Anomalous Cross-Linking by Mechlorethamine of DNA Duplexes Containing C−C Mismatch Pairs†

Rebecca M. Romero,‡ Michael Mitas,§ and Ian S. Haworth*,‡,||

Department of Biochemistry and Molecular Biology and Pharmaceutical Sciences, University of Southern California, 1985 Zonal Avenue, Los Angeles, California 90033, and Hollings Cancer Center, 866 Jonathan Lucas Street, Medical University of South Carolina, Charleston, South Carolina 29425

Received August 11, 1998; Revised Manuscript Received November 24, 1998

ABSTRACT: Nitrogen mustards such as mechlorethamine have previously been shown to covalently cross-link DNA through the N7 position of the two guanine bases of a d[GXC]·d[GYC] duplex sequence, a so-called 1,3 G−G-cross-link, when X−Y = C−G or T−A. Here, we report the formation of a new mechlorethamine cross-link with the d[GXC]·d[GYC] fragment when X−Y is a C−C mismatch pair. Mechlorethamine cross-links this fragment preferentially between the two mismatched cytosine bases, rather than between the guanine bases. The cross-link also forms when one or both of the guanine bases of the d[GCC]·d[GCC] fragment are replaced by N7-deazaguanine, and, more generally, forms with any C−C mismatch, regardless of the flanking base pairs. Piperidine cleavage of the cross-link species containing the d[GCC]·d[GCC] sequence gives DNA fragments consistent with alkylation at the mismatched cytosine bases. We also provide evidence that the cross-link reaction occurs between the N3 atoms of the two cytosine bases by showing that the formation of the C−C cross-link is pH dependent for both mechlorethamine and chlorambucil. Dimethyl sulfate (DMS) probing of the cross-linked d[GCC]·d[GCC] fragment showed that the major groove of the guanine adjacent to the C−C mismatch is still accessible to DMS. In contrast, the known minor groove binder Hoechst 33258 inhibits the cross-link formation with a C−C mismatch pair flanked by A−T base pairs. These results suggest that the C−C mismatch is cross-linked by mechlorethamine in the minor groove. Since C−C pairs may be involved in unusual secondary structures formed by the trinucleotide repeat sequence d[CCG]$_n$, and associated with triplet repeat expansion diseases, mechlorethamine may serve as a useful probe for these structures.

Mismatched base pairs in DNA can occur during replication or recombination, and have the potential to be mutagenic if not efficiently repaired. In *Escherichia coli* the repair efficiency has been reported to depend on the sequence of the mismatch (*1*), the bases flanking the mismatch (*2*), and the affinity of the MutS protein for the mismatch (*3, 4*), with pyrimidine-pyrimidine mismatches having the lowest efficiency for repair (*5, 6*). Replication fidelity is dependent on mismatch recognition, which is linked to mismatch structure, and, as a result, the structural features of mismatched base pairs have been explored using X-ray crystallography (*7*), circular dichroism (*8*), NMR spectroscopy (*9*), and ligand binding studies (*10*).

Mismatched base pairs have also been associated with secondary structures formed by DNA trinucleotide repeat sequences (*11*−*16*). The expansion and instability of such repeat sequences are associated with a number of human genetic disorders. One such disorder, Fragile-X Syndrome (*17*), is characterized by expansions of d[CGG]$_n$·d[CCG]$_n$ repeat sequences. The trinucleotide repeat expansion mechanism is unclear, but we and others have speculated that it might be due to unusual DNA conformations (*18*−*21*) that can form during replication. In the Fragile-X sequence, strand separation of the duplex can occur, and the resultant d[CGG]$_n$ and d[CCG]$_n$ single strands can fold into hairpins (*19*−*21*). For d[CCG]$_n$, the hairpin forms with a specific alignment containing repeating d[GCC]·d[GCC] 'duplex' fragments, giving a hairpin stem in which every third base pair is a C−C mismatch (*21*).

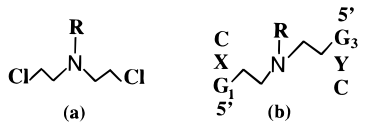

FIGURE 1: (a) Mechlorethamine (R = CH_3) or chlorambucil [R = $C_6H_4(CH_2)_3COOH$] and (b) the mechlorethamine and chlorambucil 5′-GXC···5′-GYC 1,3 G−G cross-link.

We chose to probe the structure of the d[GCC]·d[GCC] fragment with mechlorethamine, a common nitrogen mustard. Mechlorethamine [Figure 1(a), R = CH_3] and related mustards, such as chlorambucil [Figure 1(a), R = $-C_6H_4-(CH_2)_3COOH$], react with nucleophilic centers on the DNA duplex, via an aziridinium ion intermediate (*22*). This reaction occurs favorably with the guanine N7 atom (*23, 24*), but adducts with N3 of adenine have also been reported (*25*−

† This work was supported by NIH Grants CA64299-01 (I.S.H.) and DE07211-08 (R.M.R.).

* To whom correspondence should be addressed. Phone: (323) 442-3310. Fax: (323) 442-1390. E-mail: ihaworth@hsc.usc.edu.

‡ Department of Biochemistry and Molecular Biology, University of Southern California.

§ Hollings Cancer Center, Medical University of South Carolina.

|| Department of Pharmaceutical Sciences, University of Southern California.

10.1021/bi981935j CCC: $18.00
Published on Web 02/26/1999

Table 1: DNA Duplex Sequences[a]

series	sequences	
1	5′-CTCTCAGAGXCTCGTTCAG GAGAGTCTCYGAGCAAGTC-5′	X−Y = C−G C−A C−C T−G T−A T−C
2	5′-CTCTCAGAMXnTCGTTCAG GAGAGTCTmYNAGCAAGTC-5′	n−N = C−G C−G C−G C−G C−D X−Y = C−G C−G C−C C−C C−C M−m = G−C D−C G−C D−C D−C
2a	5′-CTCTCACACCGTGGTTCAG GAGAGTGTGCCACCAAGTC-5′	
3	5′-CTCTCACAMCnTGGTTCAG GAGAGTGTmCNACCAAGTC-5′	n−N = C−G T−A G−C C−G G−C T−A M−m = G−C G−C T−A C−G C−G C−G n−N = C−G T−A A−T T−A M−m = T−A T−A T−A A−T
3a	5′-CTCTCACGACTCGGTTCAG GAGAGTGCTCAGCCAAGTC-5′	
4	5′-CTCCCAATTCAATTCCCAG GAGGGTTAACTTAAGGGTC-5′	

[a] D = N7-deazaguanine.

