Thermolabile Adenine Adducts and A·T Base Pair Substitutions Induced by Nitrogen Mustard Analogues in an SV40-Based Shuttle Plasmid†

Peng Wang, Gwen B. Bauer, Richard A. O. Bennett, and Lawrence F. Povirk*

Department of Pharmacology and Toxicology, Medical College of Virginia, Virginia Commonwealth University, Richmond, Virginia 23298-0230

Received July 15, 1991; Revised Manuscript Received September 16, 1991

ABSTRACT: It was previously shown that the predominant mutations induced by melphalan (L-phenylalanine mustard) in the *supF* gene of shuttle plasmid pZ189 during replication in human cells are A·T → T·A transversions. In order to determine whether adenine adducts were formed at sequence positions corresponding to these mutations, melphalan-induced thermolabile adducts were mapped in the *supF* gene by selective depurination followed by strand cleavage in alkali. All A·T base pairs which were frequent sites for melphalan-induced A·T → T·A transversions were also prominent sites for formation of thermolabile adenine adducts. Although no mutations were detected at some prominent adduct sites, there was a significant correlation between adduct sites and mutation sites. While runs of two or more adenines were particularly prominent adduct sites, comparison of results obtained with 3′- and 5′-end-labeled DNA gave no evidence for intrastrand cross-links between adjacent adenines. Chlorambucil, another aromatic nitrogen mustard, showed sequence specificities for both mutagenesis and adenine adduct formation nearly identical to those seen with melphalan. The nonaromatic analogues mechlorethamine and phosphoramide mustard were much less efficient in inducing thermolabile adenine adducts, and mechlorethamine induced significantly fewer transversions at A·T base pairs than chlorambucil or melphalan. Formation of thermolabile adenine adducts by the aromatic nitrogen mustards was markedly reduced by blockage of the minor groove with distamycin, or by prior heat denaturation of the DNA. These results suggest that alkylation occurs primarily at the N-3 rather than N-7 position of adenine, probably as a consequence of the affinity of the aromatic rings of melphalan and chlorambucil for the minor groove. Overall, the results are consistent with the proposal that A·T → T·A transversions induced by aromatic nitrogen mustards result from adenine N-3 alkylations.

Long-term survivors of cancer chemotherapy show significantly increased risk of contracting second malignancies, which are usually unrelated to the original neoplasm and probably result from the mutagenic effects of DNA-damaging drugs (Henne & Schmähl, 1985). Patients treated with bifunctional nitrogen mustards (Figure 1), particularly melphalan (L-phenylalanine mustard, L-PAM), have the greatest incidence of secondary tumors, most commonly leukemias.

Surprisingly little work has been done on the molecular mechanisms of mutagenesis by these agents. It is known that the main site of alkylation by nitrogen mustards is N-7 of guanine (Hemminki & Kallama, 1986), but there appear to be a number of minor alkylations which have yet to be characterized. Recently, we reported that the predominant mutations induced in the shuttle plasmid pZ189, by treatment in vitro with melphalan followed by replication in human cells, were A·T → T·A transversions, and we proposed adenine N-3 adducts as possible premutagenic lesions (Wang et al., 1990). Preferential insertion of adenine (and to a lesser extent guanine) opposite this "noncoding" lesion would be expected to produce A·T → T·A (and to a lesser extent A·T → C·G) transversions (Strauss et al., 1982). Pieper and Erickson (1990) have shown that melphalan-treated DNA contains thermolabile lesions at certain adenine residues, which can be converted to strand breaks by heating at neutral pH, a result consistent with the presence of adenine N-3 adducts. In an attempt to determine whether these thermolabile lesions might be responsible for drug-induced A·T → T·A transversions, we undertook further studies of both adduct formation and mutagenesis by nitrogen mustards in the *supF* target gene of pZ189.

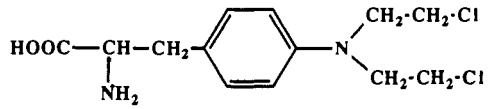
Experimental Procedures

Materials. Phosphoramide mustard, originally from Dr. V. N. Narayanan, Developmental Therapeutics Program, National Cancer Institute, was a gift of Dr. T. Kawabata. Due to the instability of this drug in aqueous solution (Hemminki, 1985), it was freshly dissolved and diluted in reaction buffer immediately before use. Other nitrogen mustards and distamycin A were from Sigma. Melphalan and mechlorethamine were dissolved in 0.1 N HCl at a concentration of 3–20 mM and diluted in water immediately before use. Chlorambucil was dissolved and diluted in dimethyl sulfoxide; the final dimethyl sulfoxide concentration in drug–DNA reactions never exceeded 4%.

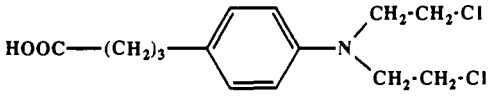
For most experiments, DNA was 5′- or 3′-end-labeled at the *Ava*I site at base 109 of *supF*, using methods described previously (Povirk et al., 1989). Following secondary cleavage with *Nci*I, a 399-bp labeled fragment, containing bases 109–183 of the *supF* tRNA sequence, was isolated (the tRNA sequence begins at base 99). To obtain a DNA fragment 5′-end-labeled at the downstream end of *supF*, 0.1 μg of pZ189 was used as a template for a polymerase chain reaction (30 cycles of 90, 37, and 72 °C, each for 1 min) with 0.2 μg of the primers GTATCACGAGGCCCT (the pBR322 *Eco*RI clockwise sequencing primer) and GCTCAATTCTTTCG-

† Supported by Grant CA40615 from the National Cancer Institute, DHHS. P.W. was supported in part by a fellowship from the Carol Jean Lipman MacFarlane and Ann Debra Lipman Fund. R.A.O.B. was supported by Training Grant ES07087 from the National Institute of Environmental Health Sciences, DHHS.

