Single amino acid mutations, but not common polymorphisms, decrease the activity of CYP1B1 against (−)-benzo[a]pyrene-7R-trans-7,8-dihydrodiol

Jennifer S. Mammen1, Gary S. Pittman2, Ying Li1, Fadi Abou-Zahr3, Bassem A. Bejjani3, Douglas A. Bell2, Paul T. Strickland1,2,5 and Thomas R. Sutter1,4,5

1 Johns Hopkins Bloomberg School of Public Health, 615 N. Wolfe St, Baltimore, MD 21205, 2 Laboratory of Computational Biology and Risk Analysis, National Institute of Environmental Health Sciences, Research Triangle Park, NC, 3 Kleberg Cytogenetics Lab, Baylor College Medical School, Houston, TX 77030 and 4 W. Harry Feinstone Center for Genomics Research, University of Memphis, Memphis, TN 38152, USA

5 To whom correspondence should be addressed
Email: strickl@jhsph.edu or tsutter@memphis.edu

Genetic differences that underlie inter-individual variation in the metabolism of common carcinogens are important potential sources of cancer susceptibility. Cytochrome P450 1B1 (CYP1B1), a central enzyme in the activation of the ubiquitous environmental carcinogen benzo[a]pyrene (B[a]P), has several genetic variants. This study investigated six rare mutations and four common polymorphisms for their effects on B[a]P metabolism. Five missense mutations associated with congenital glaucoma (Gly61Glu, Gly365Trp, Asp374Gln, Pro437Leu and Arg469Trp) dramatically decreased the capacity of CYP1B1 to convert (−)-benzo[a]pyrene-7R-trans-7,8-dihydrodiol (B[a]P-7,8-diol) to (±)-benzo[a]pyrene-r-7,t-8-dihydrodiol-9,10-epoxides. These five mutations resulted in enzymes with 3–12% of normal activity when assayed in vitro using an Saccharomyces cerevisiae microsomal expression system. A 10 bp deletion mutation produced no detectable protein or activity. In contrast, proteins containing all possible combinations of four common single nucleotide polymorphisms (Arg48Gly, Ala199Ser, Val432Leu, Asn453Ser) had modest effects on B[a]P-7,8-diol metabolism. Michaelis–Menten analysis suggested that two alleles, Arg48, Ala119, Val432, Ser453 (RAYS) and Arg48, Ala119, Leu432, Ser453 (RALS), have Km values 2-fold lower than Arg48, Ala119, Val432, Ser453 (RAYS): 1.4 ± 0.3 and 1.3 ± 0.4 μM, respectively, compared with 2.8 ± 0.8 μM (P < 0.05). However, these differences could not be confirmed with direct measurements of rate at low substrate concentration. There were no significant differences for either of two other kinetic parameters, kcat or kcat/KM, Allele frequency analysis in three populations reveals the Ser453 variant is rare among those of Asian (1%) and African ancestry (< 4%), and more common in individuals of European ancestry (16%). Haplotypes containing the Ser453 variant were uncommon; only RALS was detectable in our small populations. The RALS allele occurred between 0.5% in Asians and 15% in Europeans. Our study demonstrates that rare, disease-associated mutations in CYP1B1 significantly decrease the enzyme’s metabolism of B[a]P-7,8-diol; however, our results do not identify any major differences in this metabolism due to four common single amino acid polymorphisms.

Introduction

Benzo[a]pyrene (B[a]P) is a ubiquitous, carcinogenic polycyclic aromatic hydrocarbon (PAH) produced during combustion of organic materials. Significant human exposures occur with cigarette smoking and consuming certain foods, such as charbroiled meat (1). An individual’s risk of cancer in tissues directly exposed to this carcinogen such as the skin, lung and colon could be modulated by inter-individual variation in the metabolism of B[a]P. Genetic differences, which change the activity of enzymes that activate or detoxify B[a]P would be expected to correlate with toxicity, and therefore with cancer risk, in the relevant organs. Recently, low activity in microsomal epoxide hydrolase (mEH), an enzyme responsible for detoxifying the epoxides produced by the oxidation of many PAH compounds, was found to double the relative risk of colon cancer compared with high activity mEH (2).

