S-(2-Chloroethyl)glutathione-generated p53 Mutation Spectra Are Influenced by Differential Repair Rates More than Sites of Initial DNA Damage*

Received for publication, November 11, 2003, and in revised form, January 5, 2004 Published, JBC Papers in Press, January 7, 2004, DOI 10.1074/jbc.M312358200

J. Gerardo Valadez‡§¶ and F. Peter Guengerich¶¶

From the ¶Department of Biochemistry and Center in Molecular Toxicology, Vanderbilt University School of Medicine, Nashville, Tennessee 37232-0146 and §Department of Biotecnología y Bioingeniería, Centro de Investigación y de Estudios Avanzados del IPN, México DF 07360, Mexico

Several steps occur between the reaction of a chemical with DNA and a mutation, and each may influence the resulting mutation spectrum, i.e. nucleotides at which the mutations occur. The half-mustard S-(2-bromoethyl)glutathione is the reactive conjugate implicated in ethylene dibromide-induced mutagenesis attributed to the glutathione-dependent pathway. A human p53-driven Ade reporter system in yeast was used to study the factors involved in producing mutations. The synthetic analog S-(2-chloroethyl)glutathione was used to produce DNA damage; the damage to the p53 exons was analyzed using a new fluorescence-based modification of ligation-mediated polymerase chain reaction and an automated sequencer. The mutation spectrum was strongly dominated by the G to A transition seen in other organisms with S-(2-chloroethyl)glutathione or ethylene dibromide. The mutation spectrum clearly differed from the spontaneous spectrum or that derived from N-ethyl,N-nitrosourea. Distinct differences were seen between patterns of modification of p53 DNA exposed to the mutagen in vitro versus in vivo. In the four p53 exons in which mutants were analyzed, the major sites of mutation matched the sites with long half-lives of repair much better than the sites of initial damage. However, not all slowly repaired sites yielded mutations in part because of the lack of effect of mutations on phenotype. We conclude that the rate of DNA repair at individual nucleotides is a major factor in influencing the mutation spectra in this system. The results are consistent with a role of N7-guanyl adducts in mutagenesis.

The somatic theory of cancer holds that carcinogens are mutagens, at least with regard to what are classically termed tumor initiators (1–3). In a general model, cells are initiated by damage resulting from a DNA-alkylating agent, yielding mutations that are fixed by subsequent rounds of replication. Mutation spectra are generally observed when chemical damage to the DNA occurs; i.e. the sites of the mutation are not random and are often characteristic of the DNA-damaging agent (4, 5). Sometimes mutant spectra can be measured in tumors (in experimental animals or humans), although the existence of any “hotspots” for mutations does not necessarily indicate that the particular mutation is the cause of the tumor. The biochemical basis for where mutations occur (i.e. the mutant spectrum) is generally not well understood. Mutation resulting from chemical damage to DNA is a multistep process, and the observed mutant spectrum may be the result of events occurring during five or possibly more individual processes: (i) reaction of the chemical with DNA, (ii) further non-enzymatic reactions of the DNA adduct (e.g. opening of the imidazole ring in the case of guanyl-N7 alkyl adducts), (iii) DNA repair (or, more specifically, resistance to repair), (iv) the action of DNA polymerases, and (v) the biological effect of a particular nucleotide/amino acid change in the phenotypic assay used.

Ethylene dibromide (1,2-dibromoethane, BrCH2CH2Br) was extensively used as an insecticide, fumigant, and anti-knock agent in leaded gasoline prior to its classification as a carcinogen in the 1970s (6–8). The chemical is mutagenic in a number of bacterial and eukaryotic systems, producing primarily GC to AT transitions (9–12). Ethylene dibromide appears not to be DNA-reactive itself; enzymatic conjugation with GSH yields a half-mustard (GSCH2CH2Br) that generates an episulfonium ion that reacts with DNA (13, 14) and leads to mutations (15). Recent work has demonstrated that similar episulfonium ion chemistry appears to be involved in the reaction of the active site cysteine of the DNA repair protein O6-alkylguanine-DNA alkyltransferase with ethylene dibromide (16).

The GSH half-mustard (GSCH2CH2Cl) forms primarily S-(2-(N7-guanyl)ethyl)GSH adducts (17, 18) and also the minor adducts S-[2-(N2-guanyl)ethyl]GSH, S-[2-(O6-guanyl)ethyl]GSH, and S-[2-(N7-adenyl)ethyl]GSH (12, 19). This reactive half-mustard can be synthesized, under anhydrous conditions, and used as a model for GSCH2CH2Br, with similar results in terms of DNA alkylation and mutagenesis (15). This mustard was used in some previous work on sequence selectivity with the lacZ gene in a bacteriophage M13mp18/Salmonella typhi- murium system. The patterns of both DNA alkylation and mutation were determined in the lacZ sequence. Some similarity between the patterns of highest alkylation and mutation was seen, although the overlap was not particularly strong (12).

These results with the GSH half-mustard can be considered in context of the general question of the role of DNA adduct profiles and their influence on mutation spectra (20). Studies with other physical and chemical agents have been done, par-
p53 Mutant Spectra and DNA Repair

ticularly with UV light and PAHs1 in p53, ras, and hprt models (21–25). In the case of damage by UV light, a role for differential rates of DNA repair along the gene has been concluded to be a dominant factor in explaining the mutation spectra that appear in human tumors (26–28). With FAL1, the initial sites of DNA modification seem more important and are explained at least partly by C-5 methylation patterns at CpG sites in mammalian systems (29–31).

We considered some of the issues of the mutation spectra of ethylene dihydroxime using the half-mustard GSCH2CH2Cl in Saccharomyces cerevisiae with a plasmid containing the human tumor suppressor gene p53 and an Adr reporter responsive to the transactivation activity of p53 (32–34). This yeast system has been used previously with several physical and chemical agents (35–37). We developed a variation of an LMPCR method to identify and quantify sites of DNA damage. This mutagen has some advantages, in the context of general questions about sequence selectivity, in that DNA alkylation by GSCH2CH2Cl is largely (but not exclusively) restricted to guanine N7 atoms, in contrast to UV- and PAH-derived DNA modifications.

