

## Targeting of Lung Cancer Mutational Hotspots by Polycyclic Aromatic Hydrocarbons

Leslie E. Smith, Mikhail F. Denissenko, William P. Bennett, Haiying Li, Shantu Amin, Moon-shong Tang, Gerd P. Pfeifer

**Background:** Polycyclic aromatic hydrocarbons (PAHs) are ubiquitous in combustion products of organic matter, including cigarette smoke. Metabolically activated diol epoxides of these compounds, including benzo[*a*]pyrene diol epoxide (B[*a*]PDE), have been suggested as causative agents in the development of lung cancer. We previously mapped the distribution of B[*a*]PDE adducts within the p53 tumor suppressor gene (also known as TP53), which is mutated in 60% of human lung cancers, and found that B[*a*]PDE adducts preferentially form at lung cancer mutational hotspots (codons 154, 157, 158, 245, 248, and 273). Other PAHs may be important in lung cancer as well. **Methods:** Here we have mapped the distribution of adducts induced by diol epoxides of additional PAHs: chrysene (CDE), 5-methylchrysene (5-MCDE), 6-methylchrysene (6-MCDE), benzo[*c*]phenanthrene (B[*c*]PDE), and benzo[*g*]chrysene (B[*g*]CDE) within exons 5, 7, and 8 of the p53 gene in human bronchial epithelial cells. **Results:** CDE exposure produced only low levels of adducts. Exposure of cells to the other activated PAHs resulted in DNA damage patterns similar to those previously observed with B[*a*]PDE but with some distinct differences. 5-MCDE, 6-MCDE, B[*g*]CDE, and B[*c*]PDE efficiently induced adducts at guanines within codons 154, 156, 157, 158, and 159 of exon 5, codons 237, 245 and 248 of exon 7, and codon 273 of exon 8, but the relative levels of adducts at each site varied for each compound. B[*g*]CDE, B[*c*]PDE, and 5-MCDE induced damage at codon 158 more selectively than 6-MCDE or B[*a*]PDE. The sites most strongly involved in PAH adduct formation were also the sites of highest mutation frequency (codons 157, 158, 245, 248, and 273). **Conclusion:** The data suggest that PAHs contribute to the mutational spectrum in human lung cancer. [J Natl Cancer Inst 2000;92:803–11]

Accumulation of mutations in genes responsible for the maintenance of growth control and genomic integrity results in loss of these essential functions and ultimately leads to tumor formation. Any mutation that provides a growth advantage will be selected for during tumor progression. As the presumed guardian of the genome, the p53 tumor suppressor gene (also known as TP53) coordinates a delicate balance between arrest of the cell cycle to allow repair of damage and apoptosis if the damage is irreparable. The p53 gene is one of the most commonly mutated genes observed in human tumors. It is mutated in more than 50% of all human cancers and in about 60% of human lung cancers (1–3). While a mutation can occur anywhere along the p53 gene, most of the mutations occur in sequences that encode the DNA-binding domain of the p53 protein (2).

Lung cancer is the leading cause of cancer death for U.S. women and men (4) and is the most common type of cancer worldwide. Polycyclic aromatic hydrocarbons (PAHs), present in all products of combustion of organic matter (including tobacco smoke), have been suggested as agents responsible for the initiation and development of lung cancer (5). Because they are formed upon combustion of any organic material, they are present in notable concentrations in our environment (6) due to gasoline and diesel engine emissions and industrial sources. However, the largest concentrations of PAHs are inhaled by smokers with the mainstream smoke of cigarettes (7,8).

PAHs are activated by the cytochrome P450 enzymes to form the ultimate carcinogenic diol epoxides, although other pathways of PAH activation exist (9). The diol epoxide PAH metabolites form covalent adducts with DNA, primarily at the exocyclic N2 position of guanine residues and the N6 position of adenines (6). B[*a*]PDE [(+/-)-anti-benzo[*a*]pyrene-7,8-diol-9,10-epoxide] is the best studied of the activated PAHs. Previous studies with B[*a*]PDE have demonstrated that the DNA damage spectrum is nonrandom. B[*a*]PDE induces guanine adducts at mutational hotspots, including codons 157, 248, and 273 of the p53 gene in normal human bronchial epithelial cells (10). B[*a*]PDE-damage hotspots in exons 5, 7, and 8 correspond to methylated CpG sequences within the p53 gene (11,12). Preferential formation of B[*a*]PDE adducts is due to enhancement of adduct formation by the 5-methylcytosine base within a CpG sequence (13,14). Furthermore, B[*a*]PDE adducts in the non-transcribed strand of p53 are repaired two to four times more slowly than those in the transcribed strand (15). The slowly repaired damage hotspots correspond to mutational hotspots observed in human lung cancer. B[*a*]PDE exposure results in a mutational fingerprint with respect to the types of mutations induced. This bulky carcinogen produces predominantly G-to-T transversions (16–20), whereas spontaneous mutations tend to be G-to-A transitions probably arising through spontaneous deamination of methylated cytosines (21,22). The p53 mutation spectrum observed in human lung cancer from smokers is dominated by the presence of G-to-T transversions (approximately 30%–40%). This type of mutation is much less frequent in other

*Affiliations of authors:* L. E. Smith, M. F. Denissenko, G. P. Pfeifer (Department of Biology), W. P. Bennett (Department of Human Genetics), Beckman Research Institute of the City of Hope, Duarte, CA; H. Li, M.-s. Tang, The University of Texas M. D. Anderson Cancer Center, Smithville, TX; S. Amin, American Health Foundation, Valhalla, NY.

*Correspondence to:* Gerd P. Pfeifer, Ph.D., Department of Biology, Beckman Research Institute of the City of Hope, Duarte, CA 91010 (e-mail: gpfeifer@coh.org).

See "Notes" following "References."

© Oxford University Press

human cancers, except hepatocellular carcinoma. Of all of the G-to-T transversion mutations in lung tumors, 95% occur on the nontranscribed strand (3,23). In contrast, the p53 mutational spectrum of human lung cancer from nonsmokers demonstrates few hotspots and a much lower frequency of G-to-T transversions (3,23). These data, together with DNA adduct formation at lung cancer mutational hotspots by B[a]PDE (10) and the repair strand bias (15), suggest an etiologic link between benzo[a]pyrene (B[a]P) exposure and human lung cancer.

However, B[a]P is only one of many PAHs found in the complex mixture of chemicals in cigarette smoke (24), with concentrations reported to be as high as 10–50 ng per cigarette (6,25). Among other PAHs in tobacco smoke are chrysene, 5-methylchrysene, and 6-methylchrysene at approximately 60, 0.6, and 10 ng per cigarette, respectively (6,25,26). Benzo[c]phenanthrene has been determined to be present in cigarette smoke; however, exact concentrations have not been reported (7). Benzo[g]chrysene has not been reported in cigarette smoke but is present in coal tar and petroleum distillates (27). It is, therefore, found at appreciable levels in the environment and is also likely present in smoke condensate. The activated forms of B[a]P, chrysene, and 6-methylchrysene are planar bay region diol epoxides. The chrysene metabolite (+/-)-anti-chrysene-3,4-diol-1,2-epoxide has been reported to be only weakly mutagenic in *Salmonella typhimurium* and in Chinese hamster V-79 cells (28) and to be weakly tumorigenic in newborn mouse liver and lung model systems (29); however, the reverse diol epoxide (+/-)-anti-chrysene-1,2-diol-3,4-epoxide (CDE), used in this study, is active (28). The presence of a methyl group adjacent to the bay region in (+/-)-anti-5-methylchrysene-1,2-diol-3,4-epoxide (5-MCDE) is thought to induce steric hindrance that forces the epoxide ring away from the plane of the molecule and is responsible for its high mutagenicity and carcinogenicity in many model systems (28,30–33). In contrast, (+/-)-anti-6-methylchrysene-1,2-diol-3,4-epoxide (6-MCDE) is only weakly mutagenic in many test systems (30–32). 6-MCDE is formed in greater amounts than is 5-MCDE in mouse epidermis after exposure to the respective parent PAHs; however, 6-MCDE results in lower levels of DNA adduction (30). These differences have been attributed to decreased initial intercalation of 6-MCDE, increased repair of adducts, or increased detoxification of

6-MCDE by mechanisms such as conjugation with glutathione (32). (+/-)-Anti-benzo[g]chrysene-11,12-diol-13,14-epoxide (B[g]CDE) and (+/-)-anti-benzo[c]phenanthrene-3,4-diol-1,2-epoxide (B[c]PDE) are nonplanar fjord region diol epoxides. The nonplanar diol epoxides are more mutagenic and carcinogenic than the planar metabolites. B[c]PDE proved to be mutagenic in bacterial and mammalian assays (34) and is a potent tumor initiator in mouse skin (35). B[c]PDE adducts can occur at both adenine and guanine residues in DNA (36). Both B[c]PDE and B[g]CDE are potent mammary carcinogens in rats (37). B[g]CDE, a tumor initiator in mouse skin (38) and in lung/liver assays in newborn mice (39) and which is also highly mutagenic in the dhfr gene of Chinese hamster ovary cells (40), can likewise form adducts with both adenine and guanine residues (41).