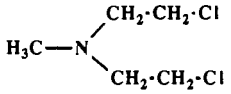
0006-2960/91/0430-11515$02.50/0



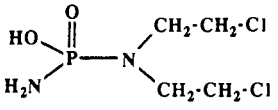
Melphalan



Chlorambucil



Mechlorethamine



Phosphoramide Mustard

FIGURE 1: Structures of nitrogen mustard analogues. The two chloroethyl groups are the alkylating moieties. Phosphoramide mustard is the active metabolite of cyclophosphamide.

GAC (corresponding to bases 212–196 of the transcribed strand of *supF*). Following phenol extraction, removal of primers by ultrafiltration, and trimming of any 3′ overhangs with the large fragment of DNA polymerase I (Hemsley et al., 1989), the resulting 238-bp product was 5′-end-labeled with T4 polynucleotide kinase and [γ-^{32}P]ATP. The product was cut with *Eco*RI, and a 211-bp fragment, 5′-end-labeled at base 212, was isolated.

Analysis of Thermolabile Sites. Reactions (25–100 μL) contained 20 μg/mL calf thymus DNA, a ^{32}P-labeled DNA fragment (<10 μg/mL), 10 mM Tris-HCl, pH 8, 1 mM EDTA, and the indicated concentrations of nitrogen mustards and were incubated at 37 °C for 1 h. Following addition of 100 μg/mL carrier tRNA, the DNA was ethanol precipitated and redissolved in 60–100 μL of 50 mM Tris-HCl, pH 7.4–1 mM EDTA, and aliquots were either reprecipitated immediately or heated at 70 °C for 1 h to depurinate thermolabile adducts and then precipitated; adduct depurination was apparently complete under these conditions since no additional drug-induced depurination was detected when samples were heated at 90 °C for 1 h. The samples were then dissolved in 100 μL of 1 M piperidine, heated at 90 °C for 30 min, lyophilized, dissolved in 80% formamide containing 10 mM EDTA, denatured for 1 min at 90 °C, and electrophoresed on sequencing gels of 8%, 12%, or 20% polyacrylamide. Gels were either dried onto plates or frozen for autoradiography.

To determine the relative cleavage at various adenine residues in *supF*, autoradiographs obtained without intensifying screens were scanned with a laser microdensitometer (Shimadzu). For quantitation, peaks were cut out and weighed on an analytical balance. Correction for multiple cleavages within a single DNA molecule was performed as described previously (Povirk & Goldberg, 1985). Breakage at each adenine residue was normalized to the average breakage at guanine residues. Only samples which sustained <0.5 break per molecule within the scanned sequence were used for quantitation. Linear regression analysis of the correlation between adduct sites and mutation sites was performed using the General Linear Models program of the SAS Statistical Package (SAS Institute, Cary, NC).

Mutation Spectra. Drug treatments were performed as described above, except that samples contained 120 μg/mL supercoiled pZ189 DNA as the substrate, and no carrier tRNA was added. Following ethanol precipitation, the plasmid was transfected into human 293 cells by the calcium phosphate procedure and allowed to replicate for 48 h. Progeny plasmid was isolated, treated with *Dpn*I to digest any unreplicated molecules, and transfected by electroporation into MBM7070 *lacZam* indicator cells. Plasmids with point mutations in *supF* were sequenced using one of the two primers described above [see Wang et al. (1990) for details].

In order to estimate the frequency of thermolabile adducts induced under these conditions at each drug dose, similar reactions were performed, but with a trace of the 211 base pair ^{32}P-end-labeled fragment added. Following ethanol precipitation, heating at neutral pH, and piperidine treatment as described above, the samples were electrophoresed on sequencing gels. Data from autoradiographs were fit to the equation $I = I_0 e^{-NkD}$, where I is the intensity of the band corresponding to the full-length labeled fragment, $N = 211$, the length of the fragment, D is the dose, k is an empirically determined constant, and kD is the average number of thermolabile adducts per nucleotide at a given dose. At least two experiments were performed for each drug, at doses comparable to those used in mutagenesis experiments.

Results

Sequence Specificity of Drug-Induced Thermolabile Sites. When a 5′-end-labeled segment of the *supF* gene was exposed to mechlorethamine, phosphoramide mustard, chlorambucil, or melphalan (Figure 2) and then treated with hot piperidine, cleavage occurred exclusively at all guanine residues, in agreement with results of Mattes et al. (1986a). Piperidine is known to induce cleavage at guanine N-7 adducts, by a mechanism involving a formamidopyrimidine intermediate (Mattes et al., 1986b). When the drug-treated DNA was first heated at neutral pH to depurinate labile adducts and then treated with hot piperidine to cleave and remove depurinated sugars, cleavage was also seen at adenine residues, as has been reported by Pieper and Erickson (1990). For the aromatic nitrogen mustards melphalan and chlorambucil, cleavage at some adenine residues was comparable to cleavage at nearby guanine residues, while cleavage at other adenine sites was barely detectable. In particular, strong cleavage sites were often found at runs of two or more adenines; for example, prominent cleavage was seen at all adenines in a TAAAT sequence (bases 134–138, Figure 2), and at all but the first adenine in a CAAAG sequence (bases 118–122), a CAAAAG sequence (bases 191–196), and two GAA sequences (bases 156–158 and 164–166). Little cleavage occurred at adenines flanked on both sides by G·C base pairs. Cleavage patterns seen with melphalan and with chlorambucil were virtually identical. Cleavage at individual adenine sites varied over a range of approximately 20-fold (see Figure 5, below).