We found differences in the damage patterns for isolated DNA and the same DNA harbored inside of yeast cells. The FLMPCR method also allowed the quantitation of piperidine-labile sites in DNA in cells as a function of time. Our results provide evidence that the mutation sites are highly localized to regions of persistent DNA adducts (i.e., regions of slow repair).

**EXPERIMENTAL PROCEDURES**

*In Vivo Exposure to S-(2-Chloroethyl)glutathione or ENU*—The yDM56p strain uses a double selection for mutant p53 transactivation activity. Both Can1 and Ade2 gene expression are regulated by a p53-responsive promoter. Cells with wild type p53 express the Can1 (arginine permease) gene, resulting in the uptake of canavanine, which is toxic to the cells and limits growth on selective media. Cells with mutant p53 that cannot drive expression of Ade2 accumulate an adenine metabolic intermediate and are red in color. *S. cerevisiae* yDM56p strain was grown to mid-log phase (A600 = 0.3, 3 × 108 cells/ml) in 250 ml of high adenine complete minimal medium as described (36). Cells were washed three times with 30 ml of PBS buffer and resuspended in 4 ml of PBS buffer (2 × 108 cells/ml). ENU (Sigma) was added to a final concentration of 100 μM as described (38). S-2-Chloroethyl(GSH) (GSCH2CH2Cl), synthesized as described (15), was freshly dissolved in chilled 0.10 M Tris-HCl buffer (pH 7.7, containing 1 mM EDTA) to a final concentration of 1 M and immediately added to yeast cells to a final concentration of 120 mM. In both cases, incubation was at 30 °C and terminated by filtration through CoolStop filters, elution buffer, and resuspended in 4 ml of PBS buffer (2 × 108 cells/ml) in the presence of 2 mM dimethylsulfate (Aldrich). The cells were run for 10 min at 25 °C. Piperidine (Fisher Scientific) was added to give a final concentration of 1 M, as well as 100 μM of Nujol mineral oil (PerkinElmer Life Sciences) and incubated for 30 min at 90 °C (39–41). After piperidine treatment, DNA fragments (a collection of DNA molecules of variable size) were purified using the Roche kit (see above). These fragments served as templates for FLMPCR.

**FLMPCR**—A general scheme of the FLMPCR technique is shown in Fig. 1. Piperidine-treated DNA fragments were used as templates in a first primer extension assay. P3 and P4 primers were employed in separate reactions for adduct detection on the transcribed and on the non-transcribed strands of p53, respectively. This procedure generated a family of variable sized, blunt-ended duplex molecules, which were substrates for T4 DNA ligase-catalyzed addition of an asymmetric linker (42). This linker provided each DNA fragment with a common defined end. The bulk of the DNA was then denatured and used as template in a second primer extension reaction with the same oligonucleotides (P3 or P4), so that the extension products included the sequence complementary to the added linker. After each member of the bulk DNA pool had two defined ends, the mixture was used as templates in an exponential PCR reaction (30 cycles) using as primers the largest oligonucleotide of the common linker, and one of four different internal (LA) primers, 5’ end-labeled with 6-FAM (Operon). Primers LA1 (5’-CCCGAGAAAACCTACCCAGGGAG/CCTACCGG) and LA2 (5’-CGGATGGAAGGGAAATTGTGCTGGAG) were used to localize the adducts on the transcribed strand of p53; primers LA3 (5’-GCCGGC- CACTGCGAACGCTTACAGCAT) and LA4 (5’-CTATCTGCGTCT CCGAACATCCCGGACG) were used for adduct detection on the non-transcribed strand. After every step of the FLMPCR assay, the DNA fragments generated were purified using the Roche kit (see above). Finally, an aliquot (1%) of the exponential PCR reactions was mixed with appropriate amounts of loading buffer, deionized formamide, and ROX-350-labeled size markers (PerkinElmer Life Sciences), loaded onto an 4% (w/v) acrylamide sequencing gel, and analyzed using an ABI Prism 377 DNA sequencer (Applied Biosystems, Foster City, CA) in the Vanderbilt facility.

**Adduct Quantitation**—The GeneScan software of the ABI Prism 377 DNA sequencer generated an image of the fluorescent bands in the acrylamide gel. The analysis of this image resulted in a profile (electropherogram) of peaks (the DNA fragments labeled with 6-FAM) and size markers (labeled with ROX-350), and the corresponding tabular display of data indicated the size (in base pairs) and the height (in arbitrary units) of each peak. The position of the peaks relative to the markers in these profiles revealed the nucleotide position where each GSCH2CH2Cl adduct was formed. Furthermore, by calculating the ratio between the heights of a given peak in the GSCH2CH2Cl-treated sample and its equivalent in the negative control, it was possible to establish the relative adduction level of every nucleotide along both p53 strands (exons 5–8).

**Adduct t1/2**—After *in vivo* exposure of yDM56p yeast cells to 200 μM GSCH2CH2Cl, washed cells were either processed immediately (time 0) or incubated at 37 °C in 10 ml of YPD broth (43). Cells were collected at 0, 3, 6, 12, and 24 h and processed for adduct quantification by FLMPCR. From the adduction level data measured at each time point, a regression curve was established for each G nucleotide in exons 5–8 of p53. Adduct t1/2 (the time at which 50% of the initial damage was removed) was then determined from a semi-logarithmic plot, i.e. log10 [adduct (versus t)] was determined for every damaged G in both strands in exons 5–8.

**RESULTS**

*Mutagenesis with Alkylating Agents*—Experiments were done with *S. cerevisiae* yDM56p to determine the mutation frequency in the p53 gene induced by two alkylating agents, ENU and GSCH2CH2Cl. Cells were exposed to 100 μM ENU or 120 μM GSCH2CH2Cl as described under “Experimental Procedures.” Survival in nonselective medium at these doses rendered similar results (~17% survival with respect to non-exposed, control cells). Exposure of yDM56p yeast to ENU

---

1 The abbreviations used are: PAH, polycyclic aromatic hydrocarbon; ENU, N-ethyl-N-nitrosourea; PBS, phosphate-buffered saline (15 mM potassium phosphate (pH 7.4) containing 150 mM NaCl); LMPCR, ligation-mediated PCR; FLMPCR, fluorescence ligation-mediated PCR; 6-FAM, 6-carboxyfluorescein; ROX-350, 5-carboxy-X-rhodamine.
produced a 20-fold increase in the induced mutant fraction. However, the mutant fraction resulting from GSCH2CH2Cl treatment was only 2.3-fold higher than the spontaneous fraction (Table I). This increase was less than the 10-fold level previously observed in the M13mp18 bacteriophage study (12), where the bacterial cells were deficient in nucleotide excision repair. It has been reported that a combined treatment with a given cytotoxic agent and the canavanine present in the selective medium kills a larger proportion of cells than the cytotoxic agent alone (36). The present results, therefore, are probably an underestimate of induced mutant fractions.