In this study, we have mapped the distribution of adducts formed by the active diol epoxide forms of chrysene, 5-methylchrysene, 6-methylchrysene, benzo[c]phenanthrene, and benzo[g]chrysene (see Fig. 1) in the nontranscribed strand of p53 for exons 5, 7, and 8 in normal human bronchial epithelial cells.

## MATERIALS AND METHODS

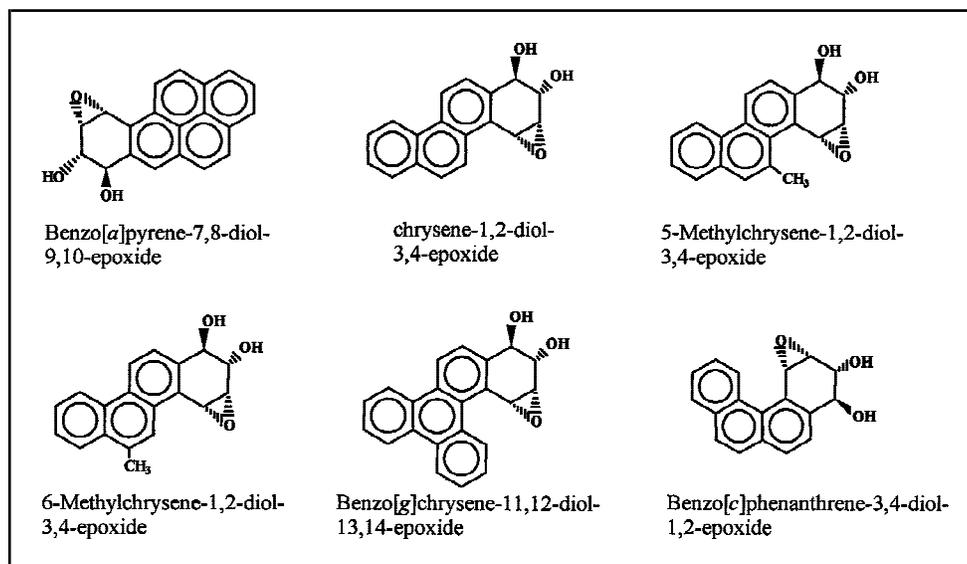
### Chemicals

Racemic diol epoxides of the following PAHs were synthesized as described in the literature: (+/-)-anti-chrysene-1,2-diol-3,4-epoxide (CDE) (6), (+/-)-anti-5-methylchrysene-1,2-diol-3,4-epoxide (5-MCDE) (42), (+/-)-anti-6-methylchrysene-1,2-diol-3,4-epoxide (6-MCDE) (43), (+/-)-anti-benzo[g]chrysene-11,12-diol-13,14-epoxide (B[g]CDE) (44,45), and (+/-)-anti-benzo[c]phenanthrene-3,4-diol-1,2-epoxide (B[c]PDE) (46) (Fig. 1). (+/-)-Anti-benzo[a]pyrene-7,8-diol-9,10-epoxide (B[a]PDE) was obtained from the National Cancer Institute repository (Midwest Research Institute, Kansas City, MO). The compounds were dissolved in dimethyl sulfoxide (DMSO), and stock concentrations were confirmed by spectrophotometry.

### Cell Culture and DNA Modification

Normal human bronchial epithelial cells were cultured in medium recommended by the supplier (Clonetics, San Diego, CA). These cells were treated with various concentrations of the PAH metabolites for 30 minutes at 37 °C in the dark. The highest concentrations used (except for those of CDE) produced similar adduct levels in total genomic DNA, as measured after cleavage with the UvrABC excision nuclease complex of *Escherichia coli* and alkaline agarose gel electrophoresis (approximately one adduct every 5 kilobases [kb]). Controls

**Fig. 1.** Structures of polycyclic aromatic hydrocarbon anti-diol epoxides investigated in this study. The structures of only the (+) optical isomers are shown.



were treated with the appropriate concentration of DMSO only. After treatment, the cells were harvested and genomic DNA was isolated as previously described (10).

### Cleavage of Damaged DNA by UvrABC

Uvr proteins, comprising the nucleotide excision repair complex of *E. coli*, were purified as described previously (47). Purified DNA was treated with a molar excess of Uvr proteins as described previously (10,47). Cleavage of adducted DNA was confirmed by running 1  $\mu$ g of treated DNA in 0.6% agarose gels after denaturation in formamide. Under the reaction conditions used, DNA cleavage by UvrABC proteins is quantitative (10,48). Efficient cleavage was achieved with a 10-fold molar excess of UvrABC subunits over DNA (10-kb size). Additional increases in enzyme concentration (up to 40-fold excess) did not result in additional cleavage. After UvrABC treatment, the proteins were removed by phenol-chloroform extraction. The cleaved DNA was precipitated and resuspended in TE buffer (i.e., 1 mM Tris-HCl [pH 8.0] and 0.1 mM EDTA).

### Ligation-Mediated Polymerase Chain Reaction and Quantitation of Data

Ligation-mediated polymerase chain reaction (LMPCR) was performed for exons 5, 7, and 8 of the human p53 gene. The conditions and oligonucleotide primers utilized for LMPCR of these exons were as previously described (49,50), with the following modifications: Longer primers having higher annealing temperatures were used during the initial primer extension step. Primer 5-4D (5'-GGGCCAGACCTAAGAGCAATCAGT) was used for exon 5, primer 7-4B (5'-CAGGGGTACGCGCAAGCAGAG) was used for exon 7, and primer 8-4D (5'-AGGCAAGGAAAGGTGATAAAAGTG) was used for exon 8.

The extension was performed with the use of a 16 : 1 combination of the Vent exo- and Vent polymerases (New England Biolabs, Inc., Beverly, MA) with a thermocycler protocol of 3 minutes at 95 °C, 3 minutes at the primer annealing temperature, and 15 minutes at 72 °C. After elongation of the primer, DNA was precipitated in the presence of glycogen. The primer-extended fragments were ligated as previously described (49-51). Ligated fragments were polymerase chain reaction (PCR) amplified after ethanol precipitation with Amplitaq Gold DNA polymerase (The Perkin-Elmer Corp., Foster City, CA) according to the manufacturer's directions. The amplified PCR products were separated by denaturing gel electrophoresis, electroblotted to a Genescreen Plus (Du Pont NEN, Boston, MA) nylon membrane, and hybridized with <sup>32</sup>P-labeled p53-specific probes as previously described (51). All LMPCR experiments were repeated at least once and gave very similar results. The nylon membranes were exposed to a PhosphorImager (Molecular Dynamics, Sunnyvale, CA). Radioactivity was measured in all PAH metabolite-specific bands of the sequencing gel that could be fully resolved. The radioactivity of each band was quantified after background subtraction with the use of lanes that were treated with UvrABC, but no carcinogen. To correct for DNA sequence-dependent LMPCR amplification efficiencies, we normalized the increases in the intensity of a radioactive signal from a treated sample by dividing this signal by the corresponding Maxam-Gilbert control band. We then calculated the relative intensity of each band in a particular exon by dividing the corrected damage signal at each nucleotide by the band of greatest intensity produced by each particular compound in a specific exon. The relative intensity was then plotted as a function of the nucleotide sequence for codons within exons 5, 7, and 8. Only one concentration of PAH diol epoxides was used for quantitation, but the LMPCR-damage patterns are quite reliable and dose independent as long as a sufficient adduct frequency of more than one adduct every 10 kb is provided.