In contrast, DNA treated with the nonaromatic nitrogen mustards mechlorethamine and phosphoramide mustard showed much less cleavage at adenine residues. However, the cleavage that was detected at high doses occurred at most of the same nucleotide positions as that seen with the aromatic analogues. Because they were relatively infrequent, adenine lesions induced by the nonaromatic analogues could only be accurately quantitated at prominent, well-separated cleavage sites such as positions 135–137. On the basis of comparison of cleavage at these sites with cleavage at nearby guanine positions, it was estimated that mechlorethamine was about 6-fold less effective, and phosphoramide mustard about 20-fold

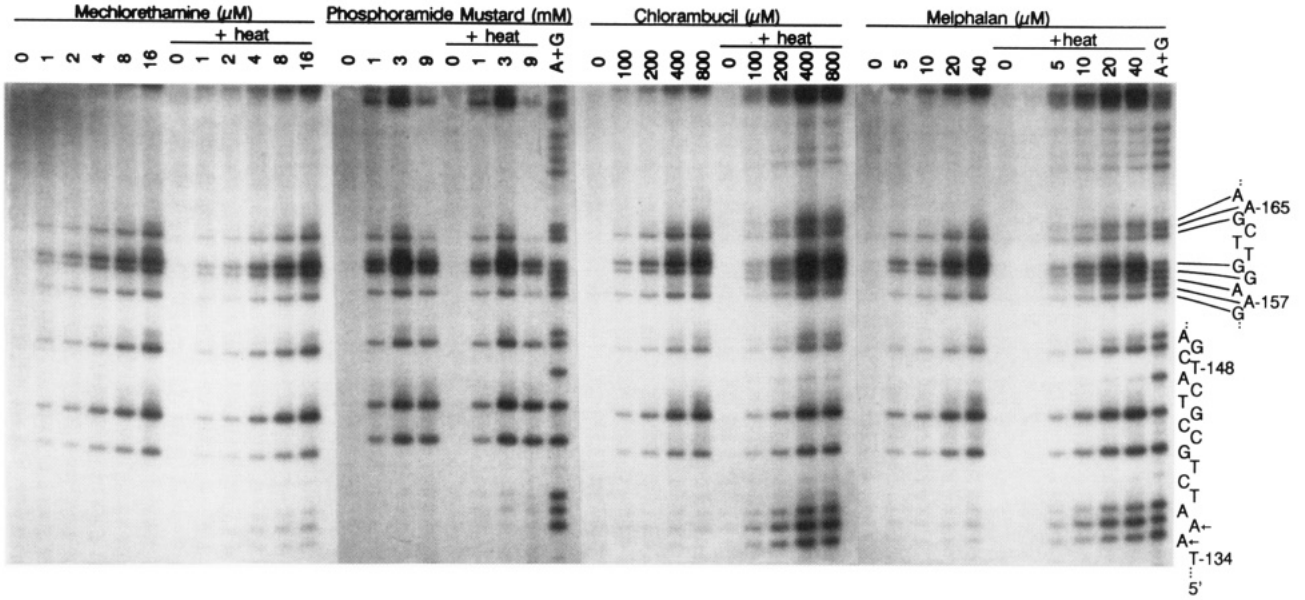


FIGURE 2: Thermolabile lesions in DNA treated with nitrogen mustard analogues. A 5′-end-labeled DNA fragment from the coding strand of *supF* was treated with the indicated concentrations of various nitrogen mustards and then either treated with piperidine or heated at neutral pH and then treated with piperidine. "A+G" indicates a Maxam–Gilbert sequencing marker. Arrows indicate positions of mutational hotspots for melphalan- and chlorambucil-treated pZ189 (see Figure 7). Electrophoresis was on a 12% polyacrylamide gel. The cluster of unresolved bands at the very top of the gels for chlorambucil and melphalan (which were well resolved on an 8% gel) correspond to cleavage at all but the first adenine in a CAAAAG sequence (bases 191–196).

less effective, than melphalan or chlorambucil in inducing thermolabile adenine adducts.

Similar patterns of melphalan- and chlorambucil-induced thermolabile adenine lesions were seen in the complementary (transcribed) strand of *supF* (not shown). In particular, at each of three GAA sequences (bases 172–170, 163–161, and 155–153), much greater cleavage was seen at the adenine nearest the 3′ end, whether the strand was 5′- or 3′-end-labeled. This specificity is in agreement with the results shown in Figure 2, and with results obtained by Pieper and Erickson (1990) using another 3′-end-labeled fragment. However, as discussed below, our results with 5′-end-labeled fragments imply that, at least at the GAA sequences, few if any of the thermolabile lesions are intrastrand adenine–adenine cross-links.

Both N-7 and N-3 alkylations render adenine labile to heat depurination (Singer, 1975). Since N-7 lies in the major groove of DNA, while N-3 is exposed in the interior of the minor groove, it should be possible to selectively block N-3 alkylation by carrying out the reaction in the presence of a drug such as distamycin, which binds in the minor groove with very high affinity (Zimmer & Wahnert, 1986). Distamycin completely eliminated the chlorambucil-induced production of thermolabile sites at adenine residues, while having little effect on cleavage at guanine residues (Figure 3). Sequences at which distamycin is expected to bind most strongly in the minor groove, such as TAAAT (Zimmer & Wahnert, 1986), were almost completely protected even at the lowest distamycin concentrations. At intermediate concentrations (5.6 μM) all sites were protected. Virtually identical results were seen with melphalan (not shown). These results strongly suggest that nearly all the thermolabile sites at adenine residues represent adenine N-3 alkylations.