Mutant Characterization—To determine the proportion of red colonies that actually contain mutations in p53, the DNA from selected red colonies was PCR-amplified and checked in the functional assay (36). PCR products that rendered red colonies were considered p53 mutants, whereas those that produced white colonies were presumed to contain mutations outside the p53 region (e.g. in the Ade2 reporter gene). All PCR products from red colonies were analyzed by agarose gel electrophoresis to detect the presence of large insertions and deletions. Two out of 10 (20%) spontaneous red colonies were found to contain large insertions (20 and 70 bp long, determined by further sequencing; Table II). In contrast, only 2 of 32 (6%) ENU-induced clones had large deletions (Table II), and only 1 of 45 (2%) GSCH2CH2Cl-induced red colonies contained a deletion (13 bp, determined by nucleotide sequencing, Table III).

Results of the p53 functional assay are presented in Table I. Of 23 spontaneous red colonies tested, 17 (74%) contained mutant p53, and the sequencing of exons 5–8 revealed that 10 of them (59%) harbored mutations in this region of the gene. Among ENU-induced colonies, 51 of 71 (72%) contained mutant p53 as judged by the functional assay, and 34 (67%) of these had detectable mutations in the DNA binding domain of the gene. Of the GSCH2CH2Cl-induced red colonies, 95 of 133 (71%) were mutants in p53, but only 46 of these 95 (48%) harbored mutations in the central domain of the gene. Spontaneous and ENU-induced mutations found in the sequenced region of the gene are listed in Table II and the GSCH2CH2Cl-induced mutations in Table III. In the spontaneous mutants (in addition to the two large insertions), there were two 2-bp deletions, three transitions, and three transversions. In the ENU- and GSCH2CH2Cl-induced mutants, the vast majority of p53 mutations (78 and 87%, respectively) were G to A transitions. Mutations were observed in a similar number of codons after exposure to ENU and GSCH2CH2Cl (25 and 24, respectively) in the region between the 3′ end of exon 4 and the 5′ end of exon 9. ENU and GSCH2CH2Cl both produced mutations preferentially in purine-rich sites (see Tables II and III). The determined consensus sequence around the mutated base for ENU was (5′) G32G45G52G54G61G62G65 whereas that for GSCH2CH2Cl was (5′) G35G36G40G52G54G56G59 (where the indicated value in each position is the percentage of cases where that nucleotide appears in the mutants and the mutated base is in the central position (in bold font)). The distribution of mutations induced by the alkylating agents tested was not random. Some codons were more frequently changed than others. Furthermore, each treatment produced a distinctive and statistically different fingerprint of hotspots (p values < 0.05 calculated with the computer program designed by Cariello and co-workers (Refs. 44 and 45)). After ENU-treatment, codons 141, 177, 194, 199, and 278 were found mutated more than once (Table II). Codons 152, 177, 244, and 279 were frequent targets for mutation by GSCH2CH2Cl (Table III). Taken together, the mutations detected were symmetrically distributed between the transcribed and non-transcribed strands (~50% in each strand for both chemicals tested). However, when individual exons were analyzed, it was evident that GSCH2CH2Cl produced mutations almost exclusively in the transcribed strand of exon 5 and in the non-transcribed strands of exons 7 and 8 (Table III). ENU treatment rendered rather diverse results; all mutations in exon 7 were in the non-transcribed strand, and mutations in other exons were rather equally distributed (Table II).

To our knowledge, the p53 mutations present in clones SPO2, SPO5, ENU1, ENU13, and ENU54 and in GSCH2CH2Cl clones 9, 43, 49, and 91 have not been reported before. They were not included in the list of 14,968 entries available on a web site collection (p53.curie.fr).

Analysis of GSCH2CH2Cl-produced DNA Adducts by LMPCR Assays—Among the different types of DNA adducts produced by GSCH2CH2Cl, the most abundant (>95%) is S-[2-(N′-guanyl)ethyl]GSH (12). Thus, GSCH2CH2Cl mimics the action of dimethylsulfate, a compound that methylates (among other positions) the N7 atom of guanine in DNA. Hot piperidine removes N7-alkylguanines from the sugar and catalyzes β-elimination of phosphates from the empty sugars to cleave the DNA (39, 40).

The strand breaks and the free 5′-phosphate termini generated by the combined action of GSCH2CH2Cl and piperidine render treated DNA a suitable substrate for analysis using LMPCR. Modifications to this technique allowed the mapping of DNA damage induced by GSCH2CH2Cl in the central domain (exons 5–8) of human p53 at nucleotide resolution. LMPCR assays have been successfully used to detect the presence of adducts in several DNA sequences at nucleotide resolution (46–48). The traditional assay is limited to a relatively short stretch of sequence (usually <200 bp) (48). Another inconvenience is the need for several (usually 4) oligonucleotide primers to analyze a genomic region of some dozens of nucleotides. Third, the common protocol requires hybridization of exponentially amplified products to radiolabeled probes to reveal the adducted nucleotides. Finally, the quantitative analysis of adducted bases requires the use of autoradiography/densitometry. We modified the LMPCR protocol, using a total of seven primers (four of them 6-FAM-labeled) and the analysis of fluorescence-amplified PCR products performed with the help of GeneScan™ software (ABI Prism Genetic Analysis instrument, PerkinElmer Life Sciences). With these modifications we analyzed in vitro and in vivo GSCH2CH2Cl-induced adducts on both strands of the complete DNA binding domain (exons 5–8, total of 1,088 bp) of the human p53 gene. We term this technique FLMPCR (Fig. 1). The complete procedure can be completed in 2 days and renders reproducible results on the localization and (relative) quantification of nucleotide adducts.