## RESULTS

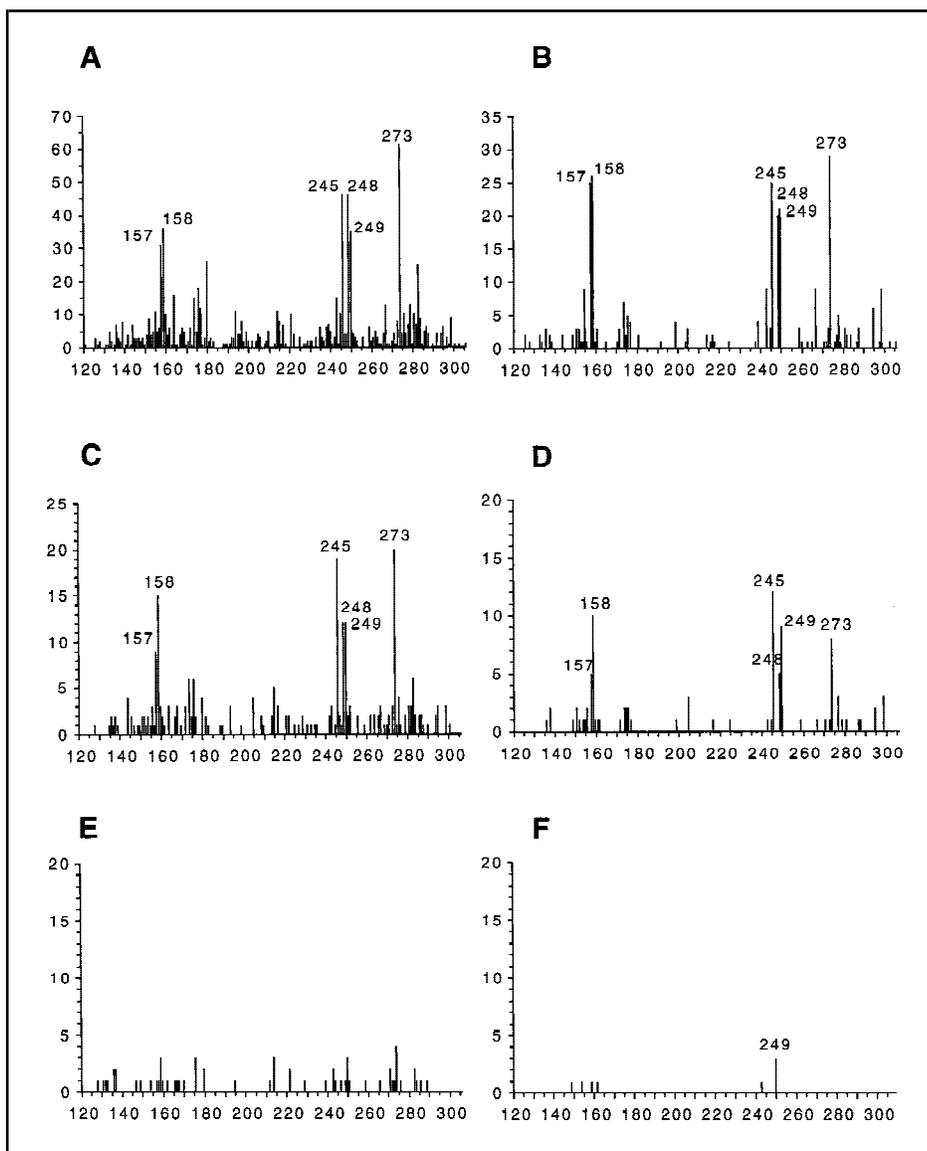
Using an online database of mutations, we have assessed the distribution of mutations along the p53 gene in lung cancers from smokers and nonsmokers (Fig. 2). We show the p53 mutational spectra for all mutations in 1) all lung cancer case patients minus nonsmokers and minus occupationally exposed individuals (given the incidence of lung cancer in smokers versus nonsmokers, 90% of these cases are expected to be smokers), 2) smokers specifically identified as such in the p53 database, and 3) identified nonsmokers. Separately, the distributions of G→T transversions only are shown. For nonsmokers, 62 mutations

could be confirmed from the literature. We have excluded mutations in individuals known to be exposed to radon and mustard gas as well as in coal miners. Fig. 2 shows that the mutational spectra are very different between smokers and nonsmokers and that G→T transversions are rare in nonsmokers (eight of 62). In smokers, the mutational spectrum is characterized by mutational hotspots at codons 157, 158, 245, 248, 249, and 273. In contrast, the p53 mutational spectrum in lung cancers from nonsmokers is relatively random, with only a few minor hotspots observed (Fig. 2, E and F). For smokers and all lung cancer case patients minus nonsmokers, the G→T transversion spectra are characterized by the same hotspots as the spectra of all mutations. It should be remembered that certain PAH adducts can induce G→A and G→C mutations in addition to G→T, depending on the sequence context (20). Importantly, the only site that shows a clustering of G→T transversions in nonsmokers is at codon 249.

We (10) previously reported that normal bronchial epithelial cells exposed *in vitro* to B[a]PDE accumulate hotspots of guanine adducts at codon positions matching several of the most prominent lung cancer mutation hotspots found in a database of somatic p53 alterations. Because B[a]P is only one of many PAH compounds found in cigarette smoke and in the environment, we have mapped at nucleotide resolution the adduct patterns induced by five additional activated PAH metabolites (Fig. 1) in normal bronchial epithelial cells. UvrABC nuclease in combination with LMPCR was used for adduct mapping. Under the conditions used, UvrABC cleavage is quantitative for B[a]PDE adducts (10,48). We believe that the other PAH adducts are also cleaved quantitatively, since increases in enzyme concentrations or reaction times did not result in additional cleavage.

Fig. 3, A, demonstrates the profile of PAH adducts produced within the nontranscribed strand of exon 5 after a single 30-minute exposure of cells to each activated PAH compound. The relative intensity of damage for each compound and previously obtained data for B[a]PDE damage within exon 5 are summarized in Fig. 3, B. Codons 154, 156, 157, 158, and 159 demonstrated the overall highest accumulation of DNA adducts after exposure to the PAH metabolites. These codons all contain 5-methylcytosines within a CpG sequence and were also sites of greatest damage induced by B[a]PDE (10) (Fig. 3, B). In addition, codons 164, 171, and 175 demonstrate substantial levels of DNA adducts, depending on the specific agent. CDE was weakly effective at producing DNA adducts at codon 156, with lesser amounts of adducts found at codons 154 and 158 (Fig. 3, A). 5-MCDE produced adducts at codons 154, 156, 157, 158 (highest), 159, 171, and 175. 6-MCDE adducts were formed primarily at codons 154 (highest), 157, 158, 159, 164, 170, 171, and 175. B[g]CDE adducts occurred primarily at codons 158 (highest), 175, 159, and 156. B[c]PDE adducts were observed primarily at codons 158 (highest), 175, 154, 156, and 159. Codon 158 was more selectively targeted by 5-MCDE, B[c]PDE, and B[g]CDE relative to B[a]PDE and 6-MCDE. The adduct profiles of the fjord region PAHs B[c]PDE and B[g]CDE were more similar to each other than to the profiles of the bay region PAHs. DNA damage at codons 154, 157, and 158 is of particular interest, since mutations at these positions are common in lung cancer (Fig. 2) but are rare in other tumors. The effective order of total DNA adduct formation at comparable concentrations of the diol epoxides within the nontranscribed

**Fig. 2.** p53 mutational spectra in lung cancer. The number of p53 mutations (y-axis) is shown by codon position (x-axis; major peaks are labeled). **Panels A and B:** all lung cancer case patients minus nonsmokers and minus occupationally exposed individuals ( $n = 989$ ); given the incidence of lung cancer in smokers versus nonsmokers, 90% of these are expected to be smokers. **Panels C and D:** smokers specifically identified as such in the p53 database ( $n = 263$ ). **Panels E and F:** nonsmokers ( $n = 62$ ). The distribution of all mutations (panels A, C, and E) and G→T transversions only (panels B, D, and F) is shown separately. For nonsmokers, all mutations were confirmed from the primary literature. Lung cancer mutations were collected by Internet (<http://www.iarc.fr/P53/index.html>) from an updated p53 database (54). Mutations were excluded from lung cancers related to occupational exposures, including exposures to chromate, mustard gas, coal, and radon in mines. In addition, 107 mutations from one unusual report (59) were removed.



strand of exon 5 was CDE < 6-MCDE < B[c]PDE < 5-MCDE = B[g]CDE (Fig. 3, A; data not shown).