B[a]P has 12 carbon molecules available for oxidation reactions and oxidative metabolism produces a number of phenols, quinones and diones (for review see ref. 3). The most carcinogenic metabolites are the sterically hindered, but highly reactive diol-epoxides (4). These compounds are formed by two sequential oxidations. In the relevant reactions for our experiments, the first oxidation yields benzo[a]pyrene-7,8-epoxide, which is hydrolyzed by mEH predominantly to the enantiomer (−)-benzo[a]pyrene-7R-trans-7,8-dihydrodiol (B[a]P-7,8-diol) (5); a second oxidation produces (±)-B[a]P-r-7,t-8-dihydrodiol-c-9,10-epoxide (DE1) and (±)-B[a]P-r-7,t-8-dihydrodiol-t-9,10-epoxide (DE2). The diol-epoxides are readily hydrolyzed in vitro to tetrox products, which can be separated by HPLC. DE2 is the more carcinogenic epoxide (4,6,7) and its production is highly favored over DE1 when B[a]P-7,8-diol is incubated with rat liver microsomes (5).

Both oxidation reactions can be catalyzed by several cytochrome P450s (8,9) including CYP1A1 and CYP1B1 (10), which are expressed in extra-hepatic tissues (11,12). Target-tissue expression of CYPs is a critical variable in metabolism, since the short half-lives of the ultimate carcinogens make these species likely to cause DNA damage primarily in the cells where they are formed. Therefore, inter-individual variation in the activity of CYP1B1 might be a cancer susceptibility factor in those organs where it is highly expressed (11,12).

Six common single nucleotide polymorphisms (SNPs) have been identified in CYP1B1, four of which change the amino acid sequence: Arg48Gly, Ala199Ser, Val432Leu and Asn453Ser (Figure 1) (11,13,14). Primary congenital glaucoma, an

Abbreviations: B[a]P, benzo[a]pyrene; B[a]P-7,8-diol, (−)-benzo[a]pyrene-7R-trans-7,8-dihydrodiol; CYP, cytochrome P450; DE1, (±)-B[a]P-r-7,t-8-dihydrodiol-c-9,10-epoxide; DE2, (±)-B[a]P-r-7,t-8-dihydrodiol-t-9,10-epoxide; EH, Estimated Haplotypes; mEH, microsomal epoxide hydrolase; RTTC, (±)-benzo[a]pyrene-r-7,t-8,t-9,c-10-tetrahydrodrotetrol.
autosomal recessive disease, has been mapped to CYP1B1 and a variety of mutations segregate exclusively with disease (14–19). The mutant haplotypes are very rare and have only been observed in those families affected by congenital glaucoma. Among these are at least 15 missense mutations, found throughout the protein, generally representing non-conservative changes in highly conserved regions of the protein. These mutations are presumed to result in a major disruption of CYP1B1 activity. In contrast, the common polymorphisms are found in not well conserved areas of the protein and do not a priori suggest dramatic effects on function. However, several epidemiological studies have already examined one or more of the four CYP1B1 common polymorphisms as disease risk factors with inconsistent results (20–25), and so an examination of the effects of these polymorphisms will help to clarify the literature. Since even small functional changes could have a large attributable risk if due to a common polymorphism or haplotype, we selected three non-diseased (control) populations to generate descriptive data on allele and haplotype frequencies for these loci.

We tested the in vitro activity of every possible haplotype generated by combinations of the four common polymorphisms in the oxidation of the proximate carcinogen B[a]P-7,8-diol. We also examined five missense substitutions that have been associated with congenital glaucoma: Gly61Glu, Gly365Trp, Asp374Asn, Pro437Leu and Arg469Trp, and a 10 bp deletion, 1192del10 that causes a premature stop (14–17).