Comparison of in Vico and in Vitro GSCH2CH2Cl-induced Damage Distribution in the p53 Gene—We used concentrations of GSCH2CH2Cl similar to those used to generate the mutants
Table II
Sequence alterations in spontaneous and ENU-induced p53 mutants

| Clone   | Mutation | Sequence contexta | Strandb | Codon | Exon | Amino acid change |
|---------|----------|-------------------|---------|-------|------|------------------|
| SPO5   | G to T   | GTGCA             | T       | 124   | 4    | Cys to Stop      |
| SPO6   | 70 bp I  | GG[T][AT]         | NT      | 124–148| 4 and 5 |                  |
| SPO21  | G to T   | TGGTA             | T       | 163   | 5    | Tyr to Stop      |
| SPO47  | 2 bp Δ   | C[A][AG]          | T       | 214   | 6    | Frameshift       |
| SPO2ε  | G to A   | TAGAT             | T       | 264   | 8    | Leu to Leu       |
| SPO3   | 20 bp I  | GG[T][AG]         | NT      | 279–286| 8    |                  |
| SPO15  | 2 bp Δ   | C[A][AG]          | T       | 285   | 8    | Frameshift       |
| SPO2ε  | G to A   | CCGAG             | T       | 289   | 8    | Leu to Leu       |
| SPO10  | G to A   | GTGAG             | T       | 296   | 8    | His to Tyr       |
| SPO19  | G to C   | GGGCA             | T       | 309   | 9    | Pro to Ala       |

a The sequence is presented 5’ to 3’ for the strand where the mutated base is a purine. Mutated bases are underlined. Insertions (I) and deletions (Δ) are indicated in brackets in the non-transcribed strand.
b The sequence context for the mutated base is shown for the transcribed (T) or non-transcribed (NT) strand.
c SPO1 Clone SPO2 has two point mutations each.
d Deletion was determined by agarose gel electrophoresis.

As indicated above, to have patterns of alkylation that could be compared with the mutant spectra. From our previous (12, 15) and current experience, the level of DNA damage is such that the average exon is damaged no more than once, avoiding a contributing fraction of short fragments resulting from multiple cleavages. Subsequent work (see below) supports the lack of multiple cleavage, in that we did not observe any consistent trends of greater apparent damage at one end of exons (Fig. 2). Examples of sites preferentially modified following in vitro treatment include nucleotides located in the transcribed strand at codons 127 and 175–178 (exon 5), 188–193 (exon 6), 243, 248, and 250–256 (exon 7), and 278, 297, and 300 (exon 8). In the non-transcribed strand, nucleotides at codons 147, 158, 160, 171, 174, and 180 (exon 5), 188, 199, and 206 (exon 6), and 265 and 287 (exon 8) were also more affected by the mutagen to the target DNA. All four bases sustained some alkylated A and G has also been reported (Refs. 49 and 50)).

(see above) and depurination (Ref. 40); limited cleavage at bases other than the cellular environment influences the accessibility of the mutagen to the target DNA. All four bases sustained some DNA damage (adenine could be a N⁵-alkyl adduct, susceptible to depurination (Ref. 40); limited cleavage at bases other than alkylated A and G has also been reported (Refs. 49 and 50)). As indicated above, the damage distribution was non-ran-
TABLE III

Sequence alterations in GSCH₂CH₂Cl-induced p53 mutants

| GSCH₂CH₂Cl clone | Mutation | Sequence context | Strand | Codon | Exon | Amino acid change |
|------------------|----------|-----------------|--------|-------|------|------------------|
| 8                | G to A   | GGGAG           | T      | 127   | 5    | Ser to Phe       |
| 129               | G to A   | AAGGC           | T      | 142   | 5    | Pro to Ser       |
| 88               | G to A   | TGCGT           | NT     | 146   | 5    | Trp to Stop      |
| 123               | G to A   | TTGGT           | NT     | 146   | 5    | Trp to Stop      |
| 102               | G to A   | GGGGG           | T      | 151   | 5    | Pro to Ser       |
| 2                | G to A   | CGGGG           | T      | 152   | 5    | Pro to Ser       |
| 19               | G to A   | CGGGG           | T      | 152   | 5    | Pro to Ser       |
| 101               | G to A   | CCGGG           | T      | 152   | 5    | Pro to Ser       |
| 67               | 13 bp Δ | GCΓGG           | T      | 152/6 |      |                  |
| 60               | G to A   | GCGCC           | T      | 175   | 5    | Arg to Cys       |
| 117               | G to A   | GCGCC           | T      | 175   | 5    | Arg to Cys       |
| 5                | G to A   | GGGGG           | T      | 177   | 5    | Pro to Leu       |
| 15               | G to A   | GGGGG           | T      | 177   | 5    | Pro to Ser       |
| 66               | del G    | TGGGG           | T      | 177   | 5    | Frameshift       |
| 71               | G to A   | GGGGG           | T      | 177   | 5    | Pro to Leu       |
| 89               | G to A   | GGGGG           | T      | 177   | 5    | Pro to Ser       |
| 116               | G to A   | GGGGG           | T      | 177   | 5    | Pro to Ser       |
| 32               | A A      | GGGAG           | T      | 190   | 6    | Frameshift       |
| 110               | G to A   | AAGGG           | T      | 190   | 6    | Pro to Ser       |
| 131               | G to A   | CTGAG           | T      | 192   | 6    | Gln to Stop      |
| 87               | G to A   | AGGAA           | NT     | 199   | 6    | Gly to Gln       |
| 124               | G to A   | GTGCA           | NT     | 238   | 7    | Cys to Tyr       |
| 4                | G to A   | GGGGG           | NT     | 244   | 7    | Gly to Asp       |
| 49               | G to A   | GGGGG           | NT     | 244   | 7    | Gly to Asp       |
| 94               | G to A   | TGGGG           | NT     | 244   | 7    | Gly to Ser       |
| 98               | G to A   | TGGGG           | NT     | 244   | 7    | Gly to Ser       |
| 105               | G to A   | TGGGG           | NT     | 244   | 7    | Gly to Ser       |
| 9               | G to CA  | ATGCC           | T      | 245   | 7    | Frameshift       |
| 18               | G to A   | CGGCA           | NT     | 245   | 7    | Gly to Asp       |
| 90               | G to A   | CGGCA           | NT     | 248   | 7    | Arg to Gln       |
| 103               | G to A   | ACCGG           | NT     | 259   | 7    | Asp to Asn       |
| 43               | Δ AAT    | GTΔXCT          |       | 263   | 8    | D (Asn)          |
| 47               | G to A   | ACCGG           | NT     | 267   | 8    | Arg to Gln       |
| 59               | G to A   | ACCGG           | NT     | 267   | 8    | Arg to Gln       |
| 118               | G to A   | CAGGA           | NT     | 278   | 8    | Pro to Leu       |
| 24               | G to A   | TGCGT           | NT     | 279   | 8    | Gly to Gln       |
| 50               | G to A   | TGCGT           | NT     | 279   | 8    | Gly to Gln       |
| 86               | G to A   | TGCGT           | NT     | 279   | 8    | Gly to Gln       |
| 128               | G to A   | TGCGT           | NT     | 279   | 8    | Gly to Gln       |
| 132               | G to A   | TGCGT           | NT     | 279   | 8    | Gly to Gln       |
| 3                | G to A   | GAGAC           | NT     | 281   | 8    | Asp to Asn       |
| 16               | G to A   | GAGAC           | NT     | 281   | 8    | Asp to Asn       |
| 6                | G to A   | AGGAA           | NT     | 286   | 8    | Gln to Lys       |
| 91               | G to A   | TGGGG           | T      | 309   | 9    | Pro to Ala       |
| 49               | G to C   | GGGAG           | T      | 316   | 9    | Pro to Ala       |
| 120               | G to A   | CTGGG           | T      | 317   | 9    | Gln to Stop      |