The differences between the aromatic and nonaromatic nitrogen mustards suggested that the aromatic moieties might have a strong affinity for the DNA minor groove, thus positioning the aromatic drugs for selective alkylation of adenine N-3. To test this hypothesis, DNA was heat-denatured for 1 min at 90 °C immediately before drug addition. As shown in Figure 4, denatured DNA showed a dramatic reduction in

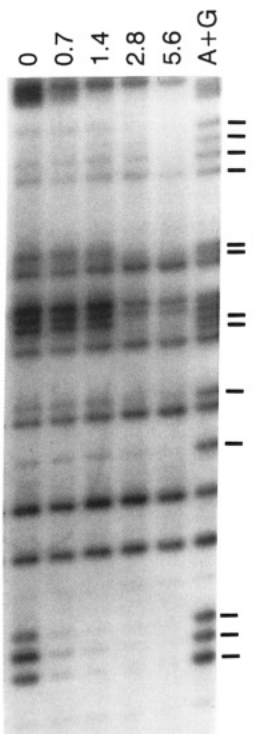


FIGURE 3: Effect of distamycin on formation of thermolabile sites by chlorambucil. The 5′-end-labeled *supF* DNA fragment (see Figure 2) was treated with 400 μM chlorambucil in the presence of the indicated concentrations of distamycin (in μM). Distamycin was added 15 min before chlorambucil. Bars show positions of adenines in the sequence. No detectable cleavage was induced by distamycin alone.

drug-induced thermolabile adenine sites, while, again, guanine sites were little affected. Thus, noncovalent interactions of the aromatic rings of melphalan and chlorambucil with the minor groove of DNA appear to promote adenine N-3 alkylation, while guanine N-7 alkylation shows much less dependence on DNA secondary structure.

Comparison of Specificity of Alkylation and Mutagenesis by Melphalan. In order to test for a possible correlation

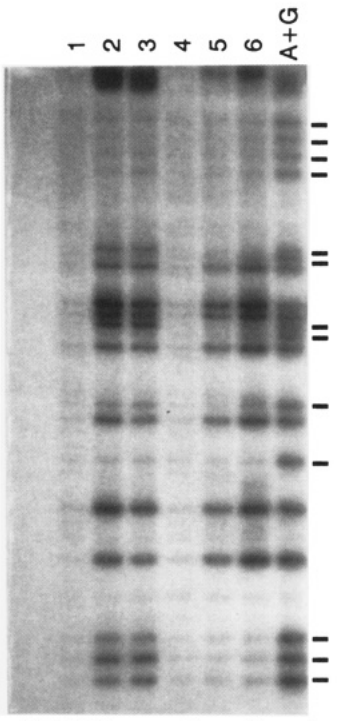


FIGURE 4: Effect of DNA denaturation on formation of thermolabile adducts. Either native DNA (lanes 1–3) or denatured DNA (lanes 4–6) was treated with no drug (lanes 1 and 4), 20 μM melphalan (lanes 2 and 5), or 400 μM chlorambucil (lanes 3 and 6), followed by heating at neutral pH and treatment with piperidine. Bars show positions of adenines in the sequence (see Figure 2).

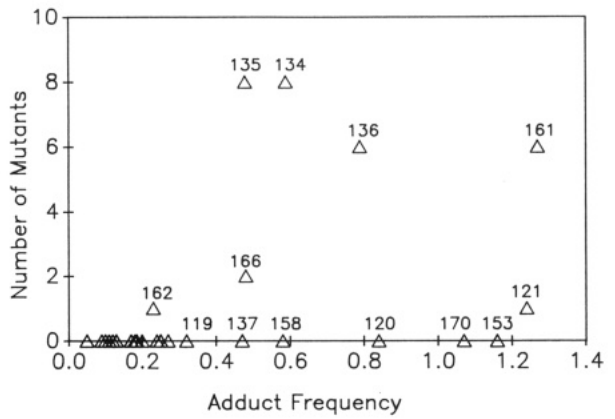


FIGURE 5: Comparison of thermolabile adduct formation and generation of single-base substitutions by melphalan at various adenine residues in the *supF* gene. Adduct frequencies, determined by densitometry of sequencing gels as described under Experimental Procedures, have been normalized to the average cleavage at guanine residues in the sequence and represent the mean of at least 2 and in most cases 3 measurements. Numbers indicate sequence positions. The points near the origin, indicating sites at which no mutations occurred and at which adduct levels were 0.05–0.27, correspond to (in order of increasing adduct frequency) base pairs 138, 183, 180, 165, 177, 157, 147, 165, 151, 131, 140, 148, 128, 145, 162, 132, and 125. Mutation data are from Wang et al. (1990).

between the specificity of base substitution mutagenesis and that of adduct formation, thermolabile adducts were quantitated at all adenines in both strands of bases 117–183 of *supF* (which includes 31 of the 35 A·T base pairs in the gene and all of those at which single-base substitutions occurred), using the three end-labeled fragments described under Experimental Procedures. Adducts at each site were compared to the numbers of single-base substitutions detected in a previous study (Wang et al., 1990). All mutational hotspots were frequent sites of thermolabile adducts (Figure 5), but there

