Carcinogenesis vol.19 no.3 pp.387–393, 1998

Human microsomal epoxide hydrolase: 5′-flanking region genetic polymorphisms

Stuart Raaka, Christopher Hassett and Curtis J.Omiecinski[1]

Department of Environmental Health, University of Washington, 4225 Roosevelt Way NE, #100, Seattle, WA 98105–6099, USA

[1]To whom correspondence should be addressed
Email: cjo@u.washington.edu

Microsomal epoxide hydrolase (mEH) catalyses the hydrolysis of xenobiotic epoxides, including various epoxide derivatives of the procarcinogenic polyaromatic hydrocarbons. Levels of mEH enzymatic activity among different cell types and between individuals within the population vary considerably. Genetic polymorphisms within the structural region of the human *mEH* gene exist and appear to contribute to the population variance in functional expression. In this study, we used single strand conformational polymorphism analysis and direct DNA sequencing approaches to identify seven additional polymorphic sites within the upstream region of the *mEH* gene, spanning –743 to +185 bp, relative to the transcription initiation site. Allelic frequencies and linkages of the polymorphic nucleotides were determined in 51 individuals using restriction fragment length polymorphism or competitive oligonucleotide priming assays. To determine the functional significance of the individual nucleotide substitutions, DNA fragments representing the variant alleles were cloned into the heterologous pBRAMScat2 reporter vector, transfected into HepG2 cells and assessed for reporter gene expression. Results indicated that certain of these polymorphic loci might differentially regulate transcription, with the maximum contribution of any of the variants modifying levels of reporter gene activity by ~ 30%. These observations establish that genetic variation in the 5′ flanking sequence of *mEH* gene is likely an additional contributing factor to the range of functional mEH expression existing in human populations.

Introduction

Microsomal epoxide hydrolase (mEH*, E.C. 3.3.2.3) catalyses, with broad substrate specificity, the conversion of highly reactive, and cytotoxic arene oxides and aliphatic epoxides to less toxic *trans*-dihydrodiols (1). Substrates for mEH are principally xenobiotic epoxide metabolites generated by Phase I oxidation reactions. Therefore, mEH biotransformation typically results in detoxification and preparation for Phase II conjugation reactions. However, in certain cases, mEH also participates in bioactivation pathways, with a well-characterized example being the metabolic conversion of benzo[*a*]pyrene (a component of cigarette smoke) to the highly mutagenic (+)-*anti*–7,8-diol–9,10-epoxide (2).

The mEH enzyme is expressed in all tissues thus far examined (3,4) with highest levels in the liver, kidney, and testis and 10–100-fold lower levels in the lung and lymphocytes (5). Within cells, mEH is localized mainly to the endoplasmic reticulum where it can transiently associate with the cytochrome P-450 mixed function oxygenase system (6). Endogenous substrates for mEH have not been readily identified. However, the mEH-catalysed hydrolysis of certain steroid epoxides (7), indicates potential involvement in steroid metabolism. Additionally, the high degree of mEH structural conservation between several mammalian species (8) and apparent ubiquitous tissue expression imply an important role in cellular metabolism.

There is evidence for polymorphic mEH expression within the human population. Interindividual differences in mEH activity ranging in scale from several to 40-fold have been reported in various human tissue types (5,9,10). Clinical studies have demonstrated associations between low mEH enzyme activity levels, and adverse drug responses or disease states. Among these are reports that low levels of mEH activity correlate with the occurrence of fetal hydantoin syndrome (11,12). Individuals displaying hepatotoxicity in response to phenytoin and acetaminophen exposure have decreased levels of mEH activity when compared to control subjects (13,14). In contrast, several other studies have not found an association between mEH expression or genotype, and adverse anticonvulsant drug responses (15–17).

Heckbert *et al*. (18) reported a consistent, but statistically non-significant association between low mEH activity and the occurrence of lung and other smoking-related cancers. Recently, McGlynn *et al*. (19) reported the association of a mEH exon 3 polymorphism, which reduces enzyme activity *in vitro*, with the occurrence of hepatocellular carcinoma in an aflatoxin and hepatitis B-exposed population. Interestingly, Lancaster *et al*. (20) observed an association between the higher activity exon 3 allele and the occurrence of ovarian cancer. These findings imply mEH involvement in both prevention and potentiation of the carcinogenic process.

The molecular basis for variation in mEH activity has not been characterized completely. As indicated above, genetic polymorphisms have been identified within the coding region of the gene (15,21) which result in amino acid substitutions at two positions (residue 113Tyr/His; residue 139Arg/His). *In vitro* expression analyses indicated that the corresponding proteins varied as much as 65% in relative activity (21). However, these alterations were more likely the result of altered protein stability, and not enzyme specific activity. While these structural polymorphisms of the enzyme are likely biologically and clinically relevant by themselves, they do not appear to account for the complete spectrum of mEH activity variation within the population.

Many human genes, including cytochrome P450 2E1 (22),

***Abbreviations:** mEH, microsomal epoxide hydrolase; SSCP, single strand conformational polymorphism; RFLP, restriction fragment length polymorphism; COP, competitive oligonucleotide priming; CAT, chloramphenicol acetyl transferase; TK, thymidine kinase; ASO, allele-specific oligomer; SEAP, secreted alkaline phosphatase.

apolipoprotein AI (23), CIII (24) and IL–4 (25), contain gene promoter region polymorphisms that modulate their respective transcriptional expression. We hypothesized that functional genetic polymorphisms are present in the mEH promoter region, which contribute to the range of variation noted previously among individuals. In this report, we describe the identification and functional characterization of polymorphisms in the promoter region of the human *mEH* gene. Seven polymorphic loci were identified by polymerase chain reaction-single strand conformational polymorphism (PCR-SSCP) analysis and by DNA sequencing. Allelic frequencies were determined either by DNA sequencing, restriction fragment length polymorphism (RFLP) or competitive oligonucleotide priming (COP) analysis. To ascertain the functional significance of nucleotide substitutions, DNA fragments containing the variant residues were cloned into heterologous pBRAMScat2 chloramphenicol acetyl transferase (CAT) reporter vectors containing the thymidine kinase (TK) promoter. These chimeric DNA constructs were transiently transfected and expressed in human hepatoma HepG2 cells. The results of these experiments demonstrate that polymorphic loci present in the 5′ region of the *mEH* gene differentially regulate transcriptional activity.