*a The sequence is presented 5’ to 3’ for the strand where the mutated base is a guanine. Mutated bases are underlined. Insertions (I) and deletions (Δ) are indicated in brackets.

*b The sequence context for the mutated base is shown for the transcribed (T) or non-transcribed (NT) strand.

The GSCH₂CH₂Cl-induced Adduct t₁/₂ Varies among Nucleotides of p53—Inspection of the initial damage caused by in vitro exposure to GSCH₂CH₂Cl revealed that more nucleotides were adducted than those that resulted in mutations (Fig. 2), implying additional factors that affect mutation fixation after the initial damage has been produced. One factor could be the slower removal of adducts at certain nucleotides. To test this possibility, the time courses for GSCH₂CH₂Cl-DNA adduct repair were determined by FLMPCR assays after in vivo exposure to the mutagen. Results of the repair kinetics for GSCH₂CH₂Cl-adducted guanine nucleotides formed along both the transcribed and non-transcribed strands of p53 gene were quantified for exons 5–8 (Fig. 3). The time period required for removal of 50% of the GSCH₂CH₂Cl adduct was non-measurable, e.g. codons 226, 233, and 245 (transcribed strand); Fig. 3). In contrast, the t₁/₂ was ≥60 h for codons 130, 244, and 245 (non-transcribed strand) (Fig. 3). Not all slowly repaired sites yielded mutations. Examples are codons 128–130 and 295 (transcribed strand, exons 5 and 8, respectively), where the adduct half-lives for their guanine residues were >50 h. On the other hand, a few mutations were found on nucleotides where the t₁/₂ of the adduct was non-measurable, e.g. codon 175 (exon 5, transcribed...
Another factor is the structural influence of bound proteins, which may be a factor in the differences in modification patterns observed in vitro and in vivo (Fig. 2) (58). (ii) A second step in some cases is non-enzymatic conversion of a DNA adduct to another type of adduct, e.g. depuration of an N^7-guanyl adduct or conversion to an imidazole ring-opened form (59). However, at the present time there is no evidence that these or other potential transformations (e.g. oxidation of 8-oxo-7,8-dihydroguanine (Ref. 60)) show sequence specificity, although the probability is likely. The product could be either less or more mutagenic (or blocking) than the initial adduct. (iii) The next step considered is DNA repair. Evidence already exists for a role of repair rates in determining the mutation spectra in tumors resulting from exposure of mammalian cells to sunlight (and corresponding to mutations in p53 in human tumors) (27, 28, 61). Some details of how the individual DNA repair systems can be influenced by individual chemical adducts in certain sequences have been published (62, 63), but few generalizations can be made at this time about these systems. (iv) Polymerization is the next step (if the DNA is not repaired). Several possibilities can be considered, depending upon the chemical adduct, the particular polymerase(s) involved, and the DNA sequence. The position at which an adduct can be found can influence the blockage of a polymerase, the qualitative and quantitative nature of misincorporation opposite the adduct, and the tendency for the polymerase(s) to extend beyond the site of insertion (opposite the adduct) (64, 65). (v) The last step is the effect on biological activity beyond these steps, i.e. phenotype. The position of the mutation can have a dramatic effect on the phenotype; obviously, some changes will not influence the biological properties of a protein and some will be critical. Further, the phenotypic effects may differ between in vitro assay systems used in the laboratory and in tumors, because different functions may be involved.

Because the yeast functional assay identifies mutant p53-expressing clones by a red/white colony color selection based on loss of transactivation ability of an integrated Ade2 gene, it can be assumed that, excepting the clones with double mutations (e.g. clones SPO2, ENU5, ENU13, and GSCH2CH2Cl 49; see Tables II and III), all the mutations obtained with this system affect the transactivating activity of p53 protein. In previous work we compared sites of S-(2-chloroethyl)GSH alkylation of the lacZ gene in bacteriophage DNA with the mutation spectrum (in the bacterium S. typhimurium) (12). That study demonstrated a dominant pattern of G to A transitions, which had previously been suggested in an Escherichia coli assay with ethylene dibromide (9). The results (G to A transitions) argue against a strong influence of depurination, where the apurinic sites that would be expected from depurination of S-(2-(N^7-guanyl)ethyl)GSH adducts should produce G to T transversions (66). The predominance of G to A transitions has now been reported for studies with ethylene dibromide in several systems, including Drosophila and Chinese hamster ovary cells (9–11). In addition to S-(2-(N^7-guanyl)ethyl)GSH, some other G adducts have been shown to be produced from S-(2-chloroethyl)GSH, including the N^2- and O^6-guanyl adducts (12). All three of the known G adducts (N^2, N^6, O^6) have been shown to block replicative polymerases (67), although the G to A transitions have not been specifically attributed to a single G modification yet.