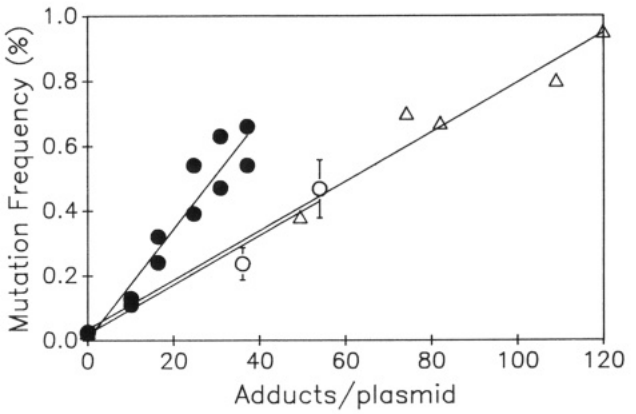


FIGURE 6: Mutagenesis of pZ189 by nitrogen mustard analogues. Frequencies of mutations in the *supF* gene were determined following in vitro treatment of pZ189 with mechlorethamine (●), melphalan (△), or chlorambucil (○), replication in human 293 cells, and rescue in indicator bacteria. Melphalan data are taken from Wang et al. (1990). A 10-fold reduction in the recovery of replicated plasmid from the human cells was seen at 9 μM mechlorethamine (28 adducts/plasmid), 33 μM melphalan (90 adducts/plasmid), and 195 μM chlorambucil (35 adducts/plasmid). For chlorambucil, mean values ± standard errors for 4 experiments are shown. Because a number of different melphalan and mechlorethamine concentrations were used, mutation frequencies of individual cultures from several experiments are shown.

were several prominent adduct sites at which no mutations were detected, such as base pairs 120, 153, and 170. Nevertheless, linear regression analysis of these data indicated a statistically significant ($p = 0.02$) correlation between adduct frequency and mutation.

Comparison of Mutational Spectra of Chlorambucil and Mechlorethamine. If thermolabile adenine adducts are primarily responsible for A·T → T·A (and A·T → C·G) transversions, aromatic nitrogen mustards would be expected to produce a larger proportion of these mutations than nonaromatic analgues. To test this prediction, spectra of mutations induced by chlorambucil and by mechlorethamine were determined for pZ189 replicated in human 293 cells (mutagenesis studies with phosphoramide mustard were precluded by the high drug concentrations required for in vitro modification and the small amount of drug available).

As might be expected from in vitro studies (Figure 2), chlorambucil treatment inhibited recovery of progeny pZ189 plasmid and induced mutagenesis only at much higher concentrations than those required for mechlorethamine or melphalan. In order to more directly compare mutagenic potencies of the three drugs, overall alkylation levels were measured for each drug under the conditions of the mutagenesis experiments, by quantitating the loss of full-length end-labeled DNA fragments in experiments similar to those shown in Figure 2. These experiments showed that adduct formation was a linear function of drug dose and that production of an average of one thermolabile adduct in one strand of a 211-bp fragment (the equivalent of 52 adducts per pZ189 plasmid) required 17 μM mechlorethamine, 19 μM melphalan, and 290 μM chlorambucil (data not shown). In the case of melphalan, these results are in reasonably good agreement with those obtained previously (Wang et al., 1990) by measuring binding of radiolabeled drug (52 adducts per plasmid at 23 μM melphalan), suggesting that thermolabile adducts accounted for the majority of DNA alkylation. When compared on the basis of induced adducts, mechlorethamine was somewhat more potent in inducing mutations than either melphalan or chlorambucil (Figure 6). However, mechlorethamine induced relatively fewer base substitutions and

Table I: Characterization of Mutants Induced by Chlorambucil and Mechlorethamine[a]

	chlorambucil	mechlorethamine
plasmid clones examined	134	172
no. with normal mobility[b]	115 (86)	92 (53)
no. sequenced[c]	83	60
single-base substitutions	62 (64)[d]	26 (23)[d]
double substitutions	13[e] (13)	1 (1)
1–2-base deletions	6 (6)	0
deletions ≥3 bases	2 (2)	31 (28)
no sequence	0	2[f] (2)

[a] Pooled results of seven transfections for each drug, with mutation frequencies between 0.25% and 0.7%. Values in parentheses are in percent (%). [b] Electrophoretic mobility on agarose indistinguishable from pZ189. [c] Only plasmids with normal mobility, i.e., without large deletions or rearrangements, were sequenced. [d] Expressed as a fraction of total clones examined, e.g., (62/83) × 86% = 64%. [e] Including one triple substitution. [f] Probably deletions of both sequencing primers; several of the sequenced deletions involved deletion of one of the primers.

more small and large deletions than chlorambucil (Table I) or melphalan (Wang et al., 1990); thus, the number of induced base substitutions per adduct was actually somewhat less for mechlorethamine than for the other drugs.

As also might be predicted from in vitro studies, the specificities of chlorambucil- and melphalan-induced substitutions were very similar. Of the single-base substitutions induced by chlorambucil, 65% (40/62) were A·T → T·A or A·T → C·G transversions, as compared with 74% for melphalan (Wang et al., 1990), and most of these occurred at the same four sites, base pairs 134, 135, 136, and 161 (Figure 7). In the case of mechlorethamine, A·T → T·A or A·T → C·G transversions accounted for only 31% (8/26) of the single-base substitutions. Most of these occurred at bases 161 and 162, which, like all other adenines in the sequence, showed only a barely detectable level of thermolabile lesions following treatment with mechlorethamine (data not shown). In summary, the mutational spectrum of mechlorethamine showed, as expected, a lower proportion of transversions at A·T base pairs, although the difference was not as great as would be expected from the more dramatic difference in the relative frequency of thermolabile adenine adducts. Both chlorambucil and mechlorethamine, like melphalan, induced few mutations at GNC sequences, which are potential sites of interstrand cross-links (Ojwang et al., 1989).