Materials and methods

PCR-SSCP analysis

Genomic DNA was obtained from a study of predominately Caucasian male workers in western Washington State (26). DNA samples were PCR amplified and radiolabeled using the primer sequences and annealing temperatures shown in Table I, panel A. The locations of the PCR products within the mEH promoter, relative to the transcription start site (27), also are indicated in Table I. Each amplification reaction contained 2.5U Taq polymerase (Promega, Madison, WI), 1.5 mM $MgCl_2$, 0.5 μg genomic DNA, 100 μM dNTPs, 0.2 μM primers and 2.0 μCi of [α-^{33}P] dATP (3000 Ci/mM, NEN) in a volume of 25 μl. PCR conditions included initial denaturation at 95°C for 3 min, addition of polymerase and dNTP mix at 85°C, followed by 30 cycles with 1 min denaturation at 94°C, 45 s annealing (see Table I for temperatures) and 72°C extension for 45 s. Reaction products were analysed by agarose gel electrophoresis to verify expected size. SSCP analyses were performed at room temperature using 6% acrylamide/7.5% glycerol gels and electrophoresed at 8–10 W for 14–18 h, as described by Orita *et al*. (28). For analysis, gels were dried and subjected to autoradiography.

DNA sequencing

Genomic DNA samples that exhibited unique patterns with SSCP analysis were PCR amplified, purified with Microcon-100 filters (Amicon, Beverly, MA) and sequenced on both strands using the *fmol*TM cycle sequencing system (Promega, Madison, WI). Sequence analysis employing Sequenase Version 2.0 T7 DNA polymerase (USB, Cleveland, OH) was used to verify plasmid identity (see below).

Genotype determination and allelic frequency

Genomic DNA samples were PCR-amplified with primers FP6 and RPS4 (Table I, panel B), ethanol precipitated, digested and analysed by agarose gel electrophoresis. Six of the identified polymorphic residues occurred within restriction enzyme recognition sites thereby allowing analyses by RFLP. Since RFLP analysis was not possible at one polymorphic site, a modification of competitive oligonucleotide priming (COP) analysis (29) was employed. The allele specific oligonucleotides (ASO), RP–362C and RP–362T (Table I, panel C) contained the variant base at the 3′ end, rather than internally. Five ng of FP6/RPS4-derived PCR product obtained from each genomic DNA was used as a template in two separate PCR reactions containing 0.1 μM of each unlabeled ASO, 5×10^5 dpm of either RP–362C or RP–362T ^{32}P-labeled ASO, 0.2 μM FPS3 and other PCR reagents as indicated above. PCR primers with perfect complementarity to the template DNA at the variable nucleotide are preferentially incorporated in the amplification reaction. Products were electrophoresed in 1.5% agarose gels which were dried and analysed by autoradiography. Hardy–Weinberg equilibrium was assessed for each polymorphic locus using the χ^2 test at significance level of $P < 0.05$.

Plasmid construction and preparation.

Fragments of mEH 5′ DNA were generated by PCR amplification of genomic DNA using the primer set shown in Table I, panel B. In several instances, other amplification primers containing mismatches relative to the template were used to create restriction sites and facilitate cloning. The PCR products were digested with appropriate restriction enzymes (see below), agarose gel purified and ligated into the reporter vector pBRAMScat2, which contains the thymidine kinase promoter (30). DH5α *E.coli* cells were transformed and resulting colonies screened either by PCR or grid hybridization for the presence of mEH DNA sequence. Plasmids were prepared by the method of Lee and Rasheed (31), and analysed by sequencing and restriction digestion to verify insert size, junctions and polymorphic sites. Positive clones were prepared in triplicate, quantified spectrophotometrically, and further assessed by agarose gel electrophoresis with quantitative and qualitative (supercoiling) comparison before use in cell transfection experiments.

Cell transfection

Human hepatoma cells (HepG2, ATCC #HB–8065) were grown in D-MEM/F–12 media (Gibco/Life Technologies, Grand Island, NY) supplemented with 5% Nu-Serum (Collaborative Research, Bedford, MA), 100 U/ml penicillin and 0.1 mg/ml streptomycin (P/S; Sigma, St Louis, MO) in 75 cm^2 flasks

Table I. Oligomer primers used in amplification reactions. Forward (sense) and reverse (antisense) orientations are designated by FP and RP, respectively, in the primer name. Annealing temperature is °C. Genomic location is with respect to the transcription start site (+1). Product size incorporated amplificiation primers. Primers were employed as follows: Panel A, SSCP analyses; Panel B, amplification of genomic DNA samples for RFLP determination; Panel C, COP analysis.