Patterns of DNA modification observed in vitro and in vivo were not identical (Fig. 2), presumably because of either the coating of the DNA or the presence of DNA-bound proteins. Cloutier et al. (58) have also reported differences in alkylation patterns inside and outside of cultured cells. Further, the initial pattern of DNA damage is not identical to that seen after

**Figure 1. Schematic representation of the FLMPCR assay.** S. cerevisiae yDM56p cells were exposed to 200 mM GSCH2CH2Cl (in vitro experiments), or the isolated p53 cDNA plasmid was incubated in the presence of 2 mM GSCH2CH2Cl (in vitro experiments). DNA adducts are formed all along the sequence (three solid circles depicted here only on one of the strands for simplification). a, hot piperidine treatment of adducted DNA produced strand breaks preferentially at N^7-alkylguanine sites, leaving 5’ and 3’ phosphates (40). b, piperidine-cleaved strands served as templates in a first primer extension reaction, using a primer external to the sequence to be analyzed. c, a common linker was added in a T4 DNA ligase-mediated reaction to define the sequences of both ends. d, a second primer extension reaction read through the longer oligonucleotide of the linker. e, the products were now suitable for exponential amplification using the longer oligonucleotide of the linker and an internal, fluorescence-labeled oligonucleotide (depicted with a star) as primers. f, the final products were resolved in a sequencing gel and analyzed by the GeneScan software of the ABI Prism sequencer.

**DISCUSSION**

The issue of the molecular basis of mutation spectra (and the relationship to tumor initiation and development) has been considered and reviewed by others (4, 5, 22–24, 28, 29, 37, 51). At least five major steps on the pathway from a carcinogen to a mutation may be involved. (i) One step is binding selectivity (of the activated carcinogen) (52). This could be the result of any of at least three factors. One is the influence of the nucleotide sequence on the chemical reactivity at a site, e.g. the enhanced nucleophilicity of a G within a run of G nucleotides (or, more properly, a G with a 5’ G neighbor (Refs. 41 and 53–56)). Another factor is the influence of the sequence on DNA structure, e.g. favoring intercalation or groove binding (57). Still another factor is the structural influence of bound proteins, which may be a factor in the differences in modification patterns observed in vitro and in vivo (Fig. 2) (58). (ii) A second step in some cases is non-enzymatic conversion of a DNA adduct to another type of adduct, e.g. depuration of an N^7-guanyl adduct or conversion to an imidazole ring-opened form (59). However, at the present time there is no evidence that these or other potential transformations (e.g. oxidation of 8-oxo-7,8-dihydroguanine (Ref. 60)) show sequence specificity, although the probability is likely. The product could be either less or more mutagenic (or blocking) than the initial adduct. (iii) The next step considered is DNA repair. Evidence already exists for a role of repair rates in determining the mutation spectra in tumors resulting from exposure of mammalian cells to sunlight (and corresponding to mutations in p53 in human tumors) (27, 28, 61). Some details of how the individual DNA repair systems can be influenced by individual chemical adducts in certain sequences have been published (62, 63), but few generalizations can be made at this time about these systems. (iv) Polymerization is the next step (if the DNA is not repaired). Several possibilities can be considered, depending upon the chemical adduct, the particular polymerase(s) involved, and the DNA sequence. The position at which an adduct can be found can influence the blockage of a polymerase, the qualitative and quantitative nature of misincorporation opposite the adduct, and the tendency for the polymerase(s) to extend beyond the site of insertion (opposite the adduct) (64, 65). (v) The last step is the effect on biological activity beyond these steps, i.e. phenotype. The position of the mutation can have a dramatic effect on the phenotype; obviously, some changes will not influence the biological properties of a protein and some will be critical. Further, the phenotypic effects may differ between in vitro assay systems used in the laboratory and in tumors, because different functions may be involved.

Because the yeast functional assay identifies mutant p53-expressing clones by a red/white colony color selection based on loss of transactivation ability of an integrated Ade2 gene, it can be assumed that, excepting the clones with double mutations (e.g. clones SPO2, ENU5, ENU13, and GSCH2CH2Cl 49; see Tables II and III), all the mutations obtained with this system affect the transactivating activity of p53 protein. In previous work we compared sites of S-(2-chloroethyl)GSH alkylation of the lacZ gene in bacteriophage DNA with the mutation spectrum (in the bacterium S. typhimurium) (12). That study demonstrated a dominant pattern of G to A transitions, which had previously been suggested in an Escherichia coli assay with ethylene dibromide (9). The results (G to A transitions) argue against a strong influence of depurination, where the apurinic sites that would be expected from depurination of S-(2-(N^7-guanyl)ethyl)GSH adducts should produce G to T transversions (66). The predominance of G to A transitions has now been reported for studies with ethylene dibromide in several systems, including Drosophila and Chinese hamster ovary cells (9–11). In addition to S-(2-(N^7-guanyl)ethyl)GSH, some other G adducts have been shown to be produced from S-(2-chloroethyl)GSH, including the N^2- and O^6-guanyl adducts (12). All three of the known G adducts (N^2, N^6, O^6) have been shown to block replicative polymerases (67), although the G to A transitions have not been specifically attributed to a single G modification yet.