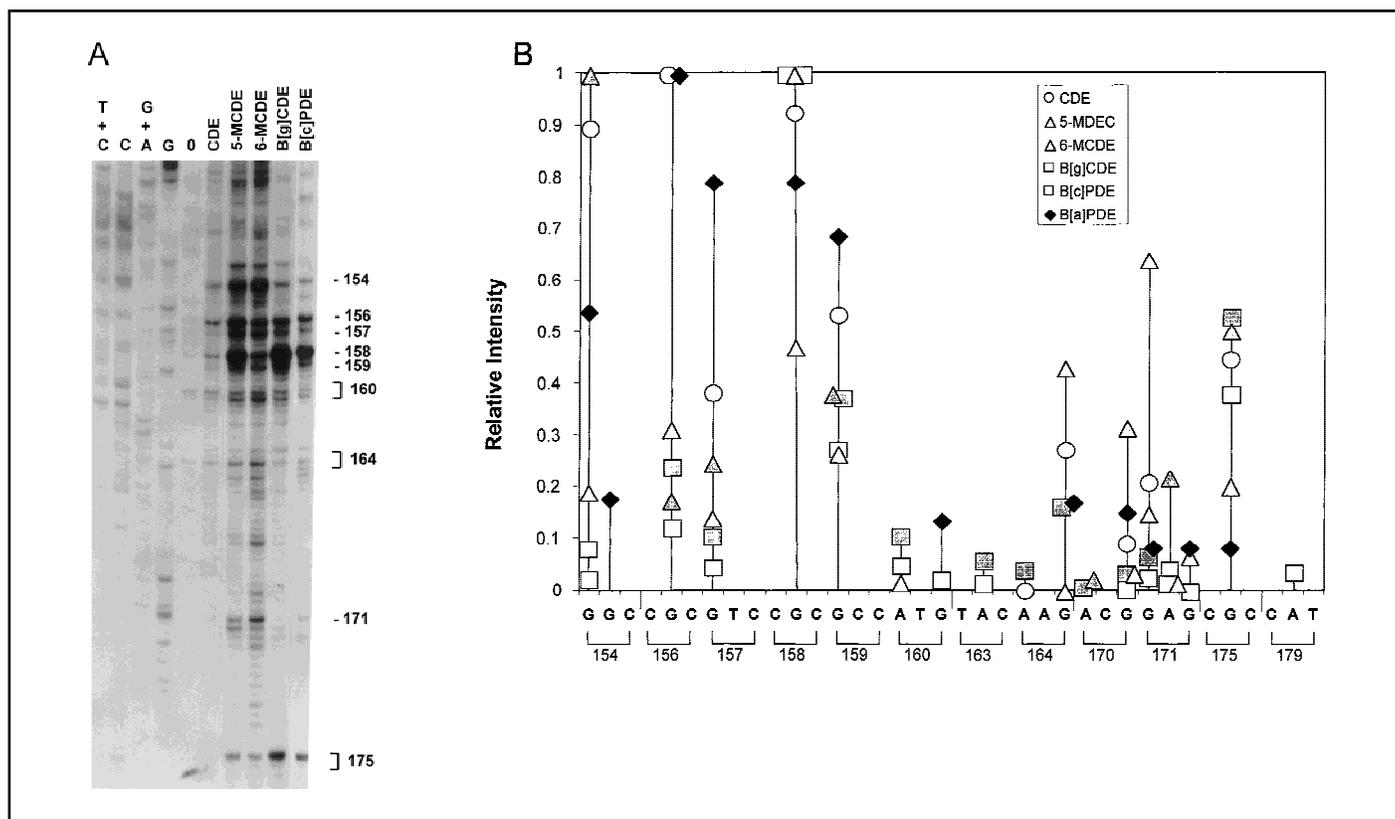
Fig. 4 shows the PAH metabolite-induced formation of adducts and the relative intensity quantitation for exon 7 of the nontranscribed strand of p53. Weak CDE-adduct bands are observed at several positions; however, after subtraction of the UvrABC-treated undamaged control, CDE produced only low levels of adducts within exon 7. The darker bands near the top of the CDE lane correspond to sequences within an intron and are, therefore, not considered in our analysis of DNA damage in the coding regions of p53. The low levels of CDE adducts are consistent with the weak carcinogenic potency of this compound (6,26). 5-MCDE produced adducts most efficiently in codons 245, 248, and 237 in exon 7. Note that the adducted guanine in codon 245 is a site that is amplified relatively inefficiently by LMPCR (15) and, thus, does not appear strongly damaged in the autoradiograph. This factor is corrected, however, by the quantitation procedure used (Fig. 4, B) and, thus, this site does appear to be an important damage site, in particular for 5-MCDE. 6-MCDE was most effective at the first guanine residue of codon 248, with lesser amounts of adducts produced at codons 245, 237, 238, and several others. B[g]CDE and B[c]PDE elicited

high amounts of adducts with the first guanine base in codon 248 as well. B[g]CDE and B[c]PDE also produced increased numbers of adducts with adenines in codons 236, 239, 240, and 246. B[a]PDE adducts were previously demonstrated to occur primarily at codons 248 and 245 (10,14). Both of these hotspots for DNA damage are CGG sequences in which the initial cytosine is methylated. Codon 237 does not contain a 5-methylcytosine within its ATG sequence; however, this codon is substantially damaged by the fjord region PAHs (Fig. 4, B).

Data for PAH adducts formed within exon 8 are shown in Fig. 5, A and B. CDE produced no appreciable levels of adducts within the nontranscribed strand of exon 8. All of the other PAHs produced adducts at a single hotspot—in codon 273 (Fig. 5, A). Lower amounts of adduction were seen at the guanines in codons 282, 283, and 290 (Fig. 5, A and B). Thus, all of the PAH metabolites studied were highly effective at producing guanine adducts within the sequence 5-methyl-CGT of codon 273, which is the major mutational hotspot in lung cancer (Fig. 2).

## DISCUSSION

PAH compounds are present in the environment as a result of combustion of organic materials. They are found as components



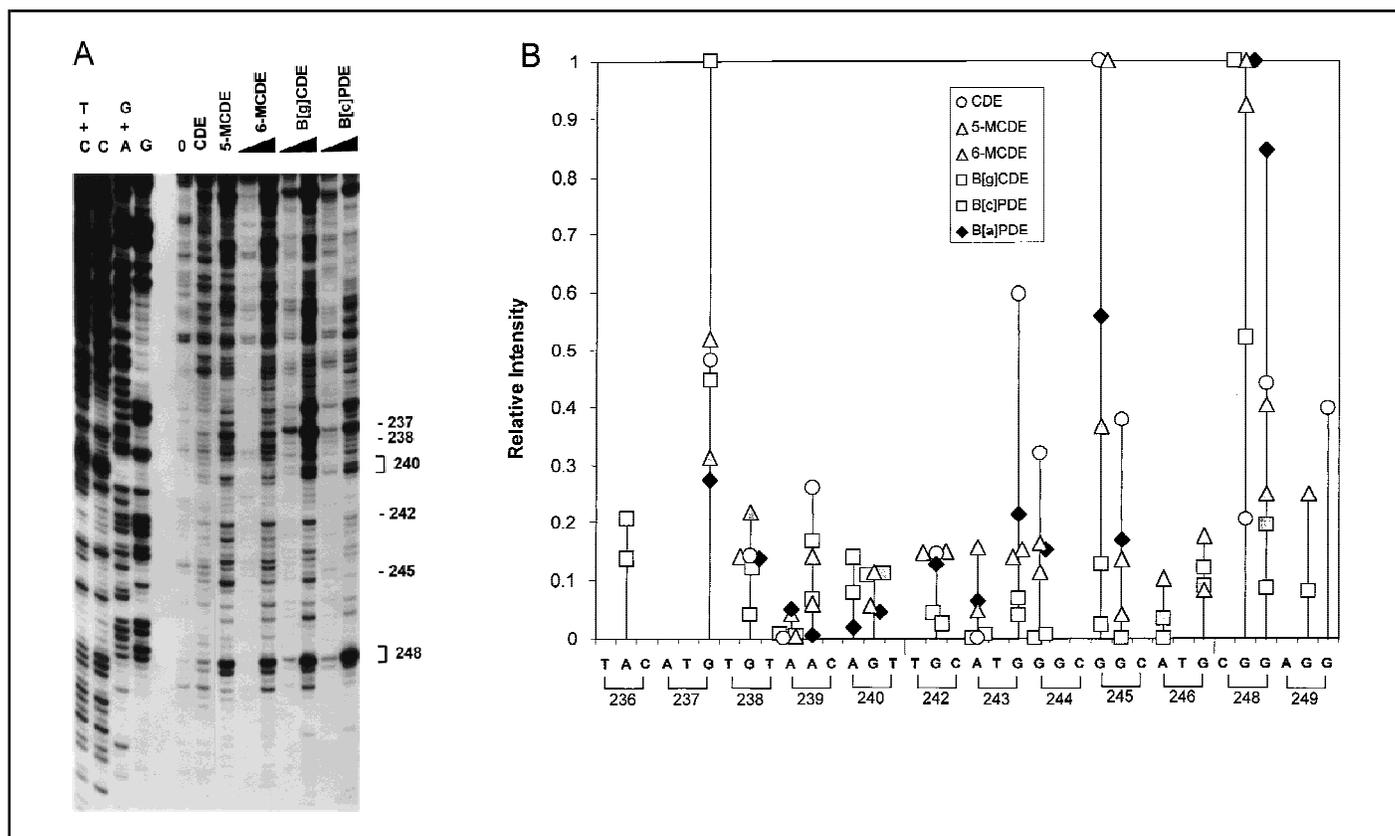
**Fig. 3. Panel A:** distribution of polycyclic aromatic hydrocarbon (PAH) diol epoxide adducts along exon 5 of the p53 gene. Concentrations of PAH compounds were 25  $\mu$ M CDE (lane 6), 10  $\mu$ M 5-MCDE (lane 7), 50  $\mu$ M 6-MCDE (lane 8), 50  $\mu$ M B[g]CDE (lane 9), and 50  $\mu$ M B[c]PDE (lane 10). Lanes 1–4: Maxam–Gilbert sequencing controls. Lane 5: UvrABC-treated DNA from cells exposed to solvent control. **Panel B:** quantitation of PAH diol epoxide damage in exon 5 of p53 by sequence and codon. Only the most substantially damaged codons are shown. Quantitative data using 4  $\mu$ M B[a]PDE were derived from the sequencing gels presented previously (10) and are included for comparison of the