	Primer set	Sequence (5′→3′)	T° anneal	Genomic location	Product size (bp)
A	FP1	CCACGCTGTGCACACATGAG	58	–25→+185	210
	RP1	CCCAGTGCTTTCACCTGTCC			
	FP2	GCCCTCCTCTTCCTG	54	–193→+57	250
	RP2	GGCACTGGCCCCCAC			
	FP5	CCTGCTGTCACACCAGGC	58	–430→–165	266
	RP5	CCAGGTGTATCACCCAGG			
	FPS3	TAGCTCTCCGACAAAGCAGACG	65	–613→–370	244
	RPS3	AATAGCAGGTGGTTAATGCACC			
	FP6	GGTCTAGGAATTGTCAAGCGCCCAG	54	–743→–538	206
	RP4	CAAGGAGCCCACCCT			
B	FP6	GGTCTAGGAATTGTCAAGCGCCCAG	65	–743→+82	825
	RPS4	TGTGAGAGAGGCAGGGCAGGCTGC			
C	FPS3	TAGCTCTCCGACAAAGCAGACG	63.5	–613→–343	271
	RP-362C	TTGCCACCTAGTAACAGCCC			
	RP-362T	TTGCCACCTAGTAACAGCCT			

(Becton Dickinson, Lincoln Park, NJ). Duplicate flasks were seeded for each transfected DNA construct, and triplicate plasmid preparations were employed for each construct in order to minimize artifactual results potentially arising from either qualitative or quantitative variation. Thus, transfection assay results were determined from the independent analyses of six flasks per plasmid construct. At ~ 40% cell confluency, media was changed to D-MEM/F–12 supplemented with insulin, transferrin, and selenium (ITS^{TM} culture supplement; Collaborative Biomedical Products, Bedford, MA) and P/S for 8 h. At that time, plasmid DNA transfections were performed using the Lipofectin^{TM} reagent (32) and Opti-MEM^{TM} media (Gibco/Life Technologies, Grand Island, NY). Each flask was co-transfected with 10 µg of chimeric mEH/pBRAMScat2 plasmid DNA and 2.5 µg of the plasmid pBC12/RSV/SEAP, which contains a secreted alkaline phosphatase (SEAP) reporter gene (33). This plasmid was included as an internal standard for normalization of transfection efficiency. Other controls included mock-transfected flasks (no DNA) and transfection with pBRAMScat2 containing no mEH inserts. Six h post-transfection, media containing ITS supplemented with BSA and linoleic acid (ITS^{+TM}; Collaborative Biomedical Products) was added. At 48 h post-transfection, aliquots of media were removed for SEAP analysis, cells were lysed and extracts were assayed for total protein content (BCA Assay, Pierce, Rockford, IL) and CAT protein content (CAT ELISA kit, Boehringer Mannheim, Indianapolis, IN). SEAP (triplicate), protein (triplicate) and CAT assays (duplicate) were analysed using a Molecular Devices 96-well plate reader with Delta Soft II software, version 4.14. Coefficients of variation of $<$10% were obtained for all assays. CAT ELISA results were normalized to total protein and SEAP activity values obtained from identical flasks.

Results

Identification and localization of polymorphic loci

PCR-SSCP screening of the 5′ region of the *mEH* gene was accomplished using the five primer sets shown in Table I, panel A. Using this approach, no polymorphisms were identified in the two-sample sets which encompassed the region spanning from –193 to +185 of the *mEH* gene (data not shown). However, analysis of 24 genomic DNA samples suggested that three of the five regions contained polymorphic loci. Figure 1A, an autoradiogram obtained from the SSCP assessment of the region spanning –430 to –165, demonstrated a complex banding pattern suggesting the presence of multiple polymorphic sites. In contrast, screening the regions –743 to –538 (Figure 1B) and –613 to –370 (Figure 1C), demonstrated simple banding patterns implying that fewer polymorphisms were present.

DNA sequence was determined for putative polymorphic regions in order to identify the exact polymorphic locations and base substitutions. Seven polymorphic sites were identified, and the positions with base substitutions (shown in Figure 2) are as follows: –200C/T, –259C/T, –290T/G, –362A/G, –399T/C, –613T/C, and –699T/C. PCR cycle sequencing enabled the identification of polymorphic sites and determination of genotype from a single reaction, as shown for four individuals at the –362A/G site (Figure 3). In this example, individual #1 is a heterozygote, #2 and #3 are homozygous –362G, and #4 is homozygous –362A.

Frequency analysis

To assess polymorphic site frequencies, genomic DNA samples from 51 individuals were genotyped by RFLP or COP analysis. RFLP characterization was employed for the polymorphic sites located at –200 (Mae III), –259 (Ssp I), –290 (Ban I), –399 (Bsl I), –613 (Nhe I) and –699 (Dde I). From this analysis we concluded that the polymorphic allele frequencies were not equal. For instance, variant residues –200T, –259T and –290G were detected in 90/102 alleles analysed (88%). Likewise, the polymorphic residues –613C and –699C were the predominant variants, observed in 69/102 alleles (68%). RFLP analysis of the DNA samples revealed a frequency of ~82% (83/102) for the –399C variant.

Because a diagnostic restriction site was not present, the –362 polymorphism was analysed by COP analysis. Characteristic results of this analysis, obtained from the DNA of 15 individuals, are presented in Figure 4. Agarose gels shown in Figure 4A and B (upper panels) demonstrate the presence of a prominent 271 bp product derived from the primers FPS3 and either RP–362C or RP–362T, respectively. Figure 4A and B (lower panels) present the autoradiograms obtained from the dried gels. The PCR bands seen in Figure 4A contained the radiolabeled RP–362C primer that is preferentially incorporated in products containing the 362G allele. Likewise, the autoradiogram in Figure 4B demonstrates the incorporation of the radiolabeled RP–362T primer, indicative of the 362A allele. From this Figure, it is concluded that individuals 1, 3,