Patterns of DNA modification observed in vitro and in vivo were not identical (Fig. 2), presumably because of either the coating of the DNA or the presence of DNA-bound proteins. Cloutier et al. (58) have also reported differences in alkylation patterns inside and outside of cultured cells. Further, the initial pattern of DNA damage is not identical to that seen after
several hours, because of the influence of variable half-lives of DNA repair at individual G sites (Figs. 4–7). The sites at which mutations were clustered generally corresponded to sites with slower repair. However, some slowly repaired (and phenotypically detectable) sites did not produce mutants with the phenotype under consideration in the assay. In the earlier work with the mutagen S-(2-chloroethyl)GSH in the bacteriophage lacZ complementation system, some overlap was noted between the patterns of piperidine-generated cleavage at Gs and mutations but was not complete (12). As in the current study, the overlap of the patterns had many exceptions, i.e. regions of mutation with low alkylation and regions of extensive alkylation but no mutations. These studies were done in an excision repair-deficient strain of bacteria (to enhance mutations); thus, adduct repair could be discounted as a reason for the lack of correlation between initial adduct sites and mutations. We hypothesized, at that time, that the sequence context of a mutation influences the ability of polymerases to extend beyond the adducted site to a “fix” a mutation in the gene. A similar explanation may contribute to the finding in this study of stably adducted sites that are not adducted and some weakly adducted sites that are mutagenic, although this hypothesis has not been addressed yet.

The method used here is focused on N7-guanyl adducts and is a variant of previously described LMPCR methods (46–48). The use of a fluorescence end label (FLMPCR) was superior to previous approaches with 32P-end-labeled DNA in that it permitted in vivo studies and direct comparisons of alkylation

---

**Fig. 2.** Comparison of *in vitro* and *in vivo* induced GSHCH2CH2Cl-specific initial damage on transcribed and non-transcribed strands along the p53 gene. Electropherograms were generated by the ABI Prism GeneScan software when exons 5–8 of p53 were assayed by FLMPCR after *in vivo* exposure of S. cerevisiae yDM56p cells to 200 mM GSHCH2CH2Cl or *in vitro* treatment of p53 cDNA with 2 mM GSHCH2CH2Cl, as described under “Experimental Procedures.” The positions of the size markers (GeneScan ROX-350, PerkinElmer Life Sciences) are indicated with arrows, and the traces correspond to the 6-FAM-labeled fragments of the indicated exon of p53. Underlined numbers indicate selected codons. A, exon 5; B, exon 6; C, exon 7; D, exon 8.
patterns inside the cells with isolated DNA (Fig. 2). The use of fluorescence methodology was also important in that it allowed the use of a commercial sequencer instead of incubations with radioactive oligonucleotides, conventional gel electrophoresis, and autoradiography. A similar method has been employed in which a near-IR fluorescence label was used (48); our 6-FAM-
labeled system was utilized with a conventional sequencer in a facility core sequencing laboratory.

The majority of the previous literature on the subject of sequence-related effects of DNA modification comes from work with UV damage (especially cyclobutane dimers and pyrimidine[6-4]pyrimidone photoproducts) and benzo[a]pyrene (dihydrodiol epoxide)-derived adducts (26–31, 61, 68–71). Comparisons of the patterns have been made with the sites of p53 mutations in tumors of people who have had exposure to these physical and chemical agents. (In the case of UV light, skin cancers are readily attributed to the agent; with lung cancers, exposure of smokers to benzo[a]pyrene is real but can be complicated by the plethora of other carcinogens.) In the case of UV light damage, the evidence suggests a strong role for DNA repair in influencing which adducts will go on to produce p53 mutations in tumors, as predicted from studies in human fibroblast cells (27). With benzo[a]pyrene-derived adducts, the situation is probably more complex, in part because of questions about the chemical origin of the lung tumors. With regard to benzo[a]pyrene (and its diol epoxides), the mutation spectra appear to be dominated by the sequence selectivity of DNA adduct formation, which is linked to the presence of 5-methyl-C at CpG sites (24, 31, 72). Rates of DNA repair have not been implicated as much as sites of adduct formation in the work by Pfeifer’s group (70, 71). However, transcription-coupled repair is an issue in terms of which strand is repaired. In other work by Wei and Conney (22, 73–75), the mutation spectrum (with other genes) resulting from benzo[a]pyrene diol epoxide treatment of mammalian cells was dependent upon the dose of the carcinogen, and saturation of DNA repair has been proposed as a likely mechanism underlying this phenomenon. Work by Tang’s group on human rasK gene codons 12 and 14 has also been interpreted in terms of effects of C methylation on binding, and rates of repair also appear to be a factor (25) (although these rasK codon 12 and 14 adducts have not been extensively evaluated in terms of producing mutations).

Epidemiological evidence for the tumorigenicity of ethylene dibromide in humans is very limited, and its classification as a carcinogen (and restricted industrial and agricultural use (Ref. 8)) is based on animal studies (6). No information is available about ethylene dibromide-generated mutation spectra of p53 or other genes in tumors. The p53 mutant spectra reported here might not predict sites of human p53 mutations in tumors, for several reasons. The influence of C methylation is not reported in the yeast system. Strand bias was not apparent in our work (Figs. 4–7) and might occur in mammalian systems. Another point is that the mutant spectra are influenced strongly in this system by differential rates of DNA repair. However, the rates may be a factor in yeast more than in mammalian liver (e.g. a $t_{1/2}$ of 70 h was found for an “average” S-[2-(N7-guanyl)ethyl]GSH adduct in rat liver (18), compare $t_{1/2}$ $\sim$20–80 h here; Figs. 3–7). One might expect slower repair in mammalian tissue to produce an even more dramatic influence of differential repair, but this cannot be assumed to be the case in the

---

**Fig. 5.** Initial damage, adduct half-life, and mutations produced on both strands along exon 6 of the p53 gene by in vivo exposure to GSCH$_2$CH$_2$Cl. Treatment and analysis were as described for Fig. 4.
absence of data. Another major issue to consider is that the phenotypes in yeast and tumors are driven by different factors, as pointed out by Yoon et al. (72). In the yeast system, only the transcription of the Ade2 gene is observed. The reasons why p53 mutations contribute to tumorigenesis are more complex (24, 72).

