Nucleic Acids Research, 1997, Vol. 25, No. 6 **1211–1218**

Specificity and kinetics of interstrand and intrastrand bifunctional alkylation by nitrogen mustards at a G-G-C sequence

Gwen B. Bauer+ and Lawrence F. Povirk*

Department of Pharmacology and Toxicology, Medical College of Virginia, Virginia Commonwealth University, Richmond, VA 23298, USA

Received November 20, 1996; Revised and Accepted January 28, 1997

ABSTRACT

Previous work showed that melphalan-induced mutations in the *aprt* gene of CHO cells are primarily transversions and occur preferentially at G-G-C sequences, which are potential sites for various bifunctional alkylations involving guanine N-7. To identify the DNA lesion(s) which may be responsible for these mutations, an end-labeled DNA duplex containing a frequent site of melphalan-induced mutation in the *aprt* gene was treated with melphalan, mechlorethamine or phosphoramide mustard. The sequence specificity and kinetics of formation of both interstrand and intrastrand crosslinks were determined. All mustards selectively formed two base-staggered interstrand crosslinks between the 5′ G and the G opposite C in the 5′ G-G-C sequence. Secondary alkylation was much slower for melphalan than for the other mustards and the resulting crosslink was more stable. Mechlorethamine and phosphoramide mustard induced intrastrand crosslinks between the two contiguous Gs in the G-G-C sequence in double-stranded DNA, but melphalan did not. Molecular dynamic simulations provided a structural explanation for this difference, in that the monofunctionally bound intermediates of mechlorethamine and phosphoramide mustard assumed thermodynamically stable conformations with the second arm in a position appropriate for intrastrand crosslink formation, while the corresponding melphalan monoadduct did not.

INTRODUCTION

The nitrogen mustards are among the oldest cancer chemotherapeutic agents and remain the drugs of choice for treatment of many human cancers, including Hodgkin's disease, multiple myeloma, Burkitt's lymphoma and, in combination chemotherapies, various carcinomas (1). Clinically useful mustard analogs include melphalan (L-phenylalanine mustard), the prototypical mustard mechlorethamine (HN2) and cyclophosphamide, which is converted to its active form, phosphoramide mustard, via the cytochrome P450 pathway (Fig. 1). In aqueous solution, each of the chloroethyl side chains of these compounds can spontaneously cyclize to form an aziridinium ion capable of adding to a nucleophilic site in DNA. The resulting monoadduct can form a second aziridinium ion which can simply react with solvent or can add to another nearby nucleophilic site, resulting in a crosslink either between DNA and protein or between two DNA bases (reviewed in 2). While DNA interstrand crosslinking has been demonstrated for several nitrogen mustards, the abundance of bifunctional adducts in mustard-treated DNA (3) suggests that intrastrand crosslinks may also be formed, however, the existence of such lesions in defined sequence DNA has not been demonstrated.

A serious complication of treatment with nitrogen mustards is the increased risk of a secondary leukemia in long-term survivors (4). The occurrence of characteristic karyotypic changes, particularly deletion of a 5q or 7q arm, in therapy-related acute myelogenous leukemias (5,6) suggests that specific genetic loci are consistently being altered. The exact mechanisms of these putative drug-induced genetic alterations remain largely unknown. If the therapeutic and mutagenic effects of these drugs result from different lesions in the genome, the two effects may be separable.

The predominant mutations induced by the aromatic nitrogen mustards melphalan and chlorambucil in the shuttle vector pZ189 replicated in human 293 cells were A·T→T·A transversions, apparently resulting from adenine N-3 alkylations (7,8). In contrast, however, most melphalan-induced base substitutions in the *aprt* gene of CHO cells occurred at G·C base pairs; G·C→T·A and G·C→C·G transversions predominating (9). Many of the base substitutions were at G-N-C (particularly G-G-C) sequences, which are potential interstrand crosslink sites, at least for the prototypical mustard mechlorethamine (10). Surprisingly, however, many of these substitutions occurred at the internal base pair (3′ G in the G-G-C), which is not expected to participate in the crosslink.

To identify possible guanine lesions which mediate G·C→T·A and G·C→C·G transversions, adducts formed by several nitrogen mustards were examined in a 20 bp DNA duplex containing a G-G-C site and resembling a frequently mutated sequence in the *aprt* gene. Separate 5′- and 3′-end-labeling experiments reveal that both interstrand and intrastrand crosslinks are formed at the G-G-C site, with significant differences between the various mustards in their propensity to form particular bifunctional adducts. Moreover, molecular dynamic simulations suggest a possible structural basis for some of these differences in specificity.

*To whom correspondence should be addressed. Tel: +1 804 828 9640; Fax: +1 804 828 8079; Email: lpovirk@gems.vcu.edu

+Present address: Department of Medicinal Chemistry, Medical College of Virginia, Virginia Commonwealth University, Richmond, VA 23298, USA

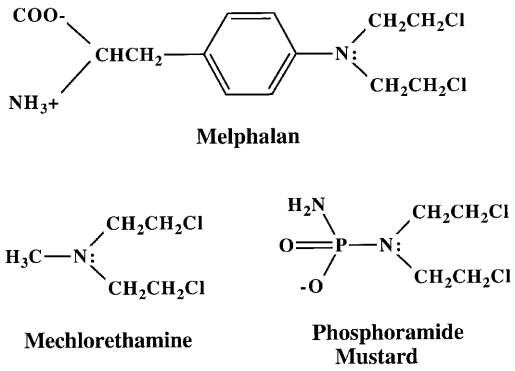

Figure 1. Stuctures of melphalan, mechlorethamine and phosphoramide mustard. Phosphoramide mustard is the active metabolite of cyclophosphamide.

MATERIALS AND METHODS

Materials

Phosphoramide mustard was obtained from Dr V.N.Narayanan (Developmental Therapeutics Program, National Cancer Institute). Since phosphoramide mustard is highly reactive in aqueous solution (11), it was freshly dissolved in reaction buffer (10 mM Tris–HCl, 0.1 mM EDTA, pH 8) at a concentration of 100 mM immediately before use. Melphalan and mechlorethamine (Sigma, St Louis, MO) were dissolved in 0.1 N HCl at a concentration of 10 mM, stored at –70°C and diluted in reaction buffer before use.