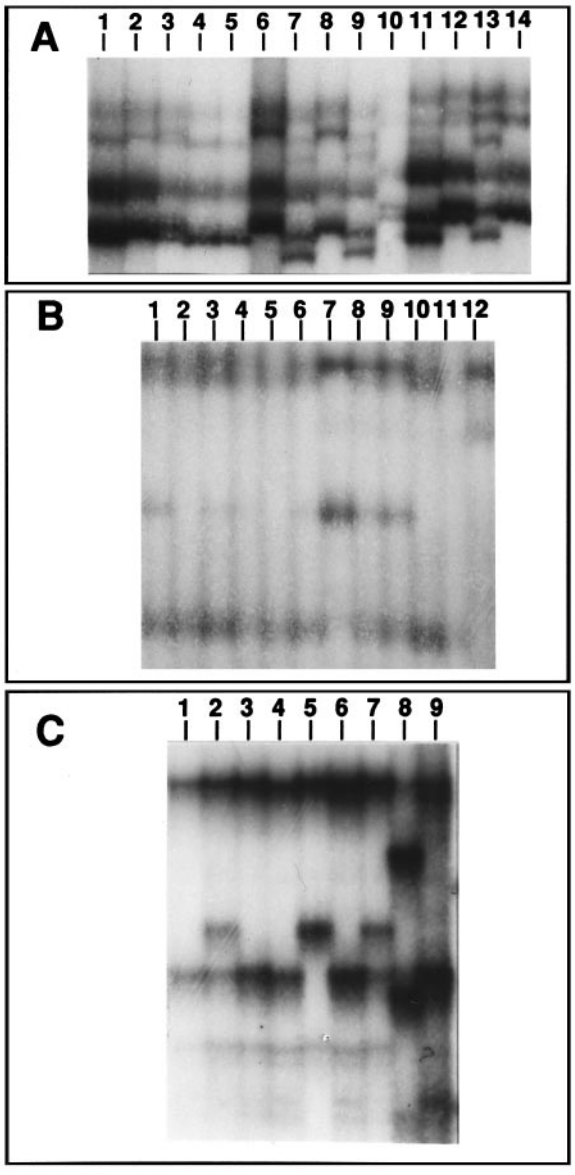

Fig. 1. SSCP analysis of three regions of the human *mEH* gene. Each lane was loaded with radiolabeled PCR product from various genomic DNA templates. Differences in the banding patterns suggest the presence of polymorphic sites within the analysed DNA fragments. Panel **A** was generated using the FP5/RP5 primer set and covers the region from –430 to –165. Panel **B** screened from –743 to –538 with the FP6/RP4 primers and panel **C** was obtained from the FPS3/RPS3 primers and covers –613 to –370. Identical lane numbers in the three panels do not represent the same DNA samples.

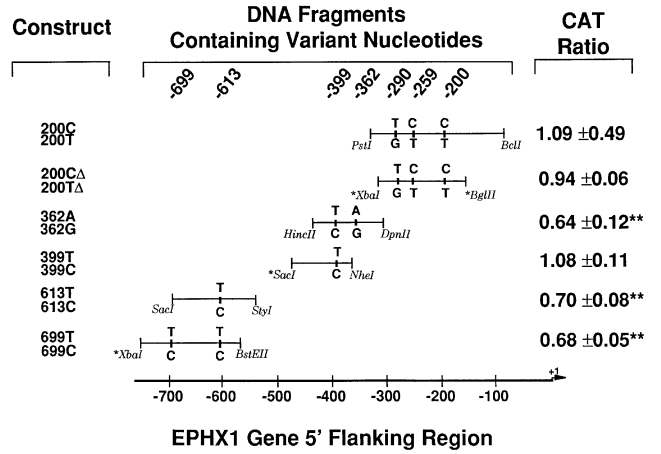

Fig. 2. Polymorphic loci in the 5′ region of the *mEH* gene. The mEH 5′ flanking region is depicted at the bottom and the polymorphic site locations are shown above at –699, –613, –399, –362, –290, –259 and –200. Each allelic construct, designated at the left, contains all of the polymorphic residues indicated. Restriction sites used for cloning into pBRAMScat2 are presented at the boundaries of each fragment containing relevant polymorphic residues. *Indicates restriction site created by PCR. The CAT ratios determined for orthologous clones (for example 200C/200T) and shown on the far right are means ± standard deviations derived from two to four experiments. Those marked with ** indicate statistical significance at $P < 0.05$, compared with ratios of unity, using standard *t*-test analyses.

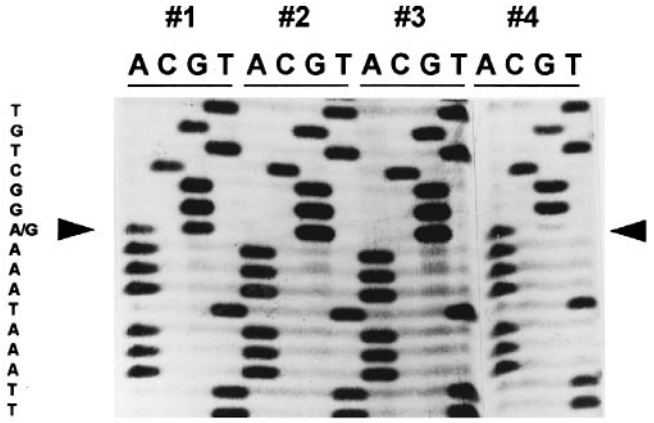

Fig. 3. PCR cycle sequence analysis of genomic DNA samples. Arrows indicate the polymorphic site at –362. Individual #1 is a heterozygote as seen by bands in both the A and G lanes. Samples #2 and #3 are homozygous G, and #4 is homozygous A.