27). Since the nitrogen mustards are bifunctional, they can form interstrand cross-links with suitable DNA sequences. Formation of guanine−guanine interstrand cross-links was originally suggested to occur between the N7 atoms of guanine bases in neighboring base pairs (that is, at a d[GC]·d[GC] site; a 1,2 G−G cross-link) (*28*, *29*), because this site provides the shortest N7 to N7 distance in a regular B-DNA duplex (*30*). However, it is now accepted that a d[GGC]·d[GCC] duplex fragment is preferentially cross-linked in the major groove of DNA between the distal guanine bases (*31*−*35*), giving a 1,3 G−G cross-link product [Figure 1(b)]. This selectivity for the d[GXC]·d[GYC] site over the d[GC]·d[GC] site is retained for X−Y = T−A (*36*).

The N7 to N7 distance in the d[GXC]·d[GYC] site is somewhat larger than the span of the bis(ethyl)amine cross-linking bridge of mechlorethamine. To accommodate the 1,3 G−G cross-link, a distortion of the DNA is necessary to reduce the interstrand guanine−guanine N7 to N7 distance from 8.9 Å in a 'B-DNA' conformation (*30*) to approximately 5.1 Å in the cross-linked species (*32*, *37*). A bend of about 15° for a mechlorethamine-cross-linked oligomer has been estimated (*31*, *32*), which could be a result of this distortion. We have previously suggested that the local distortion of the DNA and a decrease in the N7 to N7 distance can be induced by the initial, noncovalent interaction of the mustard with the DNA (*37*).

In performing the work described here, our original hypothesis was that a mismatch X−Y pair at the center of the d[GXC]·d[GYC] cross-link site would allow the DNA helix to distort more easily, and should therefore allow for more efficient cross-linking. However, for X−Y = C−C, we have observed a cross-linked species with different properties from those expected for a 1,3 G−G cross-link. We show that this species results from cross-linking of the C−C mismatch pair, and provide evidence that the cross-link forms between the cytosine N3 atoms in the minor groove of the DNA helix.

MATERIALS AND METHODS

(i) Chemicals. Mechlorethamine [*N*,*N*-bis(2-chloroethyl)-methylamine], chlorambucil (4-[*p*-[*N*,*N*-bis(2-chloroethyl)-amino]phenyl]butyric acid), Hoechst 33258, and T4 polynucleotide kinase were purchased from Sigma. [γ-^{32}P]ATP was purchased from ICN. N7-deazaguanine was purchased from Glen Research. All synthetic oligonucleotides (Table 1) were synthesized on an Applied Biosystems Model 394 automated synthesizer, deprotected, and purified with a COP cartridge at the USC Norris Cancer Center. DNA used for sequencing reactions was further purified on a 20% denaturing polyacrylamide gel. All other reagents were analytical grade.

(ii) ^{32}P-5′-End-Labeling of DNA. Approximately 10 μg of column-purified synthetic DNA was 5′-end-labeled with [γ-^{32}P]ATP (5 μL, 4500 Ci/mmol) by incubation in buffer [30 mM Tris (pH 7.8), 10 mM $MgCl_2$, 5 mM dithiothreitol] and 30 units of T4 polynucleotide kinase for 1 h at 37 °C (*38*). The reaction was stopped by the addition of 5.5 μL of 3 M sodium acetate (pH 5.2) and 150 μL of prechilled 95% ethanol. The unincorporated [γ-^{32}P]ATP was removed by precipitation in 95% ethanol at −20 °C overnight, lyophilized, and resuspended in a 0.1 M NaCl solution.

(iii) Alkylation of DNA. An equal amount of the unlabeled complementary strand was added to a 0.1 M NaCl solution of the labeled oligonucleotide, heated to 65−70 °C, and then slowly cooled to room temperature. Following annealing of the strands, a 1 μM duplex DNA solution containing 0.1 M NaCl and 10 mM Tris (pH 7.5) was incubated for 6 h at 37 °C with 100 μM mechlorethamine or chlorambucil in a total volume of 100 μL. For each experiment, a fresh solution of the nitrogen mustard (100 mM) was prepared in dimethyl sufoxide (DMSO), rapidly diluted to 10 mM, and immediately added to the DNA solution. Following incubation with the mustard, the reaction was terminated by addition of 5.5 μL of 3 M sodium acetate, 1 μL of tRNA (5 mg/mL), and 150 μL of prechilled 95% ethanol, and precipitated in 3 times the volume of prechilled 95% ethanol at −20 °C overnight, washed, and then lyophilized. The DNA was then dissolved in 2 μL of distilled water and 8 μL of tracking dye (80% formamide, 1 mM EDTA, 0.025% bromophenol blue and xylene cyanol). Controls were performed with omission of complement strand, or omission of the alkylating agent, or omission of the annealing step.

(iv) Detection of Alkylated DNA. The samples were loaded onto a 20% denaturing polyacrylamide gel, DPAGE [29:1 acrylamide/bisacrylamide, 8 M urea, 89 mM Tris−borate (pH 8.5), 2 mM EDTA (TBE buffer); 0.4 mm thick, 42 × 33