relative sequence distribution of B[a]PDE adducts. The intensity of PAH-induced bands was quantified with the use of a PhosphorImager, and the relative intensity of each compound was calculated as described in the “Materials and Methods” section. **Abbreviations used:** CDE = (+/-)-anti-chrysene-1,2-diol-3,4-epoxide; 5-MCDE = (+/-)-anti-5-methylchrysene-1,2-diol-3,4-epoxide; 6-MCDE = (+/-)-anti-6-methylchrysene-1,2-diol-3,4-epoxide; B[g]CDE = (+/-)-anti-benzo[g]chrysene-11,12-diol-13,14-epoxide; B[c]PDE = (+/-)-anti-benzo[c]phenanthrene-3,4-diol-1,2-epoxide; B[a]PDE = (+/-)-anti-benzo[a]pyrene-7,8-diol-9,10-epoxide.

of complex mixtures in tobacco smoke, polluted air, gasoline and diesel engine exhaust, and industrial waste from various manufacturing and chemical processes. Little is known about the relevant exposures, DNA binding, and repair of damage induced by a complex mixture of PAHs *in vivo*. Data from studies of cigarette smoke condensate and polluted urban air in various strains of *Salmonella typhimurium*, however, have suggested that the mutation spectrum of a complex mixture is dominated by the presence of one or two compounds within the mixture (52,53). Mutation spectra for the base substitution allele in strain TA100 by cigarette smoke condensate, for example, included primarily G-to-T transversions and were largely similar to the mutation spectrum in the TA100 strain with B[a]P (53). The data presented here further suggest that B[a]PDE may be useful as a reference compound for the DNA damage and mutation spectra induced by a variety of other activated PAHs. Strong similarities exist for the major binding sites (e.g., codons 248 and 273), although there are also distinct differences at other sequences. Some caution should be used, however, when trying to assess the overall cancer risk inherent in exposure to a complex mixture such as cigarette smoke. While the overall damage spectrum is reasonably well represented by B[a]PDE, the overall risk of exposure to a complex mixture may not be a simple additive effect of the individual compounds present. The specific carci-

nogenic risk attributable to an individual compound in a mixture is determined by the relative amount of the compound present, the efficiency of its biotransformation to ultimate active species, the efficiency of detoxification, differences in covalent DNA binding, interaction with non-DNA targets, the capacity of repair enzymes to remove individual adducts, and the mutagenicity of an adduct during DNA replication. Possible interactions between compounds at the level of biotransformation as well as a possible saturation of repair enzyme systems also need to be considered. All of the above factors may contribute to underestimation or overestimation of the cancer risk imposed by one particular compound in a complex mixture. Given the fact that p53 is a nonselectable gene, it is technically impossible to produce a mutational spectrum in the p53 gene with the use of cigarette smoke extract and *in vitro* systems. We think that the comparison of adduct profiles generated *in vitro* with mutational profiles in tumors is worthwhile. Although this is only an association, it is the only technically feasible approach available at present.

All PAH compounds demonstrated strong adduction at guanines in codons 157 and 158 of exon 5, codons 245 and 248 of exon 7, and codon 273 of exon 8, which are prominent hotspots for mutation in human lung cancer. Mutations in lung cancer are not particularly common at codon 237, another strong PAH-binding site, but G-to-T transversions at the guanine of this



**Fig. 4. Panel A:** distribution of polycyclic aromatic hydrocarbon (PAH) diol epoxide adducts along exon 7 of the p53 gene. Concentrations of PAH compounds were 25  $\mu\text{M}$  CDE (lane 7), 10  $\mu\text{M}$  5-MCDE (lane 8), 10  $\mu\text{M}$  and 50  $\mu\text{M}$  6-MCDE (lanes 9 and 10, respectively), 10  $\mu\text{M}$  and 50  $\mu\text{M}$  B[g]CDE (lanes 11 and 12, respectively), and 10  $\mu\text{M}$  and 50  $\mu\text{M}$  B[c]PDE (lanes 13 and 14, respectively). Lane 5 is blank. Lanes 1–4: Maxam–Gilbert sequencing controls. Lane 6: UvrABC-treated DNA from cells exposed to solvent control. Panel B: quantitation of PAH diol epoxide damage in exon 7 of p53 by sequence and codon. Only the most substantially damaged codons are shown. Con-