13 and 14 are homozygous for 362A, while subjects 2, 5, 6, 7, 8, 11 and 12 are homozygous for 362G. Heterozygous individuals are 4, 9, 10 and 15. COP analysis of the –362 polymorphism for all 51 individuals analysed above revealed a frequency distribution identical to the –613 and –699 variants, with the G residue present in 68% of the alleles analysed.

Linkage analysis

Initial DNA sequence characterization, as well as allele frequency determinations, suggested that some of the polymorphic loci might be linked. Two linkages, each containing three polymorphic sites, were confirmed from homozygous individuals and occur as follows. The –200/–259/–290 alleles exist as either the combination C/C/T or T/T/G, while variants at –362/–613/–699 occur as either A/T/T or G/C/C, respectively. These are subsequently referred to as the –200 linkage (–200, –259, –290) and the –600 linkage (–362, –613, –699). The –399 site, which lies within the –600 linkage, appears to

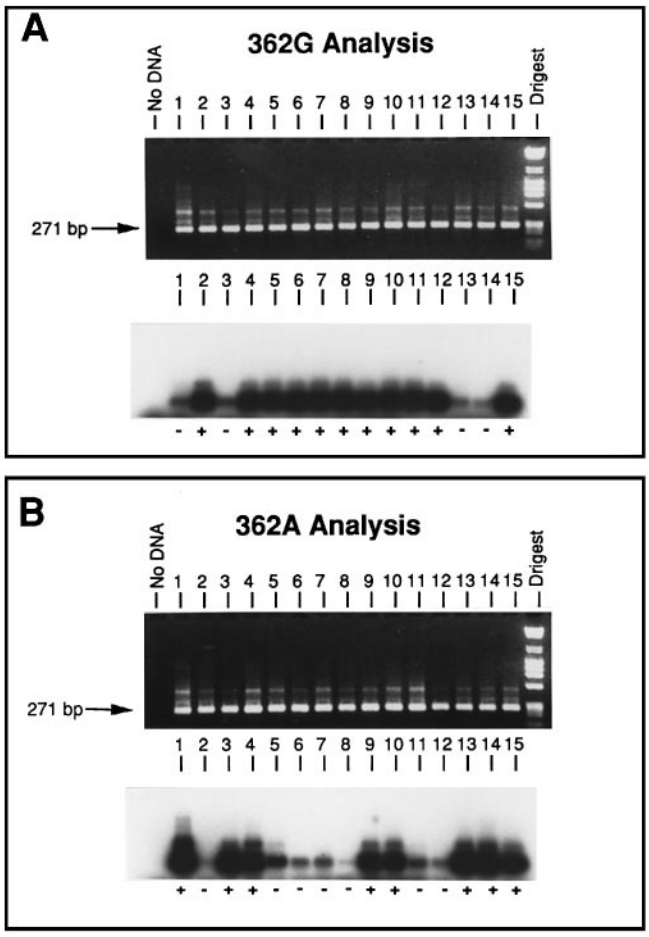

Fig. 4. Representative COP analysis. Fifteen genomic DNA samples were characterized for polymorphism at the –362 site by competitive oligonucleotide priming (COP). Panel **A** presents the results obtained from PCR amplification with the radiolabeled primer RP–362C. The upper half shows the ethidium bromide stained agarose gel with a prominent 271 bp PCR product in all samples except the negative control (no DNA). Below this is an autoradiogram of the dried gel showing preferential incorporation of radiolabel (+) in templates containing the –362G allele. Panel **B** is identical to panel A except the PCR was conducted with the radiolabeled RP–362T oligomer. Samples 4, 9, 10 and 15 are heterozygotes; 2, 5, 6, 7, 8, 11 and 12 are homozygous G; 1, 3, 13 and 14 are homozygous A.

be independent from this group as can be deduced from the allelic frequencies reported above.

The 51 genomic DNA samples used for the genotyping analyses were obtained from 42 Caucasians, five Asians, three native Americans and one Hispanic. Although the sample size was small, no marked deviation was evident for any of the races with respect to the linkages reported above. Each of the polymorphic alleles at the 200 and 600 linkages was observed in Caucasians, Asians and native Americans. Both alleles at the –399 site were detected in the Caucasians, Asians and the Hispanic. These observations demonstrate that the polymorphisms and their associated linkages, are present in several races. However, because allele frequencies often vary among races and the small sample size for some ethnic groups in this study, only the 42 Caucasians were included for population analysis.

The polymorphic sites and linkages, allele frequencies, genotype frequencies and expected frequencies at Hardy–Weinberg equilibrium are shown in Table II. Frequency determinations for the variant nucleotides within both linkages and

Table II. Polymorphic sites, allelic base substitutions (listed 3′ to 5′), allelic frequency, genotypic frequency and frequencies predicted at Hardy-Weinberg equilibrium. The data are derived from 42 Caucasian genomic DNA samples (84 alleles).

Linkage	Genotype (allele)	Allelic frequency	Genotype frequency	Expected frequency
200	–200T, –259T, –290G (TTG allele)	0.881	0.785	0.776
	–200T/C, –259T/C, –290G/T	–	0.190	0.210
	–200C, –259C, –290T (CCT allele)	0.119	0.024	0.014
600	–362G, –613C, –699C (GCC allele)	0.679	0.500	0.461
	–362G/A, –613C/T, –699C/T	–	0.357	0.436
	–362A, –613T, –699T (ATT allele)	0.321	0.143	0.103
none	–399C (C allele)	0.798	0.643	0.637
	–399C/T	–	0.309	0.322
	–399T (T allele)	0.202	0.048	0.041

the –399 site suggest each of these polymorphisms are in Hardy–Weinberg equilibrium.