Discussion

While it has been known for some time that nonaromatic nitrogen mustards induce predominantly guanine N-7 adducts (Hemminki & Kallama, 1986), very little work has been done on adducts induced by the aromatic analogues, which have been assumed to act similarly. The possibility of biologically significant alkylation of adenine residues by these agents was raised by the specific termination of transcription at adenine residues in melphalan-treated templates (Pieper et al., 1989), and by the predominance of A·T → T·A transversions in the mutational spectrum of melphalan-treated pZ189 replicated in human cells (Wang et al., 1990). In both cases, it was proposed that the lesions responsible were probably adenine N-3 adducts. It was subsequently shown that thermolabile adenine lesions are major products of DNA treatment with melphalan or chlorambucil; the fact that formation of these lesions can be blocked by distamycin, a minor groove binder, strongly suggests that they are adenine N-3 adducts. Aromatic nitrogen mustards, which induced much greater adenine alkylation than nonaromatic analogues, may initially bind noncovalently in a manner similar to that of distamycin and netropsin (though with much lower affinity), with the aromatic ring inserted edgewise into the minor groove and the flexible side chains conforming to the DNA curvature (Kopka et al., 1985). Whatever the molecular details of binding, thermolabile adduct formation clearly depends on interactions of the aromatic moieties with the minor groove, since few such adducts are formed in denatured DNA (Figure 4). A nitrogen mustard–distamycin conjugate has recently been synthesized which alkylates adenine N-3 almost exclusively (Groggini et al., 1991).

Pieper and Erickson (1990), who examined only 3′-end-labeled DNA fragments, found that, for each of several GAA sequences, only the adenine nearest the labeled 3′ end was a prominent thermolabile site in melphalan- or chlorambucil-treated DNA; we observed a similar specificity in a 3′-end-labeled fragment from *supF*. On the basis of this result, as well as results of transcription termination studies, Pieper and Erickson proposed that most of the thermolabile adducts were intrastrand cross-links between adjacent adenines; when alkylated adenines were depurinated and cleaved, only one of the two cross-linked adenines, the one nearest the labeled 3′ end, would be detected as a cleavage site. If this proposal is correct, 5′-end-labeled DNA should give just the opposite pattern; i.e., the adenine nearest the 5′ end should show the most cleavage. However, at all five GAA sites examined (bases 135–137 and 164–166, as shown in Figure 2, and bases 172–170, 163–161, and 155–153 in the transcribed strand, not shown) our studies with 5′-end-labeled fragments consistently indicated much more cleavage at the adenine nearest the 3′ end. These results are inconsistent with a predominance of intrastrand adenine–adenine cross-links and instead suggest that the 5′ adenine is simply less susceptible to alkylation. Most likely, the N-3 position of the first adenine in GAA sequences is shielded from alkylation by the steric hindrance of the N^2 amino group of the preceding guanine. It may be noted that the data do not exclude the possibility of cross-links between adenine N-3 and other potential minor groove alkylation sites such as O^2 of thymine or cytosine; however, the transcription termination results argue against significant involvement of the N-3 lesions in interstrand cross-links (Pieper & Erickson, 1990).

Although the data are not conclusive, comparisons of adduct and mutation spectra generally support a role for adenine N-3 adducts in the transversions induced by aromatic nitrogen mustards at A·T base pairs in pZ189. All hotspots for A·T → T·A transversions in *supF* were relatively frequent sites for formation of thermolabile adenine lesions, and there was a significant correlation between adduct sites and mutation sites. Three of the four A·T base substitution hotspots were absent in the spectrum of the nonaromatic analogue mechlorethamine, a drug that induces markedly less adenine alkylation than the aromatic drugs at these same sites. The persistence of the A·T → T·A transversion hotspot at base pair 161 in the spectrum of mechlorethamine-induced mutations is not readily explained on the basis of adduct frequency. However, it is clear from all the spectra that sequence-dependent factors other than adduct frequency, such as efficiency of repair or of replicative bypass, account for the majority of the variation of mutation frequency at different sites. In the case of mechlorethamine, base 161 may simply represent an extreme case of an infrequent adduct with a very high probability of replicative bypass.

The idea that a specific type of adduct such as adenine N-3 accounts for the majority of melphalan- and chlorambucil-induced mutagenesis is attractive because it raises the possibility that the cytotoxic and mutagenic effects of these drugs may be separable. For example, a drug that was sterically