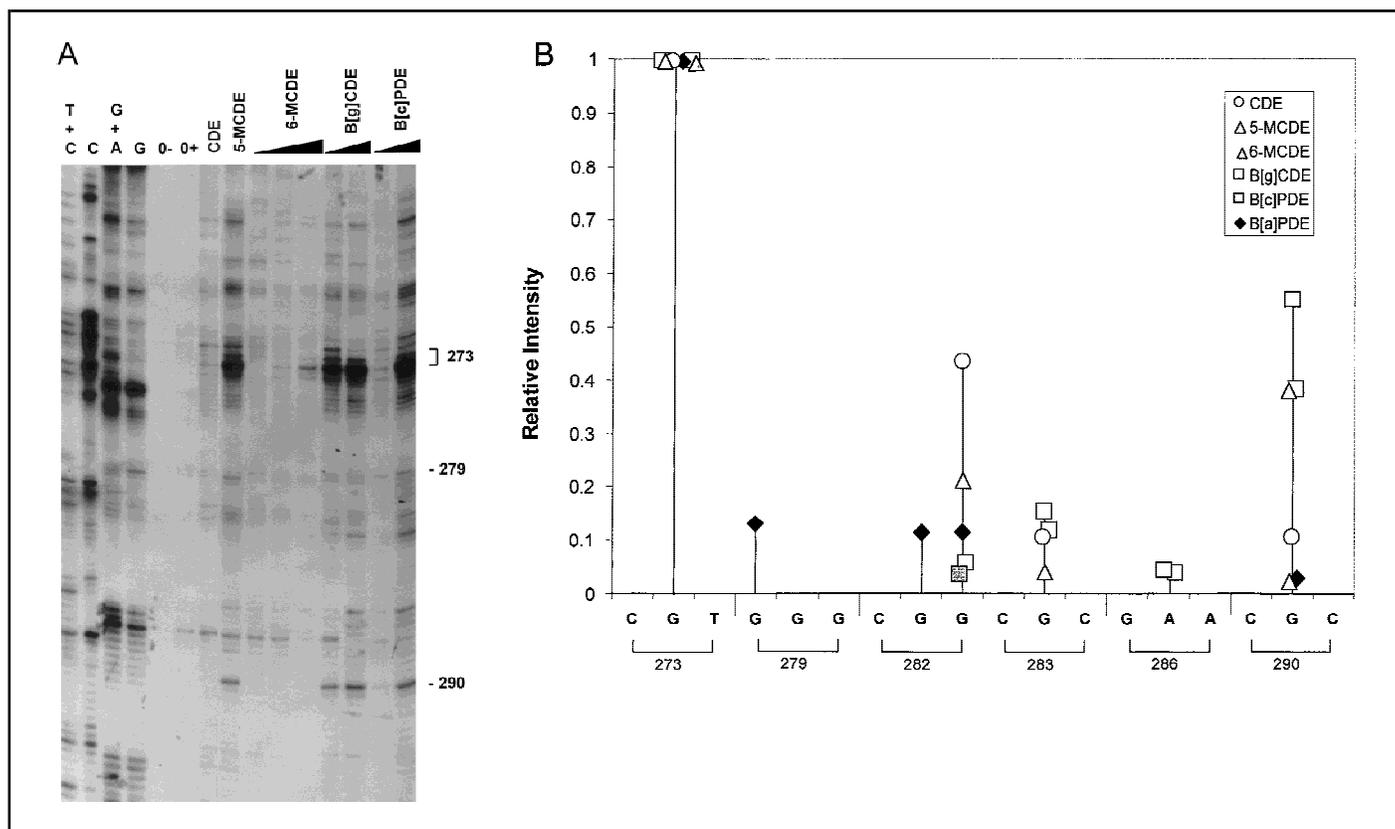
centrations of PAH compounds used for quantitation were 25  $\mu\text{M}$  CDE, 10  $\mu\text{M}$  5-MCDE, 50  $\mu\text{M}$  6-MCDE, 50  $\mu\text{M}$  B[g]CDE, 50  $\mu\text{M}$  B[c]PDE, and 4  $\mu\text{M}$  B[a]PDE (10). **Abbreviations used:** CDE = (+/-)-anti-chrysene-1,2-diol-3,4-epoxide; 5-MCDE = (+/-)-anti-5-methylchrysene-1,2-diol-3,4-epoxide; 6-MCDE = (+/-)-anti-6-methylchrysene-1,2-diol-3,4-epoxide; B[g]CDE = (+/-)-anti-benzo[g]chrysene-11,12-diol-13,14-epoxide; B[c]PDE = (+/-)-anti-benzo[c]phenanthrene-3,4-diol-1,2-epoxide; B[a]PDE = (+/-)-anti-benzo[a]pyrene-7,8-diol-9,10-epoxide.

codon are seen frequently in breast tumors (54). The reason could be that the PAHs that damage codon 237 are not abundant in cigarette smoke or that this mutation is not selected in lung cancer. The damaged guanines in codons 157, 158, 245, 248, and 273 are all within methylated CpG dinucleotides. The effect of cytosine methylation on the targeting of DNA damage by the PAHs investigated in this study is not yet well defined. Previous data from our studies (13), using methylated and unmethylated plasmids containing exons 5, 7, or 8 of p53, demonstrated that the presence of 5-methylcytosine in a CpG sequence enhanced B[a]PDE binding. A subsequent study (14) demonstrated that cytosine methylation resulted in increased binding of B[g]CDE to guanines in CpG sequences. The data presented here suggest that the presence of methylated cytosines in DNA-damage hotspot sequences may contribute to the binding of other activated PAHs. No specific sequence rules, other than a preference for methylated CpG sequences, could be identified.

Of the six most dominant lung cancer mutational hotspots that contain guanine bases (codons 157, 158, 245, 248, 249, and 273; Fig. 2), five are prominent PAH adduct-binding sites. The only exception is codon 249. This codon may represent a special case, where mutations may be frequent as a consequence of strong selection rather than preferential adduct formation [see also (55)]. Of the five strongest adduct-binding sites in exon 5

(codons 154, 156, 157, 158, and 159), two (codons 157 and 158) are prominent mutational hotspots in lung tumors and one (codon 154) is a moderate hotspot. Although they form high levels of adducts, codons 156 and 159 are not commonly mutated in lung tumors. At codons 156 and 159, a G-to-T transversion was described only once, each in a glioblastoma and a breast cancer, respectively (54), indicating that this type of mutation rarely may be selected at these codons. In exon 7, the three codons showing the highest levels of adduction (codons 237, 245, and 248) are mutational hotspots in breast cancer (codon 237) or lung cancer (codons 245 and 248). In exon 8, extremely high levels of adducts are seen at codon 273 (Fig. 5, A). This adduct distribution is largely similar to the lung cancer mutational spectrum, where codon 273 is the most commonly mutated site along the entire p53 sequence (Fig. 2). Moderate levels of adducts also form at codon 290, which is at the 3' end of the DNA-binding domain of the p53 protein, and mutations here probably are not selected. In summary, it appears that the p53 mutational spectrum in lung cancer is largely adduct driven, but selection determines which of the preferentially targeted codons are sampled into the tumor mutation database.

The sterically hindered fjord region diol epoxides, B[g]CDE and B[c]PDE, also bound substantially to adenines in codons 236, 239, 240, and 246 in exon 7, although the levels of adduc-



**Fig. 5. Panel A:** distribution of polycyclic aromatic hydrocarbon (PAH) diol epoxide adducts along exon 8 of the p53 gene. Concentrations of PAH compounds were 25  $\mu$ M CDE (lane 7), 10  $\mu$ M 5-MCDE (lane 8), 5  $\mu$ M, 10  $\mu$ M, and 50  $\mu$ M 6-MCDE (lanes 9, 10, and 11, respectively), 10  $\mu$ M and 50  $\mu$ M B[g]CDE (lanes 12 and 13, respectively), and 10  $\mu$ M and 50  $\mu$ M B[c]PDE (lanes 14 and 15, respectively). Lanes 1–4: Maxam–Gilbert sequencing controls. Lane 5: DNA from solvent control cells not treated with UvrABC. Lane 6: UvrABC-treated DNA from cells exposed to solvent control. Less DNA was loaded in the 50  $\mu$ M 6-MCDE lane. **Panel B:** quantitation of PAH diol epoxide damage in

exon 8 of p53 by sequence and codon. Codons that were not substantially damaged were excluded. Concentrations of PAH compounds used were 25  $\mu$ M CDE, 10  $\mu$ M 5-MCDE, 50  $\mu$ M 6-MCDE, 50  $\mu$ M B[g]CDE, 50  $\mu$ M B[c]PDE, and 4  $\mu$ M B[a]PDE (10). **Abbreviations used:** CDE = (+/-)-anti-chrysene-1,2-diol-3,4-epoxide; 5-MCDE = (+/-)-anti-5-methylchrysene-1,2-diol-3,4-epoxide; 6-MCDE = (+/-)-anti-6-methylchrysene-1,2-diol-3,4-epoxide; B[g]CDE = (+/-)-anti-benzo[g]chrysene-11,12-diol-13,14-epoxide; B[c]PDE = (+/-)-anti-benzo[c]phenanthrene-3,4-diol-1,2-epoxide; B[a]PDE = (+/-)-anti-benzo[a]pyrene-7,8-diol-9,10-epoxide.

tion were less than the quantities achieved at heavily damaged guanines. The majority of the mutations in the lung tumor database occur at guanines and are most frequently G-to-T transversions. Codons that have been reported to have frequent mutations at adenines in the human lung cancer database are 163 and 179. Codon 179, a prominent mutational hotspot in lung cancers (Fig. 2), does not contain a guanine on the nontranscribed strand and, therefore, is not an important target for DNA damage induced by B[a]PDE and chrysene metabolites. We observed little binding by B[g]CDE and B[c]PDE at the codon 179 sequence. The mutations observed at codon 179 are A-to-G transversions, a type of mutation that can be induced by fjord region PAHs (56,57). However, A-to-T transversion is a more common type of mutation induced by these compounds (56,57), which is rare (5% of all mutations) in the lung cancer database (3,54). An alternative hypothesis regarding the production of A-to-G transversions is exposure to nitric oxide (58). The adenine nucleotide most substantially damaged by fjord region PAHs occurs in the second position of codon 236. This lesion most likely would produce an A-to-T transversion, TAC (Tyr) to TTC (Phe), but this mutation has never been reported in the p53 database of more than 10 000 entries; therefore, it is likely that this event is not selected during tumorigenesis.