Transfection experiments

To assess the functional significance of these sequence variations, 12 mEH DNA fragments containing the polymorphic sites were cloned into pBRAMScat2 reporter vectors using the restriction sites shown in Figure 2. The CAT ratios shown at the right of the figure were derived using the normalized CAT expression of clones from two to four experiments. CAT protein or SEAP activity was not detected in any of the mock transfections. Chimeric constructs containing either of the variant nucleotides at positions –200, –259, –290 and –399 demonstrated no consistent effect on CAT reporter expression. In contrast, constructs, which contain allelic variants at positions –362, –613 and –699, appeared to modulate CAT expression by ~ 30% in an allele specific manner.

Discussion

We have identified and characterized seven polymorphic sites in the mEH proximal promoter region. Six of these base changes are transitions: five of the C/T class and one of the A/G type. The –290 G/T polymorphism was the only transversion identified in this study. These findings are in general agreement with previous estimates showing that transitions occur more frequently than transversions, with pyrimidines favored over purines (34). The frequency of polymorphisms within this region of the human *mEH* gene (~ 1 per 100 bp) is slightly higher than the mean heterozygosity in the human genome of 1 per 250–300 bp estimated by Cooper *et al*. (35). Four of the *mEH* gene polymorphic sites (–362, –399, –613 and –699) are apparent by comparing recently published sequences (15,36). The three polymorphic loci in the 200 linkage are not evident by comparative sequence analysis, probably due to the relatively low frequency of the CCT allele. An additional non-coding polymorphism, recently identified at residue +66 (16), was not detected by the conditions employed in our SSCP analysis.

To our knowledge, the presence of two separate linkage groups within a small promoter region, each containing three variant bases, represents an unusual finding. Promoters of numerous genes contain single polymorphic linkages. For instance, the cytochrome P450 CYP2E1 promoter contains five variant sites between –1259 and –771 bp, which were reported by Hayashi *et al*. (22) to exist in complete linkage disequilibrium. Linkage such as this is consistent with a gene conversion event (37). Powers and Smithie (38) reported that multiple gene conversions between fetal γ and α-globin genes involve sequences <300 bp in length, and the linkages reported here are within this size range. To consider gene conversion as a mechanistic explanation for our observations in the human *mEH* gene, homologous sequence must be present on the same chromosome. However, computer analyses of the DNA sequence currently deposited in Genbank (Release 102.0) reveals no unexpected sequence homologies to the mEH 5′ upstream region.

The distal boundaries for the *mEH* gene linkages are not known. We have determined by DNA sequencing that the 200 linkage is not associated in a predictable manner with the variants at +66 (16), nor the exons 3 and 4 coding polymorphisms (21) which reside >6500 bp downstream. However, as exemplified by the –399 site within the 600 linkage, this does not preclude continued linkage downstream of the +66 site. The 5′ boundary of the 600 linkage has not been determined.

The CAT expression data summarized in Figure 2 suggest that two polymorphic loci in the human mEH 5′ region (–362 and –613) are each capable of modulating transcriptional activity by an average of ~ 30%. The importance of the –613 site (clones 613T/613C, CAT ratio = 0.70) is implied by the finding that no difference in the CAT expression ratio resulted when the –699 site was included in a similar construct (clones 699T/699C, CAT ratio = 0.68). Inclusion of the –399 site within the –362 clones (see Figure 2) is also not likely to contribute to the observed variation because its presence alone did not alter the ratio of expression activities (clones 399C/399T, CAT ratio = 1.08). It is possible that functional assessment of the entire linkage (–362 to –699) might reveal more significant differences in activity between the two alleles. Of interest, the 600 linkage is potentially not in Hardy–Weinberg equilibrium ($\chi^2 = 2.96$, $0.05 < P < 0.10$). The analysis of a larger population may reveal statistically significant differences between observed and expected frequencies.

The molecular mechanisms responsible for the observed differences in transcriptional activity are unknown. We have analysed the mEH promoter and the polymorphic positions for the presence of transcription factor binding sites using the compilation of Boulikas (39) and the TransFac/MatInspector database (40). The –290 site lies within a near consensus (12/13 bp) NF–1 recognition element (41). However, no consistent effect on transcriptional activity was detected between the variant 200 linkage alleles containing this site. We were not able to identify putative transcription factor recognition elements within or near the remaining polymorphic positions characterized in this report. The higher measured transcriptional activities of polymorphic fragments containing G or C at the variant positions, compared to A or T in the respective positions, imply that the increase in hydrogen bonding potential at these sites may be a mechanistic factor. Future analysis of

the mEH promoter region may help elucidate the molecular basis for these variations.

The *in vivo* effects of the 5′ polymorphisms identified in the current study were examined in human liver samples that were previously characterized for mEH activity level using benzo[*a*]pyrene–4,5-oxide as substrate (42). Each of the 5′ polymorphisms described above was determined in five low (256 ± 38 pmol/min/mg protein) and five high (1065 ± 215 pmol/min/mg protein) activity liver samples and comparatively analysed. No significant correlation between genotype and mEH activity was demonstrated (data not shown). However, when the genotypes were given a numerical ranking based on the *in vitro* experimental data and plotted against liver activities, the slope of the regressed line was positive, as would be expected if there was a correlation. It is possible that analyses of a larger data set might yield more insight into a potential genotype/phenotype correlation.

Taken together, the functional effects of 5′ upstream (this study) and coding region (21) *mEH* gene polymorphisms indicates that no single site accounts for the mEH activity variations observed in the general human population. However, combinations of 'low activity' alleles present in the *mEH* gene could account for the spectrum of phenotypic variation observed in humans. A working hypothesis supported with these findings is that allelic variation in the regulatory and structural regions of the *mEH* gene, and other factors function co-operatively to determine an individual's activity status. Possible undetermined factors include additional 5′ polymorphisms, post-transcriptional regulation, gene induction by environmental agents and variation in the levels of transcriptional proteins required for mEH expression. The participation of these factors must be characterized more definitively in order to allow prediction of individual activities on the basis of these integrated phenomena. Genetic polymorphisms within the *mEH* gene locus likely represent an important mechanism for the modulation of carcinogen disposition that may ultimately influence individual susceptibility to cancer development.