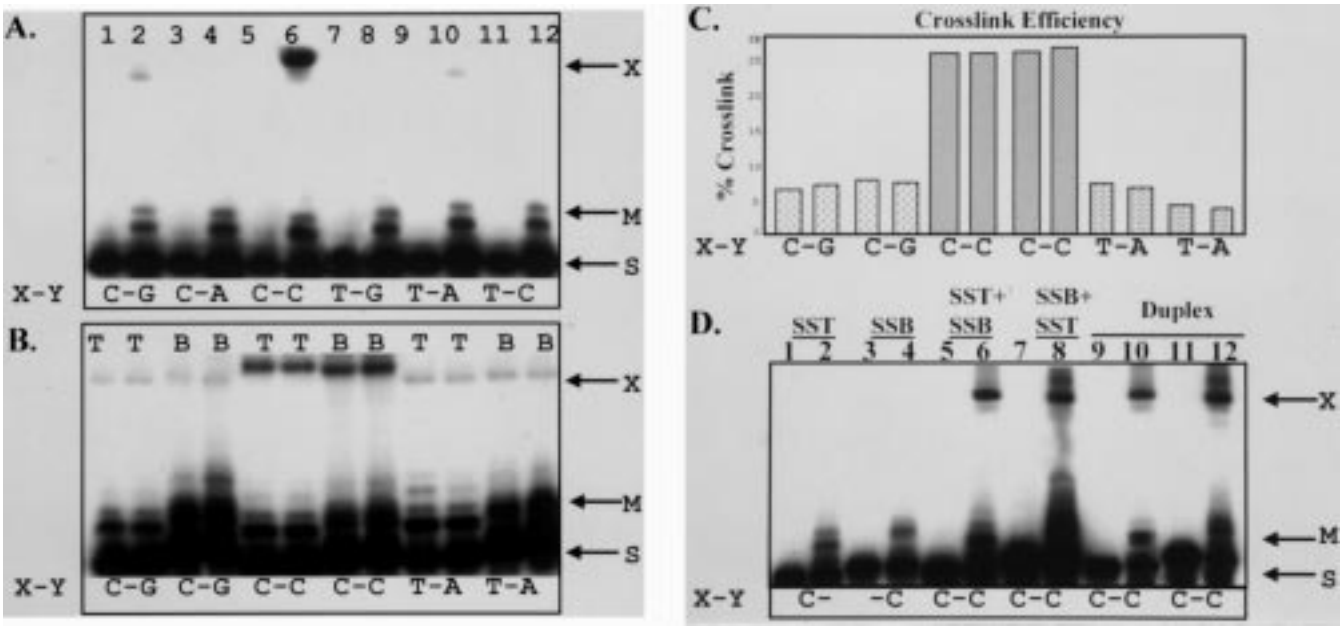

FIGURE 2: Mechlorethamine cross-linking of DNA duplexes of the sequence d[CTCTCAGAGXCTCGTTCAG]·d[CTGAACGAGYCTCTGAGAG] (Table 1, series 1). Bands are identified as X (cross-link), M (monoadduct), and S (unreacted single strands) on the right side of the figures. (A) Autoradiography of a 20% DPAGE gel of DNA duplexes where X−Y = C−G, C−A, C−C, T−G, T−A, and T−C, following incubation with 100 μM mechlorethamine, in lanes 2, 4, 6, 8, 10, and 12, respectively. Lanes 1, 3, 5, 7, 9, and 11 are controls (no mechlorethamine). In all experiments, only the top strand (X = C or T) was labeled. (B) Autoradiography of a 20% DPAGE gel for duplexes X−Y = C−G, C−C, and T−A in which either the top strand (lane T) or the bottom strand (lane B) was labeled, following incubation with 100 μM mechlorethamine. Four experiments were run for each duplex, two with the top strand labeled and two with the bottom strand labeled. The exposure time results in weak bands for the Watson−Crick duplexes, to avoid overexposing the X−Y = C−C cross-link band. (C) Quantification of the mechlorethamine-cross-linked bands in Figure 2B (bands labeled X) using densitometry. The intensity of the cross-link band is expressed as a percentage of the total DNA in each lane of the gel. (D) Autoradiography of a 20% DPAGE gel for single strands or duplexes where X = C, Y = C, following incubation with 100 μM mechlorethamine. Lanes 2 and 4 contain only top strand and only bottom strand, respectively. Lanes 6 and 8 contain both strands, but the experiment was performed with omission of the annealing step, and with the top strand labeled (lane 6) and bottom strand labeled (lane 8). Lanes 10 and 12 show the results of experiments peformed as in Figure 2B, with the top strand labeled (lane 10) and the bottom strand labeled (lane 12). Lanes 1, 3, 5, 7, 9, and 11 are controls (no mechlorethamine).

cm, 2500 V, 45 W], until the xylene cyanol marker migrated 15 cm.

(v) Determination of Cross-Linking Site. The band thought to be due to the mechlorethamine-cross-linked DNA was recovered from the gel using the crush-and-soak procedure (*38*). The DNA was then ethanol-precipitated, washed, lyophilized, and resuspended in 10% aqueous piperidine in a total volume of 100 μL. To ensure complete cleavage of all alkylated bases, samples were heated for 1 h at 90 °C. To determine the cross-linking site, control Maxam−Gilbert G and C reactions were performed in parallel *(39)*. Following this reaction, the control DNA was cleaved using 10% aqueous piperidine, in a total volume of 100 μL, for 30 min at 90 °C (*39*). All samples were lyophilized overnight, resuspended in 2 μL of distilled water and 8 μL of tracking dye (80% formamide, 1 mM EDTA, 0.025% bromophenol blue and xylene cyanol), heated at 90 °C for 2 min, chilled on an ice bath, and then loaded onto a 20% denaturing polyacrylamide gel [29:1 acrylamide/bisacrylamide, 8 M urea, 89 mM Tris−borate (pH 8.5), 2 mM EDTA (TBE buffer); 0.4 mm thick, 42 × 33 cm, 2900 V, 50 W] until the xylene cyanol marker had migrated 10 cm. Bands were assigned by reference to the Maxam−Gilbert G and C lanes.

(vi) pH-Dependent Alkylation Reaction. The DNA was alkylated with chlorambucil or mechlorethamine by the same method as in *(iii)* except that the buffer used was 10 mM potassium phosphate prepared at pH 4.0, 5.8, and 8.0. In these experiments, incubation with the mustard was for 24 h at 37 °C, to allow sufficient reaction time for the slower reacting chlormbucil.

(vii) Dimethyl Sulfate (DMS) Probing Reaction. Mechlorethamine- and chlorambucil-cross-linked DNA duplexes were recovered from the gel, purified by precipitation, and then incubated with 1% DMS, in a total volume of 50 μL, at 37 °C for 30 min. The DNA was then ethanol-precipitated, washed, lyophilized, and cleaved with 10% aqueous piperidine, in a total volume of 100 μL at 90 °C for 30 min to convert sites of alkylation into strand breaks.

(viii) Hoechst 33258 Inhibition of Cross-Linking. Experiments were performed similarly to those described in steps *(ii)*, *(iii)*, and *(iv)*, except that, prior to alkylation with mechlorethamine, the annealed DNA duplexes were preincubated at room temperature for 30 min with varying concentrations of Hoechst 33258. The incubation time for the alkylation reaction was 1 h.

RESULTS

Anomalous Cross-Linking of a DNA Duplex Containing a C−C Mismatch Pair. DNA duplexes of the sequence shown in Table 1, series 1, were incubated with mechlorethamine, and the products were electrophoresed on a 20% polyacrylamide denaturing gel. In Figure 2A, the bands with slower mobility (labeled X) are due to mechlorethamine-cross-linked duplexes (*32, 35, 40*). These bands are observed for duplexes in which X−Y = C−G and T−A, Watson−