The oligonucleotides 5′-$_{OH}$CCACACTCGCCTGCGATCG-G$_{OH}$-3′ (oligonucleotide 1), 5′-$_{OH}$CCGATCGCAGGCGAGTGT-GG$_{OH}$-3′ (oligonucleotide 2) and 5′-$_{OH}$TCCACACTCGCCT-GCGATCGG$_{OH}$-3′ (oligonucleotide 3) were synthesized by the phosphoramidite method and purified by polyacrylamide gel electrophoresis. Oligonucleotides 1 and 2 were 5′-end-labeled, using [γ-^{32}P]ATP (3000 Ci/mmol; New England Nuclear) and polynucleotide kinase (New England Biolabs) according to standard procedures (12). Each labeled oligonucleotide (1 µg) was either repurified on a denaturing gel or annealed to a 2-fold excess of its complement at 10°C for 3 h in 75 µl 0.5 M NaCl, 25 mM Tris–HCl, pH 8.0, and purified on a non-denaturing gel run at 4°C.

To obtain a 3′-end-labeled duplex, oligonucleotide 2 (0.64 µg) was incubated for 3 h at 10°C in the presence of 0.96 µg complementary oligonucleotide 3, 200 µCi [α-^{32}P]dATP (3000 Ci/mmol; New England Nuclear), 0.2 mM each dGTP, dCTP and dTTP and 10 U Klenow fragment (New England Biolabs) in the buffer provided plus 0.1 M NaCl and the labeled duplex was purified on a non-denaturing gel.

To obtain a single-stranded 3′-end-labeled substrate, oligonucleotide 2 was treated with terminal transferase (US Biochemical Co.) and 100 µCi [α-^{32}P]ddATP (3000 Ci/mmol; Amersham) according to the manufacturer's instructions, ethanol precipitated and purified on a denaturing gel.

Drug treatment and quantitative analysis of adducts

Reaction mixtures (10 µl) contained 80 µg/ml calf thymus DNA, single- or double-stranded ^{32}P-end-labeled DNA oligonucleotide (<10 µg/ml), reaction buffer (10 mM Tris–HCl, pH 8, 0.1 mM EDTA) and the indicated concentrations of nitrogen mustards. After an initial 30 min incubation at 37°C, the DNA was ethanol precipitated in the presence of 100 µg/ml carrier tRNA, in order to remove free drug. The DNA was redissolved in 50 µl reaction buffer and split into five 10 µl aliquots, which were incubated for 0, 3, 6, 12 or 24 h at 37°C and then frozen at –70°C. At 24 h, the samples were thawed, ethanol precipitated, dried, dissolved in 0.1 N sodium hydroxide and incubated for 1 h at 37°C in order to convert N-7 adducts to the more stable formamidopyrimidine (FAPY) derivatives (13). After a final precipitation to remove NaOH, each sample was dissolved in 20 µl reaction buffer and split into two aliquots to be either frozen at –70°C or treated with 1 M piperidine at 90°C for 30 min in a volume of 100 µl in order to cleave both N-7 adducts and their FAPY derivatives (14). Aliquots of 10 µl loading solution (80% formamide plus 20 mM EDTA) were added to untreated and lyophilized piperidine-treated samples and DNA was denatured for 1 min at 90°C immediately before loading. The samples were electrophoresed in a 20% denaturing polyacrylamide/bisacrylamide (20:1) gel. The wet gels were autoradiographed at –70°C for 10 h.

β-Emissions from bands in the gel corresponding to radioactive crosslinked DNA or broken DNA fragments were quantitated using a Betascope 603 Blot Analyzer (Betagen Inc.). Radioactivity in each band was normalized to the total radioactivity in the lane to give the percent DNA in each band. Each assay was reproduced a minimum of three times; the mean and standard error of each time point was calculated and plotted. An ANOVA (15) was performed to assess the statistical significance of time-dependent changes in alkylation at each site.

Determining the sequence positions of interstrand crosslinks

Following Betascope analysis, the slower migrating crosslinked DNA was excised from the gel. Crosslinked DNA bands for all the time points for each concentration were pooled and eluted for 16 h. After removal of gel fragments by centrifugation, the DNA was ethanol precipitated and heated in 100 µl 1 M piperidine for 1 h at 90°C to cleave the stabilized FAPY interstrand crosslinks. Following lyophilization, the samples were analyzed on denaturing gels as above.

Quantitative determination of intrastrand crosslinks

The extent of intrastrand crosslinking between the two adjacent Gs in the G-G-C sequence was calculated by comparison of cleavage patterns of 5′- and 3′-end-labeled DNA as follows.

Let *A* be the the apparent ratio of 5′-G/3′-G cleavage when oligonucleotide 2 is 5′-end-labeled, *B* be the the apparent ratio of 3′-G/5′-G cleavage when oligonucleotide 2 is 3′-end-labeled, *x* be the fraction of adducts at the G-G site which are either monoadducts at the 5′ G or interstrand crosslinks involving the 5′ G, *y* be the fraction of adducts at the G-G site which are monoadducts at the 3′ G and *z* be the fraction of adducts at the G-G site which are intrastrand crosslinks between the 5′ and 3′ Gs. The total of all adduct fractions, 1, is equal to the sum of the fractions of monofunctional adducts, interstrand crosslinks and intrastrand crosslinks at the G-G site.