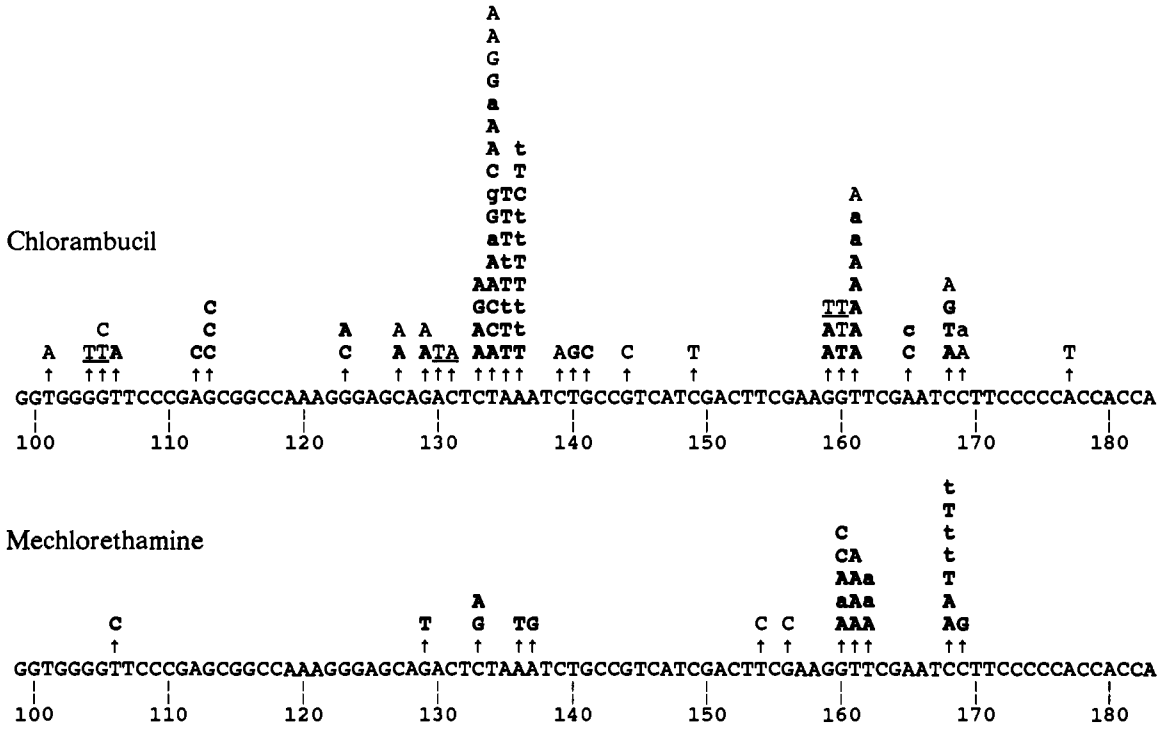


FIGURE 7: Spectra of base substitutions induced by chlorambucil or mechlorethamine in pZ189. Boldface letters indicate single-base substitutions, while lightface letters indicate those that occurred as part of a double (or triple) substitution. Underlines indicate tandem double substitutions. Lower-case letters indicate multiple identical mutants recovered from the same mammalian cell culture; i.e., only the first such mutant from each culture was assigned a capital letter. All mutants were taken from cultures with mutation frequencies between 0.25% and 0.7%. Not shown in the chlorambucil spectrum are second substitutions lying outside the tRNA sequence, including three A·T → T·A transversions and one A·T → G·C transition.

incapable of minor groove binding might induce far fewer mutagenic adenine N-3 alkylations, yet still efficiently induce guanine N-7 alkylations (including interstrand cross-links), which may be the lesions responsible for cytotoxicity. If so, such a drug might conceivably be less prone to produce second malignancies. However, the significance of adenine N-3 adducts and A·T → T·A transversions in carcinogenesis is far from clear. Base substitution (*ras*) mutations have been detected only sporadically in putative secondary leukemias, and most of the substitutions have been found at G·C base pairs (Greenberger, 1989; Bartram et al., 1989). Base substitution mutations could also result in inactivation of tumor suppressor genes, but although the frequent loss of the q arm of chromosome 5 in secondary leukemias (Henne & Schmähl, 1985) raises the possibility of a tumor suppressor gene at this locus, no candidate gene has yet been identified. Furthermore, preliminary data on mechlorethamine-induced mutations at the hamster HGPRT locus indicate that most of them are not base substitutions but deletions (Henner et al., 1991). Although base substitutions do appear to predominate among melphalan-induced mutations at the much smaller APRT locus, only a small fraction of these were A·T → T·A or A·T → C·G transversions; again, most mutations were at G·C base pairs (Austin et al., 1991). Substitutions at G·C base pairs also dominated the spectrum of mutations induced by mechlorethamine in a yeast shuttle vector system (Kunz & Mis, 1989).

In short, the nitrogen mustards appear to be very complex mutagens, capable of inducing a variety of base substitutions as well as deletions. Individual mutation spectra generated in various biological systems show wide divergence, for reasons that are not readily apparent. Given this complexity, it seems clear that elucidation of the role of specific drug-induced DNA alterations in the etiology of secondary malignancies, which is a necessary prerequisite to the rational design of potentially less carcinogenic analogues, will be no easy task.

References

Austin, M. J. F., Han, Y.-H., & Povirk, L. F. (1991) *Proc. Am. Assoc. Cancer Res. 32*, 101.

Bartram, C. R., Ludwig, W. D., Hiddemann, W., Lyons, J., Buschle, M., Ritter, J., Harbott, J., Frohlich, A., & Janssen, J. W. (1989) *Leukemia 3*, 247.

Greenberger, J. S. (1989) *Intl. J. Cell Cloning 7*, 343.

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

Hemminki, K. (1985) *Cancer Res. 45*, 4237.

Hemminki, K., & Kallama, S. (1986) in *Carcinogenicity of alkylating cytostatic drugs* (Schmähl, D., & Kaldor, J. M., Eds.) pp 55–70, International Agency for Research on Cancer, Lyon, France.

Hemsley, A., Arnheim, N., Toney, M. D., Cortopassi, G., & Galas, D. J. (1989) *Nucleic Acids Res. 17*, 6545.

Henne, T., & Schmähl, D. (1985) *Cancer Treat. Rev. 12*, 77.

Henner, W. D., Sanderson, B. J. S., Johnson, K. J., & Walker, K. K. (1991) *Proc. Am. Assoc. Cancer Res. 32*, 104.

Kopka, M. L., Yoon, C., Goodsell, D., Pjura, P., & Dickerson, R. E. (1985) *J. Mol. Biol. 183*, 553.

Kunz, B. A., & Mis, J. R. (1989) *Cancer Res. 15*, 279.

Mattes, W. B., Hartley, J. A., & Kohn, K. W. (1986a) *Nucleic Acids Res. 14*, 2971.

Mattes, W. B., Hartley, J. A., & Kohn, K. W. (1986b) *Biochim. Biophys. Acta 868*, 71.

Ojwang, J. O., Grueneberg, D. A., & Loechler, E. L. (1989) *Cancer Res. 49*, 6529.