Although this study has focused on PAHs as a major class of

carcinogenic components of cigarette smoke, it does not exclude the possibility that other carcinogens present in smoke, such as NNK [i.e., 4-(methylnitrosamino)-1-(3-pyridyl)-1-butanone] (via pyridyloxobutyl adducts) or aromatic and heterocyclic amines, may have a similar sequence specificity.

We (10) previously reported that B[a]PDE-induced DNA damage occurred at several lung cancer mutational hotspots in the p53 gene. In this article, DNA damage was mapped for the diol epoxide forms of other PAHs, which may be present in the complex mixture of cigarette smoke. The DNA damage induced by these compounds maps, at least in part, to the same codons previously reported for B[a]PDE. The data show that B[a]PDE is a reasonable model compound for the strongest sites of adduction; however, other PAH metabolites probably contribute to the total load of adducts in the p53 gene of smokers and, therefore, may contribute to the mutational spectrum observed in human lung tumors.

## REFERENCES

- (1) Harris CC. p53 tumor suppressor gene: from the basic research laboratory to the clinic—an abridged historical perspective. *Carcinogenesis* 1996;17:1187–98.
- (2) Greenblatt MS, Bennett WP, Hollstein M, Harris CC. Mutations in the p53 tumor suppressor gene: clues to cancer etiology and molecular pathogenesis. *Cancer Res* 1994;54:4855–78.

- (3) Bennett WP, Hussain SP, Vahakangas KH, Khan MA, Shields PG, Harris CC. Molecular epidemiology of human cancer risk: gene-environment interactions and p53 mutation spectrum in human lung cancer. *J Pathol* 1999; 187:8-18.
- (4) Devesa SS, Blot WJ, Stone BJ, Miller BA, Tarone RE, Fraumeni JF Jr. Recent cancer trends in the United States. *J Natl Cancer Inst* 1995;87: 175-82.
- (5) Hecht SS, Carmella SG, Murphy SE, Foiles PG, Chung FL. Carcinogen biomarkers related to smoking and upper aerodigestive tract cancer. *J Cell Biochem Suppl* 1993;17F:27-35.
- (6) Harvey RG. Polycyclic aromatic hydrocarbons: chemistry and carcinogenicity. Cambridge (U.K.): Cambridge University Press; 1991.
- (7) Wynder EL, Hoffmann D. A study of tobacco carcinogenesis. VII. The role of higher polycyclic hydrocarbons. *Cancer* 1959;12:1079-86.
- (8) Van Duuren BL. The polynuclear aromatic hydrocarbons in cigarette smoke condensate II. *J Natl Cancer Inst* 1958;21:623-30.
- (9) Cavalieri EL, Rogan EG. Central role of radical cations in metabolic activation of polycyclic aromatic hydrocarbons. *Xenobiotica* 1995;25: 677-88.
- (10) Denissenko MF, Pao A, Tang M, Pfeifer GP. Preferential formation of benzo[a]pyrene adducts at lung cancer mutational hotspots in P53. *Science* 1996;274:430-2.
- (11) Tornaletti S, Pfeifer GP. Complete and tissue-independent methylation of CpG sites in the p53 gene: implications for mutations in human cancers. *Oncogene* 1995;10:1493-9.
- (12) Rideout WM 3d, Coetzee GA, Olumi AF, Jones PA. 5-Methylcytosine as an endogenous mutagen in the human LDL receptor and p53 genes. *Science* 1990;249:1288-90.
- (13) Denissenko MF, Chen JX, Tang MS, Pfeifer GP. Cytosine methylation determines hot spots of DNA damage in the human P53 gene. *Proc Natl Acad Sci U S A* 1997;94:3893-8.
- (14) Chen JX, Zeng Y, West M, Tang MS. Carcinogens preferentially bind at methylated CpG in the p53 mutational hot spots. *Cancer Res* 1998;58: 2070-5.
- (15) Denissenko MF, Pao A, Pfeifer GP, Tang M. Slow repair of bulky DNA adducts along the nontranscribed strand of the human p53 gene may explain the strand bias of transversion mutations in cancers. *Oncogene* 1998; 16:1241-7.
- (16) Eisenstadt E, Warren AJ, Porter J, Atkins D, Miller JH. Carcinogenic epoxides of benzo[a]pyrene and cyclopenta[cd]pyrene induce base substitutions via specific transversions. *Proc Natl Acad Sci U S A* 1982;79: 1945-9.
- (17) Mazur M, Glickman B. Sequence specificity of mutations induced by benzo[a]pyrene-7,8-diol-9,10-epoxide at endogenous aprt gene in CHO cells. *Somat Cell Mol Genet* 1988;14:393-400.
- (18) Yang JL, Maher VM, McCormick JJ. Kinds of mutations formed when a shuttle vector containing adducts of (+/-)-7 beta, 8 alpha-dihydroxy-9 alpha, 10 alpha-epoxy-7,8,9,10-tetrahydrobenzo[a]pyrene replicates in human cells. *Proc Natl Acad Sci U S A* 1987;84: 3787-91.
- (19) Chen RH, Maher VM, McCormick JJ. Effect of excision repair by diploid human fibroblasts on the kinds and locations of mutations induced by (+/-)-7beta,8alpha-dihydroxy-9alpha,10alpha-epoxy-7,8,9,10-tetrahydrobenzo[a]pyrene in the coding region of the HPRT gene. *Proc Natl Acad Sci U S A* 1990; 87:8680-4.
- (20) Jelinsky SA, Liu T, Geacintov NE, Loechler EL. The major N2-Gua adduct of (+)-anti-benzo[a]pyrene diol epoxide is capable of inducing G→A and G→C, in addition to G→T mutations. *Biochemistry* 1995;34: 13545-53.
- (21) Coulondre C, Miller JH, Farabaugh PJ, Gilbert W. Molecular basis of base substitution hotspots in *Escherichia coli*. *Nature* 1978;274:775-80.
- (22) Jones PA. DNA methylation errors and cancer. *Cancer Res* 1996;56: 2463-7.
- (23) Hernandez-Boussard T, Hainaut P. A specific spectrum of p53 mutations in lung cancer from smokers: review of mutations compiled in the IARC p53 database [published erratum appears in *Environ Health Perspect* 1998;106: A421]. *Environ Health Perspect* 1998;106:385-91.
- (24) International Agency on the Research of Cancer (IARC). Polynuclear aromatic compounds, part 1, chemical, environmental, and experimental data. IARC Monogr Eval Carcinog Risk Chem Hum, vol 32. Lyon (France): IARC; 1983. p. 37-8.
- (25) Hoffmann D, Hoffmann I. The changing cigarette, 1950-1995. *J Toxicol Environ Health* 1995;50:307-64.
- (26) Hecht SS, Bondinell WE, Hoffmann D. Chrysene and 5-methylchrysenes: presence in tobacco smoke and carcinogenicity. *J Natl Cancer Inst* 1974; 53:1121-33.
- (27) McKay JF, Latham DR. Polyaromatic hydrocarbons in high boiling petroleum distillates: isolation by gel permeation chromatography and identification by fluorescence spectrometry. *Anal Chem* 1973;45:1050-5.
- (28) Glatt H, Wameling C, Elsberg S, Thomas H, Marquardt H, Hewer A, et al. Genotoxicity characteristics of reverse diol-epoxides of chrysene. *Carcinogenesis* 1993;14:11-9.
- (29) Buening MK, Levin W, Karle JM, Yagi H, Jerina DM, Conney AH. Tumorigenicity of bay-region epoxides and other derivatives of chrysene and phenanthrene in newborn mice. *Cancer Res* 1979;39:5063-8.
- (30) Amin S, Huie K, Melikian AA, Leszczynska JM, Hecht SS. Comparative metabolic activation in mouse skin of the weak carcinogen 6-methylchrysene and the strong carcinogen 5-methylchrysene. *Cancer Res* 1985; 45:6406-12.
- (31) Hecht SS, Amin S, Huie K, Melikian AA, Harvey RG. Enhancing effect of a bay region methyl group on tumorigenicity in newborn mice and mouse skin of enantiomeric bay region diol epoxides formed stereoselectively from methylchrysenes in mouse epidermis. *Cancer Res* 1987;47: 5310-5.
- (32) Melikian AA, Prahalad KA, Amin S, Hecht SS. Comparative DNA binding of polynuclear aromatic hydrocarbons and their dihydrodiol and bay region diolepoxide metabolites in newborn mouse lung and liver. *Carcinogenesis* 1991;12:1665-70.
- (33) You L, Wang D, Galati AJ, Ross JA, Mass MJ, Nelson GB, et al. Tumor multiplicity, DNA adducts and K-ras mutation patterns of 5-methylchrysene in strain A/J mouse lung. *Carcinogenesis* 1994;15:2613-8.
- (34) Glatt H, Piee A, Pauly K, Steinbrecher T, Schrode R, Oesch F, et al. Fjord- and bay-region diol-epoxides investigated for stability, SOS induction in *Escherichia coli*, and mutagenicity in *Salmonella typhimurium* and mammalian cells. *Cancer Res* 1991;51:1659-67.
- (35) Levin W, Wood AW, Chang RL, Ittah Y, Croisy-Delcey M, Yagi H, et al. Exceptionally high tumor-initiating activity of benzo(c)phenanthrene bay-region diol-epoxides on mouse skin. *Cancer Res* 1980;40:3910-4.
- (36) Dipple A, Pigott MA, Agarwal SK, Yagi H, Sayer JM, Jerina DM. Optically active benzo[c]phenanthrene diol epoxides binding extensively to adenine in DNA. *Nature* 1987;327:535-6.
- (37) Amin S, Krzeminski J, Rivenson A, Kurtzke C, Hecht SS, el-Bayoumy K. Mammary carcinogenicity in female CD rats of fjord region diol epoxides of benzo[c]phenanthrene, benzo[g]chrysene and dibenzo[a,l]pyrene. *Carcinogenesis* 1995;16:1971-4.
- (38) Giles AS, Seidel A, Phillips DH. Covalent DNA adducts formed in mouse epidermis by benzo[g]chrysene. *Carcinogenesis* 1996;17:1331-6.
- (39) Amin S, Desai D, Dai W, Harvey RG, Hecht SS. Tumorigenicity in newborn mice of fjord region and other sterically hindered diol epoxides of benzo[g]chrysene, dibenzo[a,l]pyrene (dibenzo[def,p]chrysene), 4H-cyclopenta[def]chrysene and fluoranthene. *Carcinogenesis* 1995;16: 2813-7.
- (40) Yuan W, Kiselyov AS, Harvey RG, Carothers AM. Mutagenic specificity of syn-benzo[g]chrysene 11,12-dihydrodiol 13,14-epoxide in the dihydrofolate reductase gene of Chinese hamster ovary cells. *Carcinogenesis* 1995; 16:2869-73.
- (41) Giles AS, Seidel A, Phillips DH. *In vitro* reaction with DNA of the fjord-region diol epoxides of benzo[g]chrysene and benzo[c]phenanthrene as studied by <sup>32</sup>P-postlabeling. *Chem Res Toxicol* 1995;8:591-9.
- (42) Pataki J, Lee H, Harvey RG. Carcinogenic metabolites of 5-methylchrysene. *Carcinogenesis* 1983;4:399-402.
- (43) Amin S, Huie K, Hecht SS, Harvey RG. Synthesis of 6-methylchrysene-1,2-diol-3,4-epoxides and comparison of their mutagenicity to 5-methylchrysene-1,2-diol-3,4-epoxides. *Carcinogenesis* 1986;7:2067-70.
- (44) Kiselyou AA, Lee HM, Harvey RG. Efficient syntheses of the anti-diol and syn-diol epoxide metabolites of the carcinogenic polycyclic aromatic hydrocarbon benzo[g]chrysene. *J Org Chem* 1995;60:6123-8.
- (45) Krzeminski J, Lin JM, Amin S, Hecht SS. Synthesis of Fjord region diol epoxides as potential ultimate carcinogens of dibenzo[a,l]pyrene. *Chem Res Toxicol* 1994;7:125-9.
- (46) Misra B, Amin S. An improved synthesis of anti-benzo[c]phenanthrene-