$$1 = x + y + z$$

When an intrastrand crosslink occurs between the 5′ and 3′ Gs and the backbone is cleaved at both sites by piperidine, the lesion will

appear on a sequencing gel as a single break at the G nearest the labeled end. Thus, the apparent ratio of 5′-G/3′-G cleavage when oligonucleotide 2 is 5′-end-labeled is:

$$A = (x + z)/y = (1 - y)/y$$

or

$$y = 1/(A + 1)$$

Similarly, the apparent ratio of 3′-G/5′-G cleavage when oligonucleotide 2 is 3′-end-labeled is:

$$B = (y + z)/x = (1 - x)/x$$

or

$$x = 1/(B + 1)$$

Thus,

$$z = 1 - (y + x) = 1 - [1/(A + 1) + 1/(B + 1)]$$

Modeling

Molecular modeling studies were performed on a Silicon Graphics Crimson VGXT workstation. The 9mer GCAGGCGAG with its complement was constructed as B-DNA in Sybyl 6.1 (Tripos Associates, St Louis, MO), with charges calculated by the method of Gasteiger-Hückel. The DNA was solvated with one layer of water and relaxed by steepest descent for 300 iterations and then by conjugate gradient to an energy change of 0.05 kcal/mol between iterations. The aziridinium forms of melphalan (zwitterion), mechlorethamine and phosphoramide mustard (anion) were each minimized and then attached to the first G in the G-G-C sequence and rotated around the N7-CH_2 DNA–drug bond in 13° increments using the gridsearch option. Each rotamer was minimized with DNA constrained and an initial rotamer was chosen based on a favorable energy minimum. In general, this resulted in the adduct being centered in the major groove approximately equidistant from the interstrand or intrastrand crosslink sites. The water was then removed and each adduct was subjected to a 100 ps molecular dynamics simulation using the TRIPOS force field, with a step length of 1 fs and a temperature of 300 K, with DNA constrained. The dielectric function was set to 'Distance', which simulates solvent screening of electrostatic charges by substituting a distance-dependent function for the dielectric constant. The simulations were saved every 100 fs, giving a total of 1000 configurations.

RESULTS

Sequence specificity of interstrand crosslinks in double-stranded DNA

To determine the positions of interstrand crosslinks, the double-stranded DNA duplex was 5′- or 3′-end-labeled in the G-G-C-containing strand or was 5′-end-labeled in the complementary strand. Each labeled duplex was exposed to melphalan, mechlorethamine or phosphoramide mustard. Following drug removal by precipitation, the samples were incubated for 24 h to allow formation of interstrand crosslinks. It is generally accepted that guanine N-7 adducts are responsible for interstrand crosslinks, however, these alkylated adducts easily depurinate (16). Therefore, following the 24 h incubation, samples were treated with alkali to convert N-7 adducts, including interstrand crosslinks, to the more stable, imidazole ring-opened FAPY derivatives (13). When either strand was end-labeled, each drug produced interstrand

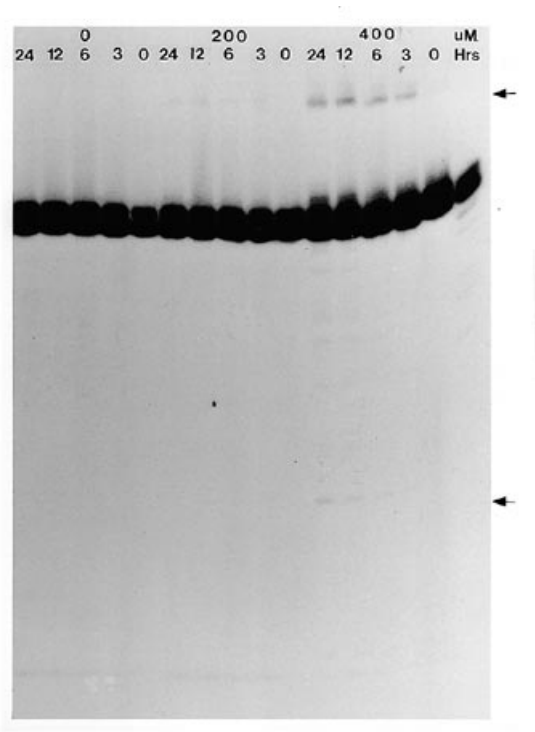

Figure 2. Slow formation of interstrand crosslinks in melphalan-treated DNA. 5′-End-labeled oligonucleotide 1 annealed to oligonucleotide 2 was treated with 0, 200 or 400 μM melphalan, then incubated at 37°C for 0–24 h and treated with 0.1 N NaOH. Analysis on a denaturing gel revealed the slow formation of alkali-stabilized crosslinked DNA (top arrow), which had slower mobility than full-length single-stranded DNA. The bottom arrow shows a band corresponding to cleavage at the crosslink site, presumably due to spontaneous depurination followed by alkali-induced cleavage.

crosslinks in a dose-dependent manner, as evidenced by the presence on the denaturing gel of a slower migrating band (shown for melphalan in Fig. 2). This band was not seen with samples of either untreated DNA or drug-treated single-stranded DNA. Although the samples were heat denatured for 1 min at 90°C and electrophoresed on a denaturing gel, the crosslink was apparently stable under these conditions. The location of the interstrand crosslink was determined by isolating the slower migrating fragments and heating them for 1 h at 90°C in piperidine, conditions known to hydrolyze most guanine N-7 adducts and their FAPY derivatives (14). The products were resolved on a denaturing gel (Fig. 3). For all three drugs, the major alkylated base in the crosslinked DNA was the 5′ G of the G-G-C sequence. Low levels of cleavage were also detected at other guanines in the fragment, but it is not known whether these represented interstrand crosslinks or independent monofunctional alkylations in the same molecule. Crosslink formation at the 5′ G apparently precluded alkylation at the adjacent 3′ G, as cleavage at this position was conspicuously absent. Surprisingly, a fraction of the melphalan-induced crosslinks apparently remained intact following piperidine treatment, as ~17% of the radioactivity still migrated in a position characteristic of crosslinked DNA. This result suggests that one or more isomers (13) of the ring-opened melphalan–FAPY adducts is unusually stable, although the possibility of some very stable non-FAPY adduct (e.g. an O^6 adduct) cannot be ruled out.