Do these results mean that the S-(2-(N7-guanyl)ethyl)GSH adducts are the ones involved in producing mutations? In earlier work (12) and here, most of the mutations are at G nucleotides (specifically GC to AT transitions) and most of the mutations correspond to sites of apparent guanyl N7-alkylation and slower repair (Figs. 4–7). However, minor amounts of two other G adducts can be formed from S-(2-chloroethyl)GSH, the N2 and O6 adducts (12). We have not specifically ruled out the possibility that these adducts could contribute to the FLMPCR results on sequence specificity of alkylation, although neither would be expected to on the basis of general chemical principles (39, 40). (However, hot piperidine treatment can lead to some other reactions other than cleavage at N7-guanyl and N6-adenyl alkyl adducts (Refs. 49 and 50), at slower rates.) We have not evaluated the relative efficiency of these reactions relative to N7-guanyl adducts but could not rule out small contributions. A fraction of the mutations produced by S-(2-chloroethyl)GSH (16).2 Mutagenicity of O6-guanyl adducts (which presumably would not be detected in the piperidine cleavage assay) would contribute to a lack of overlap between DNA adduct and mutation spectra (Figs. 4–7), as would any residual spontaneous mutations (Table II).

The results are consistent with the mutagenicity of N7-guanyl adducts in producing G to A mutations. How could this process occur? Depurination (which would produce G to T transversions) and imidazole ring-opening (not detected at a limit of 1% (Ref. 12)) can be ruled out. The literature generally argues against the mutagenicity of N7-alkylguanine adducts (77). However, closer inspection of the literature and the cited review (77) indicate that most of the evidence, which is rather indirect, was developed with N7-methylguanine. Little direct analysis of misincorporation at N7-alkylguanine adducts has

2 In principle, site-specific mutagenesis methods could be used to resolve the issue of relative mutagenicities of the three G adducts. Precedent now exists with aflatoxin B1 for introducing an unstable guanine N7-alkyl adduct into a cellular system (76). However, a major obstacle is the need to synthesize the adduct by chemical treatment of a G-containing oligonucleotide, followed by separation of all possible products. An issue arises in that the choice of an oligonucleotide target should reflect a site where mutations are known to occur. As shown in Table II, even among pentameric base sequences in mutation sites, only one of the sites identified here (codon 259) has a single G (and that mutation was only found once). Thus, synthesis of an appropriate N7-guanyl adduct for site-specific mutagenesis remains problematic.
been done and none with the more recently discovered “translesional” polymerases. There is precedent for mutation at an \( N^7 \)-aflatoxin \( B_1 \) adduct (76). In our earlier studies, we calculated that the frequency of G to A transitions (caused by GSCH\(_2\)CH\(_2\)Cl) in bacterial systems could not be attributed to adducts other than the \( N^7 \)-alkylguanine adduct, S-[2-\( N^7 \)-guanyl]ethylglutathione (12, 15). S-[2-\( N^7 \)-Guanyl]ethylglutathione was misincorporating with \( E. \) coli polymerase II (but not polymerase I), with dTTP being inserted (67) (other polymerases have not been examined). We have discussed possible mechanisms for misincorporation previously (12, 15, 78), primarily in the context of tautomerism and unusual ionization properties. The concept of base tautomerism as a basis for mispairing is not new (79–80) and has theoretical (81) and experimental (82–84) backing. The lowered \( pK_a \) of the \( N1 \) atom of guanine resulting from \( N^7 \)-alkylation is well established both with the free base (82) and in double-stranded oligonucleotides (78), lowering the \( pK_a \) to physiological pH (pH 7–8). A zwitterionic residue might indeed mispair, particularly if the polymerase is strongly blocked by the \( N7 \) substitution, as has been shown (67). These possibilities are being considered in experimental studies.2

In summary, we have characterized a system used as a model for the GSH-dependent mutagenicity of the carcinogen ethylene dibromide. The pattern of DNA alkylation observed inside cells differs from that with isolated DNA, even in the same coding sequence. The mutant spectrum did not correlate well with the pattern of initial DNA alkylation, but the mutations were generally at sites most resistant to repair. These results provide an example of the role of sequence-selective repair in determining the mutation spectrum in this system. The molecular basis of sequence-selective profiles of DNA repair is poorly understood. Our earlier research in this area suggested that nucleotide excision repair is important in preventing the mutagenicity of \( S-(2\text{-chloroethyl}) \text{GSH} \) (12), although more details of the process are unknown (or the potential involvement of other types of repair). Other evidence suggests at least partial roles for sequence selectivity of DNA repair in some other systems, e.g. UV light and PAH damage.

Acknowledgments—We thank N. F. Cariello (University of North Carolina, Chapel Hill, NC) for the statistical program we used to compare patterns and J. A. Krauser (Vanderbilt University, Nashville, TN) for assistance in using this software, K. Bhat of the Vanderbilt sequencing facility for technical contributions, and particularly J. L. Cmarik (Frederick Cancer Center, Frederick, MD) for comments on the manuscript.

REFERENCES
1. Bauer, K. H. (1928) *Mutationstheorie der Geschwulstentstehung*, Springer Verlag, Berlin
2. Heidelberger, C. (1975) *Annu. Rev. Biochem.* 44, 79–121
3. Ames, B. N. (1979) *Science* 204, 587–593
4. Benzer, S. (1961) *Proc. Natl. Acad. Sci. U. S. A.* 47, 403–416
5. Rogozin, I. B., and Pavlov, Y. I. (2005) *Mutat. Res.* 544, 65–85
6. Olson, W. A., Habermann, R. T., Weisburger, E. K., Ward, J. M., and Weisburger, J. H. (1973) *J. Natl. Cancer Inst.* 51, 1993–1995
7. Rannug, U. (1980) *Mutat. Res.* 76, 269–295

Fig. 7. Initial damage, adduct halflife, and mutations produced on both strands along exon 8 of the \( p53 \) gene by \textit{in vivo} exposure to GSCH\(_2\)CH\(_2\)Cl. Treatment and analysis were as described for Fig. 4.
S-(2-Chloroethyl)glutathione-generated p53 Mutation Spectra Are Influenced by Differential Repair Rates More than Sites of Initial DNA Damage

J. Gerardo Valadez and F. Peter Guengerich

J. Biol. Chem. 2004, 279:13435-13446. doi: 10.1074/jbc.M312358200 originally published online January 7, 2004

Access the most updated version of this article at doi: 10.1074/jbc.M312358200

Alerts:
- When this article is cited
- When a correction for this article is posted

Click here to choose from all of JBC’s e-mail alerts

This article cites 80 references, 24 of which can be accessed free at http://www.jbc.org/content/279/14/13435.full.html#ref-list-1