Crick pairs (Figure 2A, lanes 2 and 10), and when X−Y = C−C, a mismatch pair (Figure 2A, lane 6). The duplexes containing mismatch pairs X−Y = C−A and T−C (Figure 2A, lanes 4 and 12) do not give observable cross-link bands. A very weak band was discernible for duplex X−Y = T−G, a mismatch pair which is relatively stable compared to other mismatch pairs. The identity of the bands with slower mobility for duplexes X−Y = C−G, C−C, and T−A was further confirmed in a separate experiment by the formation of species with similar mobility in which either the top or the bottom strand was radiolabeled (Figure 2B). The cross-linking efficiency is highest for the duplex containing the C−C mismatch in both Figure 2A and Figure 2B. Quantification by densitometry of the slower mobility bands in Figure 2B shows that the cross-link band for X−Y = C−C is about 25% of the total DNA, compared to an average of 6% for the Watson−Crick-paired duplexes (Figure 2C). In addition, the cross-link species for duplex X−Y = C−C appears to have a slightly slower mobility in the denaturing gel, compared to the cross-linked Watson−Crick duplexes. To show that the band labeled X (Figure 2A,B) for X−Y = C−C was due to a cross-link between the two complementary strands, control experiments were performed with only one strand. Incubation of either top or bottom strands with mechlorethamine (Figure 2D, lanes 2 and 4) did not give a cross-link band. Further, to show that the cross-linked species for X−Y = C−C in Figure 2A,B did not result from nonduplex conformations created by the annealing conditions, experiments were performed with elimination of the annealing step. Hence, cold complementary strands were added to the reaction mixture at room temperature. As shown in Figure 2D (lanes 6 and 8), mechlorethamine was still able to cross-link under these conditions.

Piperidine Cleavage of the Cross-Linked C−C Mismatch Duplex Gives Fragments Consistent with Alkylation of the Mismatched Bases. To determine the bases through which the mechlorethamine cross-link forms, the cross-link bands for X−Y = C−G and C−C (labeled X in Figure 2B) were subjected to piperidine cleavage. This reaction results in strand cleavage to the 5′ side of the alkylated base. The results are shown in Figure 3, and suggest that the cross-link in the two duplexes is different. For X−Y = C−G, piperidine cleavage gives bands corresponding to cleavage at G9 and G28, consistent with a 1,3 G−G cross-link. However, for X−Y = C−C, piperidine treatment produces bands with mobilities consistent with cleavage at bases X and Y. This suggests that the mechlorethamine cross-link forms at the mismatched C−C pair.

N7-Deazaguanine Substitution Does Not Influence Formation of the Cross-Link in the C−C Mismatched Duplex. To prove that the cross-link of the duplex containing a C−C mismatch pair does not form through guanine N7, mechlorethamine cross-linking of duplexes containing N7-deazaguanine (Table 1, series 2, where D = N7-deazaguanine) was examined. N7-deazaguanine substitution of one or both of the guanine bases of the d[GXC]·d[GYC] fragment prevents the formation of an N7−N7 1,3 G−G cross-link. Hence, no cross-link band is observed for the N7-deazaguanine-substituted duplex when X−Y= C−G (Figure 4, lane 4). The retention of a cross-link in N7-deazaguanine-substituted duplexes containing a C−C mismatch (Figure 4, lanes 8 and 10, bands labeled X) further supports the

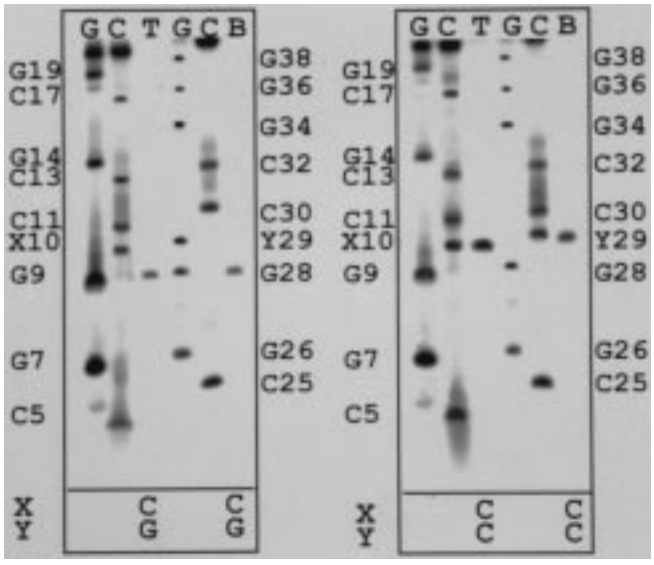

FIGURE 3: Piperidine cleavage of the cross-link bands from lanes T and B of Figure 2B for the cross-linked DNA duplexes d[$C_1T_2C_3T_4C_5A_6G_7A_8G_9X_{10}C_{11}T_{12}C_{13}G_{14}T_{15}T_{16}C_{17}A_{18}G_{19}$]· d[$C_{20}T_{21}$-$G_{22}A_{23}A_{24}C_{25}G_{26}A_{27}$ $G_{28}Y_{29}C_{30}T_{31}$ C_{32} T_{33} $G_{34}A_{35}G_{36}A_{37}G_{38}$] (X−Y = C−G or C−C, Table 1, series 1). Lanes G and C are Maxam−Gilbert G and C reactions. Lanes T and B show the results of piperidine cleavage of the top strand (that containing X) and the bottom strand (that containing Y).

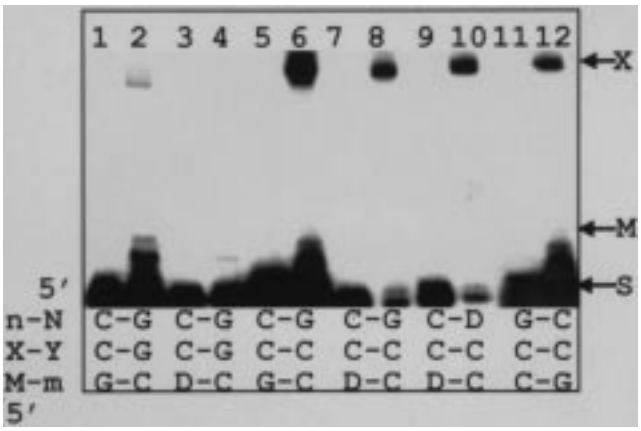

FIGURE 4: Autoradiography of a 20% DPAGE gel following incubation with mechlorethamine of duplexes of sequence d[CTCTCAGAMXnTCGTTCAG]·d[CTGAACGANYmTCTGAGAG] [where X−Y = C−G or C−C, n and m = C, and M and/or N = G or D (N7-deazaguanine) (Table 1, series 2) (lanes 1−10)] and a duplex (Table 1, series 2a) of sequence d[CTCTCACACCGTGGTTCAG]·d[CTGAACCACCGTGTGAGAG] (lanes 11 and 12). In all experiments only the top strand (that containing X) was labeled. Lanes 2, 4, 6, 8, 10, and 12 include 100 μM mechlorethamine, and lanes 1, 3, 5, 7, 9, and 11 are controls (no mechlorethamine). Bands are identified as X (cross-link), M (monoadduct), and S (unreacted single strands).

conclusion that the cross-linked species does not contain the normal 1,3 G−G cross-link. Further, a cross-link also forms for a duplex (Table 1, series 2a) which has a central d[CCG]·d[CCG] sequence and no guanine−guanine 1,3 G−G cross-link site (Figure 4, lane 12). Piperidine cleavage of each of these cross-links results in products of mobility consistent with alkylation at the C−C mismatch pair (Figure 5).