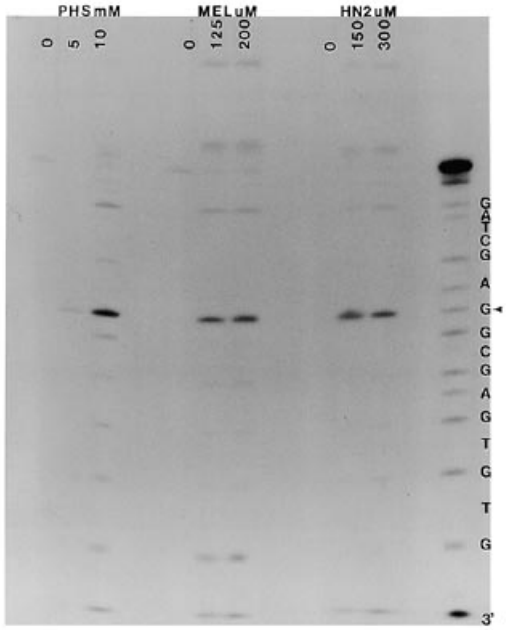

Figure 3. Determination of the site of interstrand crosslinking by nitrogen mustards in the 3′-end-labeled G-G-C-containing oligonucleotide. The more slowly migrating crosslinked species formed by phosphoramide mustard (PHS), melphalan (MEL) and mechlorethamine (HN2) was isolated from a denaturing gel and treated with hot piperidine to induce cleavage at the alkylation site. For all three drugs, the major alkylated site in crosslinked DNA was at the 5′ G in the G-G-C sequence (arrow).

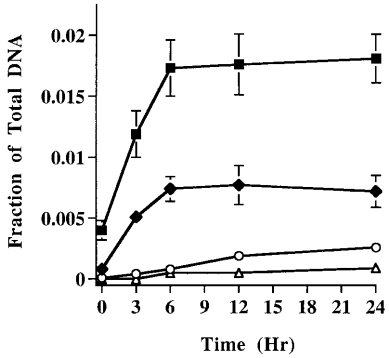

Figure 4. Time course of interstrand crosslink formation by melphalan. Interstrand crosslinking of 5′-end-labeled oligonucleotide 1 annealed to oligonucleotide 2 was determined at various times following an initial treatment with 200 (♦) or 400 µM (■) melphalan, as described in Figure 3. Spontaneous depurination at the crosslink site (5′ G in the G-C-C sequence) was also determined at various times following treatment with 200 (Δ) or 400 µM (○) melphalan. Alkylation sites which were cleaved rather than stabilized by treatment with 0.1 N NaOH were assumed to have undergone depurination.

Kinetics of interstrand crosslinking

When the 20 bp duplex was treated with melphalan, very little crosslinking occurred during the initial 30 min drug treatment. Rather, following drug removal, crosslinks formed very slowly, reaching a maximum between 6 and 12 h ($t_{1/2}$ = 3 h) (Fig. 4). When total alkylation at individual guanines in each strand (as judged by piperidine-mediated cleavage) was examined as a function of time (Fig. 5), it was apparent that alkylation of G in G-C-C in oligonucleotide 1 increased substantially over time, while alkylation of G in G-G-C in oligonucleotide 2 did not. Taken together, these data suggest that the 5′ G in the G-G-C sequence usually becomes alkylated first and, over a period of several hours, the second arm of the bifunctional alkylator spans the major groove and attaches to the 5′ G in the complementary G-C-C sequence. Preferential alkylation of guanines flanked by adjacent guanines has also been reported by others (17).

In contrast, mechlorethamine-induced crosslinks reached their maximum level during the initial 30 min incubation (Fig. 6A) and then decayed substantially over time, presumably due to spontaneous depurination. This decay did not conform to a simple exponential, suggesting two populations of adducts, perhaps a mixture of guanine N-7 and the more stable FAPY interstrand crosslinks. All mechlorethamine-induced guanine adducts were much more labile than melphalan-induced adducts, as indicated by the much higher levels of strand breaks in mechlorethamine-treated samples incubated for long time periods and then exposed to alkali; presumably, the adducts at these sites had spontaneously depurinated and could no longer be alkali stabilized. The lack of increase in piperidine-mediated cleavage at longer incubation times at sites associated with interstrand crosslinks confirmed that all mechlorethamine-induced crosslinks were formed within the first 30 min.

Phosphoramide mustard, which on a molar basis was 50–100 times less effective in forming interstrand crosslinks than mechlorethamine or melphalan, exhibited unique kinetics of interstrand crosslinking (Fig. 6B). Although the majority of crosslinks were formed within the first 30 min of drug treatment, crosslink formation continued for at least 3 h following drug removal. After 3 h, however, the level of crosslinked DNA began to slowly decline. These data, combined with the appearance of alkali-labile sites at the position of the crosslink and the known lability of phosphoramide mustard adducts (11), suggest that the crosslinked bases slowly depurinated. However, the rate of crosslink decline was slower than that seen for mechlorethamine, a difference which was especially noticeable within the first 6 h incubation. In contrast, phosphoramide mustard produced monofunctional adducts (i.e. not in the G-G-C sequence) which depurinated more rapidly than mechlorethamine-induced monofunctional adducts (data not shown). Spontaneous conversion of phosphoramide mustard adducts to FAPY derivatives has been reported (19). Thus, one possible explanation of the data is that this conversion to the more stable FAPY form occurred more readily for the interstrand crosslinks than for monoadducts. It is known that the rate of FAPY conversion can vary dramatically depending on the structure of the substituent adducted to N-7 (20).