Acknowledgements

We thank Ms Lin Jing for expert technical assistance. This research was supported by NIH grant ES–04978 (C.J.O.) and Center Grant ES–07033. C.J.O. is a Burroughs Wellcome Fund Toxicology Scholar.

References

1. Oesch,F. (1973) Mammalian epoxide hydrases: Inducible enzymes catalyzing the inactivation of carcinogenic and cytotoxic metabolites derived from aromatic and olefinic compounds. *Xenobiotica*, **3**, 305–340.
2. Sims,P., Grover,P.L., Swaisland,A., Pal,K. and Hewer,A. (1974) Metabolic activation of benzo[*a*]pyrene proceeds by a diol-epoxide. *Nature*, **252**, 326–328.
3. Oesch,F., Glatt,H. and Schimassmann,H. (1977) The apparent ubiquity of epoxide hydratase in rat organs. *Biochem. Pharmacol*., **26**, 603–607.
4. Seidegard,J. and DePierre,J.W. (1983) Microsomal epoxide hydrolase. Properties, regulation and function. *Biochim. Biophys. Acta*, **695**, 251–270.
5. Omiecinski,C.J., Aicher,L., Holubkov,R. and Checkoway,H. (1993) Human peripheral lymphocytes as indicators of microsomal epoxide hydrolase activity in liver and lung. *Pharmacogenetics*, **3**, 150–158.
6. Etter,H., Richter,C., Ohta,Y., Winterhalter,K.H., Sasabe,H. and Kawato,S. (1991) Rotation and interaction with epoxide hydrase of cytochrome P-450 in proteoliposomes. *J. Biol. Chem*., **266**, 18600–18605.
7. Vogel-Bindel,U., Bentley,P. and Oesch,F. (1982) Endogenous role of microsomal epoxide hydrolase. Ontogenesis, induction inhibition, tissue distribution, immunological behaviour and purification of microsomal epoxide hydrolase with 16 alpha, 17 alpha-epoxyandrostene–3-one as substrate. *Eur. J. Biochem*., **126**, 425–431.
8. Hassett,C., Turnblom,S.M., DeAngeles,A. and Omiecinski,C.J. (1989) Rabbit microsomal epoxide hydrolase: isolation and characterization of the xenobiotic metabolizing enzyme cDNA. *Arch. Biochem. Biophys*., **271**, 380–389.
9. Mertes,I., Fleischmann,R., Glatt,H.R. and Oesch,F. (1985) Interindividual variations in the activities of cytosolic and microsomal epoxide hydrolase in human liver. *Carcinogenesis*, **6**, 219–223.
10. Seidegard,J., DePierre,J.W. and Pero,R.W. (1984) Measurement and characterisation of membrane-bound and soluble epoxide hydrolase activities in resting mononuclear leucocytes from human blood. *Cancer Res*., **44**, 3654–3660.
11. Strickler,S.M., Dansky,L.V., Miller,M.A. *et al.* (1985) Genetic predisposition to phenytoin-induced birth defects. *Lancet*, **2**, 746–749.
12. Buehler,B.A., Delimont,D., van-Waes,M. and Finnell,R.H. (1990) Prenatal prediction of risk of the fetal hydantoin syndrome. *N. Engl. J. Med*., **322**, 1567–1572.
13. Spielberg,S.P. (1980) Acetaminophen toxicity in human lymphocytes *in vitro*. *J. Pharmacol. Exp. Ther*., **213**, 395–398.
14. Spielberg,S.P., Gordon,G.B., Blake,D.A., Goldstein,D.A. and Herlong,H.F. (1981) Predisposition to phenytoin hepatotoxicity assessed *in vitro*. *N. Engl. J. Med*., **305**, 722–729.
15. Gaedigk,A., Spielberg,S.P. and Grant,D.M. (1994) Characterization of the microsomal epoxide hydrolase gene in patients with anticonvulsant adverse drug reactions. *Pharmacogenetics*, **4**, 142–153.
16. Green,V.J., Pirmohamed,M., Kitteringham,N.R., Gaedigk,A., Grant,D.M., Boxer,M., Burchell,B. and Park,B.K. (1995) Genetic analysis of microsomal epoxide hydrolase in patients with carbamazepine hypersensitivity. *Biochem. Pharmacol*., **50**, 1353–1359.
17. Davis,C.D., Pirmohamed,M., Kitteringham,N.R., Allott,R.L., Smith,D. and Park,B.K. (1995) Kinetic parameters of lymphocyte microsomal epoxide hydrolase in carbamazepine hypersensitive patients. Assessment by radiometric HPLC. *Biochem. Pharmacol*., **50**, 1361–1366.
18. Heckbert,S.R., Weiss,N.S., Hornung,S.K., Eaton,D.L. and Motulsky,A.G. (1992) Glutathione S-transferase and epoxide hydrolase activity in human leukocytes in relation to risk of lung cancer and other smoking-related cancers. *J. Natl Cancer Inst*., **84**, 414–422.
19. McGlynn,K.A., Rosvold,E.A., Lustbader,E.D., *et al.* (1995) Susceptibility to hepatocellular carcinoma is associated with genetic variation in the enzymatic detoxification of aflatoxin B1. *Proc. Natl. Acad. Sci*., **92**, 2384–2387.
20. Lancaster,J.M., Brownlee,H.A., Bell,D.A., Futreal,P.A., Marks,J.R., Berchuck,A., Wiseman,R.W. and Taylor,J.A. (1996) Microsomal epoxide hydrolase polymorphism as a risk factor for ovarian cancer. *Mol. Carcinogen*., **17**, 160–162.
21. Hassett,C., Aicher,L., Sidhu,J.S. and Omiecinski,C.J. (1994) Human microsomal epoxide hydrolase: genetic polymorphism and functional expression *in vitro* of amino acid variants. *Hum. Mol. Genet*., **3**, 421–428.
22. Hayashi,S., Watanabe,J. and Kawajiri,K. (1991) Genetic polymorphisms in the 5′-flanking region change transcriptional regulation of the human cytochrome P-450 IIE1 gene. *J. Biochem*., **110**, 559–565.
23. Jeenah,M., Kessling,A., Miller,N. and Humphries,S. (1990) G to A substitution in the promoter region of the apolipoprotein AI gene is associated with elevated serum apolipoprotein AI and high density lipoprotein cholesterol concentrations. *Mol. Biol. Med*., **7**, 233–241.
24. Dammerman,M., Sandkuijl,L.A., Halaas,J.L., Chung,W. and Breslow,J.L. (1993) An apolipoprotein CIII haplotype protective against hypertriglyceridemia is specified by promoter and 3′ untranslated region polymorphisms. *Proc. Natl Acad. Sci*., **90**, 4562–4566.
25. Song,Z., Casolaro,V., Chen,R., Georas,S., Monos,D. and Ono,S. (1996) Polymorphic nucleotides within the human IL–4 promoter that mediate overexpression of the gene. *J. Immunol*., **156**, 424–429.
26. Checkoway,H., Costa,L.G., Camp,J., Coccini,T., Daniell,W.E. and Dills,R.L. (1992) Peripheral markers of neurochemical function among workers exposed to styrene. *Br. J. Ind. Med*., **49**, 560–565.
27. Wilson,N.M. and Omiecinski,C.J. (1989) Xenobiotic microsomal epoxide hydrolase: 5′ sequence of the human gene. *Biochim. Biophys. Acta*, **1008**, 357–358.
28. Orita,M., Iwahana,H., Kanazawa,H., Hayashi,K. and Sekiya,T. (1989) Detection of polymorphisms of human DNA by gel electrophoresis as single-strand conformation polymorphisms. *Proc. Natl Acad. Sci*., **86**, 2766–2270.
29. Gibbs,R., Nguyen,P. and Caskey,C. (1989) Detection of single base differences by competitive oligonucleotide priming. *Nucl. Acids Res*., **17**, 2437–2448.
30. Iyer,R.R. (1991) A matched set of cat vectors for rapid mutational analysis of eukaryotic promoters and enhancers. *Gene*, **105**, 97–100.
31. Lee,S.Y. and Rasheed,S. (1990) A simple procedure for maximum yield of high quality plasmid DNA. *Biotechniques*, **9**, 676–679.