A Mechlorethamine Cross-Link Forms with Any DNA Duplex Containing a Single C−C Mismatch Pair. To determine if mechlorethamine cross-linking of a C−C mismatch is sequence-dependent, cross-linking experiments

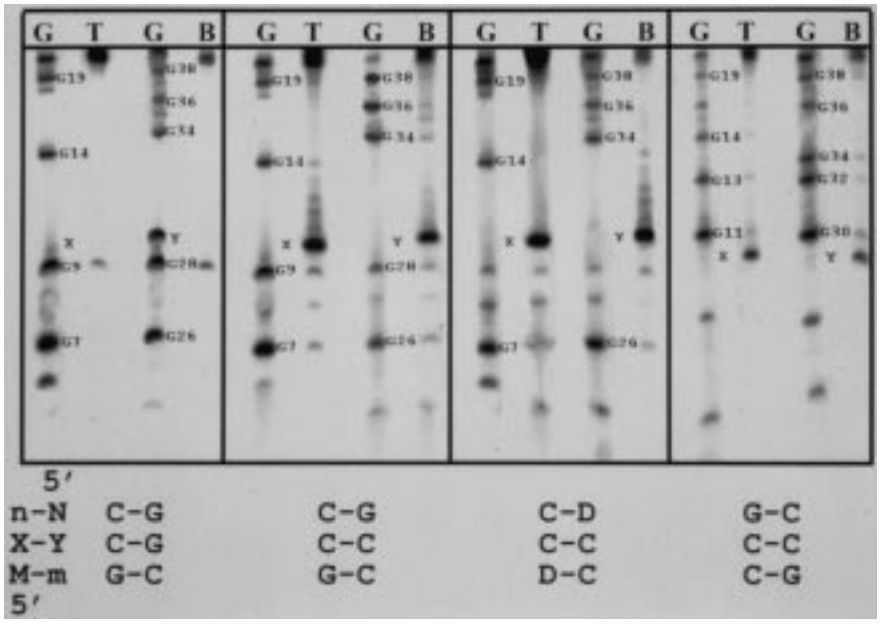

FIGURE 5: Piperidine cleavage of the cross-link bands from lanes 2, 6, 10, and 12 of Figure 4 for the mechlorethamine-cross-linked duplexes of sequence d[CTCTCAGAM$_9$X$_{10}$n$_{11}$TCGTTCAG]·d[CTGAACGAN$_{28}$Y$_{29}$m$_{30}$TCTGAGAG] [X−Y = C−G or C−C, n and m = C, and M and/or N = G or D (N7-deazaguanine) (Table 1, series 2)] and a fourth duplex (Table 1, series 2a) of sequence d[CTCTCACAC$_9$C$_{10}$G$_{11}$-TGGTTCAG]·d[CTGAACCAC$_{28}$C$_{29}$G$_{30}$TGTGAGAG]. In the figure, the duplexes are identified by their central three base pair sequence. The lanes marked with G are the Maxam−Gilbert G reactions. Lanes T and B show the results of piperidine cleavage of the top strand (that containing X) and the bottom strand (that containing Y).

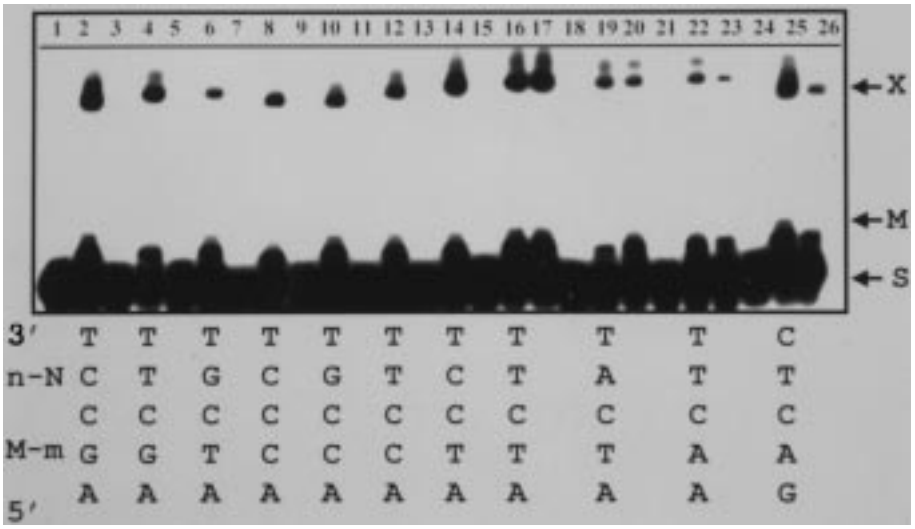

FIGURE 6: Autoradiogram of a 20% DPAGE gel following incubation with mechlorethamine of duplexes d[CTCTCACAMCnTGGTTCAG]·d[CTGAACCANCmTGTGAGAG] (Table 1, series 3) and d[CTCTCACGACTCGGTTCAG]·d[CTGAACCGACTCGTGAGAG] (Table 1, series 3a). The top strands of each duplex have the central sequences shown in the figure. In each sequence, the central C is part of a mismatch pair. In lanes 1−14, the odd-numbered lanes are controls (no mechlorethamine), and the even-numbered lanes include 100 μM mechlorethamine. Lanes 15, 18, 21, and 24 are also controls. Lanes 16, 17, 19, 20, 22, 23, 25, and 26, include 100 μM mechlorethamine. Lanes 16, 19, 22, and 25 are from experiments at 25 °C, and lanes 17, 20, 23, and 26 are from experiments at 37 °C. Bands are identified as X (cross-link), M (monoadduct), and S (unreacted single strands).

on the DNA duplexes of sequences shown in Table 1 (series 3 and 3a) were performed. The results (Figure 6) show that any C−C mismatch pair can be cross-linked by mechlorethamine, regardless of the sequence context. Piperidine cleavage of these cross-links also resulted in products that were consistent with alkylation at the C−C mismatch pair (data not shown). We note that the cross-link forms with variable efficiency, and has slightly different electrophoretic mobility, with different duplexes (Figure 6, bands labeled X). The cross-link efficiency may be related to the stability of the duplex, and the variable mobility may reflect slightly different conformations of the cross-linked species in the denaturing gel (*41*).