Intrastrand crosslinks in double- and single-stranded DNA

Melphalan-induced mutations in the *aprt* gene of CHO cells were found preferentially at potential interstrand crosslink sites (9), but many of these mutations occurred at the internal G in G-G-C sequences, a position which clearly does not directly participate in interstrand crosslinking. One possible explanation is that the mutagenic lesion may be an intrastrand crosslink between the adjacent Gs, as has been reported with cisplatin, another bifunctional chemotherapeutic agent (21). To test this possibility, melphalan-adducted DNA was treated with hot piperidine to convert N-7 alkylations to strand breaks and cleavage patterns

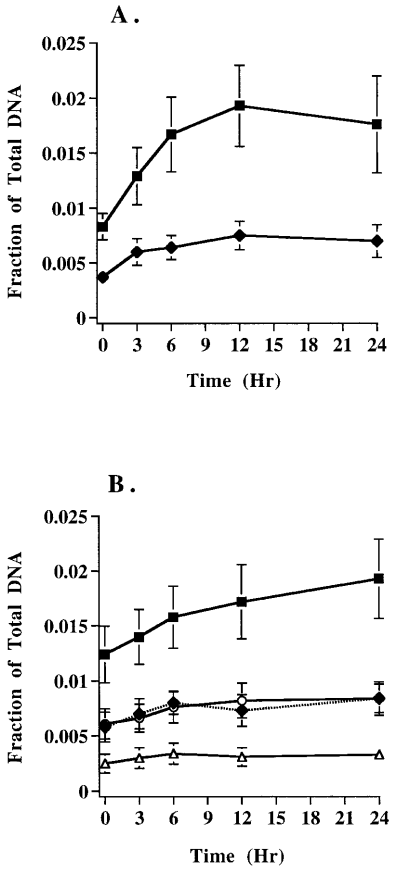

Figure 5. Time course of alkylation at interstrand crosslink sites in each strand following melphalan removal. Alkylation was determined as the fraction of piperidine-cleavable sites at each nucleotide position in the total DNA sample. (**A**) Alkylation at G opposite the C in the G-G-C sequence following treatment with 200 (♦) or 400 μM (■) melphalan. The time-dependent increase, which was statistically significant by ANOVA for both doses, suggests slow crosslinking at this site by drug already attached to the opposite strand (at 5′ G of G-G-C). (**B**) Alkylation at the 5′ G in the G-G-C sequence following treatment with 125 (♦) or 200 μM (■) melphalan or at the 3′ G following treatment at 125 (Δ) or 200 μM (○) melphalan. A relatively constant alkylation level at the 5′ position (no significant change by ANOVA at either dose) confirms that this site was the target of primary rather than secondary alkylation during the crosslinking reaction.

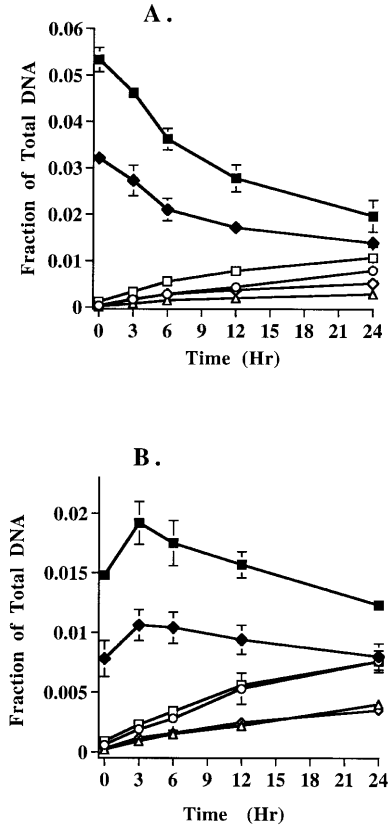

Figure 6. Time course of interstrand crosslink formation by mechlorethamine (**A**) and phosphoramide mustard (**B**) in the G-G-C sequence. (A) Crosslinking following treatment with 150 (♦) or 300 μM (■) mechlorethamine was assessed as in Figure 3. Depurination was determined (as in Fig. 4) at the 5′ G following treatment at 150 (Δ) or 300 μM (○) or at the 3′ G following treatment at 150 (♦) or 300 μM (□). (B) Crosslinking was assessed following treatment with 5 (♦) or 10 mM (■) phosphoramide mustard. Depurination was assessed at the 5′ G following treatment at 5 (Δ) or 10 mM (○) or at the 3′ G following treatment at 5 (◇) or 10 mM (□).

were compared for 5′- versus 3′-end-labeled fragments. Since intrastrand crosslinks are detected as strand breaks only at the base nearest the labeled end, these data can be used to calculate the extent of intrastrand crosslinking, expressed as a fraction of the total alkylation at the G-G site. However, in melphalan-treated double-stranded DNA there was no evidence of intrastrand crosslinking, as the cleavage pattern was essentially identical whether the strand was labeled at the 3′- or the 5′-end (Table 1). Thus, the adducts involving the 3′ G were apparently monofunctional and those at the 5′ G were either interstrand crosslinks or monofunctional adducts. Although melphalan treatment of single-stranded DNA did generate slightly different cleavage patterns for 5′- and 3′-end-labeled fragments, the amount of cleavage remained unchanged during a 24 h time period, indicating that intrastrand crosslinks (5%) were formed within the first 30 min of drug incubation.

The highly reactive mechlorethamine, which produced the largest number of interstrand crosslinks, also generated the most intrastrand crosslinks in both single-stranded (13%) and double-stranded (28%) DNA. No increase in the number of 3′ or 5′ breaks at the G-G site during a 24 h time period was detected, indicating rapid crosslink formation. Phosphoramide mustard produced detectable intrastrand crosslinking (8%) only in double-stranded DNA.

Molecular dynamics simulations

Mechlorethamine, melphalan and phosphoramide mustard adducts in double-stranded DNA were subjected to molecular dynamic simulations in an attempt to determine a structural basis for differences in intrastrand crosslink formation. The movement of the aziridinium ion intermediate was evaluated in terms of the distance from each aziridinium carbon to either the guanine N-7 in the opposite strand, corresponding to a potential G-N-C·G-N-C interstrand crosslink, or the adjacent guanine N-7 in the same strand, corresponding to a potential intrastrand crosslink. The potential energy of the drug–DNA complex was also calculated for each conformation.