Pieper, R. O., & Erickson, L. C. (1990) *Carcinogenesis 11*, 1739.

Pieper, R. O., Futscher, B. W., & Erickson, L. C. (1989) *Carcinogenesis 10*, 1307.

Povirk, L. F., & Goldberg, I. H. (1985) *Proc. Natl. Acad. Sci.*

U.S.A. 82, 3182.
Povirk, L. F., Han, Y.-H., & Steighner, R. J. (1989) Biochemistry 28, 8508.
Singer, B. (1975) Prog. Nucleic Acids Res. Mol. Biol. 15, 219.
Strauss, B., Rabkin, S., Sagher, D., & Moore, P. (1982) Biochimie 64, 829.
Wang, P., Bennett, R. A. O., & Povirk, L. F. (1990) Cancer Res. 50, 7527.
Zimmer, C., & Wahnert, U. (1986) Prog. Biophys. Mol. Biol. 47, 31.

Structural and Thermodynamic Consequences of Burying a Charged Residue within the Hydrophobic Core of T4 Lysozyme†,‡

S. Dao-pin,§ D. E. Anderson, W. A. Baase, F. W. Dahlquist, and B. W. Matthews*

Institute of Molecular Biology, Howard Hughes Medical Institute, and Departments of Physics and Chemistry, University of Oregon, Eugene, Oregon 97403

Received July 22, 1991; Revised Manuscript Received September 16, 1991

ABSTRACT: To determine the energetic and structural consequences of placing a charged group within the core of a protein, two "buried charge" mutants, Met 102 → Lys (M102K) and Leu 133 → Asp (L133D) were constructed in phage T4 lysozyme. Both proteins fold at neutral pH, although they are substantially less stable than wild type. The activity of M102K is about 35% that of wild type, while that of L133D is about 4%. M102K could be crystallized, and its structure was determined at high resolution. The crystal structure (at pH 6.8) of the mutant is very similar to that of wild type except for the α-helix that includes residues 108–113. In wild-type lysozyme, one side of this helix is exposed to solvent and the other contacts Met 102. In the M102K structure this α-helix becomes much more mobile, possibly allowing partial access of Lys 102 to solvent. The stability of M102K, determined by monitoring the unfolding of the protein with CD, is pH-dependent, consistent with the charged form of the substituted amino acid being more destabilizing than the uncharged form. The pK_a of Lys 102 was estimated to be 6.5 both by differential titration and also by NMR analysis of isotopically labeled protein with ^{13}C incorporated at the C^{ϵ} position of all lysines. As the pH is lowered below pH 6.5, the overall three-dimensional structure of M102K at room temperature appears to be maintained to pH 3 or so, although there is evidence for some structural adjustment possibly allowing solvent accessibility to the protonated form of Lys 102.

Charged residues in protein molecules are usually located on the surface (Rashin & Honig, 1984; Wada et al., 1985). When such residues are mobile and solvent-exposed, they appear to contribute very little to protein stability (Dao-pin et al., 1991a). Charged residues that are buried or largely buried are often found to be catalytically or functionally important [e.g., Gelin and Karplus (1977), Kraut (1977), Perutz (1978), Baldwin and Chothia (1979), and Kossiakoff (1983)]. When a charged amino acid is buried, it is usually paired with a residue or ion of the opposite charge and, in addition, has most if not all of its hydrogen-bonding potential satisfied (Rashin & Honig, 1984). This suggests that the burial of an isolated charge in the nonpolar protein interior would be energetically costly. If so, can a protein tolerate such a replacement and still fold? The description of the dielectric environment inside a protein molecule also remains problematic. Different theoretical calculations (Gilson et al., 1985, 1987; Klapper et al., 1986; Dao-pin et al., 1989; Gilson & Honig, 1987; Sternberg et al., 1987) have used a dielectric constant ranging from 2 to 8 and have predicted that it would cost 10–40 kcal/mol to bury a charge in a protein interior (Gilson et al., 1985).

To experimentally probe the nature of the dielectric constant, to measure the energetics of burying a charge, and to observe the tolerance of a protein molecule toward a potentially catastrophic amino acid replacement, two "buried charge" mutations, Met 102 → Lys (M102K) and Leu 133 → Asp (L133D), have been constructed in phage T4 lysozyme. Both sites of substitution are buried in the native protein structure, as shown in Figure 1. The thermal stabilities of both M102K and L133D have been measured by CD-monitored thermal denaturation. Attempts were also made to determine the pK_a values of the substituted residues by differential titration and by NMR. The structure of one of the mutants, M102K, has been determined by X-ray crystallography.

Materials and Methods

(1) Choice and Construction of Mutants. The two mutations M102K and L133D were chosen as representative replacements for two principal reasons. (a) The mutation sites are located near the center of the C-terminal lobe of T4 lysozyme. This is the largest hydrophobic region in the molecule. Residue 102 is located in a buried α-helix which extends from residue 93 to 106. Residue 133 is located within an α-helix that includes residues 127–136 and is exposed to solvent on one side and buried on the other. Both the main-chain and

† This work was supported in part by grants from the NIH (GM21967 to B.W.M. and 1S10RR04926 to F.W.D.), the NSF (DMB8905322 and DIR8815047 to F.W.D.), and the Lucille P. Markey Charitable Trust.

‡ The coordinates of the refined structure of M102K have been deposited in the Brookhaven Protein Data Bank (Ref. 1L54).

* To whom correspondence should be addressed.

§ Present address: Laboratory of Molecular Biology, NIDDKD, National Institutes of Health, Bethesda, MD 20892.

0006-2960/91/0430-11521$02.50/0