32. Felgner,P.L., Gadek,T.R., Holm,M., Roman,R., Chan,H.W., Wenz,M., Northrop,J.P., Ringold,G.M. and Danielson,M. (1987) Lipofectin: a highly efficient lipid mediated DNA-transfection procedure. *Proc. Natl Acad. Sci. USA*, **84**, 7413–7417.
33. Berger,J., Hauber,J., Hauber,R., Geiger,R. and Cullen,B.R. (1988) Secreted placental alkaline phosphatase: a powerful new quantitative indicator of gene expression in eukaryotic cells. *Gene*, **66**, 1–10.
34. Vogel,F. and Kopun,M. (1977) Higher frequencies of transitions among point mutations. *J. Mol. Evol.*, **9**, 159–180.
35. Cooper,D.N., Smith,B.A., Cooke,H.J., Niemann,S. and Schmidtke,J. (1985) An estimate of unique DNA sequence heterozygosity in the human genome. *Hum. Genet.*, **69**, 201–205.
36. Hassett,C., Robinson,K.B., Beck,N.B. and Omiecinski,C.J. (1994) The human microsomal epoxide hydrolase gene (EPHX1): complete nucleotide sequence and structural characterization. *Genomics*, **23**, 433–442.
37. Vogel,F. and Motulsky,A.G. (1986) *Human Genetics: Problems and Approaches*, 2nd edn. Springer Verlag, Berlin.
38. Powers,P.A. and Smithie,O. (1986) Short gene conversion in the human fetal globin gene region: a by-product of chromosome pairing during meiosis? *Genetics*, **112**, 343–358.
39. Boulikas,T. (1994) A compilation and classification of DNA binding sites for protein transcription factors from vertebrates. *Crit. Rev. Eukaryot. Gene Expr.*, **4**, 117–321.
40. Quandt,K., Frech,K., Karas,H., Wingender,E. and Werner,T. (1995) MatInd and MatInspector- New, fast, and versatile tools for detection of consensus matches in nucleotide sequence data. *Nucl. Acids Res.*, **23**, 4878–4884.
41. Courtois,S.J., Lafontaine,D.A., Lemaigre,F.P., Durviaux,S.M. and Rousseau,G.G. (1990) Nuclear factor–1 and activator protein–2 bind in a mutually exclusive way to overlapping promoter sequences and trans-activate the human growth hormone gene. *Nucl. Acids Res.*, **18**, 57–64.
42. Hassett,C., Lin,J., Carty,C.L., Laurenzana,E.M. and Omiecinski,C.J. (1996) Human hepatic microsomal epoxide hydrolase: Comparative analysis of polymorphic expression. *Arch. Biochem. Biophys.*, **337**, 275–283.

Received on September 12, 1997; revised on November 3, 1997; accepted on November 19, 1997