Table 1. Fraction of interstrand crosslinks formed at a G-G-C site in double- and single-stranded DNA[a]

	5′-G/3′-G[b] 5′-end-label	3′-G/5′-G[b] 3′-end-label	Intrastrand crosslinks
Double-stranded DNA			
Melphalan	2.30 ± 0.030	0.45 ± 0.025	0.01 ± 0.011
Mechlorethamine	2.54 ± 0.058	1.28 ± 0.044	0.28 ± 0.013
Phosphoramidemustard	1.98 ± 0.026	0.70 ± 0.012	0.08 ± 0.007
Single-stranded DNA			
Melphalan	0.95 ± 0.026	1.23 ± 0.020	0.05 ± 0.011
Mechlorethamine	1.07 ± 0.025	1.56 ± 0.070	0.13 ± 0.011
Phosphoramide mustard	0.78 ± 0.026	1.34 ± 00.076	0.01 ± 0.021

[a]Values are presented as the mean ± standard error of the mean (SEM) for a minimum of four experiments.
[b]Apparent ratio of cleavage at the two sites in samples incubated for 24 h following drug exposure and then treated with piperidine.

In each of these simulations, equilibrium was reached within 10 ps, after which both the energy and the adduct orientation remained essentially constant (data not shown). Prior to reaching equilibrium, all three adducts briefly sampled conformational space to within ~4.5–5.3 Å of the guanine N-7 corresponding to the interstrand crosslink site. Although this distance has been considered sufficiently small to allow interstrand crosslinking (22,23), the equilibrium conformations of both mechlorethamine and phosphoramide mustard adducts appeared to favor intrastrand crosslink formation. The average distance from the nearest aziridinium carbon atom to the adjacent guanine N-7 was 3.14 (mechlorethamine, Fig. 7, left) or 3.36 Å (phosphoramide mustard, not shown), with ~0.5 Å deviation from these positions and direct van der Waals contacts between these two moieties during the simulation. Thus, modeling studies predict that mechlorethamine and phosphoramide mustard, when adducted to DNA, assume relatively stable conformations wherein the second mustard arm is appropriately positioned for intrastrand crosslink formation at G-G sequences.

In contrast, the melphalan–DNA adduct was stably oriented in the major groove with an average distance of 6.15 ± 0.22 and 6.44 ± 0.20 Å from each carbon atom in the aziridinium ion to the adjacent guanine N-7. A major factor in stabilizing this conformation appeared to be hydrogen bonding of the positively charged terminal amine group of melphalan to the 5′ oxygen (O5′) and the corresponding phosphate oxygens (O1P or O2P) of the deoxycytidylate residue located 2 bases 5′ of the adduct site (Fig. 7, right). These hydrogen bonds tended to draw the aziridinium away from the adjacent guanine, thus providing a possible structural basis for the lack of intrastrand crosslinking by melphalan.

DISCUSSION

Nitrogen mustards produce a variety of monofunctional and bifunctional DNA lesions and a diverse spectrum of mutations, including substitutions at both G·C and A·T base pairs, as well as large deletions (2). In the endogenous *aprt* gene of CHO cells, the most frequent mutations induced by the aromatic mustard melphalan were G·C→T·A and G·C→C·G transversions and these occurred preferentially at G-G-C, G-G-C-C and to a lesser extent G-A-C sequences. Stereochemical considerations suggest that G-G-C sequences could be potential sites for formation of both interstrand and intrastrand crosslinks involving guanine N-7. Since such crosslinks are likely to be more difficult to repair than monofunctional alkylations, they represent prime candidates for premutagenic DNA lesions. The present study was undertaken to determine the propensity of several nitrogen mustards to crosslink specific guanine residues in a G-G-C·G-C-C sequence.

The results show that both melphalan and phosphoramide mustard induce interstrand guanine–guanine crosslinks with the same two base-staggered G-N-C·G-N-C specificity previously reported for mechlorethamine (10,13), but that melphalan-induced crosslinks were formed much more slowly and were more stable than mechlorethamine-induced crosslinks. The slow kinetics of formation, which were detected previously in an agarose gel assay (24), appear to be due to the very slow second arm reaction of monofunctionally bound melphalan. The greater persistence of melphalan-induced interstrand crosslinks *in vivo* ($t_{1/2}$ > 24 h) (25) as compared with mechlorethamine-induced crosslinks is probably due at least in part to differences in rates of spontaneous depurination. Conversely, single- and double-strand cleavage resulting from spontaneous depurination of mechlorethamine-induced adducts and crosslinks may account for the high proportion of deletion mutations induced by this drug in the pZ189 shuttle vector system, as well as in the *HPRT* gene in human lymphoblastoid cells (26). In contrast, aromatic mustards, which form more stable adducts, induce predominantly point mutations, both in shuttle vectors and in endogenous genes (2).

At a G-G-C sequence in double-stranded DNA, the proportion of intrastrand crosslinks formed by various mustards differed by a factor of >20-fold, with mechlorethamine > phosphoramide mustard > melphalan. These differences appear to be due at least in part to differences in the geometry of the monofunctionally bound drug in the major DNA groove, since they were less pronounced in single-stranded DNA. Moreover, molecular modeling predicted that mechlorethamine and phosphoramide mustard would assume monoadduct conformations consistent with intrastrand crosslinking, while melphalan, due to hydrogen bonding of its phenylalanine moiety to the DNA backbone, would not. It should be noted that the lack of intrastrand crosslinking by melphalan cannot be attributed to the mere bulkiness of the phenylalanine side chain, as additional modeling studies (not shown) showed that a melphalan molecule bifunctionally linked to N-7 positions of adjacent guanines can easily fit into the major groove with essentially no distortion in DNA structure.