The Formation of the C−C Cross-Link Is pH-Dependent. The reported pK_a of cytosine N3 in a C−C mismatch pair is 6.95 (*9*). Hence, if the C−C cross-link occurs through N3 of cytosine, then its formation should be blocked by protonation of N3 at low pH. However, for mechlorethamine, this is complicated by protonation of the mustard itself [mechlorethamine has a pK_a of 6.45 *(42)*], which would be expected to reduce the reactivity of the cross-linking agent. Because of this, we also carried out the pH-dependent cross-

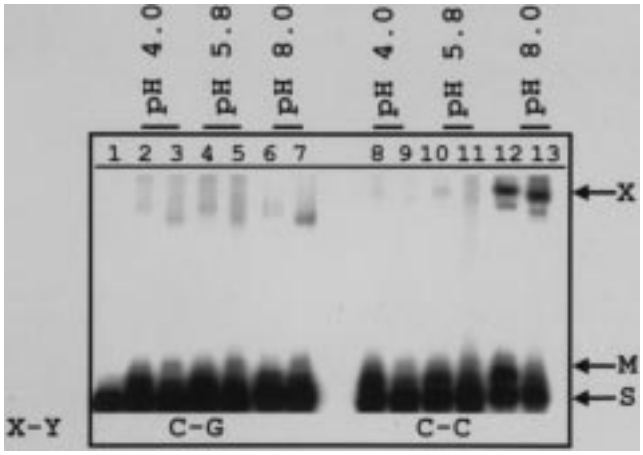

FIGURE 7: Autoradiogram of 20% DPAGE gels following incubation with mechlorethamine or chlorambucil of duplexes of sequence d[CTCTCAGAGXCTCGTTCAG]·d[CTGAACGAGYCTCTGAGAG] (Table 1, series 1), where X−Y = C−G, or C−C. Lane 1 is a control, zero nitrogen mustard. Lanes 2−7 show the products of the reaction of duplex X−Y = C−G with 100 μM chlorambucil (lanes 2, 4, and 6) or 100 μM mechlorethamine (lanes 3, 5, and 7) at the indicated pH. Lanes 8−13 show the products of the reaction of duplex X−Y = C−C with 100 μM chlorambucil (lanes 8, 10, and 12) or 100 μM mechlorethamine (lanes 9, 11, and 13) at the indicated pH. Bands are identified as X (cross-link), M (monoadduct), and S (unreacted single strands).

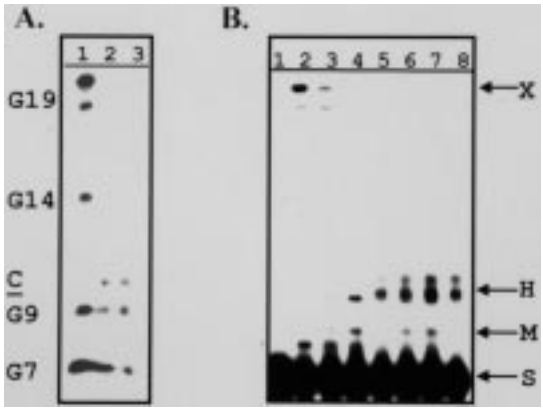

FIGURE 8: (A) Maxam−Gilbert sequencing gel of the piperidine cleavage products resulting from incubation of the chlorambucil- and mechlorethamine-cross-linked C−C mismatch duplexes with DMS. The cross-linked DNA was excised from Figure 7, lanes 12 and 13, respectively (bands labeled X). Lane 1 is the standard Maxam−Gilbert guanine reaction with sequence d[CTCTCAG$_7$-AG$_9$XCTCG$_{14}$TTCAG$_{19}$] (Table 1, series 1, X = C). Lane 2 shows the products of DMS probing of the chlorambucil-cross-linked duplex from lane 12 of Figure 7. Lane 3 shows the products of DMS probing of the mechlorethamine-cross-linked duplex from lane 13 of Figure 7. (B) Autoradiography of a 20% DPAGE gel of the duplex of sequence d[CTCCCAATTCAATTCCCAG]·d[CTGGGAATTCAATTGGGAG] (Table 1, series 4), following preincubation with Hoechst 33258, and then incubation with 100 μM mechlorethamine. Lane 1 is a control (no Hoechst 33258 and no mechlorethamine), and lanes 2−8 contain 0, 1, 2, 5, 10, 25, and 50 μM Hoechst 33258. Bands are identified as X (cross-link), H (Hoechst 33258 concentration-dependent band), M (monoadduct), and S (unreacted single strands).

linking experiments using chlorambucil (Figure 1a), which, due to the aromatic ring, has an amine with a much lower pK_a of 2.49 *(43)*, and is essentially unprotonated at the pH corresponding to the pK_a of the C−C mismatch. These experiments were performed on the duplexes shown in Table 1 (series 1) where X−Y = C−G and C−C. In Figure 7 (lanes 8, 10, and 12, bands labeled X), we show that formation of the C−C cross-link is pH-dependent for chlorambucil. The chlorambucil cross-link band is intense at pH 8, but is barely visible at pH 5.8 or pH 4.0. The mechlorethamine cross-link also forms efficiently at pH 8, but essentially no reaction occurs at lower pH. For X−Y = C−G, lanes 2−7, pH does not strongly influence the formation of the conventional 1,3 G−G cross-link with chlorambucil or mechlorethamine, providing a good control for the X−Y = C−C results. We note that protonation of cytosine N3 could influence alkylation at other atoms, either directly or through a protonation-induced DNA conformational change, and that the pH dependence does not, therefore, prove that the cross-link forms through N3, although it is suggestive of this. We also note that these experiments were conducted with longer incubation periods than those in Figures 2, 4, and 6 to allow sufficient time for the slower cross-linking reaction of chlorambucil. This results in additional cross-links forming for the C−C duplex (lanes 12 and 13) which have similar mobility to the bands observed for X−Y = C−G. This suggests that the C−C cross-link forms more rapidly than the 1,3 G−G cross-link (*41*).

The C−C Cross-Linked Species Is Reactive with DMS at Guanine Bases Adjacent to the C−C Mismatch Pair. To establish the groove in which the C−C cross-link forms, we probed the chlorambucil- and mechlorethamine-cross-linked DNA, using dimethyl sulfate (DMS), which reacts with guanine N7 atoms that are solvent-accessible. In these experiments, only cross-linked DNA species (excised from Figure 7, lanes 12 and 13, bands labeled X) were incubated with DMS. If the cross-link is formed in the major groove, one would expect the guanine bases in the d[GCC]·d[GCC] fragment to be unreactive with DMS, particularly for chlorambucil, where the additional bulk of the cross-linking agent might be expected to block the approach of DMS. In Figure 8A we show these guanines are still reactive with DMS in the chlorambucil- and mechlorethamine-cross-linked duplexes, which suggests that the cross-link is occurring in the minor groove. We note that guanines in the cross-linked species to the 3′ side of the cross-link probably react with DMS, but do not appear to do so (Figure 8A), because the conditions used for piperidine cleavage of the DMS adducts also result in cleavage of the cross-link itself (leading to the appearance of bands corresponding to fragments cleaved at a cytosine of the C−C mismatch). Hence, bands resulting from cleavage at guanine bases 3′ to the mismatch are not observable.