- 3,4-diol-1,2-epoxide via 4-methoxybenzo[*c*]phenanthrene. *J Org Chem* 1990;55:4478–80.
- (47) Tang M-s. Mapping and quantification of bulky chemical-induced DNA damage using UvrABC nucleases. In: Pfeifer GP, editor. *Technologies for detection of DNA damage and mutations*. New York (NY): Plenum Press; 1996. p. 139–53.
- (48) Tang M-s, Zheng JB, Denissenko MF, Pfeifer GP, Zheng Y. Use of UvrABC nuclease to quantify benzo[*a*]pyrene diol epoxide–DNA adduct formation at methylated versus unmethylated CpG sites in the p53 gene. *Carcinogenesis* 1999;20:1085–9.
- (49) Tornaletti S, Rozek D, Pfeifer GP. The distribution of UV photoproducts along the human p53 gene and its relation to mutations in skin cancer [published erratum appears in *Oncogene* 1993;8:3469]. *Oncogene* 1993;8:2051–7.
- (50) Tornaletti S, Pfeifer GP. Slow repair of pyrimidine dimers at p53 mutation hotspots in skin cancer. *Science* 1994;263:1436–8.
- (51) Tornaletti S, Pfeifer GP. Ligation-mediated PCR for analysis of UV damage. In: Pfeifer GP, editor. *Technologies for detection of DNA damage and mutations*. New York (NY): Plenum Press; 1996. p. 199–209.
- (52) DeMarini DM, Shelton ML, Bell DA. Mutation spectra in *Salmonella* of complex mixtures: comparison of urban air to benzo[*a*]pyrene. *Environ Mol Mutagen* 1994;24:262–75.
- (53) DeMarini DM, Shelton ML, Levine JG. Mutation spectrum of cigarette smoke condensate in *Salmonella*: comparison to mutations in smoking-associated tumors. *Carcinogenesis* 1995;16:2535–42.
- (54) Hainaut P, Hernandez T, Robinson A, Rodriguez-Tome P, Flores T, Hollstein M, et al. IARC database of p53 gene mutations in human tumors and cell lines: updated compilation, revised formats and new visualisation tools. *Nucleic Acids Res* 1998;26:205–13.
- (55) Denissenko MF, Koudriakova TB, Smith L, O'Connor TR, Riggs AD, Pfeifer GP. The p53 codon 249 mutational hotspot in hepatocellular carcinoma is not related to selective formation or persistence of aflatoxin B1 adducts. *Oncogene* 1998;17:3007–14.
- (56) Wei SJ, Chang RL, Cui XX, Merkler KA, Wong CQ, Yagi H, et al. Dose-dependent differences in the mutational profiles of (–)-(1*R*,2*S*,3*S*,4*R*)-3,4-dihydroxy-1,2-epoxy-1,2,3,4-tetrahydrobenzo[*c*]phenanthrene and its less carcinogenic enantiomer. *Cancer Res* 1996;56:3695–703.
- (57) Bigger CA, St John J, Yagi H, Jerina DM, Dipple A. Mutagenic specificities of four stereoisomeric benzo[*c*]phenanthrene dihydrodiol epoxides. *Proc Natl Acad Sci U S A* 1992;89:368–72.
- (58) Routledge MN, Wink DA, Keefer LK, Dipple A. Mutations induced by saturated aqueous nitric oxide in the pSP189 supF gene in human Ad293 and *E. coli* MBM7070 cells. *Carcinogenesis* 1993;14:1251–4.
- (59) Gao HG, Chen JK, Stewart J, Song B, Rayappa C, Whong WZ, et al. Distribution of p53 and K-ras mutations in human lung cancer tissues. *Carcinogenesis* 1997;18:473–8.

## NOTES

*Present address:* M. F. Denissenko, PharMingen, San Diego, CA.

*Present address:* M.-s. Tang, Department of Environmental Medicine, New York University, Tuxedo.

Supported by grant 6RT-0361 (to G. P. Pfeifer) from the University of California Tobacco Related Disease Research Program and by Public Health Service grant ES03124 (to M.-s.Tang) from the National Institute of Environmental Health Sciences, National Institutes of Health, Department of Health and Human Services.

Manuscript received October 19, 1999; revised February 22, 2000; accepted March 7, 2000.