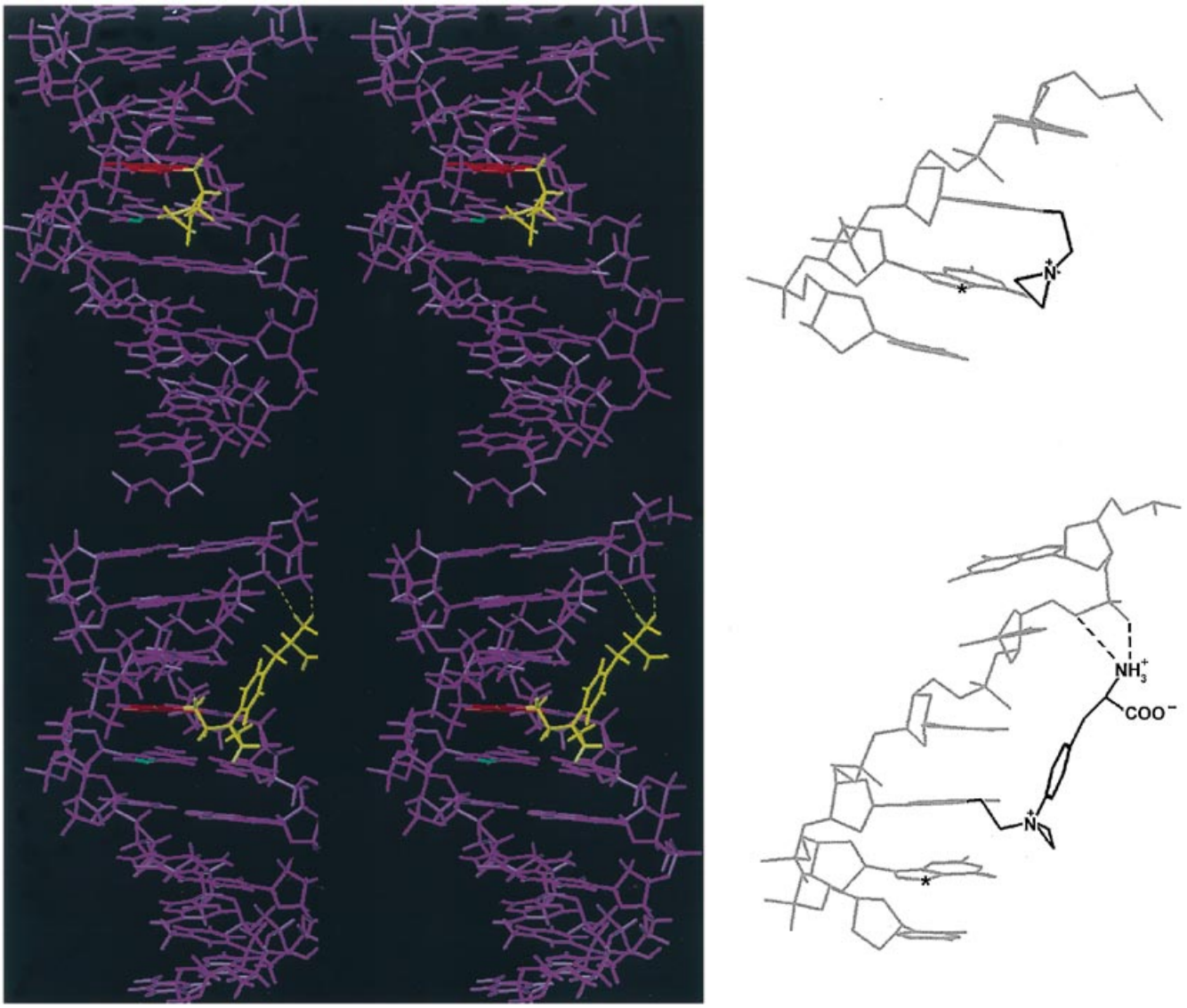

Figure 7. Calculated average structure of the mechlorethamine and melphalan aziridinium ion intermediates taken from the 100 ps dynamics. Each monoadduct is attached to the N-7 of G in the 5′-pGCAGGCGAG $_{OH}$-3′ sequence and is shown with a view into the major groove. Mechlorethamine (top) adopts an energetically stable conformation in which the aziridinium carbons are positioned near the N-7 of the adjacent G (just below the adducted base), a potential intrastrand crosslink site. Melphalan (bottom) is positioned with the aziridinium >6 Å away from the N-7 of the adjacent G, in a conformation stabilized by hydrogen bonding (dotted lines) of the terminal amine of melphalan to the 5′ phosphodiester oxygen and the 5′ phosphate oxygen of the deoxycytidylate residue located two bases 5′ of the adducted base. The adducted mustards are shown in yellow and the DNA in purple, except for the guanine (red), which is the monoaduct site, and the N-7 (green) of the adjacent guanine, which is the target for intrastrand crosslink formation. The left panel shows stereoviews of the monoadducts. The right panel shows a simplified monochrome view of the same adducts, but with mustard structures indicated and with the complementary DNA strand and all hydrogens eliminated for clarity. The asterisk indicates the target N-7 for intrastrand crosslinking.

The relatively high frequency of intrastrand crosslinking induced by mechlorethamine supports the proposal of Pieper *et al.* (27) that selective transcription termination at G-G sequences in mechlorethamine-treated DNA may be attributable to intrastrand crosslinks. Intrastrand crosslinks could also account for at least some mechlorethamine-induced base substitutions in pZ189, most of which occur at G-G sites (7). However, there is no evidence that intrastrand crosslinks correlate with cytotoxicity, since the ratio of intrastrand to interstrand crosslinks is much higher for mechlorethamine than for melphalan, yet, at equivalent levels of interstrand crosslinking, melphalan is somewhat more cytotoxic (25).

The present data do not clearly implicate any one DNA lesion as accounting for the relatively high frequency of melphalan-induced mutations at G-G-C sequences in *aprt*. Interstrand crosslinks could account for substitutions at the first and last positions in the sequence, but not at the internal G, which is mutated just as frequently (9). Intrastrand crosslinks, which have been implicated in mutagenesis by cisplatin (28,29), remain a possible candidate lesion for melphalan-induced mutations

involving the internal G. Although intrastrand crosslinks are apparently produced by melphalan only in single-stranded DNA, the very slow kinetics of the second arm alkylation reaction raise the possibility that they could be formed when monoadducted DNA becomes transiently single stranded during transcription or replication. Of course, the mutations could be the result of N-7 monoadducts, but since G-G-C sequences are not particularly strong sites for monoadduct formation, other factors must then be invoked to account for the sequence specificity of mutagenesis. Although all guanine adducts detected thus far in melphalan-treated DNA appear to involve N-7 alkylation (2), it is difficult to eliminate the possibility that other rare but highly mutagenic adducts might also be induced at these sites (e.g. guanine N^2 adducts, which have been implicated in mutagenesis by benzo[*a*]pyrene; 30). Finally, the possibility of untargeted or 'locally targeted' mutagenesis (29) or of mutagenesis by some endogenous species generated in response to malphalan treatment cannot be ruled out.