The Mechlorethamine C−C Cross-Link Is Inhibited by the DNA Minor Groove Binder Hoechst 33258. To define further the groove in which the C−C cross-link forms, we carried out mechlorethamine cross-linking experiments on the duplex shown in Table 1 (series 4), following preincubation with Hoechst 33258. The duplex sequence was designed so that two favorable Hoechst binding sites, d[AATT]·d[AATT] (*44*), flank the C−C mismatch pair. Hoechst 33258 has been shown to be a DNA minor groove binder in two X-ray structures *(45, 46)* of its complex with a duplex containing a d[AATTC]·d[GAATT] sequence. The location of the ligand is somewhat different in these two structures [associated either with the minor groove of d[AATT]·d[AATT] (*46*) or with d[ATTC]·d[GAAT] *(45)* fragments]. Hence, for the duplex (Table 1, series 4), Hoechst 33258 should bind in

the d[AATT]·d[AATT] sequences flanking the C−C mismatch, or perhaps may interact directly with the C−C mismatch minor groove, in a manner analogous to that seen for the d[ATTC]·d[GAAT] sequence. In either case, Hoechst 33258 binding should be effective in blocking cross-linking of the C−C mismatch pair, if this cross-linking occurs through the minor groove. In Figure 8B (bands labeled X), we show that the C−C cross-link is sensitive to Hoechst 33258, being significantly inhibited by 1 μM ligand, and eliminated at 2 μM ligand. We note that this effect could be due to indirect conformational changes induced by Hoecsht 33258 binding, rather than a direct blocking effect, and, indeed, we do observe a second Hoechst 33258-dependent band (labeled H in Figure 8B), the origin of which we have not yet determined. However, the combined DMS and Hoechst 33258 results do suggest that the mechorethamine C−C cross-link forms through the minor groove.

DISCUSSION

We have presented evidence for a new nitrogen mustard cross-link between the cytosine bases of a C−C mismatch pair. For the cross-linked d[GCC]·d[GCC] duplex fragment, piperidine cleavage data are consistent with alkylation at the mismatched cytosine base, and the retention of the cross-link in N7-deazaguanine-substituted duplexes excludes the possibility of the expected N7 to N7 1,3 G−G cross-link. We have also ruled out the possibility of cross-link formation at an alternative site (perhaps purine N3), because a cross-link still occurs with different base pairs adjacent to the C−C mismatch.

It seems likely that the cross-link is formed between the N3 atoms of the cytosine bases, since these are the most nucleophilic atoms of the mismatch pair. This is supported by the observation that C−C cross-link formation is pH-dependent, and cross-linking is much more efficient at a pH above the pK_a of cytosine N3. However, as explained in the previous section, alternative explanations of the pH dependence are possible. In particular, N3 protonation could induce a DNA conformational change that could prevent cross-link formation at another atom of cytosine. N3 and O2 are the only likely sites for alkylation, and in a fully Watson−Crick-paired duplex (in which N3 and O2 are involved in base pair formation), no adduct was observed between chlorambucil and a cytosine base (*47*). Methylation of cytosine N3 by dimethyl sulfate is favored over other nucleophilic centers (*48*), although reactions at O2 can also occur for other alkylating agents (*49*). We also note that N3-alkylcytosines are heat-labile (*50*), consistent with the heat-induced cleavage of the C−C cross-link.

Based on the reactivity with DMS of the N7 position of guanines adjacent to the cross-linked cytosines and on the inhibition of the cross-link by Hoechst 33258, a known DNA minor groove binder, we believe that the cross-linking reaction occurs through the minor groove, in contrast to the major groove guanine−guanine 1,3 G−G cross-link. The minor groove is also the most likely site because the 4-amino group of the cytosine bases effectively blocks approach from the major groove, although this might be less of a problem in a distorted duplex.

Various geometries have been reported for C−C mismatch base pairs in different sequence contexts. The antiparallel C−C mismatch is intrahelical in several duplex sequences (*9, 51*). In contrast, Gao et al. *(52)* have shown that the duplex d[CCGCCG]$_2$ has a central d[GCC]·d[GCC] sequence in which the cytosines of the C−C mismatch adopt an extrahelical location in the minor groove, and the two G−C pairs stack within the helix. In molecular dynamics simulations *(53)* of a hairpin conformation of the single-strand d[GCC]$_{11}$ trinucleotide repeat sequence, in which the hairpin stem has four d[GCC]·d[GCC] repeats and four C−C mismatches, we observed motion of the mismatched cytosines toward the minor groove. This is in contrast to simulations of duplexes containing single C−C mismatches *(53)*, in which the C−C pair remained essentially intrahelical. The motion of the mismatched cytosine bases in the repeating hairpin could be a prelude to the formation of an e-motif-like structure, or some other nonstandard hairpin conformation (*21*), consistent with the anomalous electrophoretic and pH-dependent behavior we and others have previously observed for d[CCG]$_n$ hairpins (*20, 21*). Probing of such structures with nitrogen mustards may reveal more details of the conformations of single-stranded d[CCG]$_n$, given the concentration of potential 1,3 G−G and C−C mismatch cross-link sites in these structures.