In conclusion, although interstrand crosslinks were detected at the 5′ G in the G-G-C sequences, there is little indication of what lesions might be involved in melphalan-induced mutations at the 3′ G, which occur just as frequently. Similar considerations apply to the spectrum of mechlorethamine-induced mutations in the yeast *SUP4*-o gene (31), where several of the most frequently mutated sites are neither expected alkylation hotspots (17) nor potential sites of inter- or intrastrand crosslinks. The disparate mutation spectra produced by nitrogen mustards in various experimental systems (2) suggest that rather than being attributable to any single highly mutagenic lesion, mustard-induced mutagenesis represents the cumulative effects of several types of adenine and guanine adducts, both monofunctional and bifunctional. This complexity would appear to preclude, for the present, any rational attempts to devise less mutagenic but equally cytotoxic analogs of these drugs.

ACKNOWLEDGEMENTS

This work was supported by grant CA40615 from the National Cancer Institute. G.B.B. was supported in part by Training Grant HL07110 from the National Heart, Lung and Blood Institute.

REFERENCES

1 Medical Letter Inc. (1993) *Med. Lett.*, **35**, 43–50.
2 Povirk,L.F. and Shuker,D.E. (1994) *Mutat. Res.*, **318**, 205–226.
3 Osborn,M.R., Wilman,D.E.V. and Lawley,P.D., (1995) *Chem. Res. Toxicol.*, **8**, 316–320.
4 Henne,T. and Schmähl,D. (1985) *Cancer Treatment Rev.*, **12**, 77–94.
5 Nowell,P.C. (1982) *Cancer Genet. Cytogenet.*, **5**, 265–278.
6 Rowley,J.D., Golomb,H.M. and Vardiman,J.W. (1981) *Blood*, **58**, 759–768.
7 Wang,P., Bauer,G.B., Bennett,R.A.O. and Povirk,L.F. (1991) *Biochemistry*, **30**, 11515–11521.
8 Wang,P., Bauer,G.B., Kellogg,G.E., Abraham,D.J. and Povirk,L.F. (1994) *Mutagenesis*, **9**, 133–139.
9 Austin,M.J.F., Han,Y.-H. and Povirk,L.F. (1992) *Cancer Genet. Cytogenet.*, **64**, 69–74.
10 Millard,J.T., Raucher,S. and Hopkins,P.B. (1990) *J. Am. Chem. Soc.*, **112**, 2459–2460.
11 Hemminki,K. (1985) *Cancer Res.*, **45**, 4237–4243.
12 Maniatis,T., Fritsch,E.F. and Sambrook,J. (1982) *Molecular Cloning: A Laboratory Manual.* Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
13 Ojwang,J.O., Grueneberg,D.A. and Loechler,E.L. (1989) *Cancer Res.*, **49**, 6529–6537.
14 Maxam,A.M. and Gilbert,W. (1980) *Methods Enzymol.*, **65**, 499–560.
15 Sokal,R.B. and Rohlf,F.J. (1981) *Biometry.* W.H.Freeman, San Francisco, CA.
16 Singer,B. (1975) *Prog. Nucleic Acids Res. Mol. Biol.*, **15**, 219–285.
17 Mattes,W.B., Hartley,J.A. and Kohn,K.W. (1986) *Nucleic Acids Res.*, **14**, 2971–2987.
18 Mehta,J.R. and Ludlum,D.B. (1980) *Cancer Res.*, **40**, 4183–4186.
19 Kallama,S. and Hemminki,K. (1986) *Chemico-Biol. Interact.*, **57**, 85–87.
20 Müller,N. and Eisenbrand,G. (1985) *Chemico-Biol. Interact.*, **53**, 173–181.
21 Eastman,A. (1983) *Biochemistry*, **22**, 3927–3933.
22 Haworth,I.S., Lee.C-S., Yuki,M. and Gibson,N.W. (1993) *Biochemistry*, **32**, 12857–12863.
23 Dong,Q., Barsky,D., Colvin,M.E., Melius,C.F., Ludeman,S.M. Moravek,J.F., Colvin,O.M., Bigner,D.D., Modrich,P. and Friedman,H.S. (1995) *Proc. Natl. Acad. Sci. USA*, **92**, 12170–12174
24 Sunters,A., Springer,C.J., Bagshawe,K.D., Souhami,R.L. and Hartley,J.A. (1992) *Biochem. Pharmacol.*, **44**, 59–64
25 Ross,W.E., Ewig,R.A.G. and Kohn,K.W. (1978) *Cancer Res.*, **38**, 1502–1506.
26 Henner,W.D., Sanderson,B.J.S., Johnson,K.J. and Walker,K.K. (1991) *Proc. Am. Ass. Cancer Res.*, **32**, 104.
27 Pieper,R.O., Futscher,B.W. and Erickson,L.C. (1989) *Carcinogenesis*, **10**, 1307–1314.
28 Burnouf,D., Daune,M. and Fuchs,R.P.P. (1987) *Proc. Natl. Acad. Sci. USA*, **84**, 3758–3762.
29 de Boer,J.G. and Glickman,B.W. (1989) *Carcinogenesis*, **10**, 1363–1367.
30 Mazur,M. and Glickman,B.W. (1988) *Somatic Cell Mol. Genet.*, **14**, 393–400.
31 Kunz,B.A. and Mis,J.R. (1989) *Cancer Res.*, **49**, 279–283.