REFERENCES

1. Kramer, B., Kramer, W., and Fritz, H. J. (1984) *Cell 38*, 879−887.
2. Jones, M., Wagner, R., and Radman M. (1987) *Genetics 115*, 605−610.
3. Su, S. S., Lahue, R. S., Au, K. O., and Modrich, P. (1988) *J. Biol. Chem. 263*, 6829−6835.
4. Wagner, R., Debbie, P., and Radman, M. (1995) *Nucleic Acids Res. 23*, 3944−3948.
5. Modrich, P. (1991) *Annu. Rev. Genet. 25*, 229−253.
6. Thomas, D. C., Roberts, J. D., and Kunkel, T. A. (1991) *J. Biol. Chem. 266*, 3744−3751.
7. Holbrook, S. R., Cheong, C., Tinoco, I., and Kim, S. H. (1991) *Nature 353*, 579−581.
8. Gray, D. M., Cui, T., and Ratliff, R. L. (1984) *Nucleic Acids Res. 12*, 7565−7580.
9. Boulard, Y., Cognet, J. A. H., and Fazakerley, G. V. (1997) *J. Mol. Biol. 268*, 331−347.
10. Chen, F.-M. (1998) *Biochemistry 37*, 3955−3964.
11. Mitas, M. (1997) *Nucleic Acids Res. 25*, 2245−2253.
12. Chen, X. A., Mariappan, S. V. S., Catasti, P., Ratliff, R., Moyzis, R. K., Laayoun, A., Smith, S. S., Bradbury, E. M., and Gupta, G. (1995) *Proc. Natl. Acad. Sci. U.S.A. 92*, 5199−5203.
13. Gacy, A. M., Goellner, G., Juranic, N., Macura, S., and McMurray, C. T. (1995) *Cell 81*, 533−540.
14. Mariappan, S. V. S., Catasti, P., Chen, X., Ratliff, R., Moyzis, R. K., Bradbury, E. M., and Gupta, G. (1996) *Nucleic Acids Res. 24*, 784−792.
15. Mariappan, S. V. S., Garcia, A. E., and Gupta, G. (1996) *Nucleic Acids Res. 24*, 775−783.
16. Petruska J., Arnheim, N., and Goodman, M. F. (1996) *Nucleic Acids Res. 24*, 1992−1998.
17. Fu, Y.-H., Kuhl, D. P. A., Pizzuti, A., Pieretti, M., Sutcliffe, J. S., Richards, S., Verkerk, A. J. M. H., Holden, J. J. A., Fenwick, J. R. G., Warren, S. T., Oostra, B. A., Nelson, D. L., and Caskey, C. T. (1991) *Cell 67*, 1047−1058.
18. Kuryavyi, V. V., and Jovin, T. M. (1995) *Nat. Genet. 9*, 339−341.
19. Mitas, M., Yu, A., Dill, J., and Haworth, I. S. (1995) *Biochemistry 34*, 12803−12811.
20. Gacy, A. M., and McMurray, C. T. (1998) *Biochemistry 37*, 9426−9434.

21. Yu, A., Romero, R. M., Barron, M. D., Dill, J., Christy, M., Gold, B., Gray, D. M., Haworth, I. S., and Mitas., M. (1997) *Biochemistry 36*, 3687−3699.
22. Rutman, R. J., Chun, E. H. L., and Jones, J. (1969) *Biochim. Biophys. Acta 174*, 663−673.
23. Mattes, W. B., Hartley, J. A., and Kohn, K. W. (1986) *Nucleic Acids Res. 14*, 2971−2987.
24. Kohn, K. W., Hartley, J. A., and Mattes, W. B. (1987) *Nucleic Acids Res. 15*, 10531−10549.
25. Pieper, R. O., and Erickson, L. C. (1990) *Carcinogenesis 11*, 1739−1746.
26. Wang, P., Bauer, G. B., Kellogg, G. E., Abraham, D. J., and Povirk, L. F. (1994) *Mutagenesis 9*, 133−139.
27. Wang, P., Bauer, G. B., Bennett, R. A. O., and Povirk, L. F. (1991) *Biochemistry 30*, 11515−11521.
28. Brookes, P., and Lawley, P. D. (1961) *J. Chem. Soc.*, 3923−3928.
29. Brookes, P., and Lawley, P. D. (1961) *Biochem. J. 80*, 496−503.
30. Arnott, S., Campbell-Smith, P., and Chandreskharan, P. (1976) in *CRC Handbook of Biochemistry*, Vol. 2, pp 411−422, CRC Press, Inc., Boca Raton, FL.
31. Rink, S. M., and Hopkins, P. B. (1995) *Biochemistry 34*, 1439−1445.
32. Rink, S. M., Solomon, M. S., Taylor, M. J., Raja, S. B., Mclaughlin, L. W., and Hopkins, P. B. (1993) *J. Am. Chem. Soc. 115*, 2551−2557.
33. Hopkins, P. B., Millard, J. T., Woo, J., Weidner, M. F., Kirchner, J. J., Sigurdsson, S. T., and Raucher, S. (1991) *Tetrahedron 47*, 2475−2489.
34. Millard, J. T., Weidner, M. F., Kirchner, J. J., Ribeiro, S., and Hopkins, P. B. (1991) *Nucleic Acids Res. 19*, 1885−1891.
35. Millard, J. T., Raucher, S., and Hopkins, P. B. (1990) *J. Am. Chem. Soc. 112*, 2459−2460.
36. Ojwang, J. O., Grueneberg, D. A., and Loechler, E. L. (1989) *Cancer Res. 49*, 6529−6537.
37. Remias, M. G., Lee, C.-S., and Haworth, I. S. (1995) *J. Biomol. Struct. Dyn. 12*, 911−936.
38. Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989) in *Molecular Cloning*: *A Laboratory Manual*, *2nd ed*., Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
39. Maxam, A. M., and Gilbert, W. (1980) *Methods Enzymol. 65*, 499−560.
40. Hartley, J. A., Berardini, M. D., and Souhami, R. L. (1991) *Anal. Biochem. 193*, 131−134.
41. Romero, R. M., Rodesittisuk, P., and Haworth, I. S., unpublished experiments.
42. Cohen, B., Van Artsdalen, E. R., and Harris, J. (1948) *J. Am. Chem. Soc. 70*, 281−285.
43. Stewart, P. J., and Owen, W. R. (1980) *Aust. J. Pharm. Sci. 9*, 15−18.
44. Abu-Daya, A., Brown, P. M., and Fox, K. R. (1995) *Nucleic Acids Res. 23*, 3385−3392.
45. Pjura, P. E., Grzeskowiak, K., and Dickerson, R. E. (1987) *J. Mol. Biol. 197*, 257−271.
46. Teng, M.-K., Usman, N., Frederick, C. A., and Wang, A. H.-J. (1988) *Nucleic Acids Res. 16*, 2671−2690.
47. Bank, B. B. (1992) *Biochem. Pharmacol. 44*, 571−575.
48. Brookes, P., and Lawley, P. D. (1962) *J. Chem. Soc.*, 1348−1351.
49. Ford, G. P., and Wang, B. (1993) *Carcinogenesis 14*, 1465−1467.
50. Liang, G., Gannett, P., Shi, X., Zhang, Y., Chen, F.-X., and Gold, B. (1994) *J. Am. Chem. Soc. 116*, 1131−1132
51. Brown, T., Leonard, G. A., Booth, E. D., and Kneale, G. (1990) *J. Mol. Biol. 212*, 437−440.
52. Gao, X., Huang, X., Smith, K. G., Zheng, M., and Liu, H. (1995) *J. Am. Chem. Soc. 117*, 8883−8884.
53. Romero, R. M., Cheng, H. Y., Mitas, M., and Haworth, I. S. (1998) *Structure, Motion, Interaction and Expression of Biological Macromolecules*, Volume 2. Proceedings of the 10th Conversation in Biomolecular Stereodynamics (Sarma, R. H., and Sarma, M. H., Eds.) Adenine Press, pp 215−220, Albany, NY.

BI981935J