Materials and methods

Yeast constructs of CYP1B1 polymorphisms

Starting with a cDNA clone described previously (11,26), a HindIII–SphI 1.9 kb segment encoding unmodified, full-length CYP1B1 was subcloned into the pCDNAII vector (Invitrogen, Carlsbad, CA) for site-directed mutagenesis using the QuikChange kit (Stratagene, La Jolla, CA). Primers containing each of the single base pair changes and their reverse complements were purchased from Integrated DNA Technologies (Coralville, IA). Coding strand primers are as follows with base pair changes shown in bold and underlined: R48G, 5'-CGGAGGCCAGCTCGGGTGCCCGCCC-3'; A119S, 5'-GCCGAGGCACGTCCTCGCCGCCGCCC-3'; G61E, 5'-CCGTGAGCCACTGATCGAAACGCCGCGCGTTGG-3'; 1192del10; 5'-CCGAGACCTGGAAGGTGC...CAGCCCAACGCATGATGCCACAC-3'; G365W, 5'-GGATCAGGTCGTGGAAGGGACCGTCTG-3'; D374N, 5'-CCGCTCCTGCTTTAGTTGGAACCGTCTGCCGTCTGTATGGG-3'; D374N; 5'-CCGCTCCTGCTTTAGTTGGAACCGTCTGCCGTCTGTATGGG-3'; E469W, 5'-CAGTGGGCCAAAAGGTGGTGCAATTGCCG-3'. Constructs containing more than one polymorphic change were generated sequentially by additional rounds of site-directed mutagenesis with the appropriate primer pairs. All vectors were verified by DNA sequence analysis and then the HindIII–SphI insert was subcloned into pYES2 (Invitrogen, Carlsbad, CA) followed by transformation into S.cerevisiae strain JL20 as described previously (26). Single colony stocks were grown in uracil deficient synthetic dextrose media (SD, ura, 2% glucose, supplemented with leucine, histidine, methionine, tryptophan and adenine) and aliquots were frozen with 20% glycerol at –80 ºC and re-streaked to single colonies as necessary.

In order to distinguish the 16 polymorphic alleles with simple and meaningful shorthand, we used the single letter amino acid designations to give each allele a four-letter code corresponding to the amino acids found at each possible polymorphic site. For example, the initial cloned cDNA sequence (11), with amino acids Arg48, Ala119, Val432 and Asn453, is called RAVN, which is the RAVN variant, as this was the background for all of the mutants called wild-type in the original linkage analysis work (14–17). The RAVN variant is defined as the protein sequence (1,2) with amino acids Arg48, Ala119, Val432 and Asn453.

Microsomal preparations

Microsomal preparations were made from expanded single colonies as described previously (26). Small aliquots of microsomes in freezing buffer (10 mM Tris pH 7.5, 1 mM EDTA, 20% glycerol) were snap-frozen in a dry ice:ethanol bath and stored at –80 ºC until needed. Total protein concentration was assayed by the Bradford reagent (Bio-Rad, Richmond, CA), which is less sensitive to the concentration of glycerol in the freezing buffer. Each sample was measured in triplicate at two different concentrations, and the measurements had <10% variation. Specific CYP1B1 protein content was estimated by quantitative western blot, in duplicate at two concentrations of total protein, using a previously described antibody (27). Standard curves were generated using CYP1B1 lymphoblastoid cell microsomes of known concentration determined spectrophotometrically (Gentest, Woburn, MA). Blots were scanned and optical density measured using MacBas (Fuji Films, Japan).
The variation in this method within a blot and between blots averaged 15% (CV; range 4–30%).

**Cytochrome c reductase assay**

To determine endogenous cytochrome P450 reductase activity, 25 μg of total protein from each microsomal preparation was diluted in 910 μl of 300 mM KPO4, pH 7.7, to which 80 μl of 0.5 mM cytochrome c (in 10 mM KPO4, pH 7.7) was added. A baseline scan at 430 nm was obtained for 3 min in a cuvette with 1 cm path length. Ten microliters of 10 mM NADPH was added and mixed quickly and the absorbance was measured for 3 min. Cytochrome c reductase activity was calculated for the linear phase of the reaction as follows: nmole reduced/(μg x min) = ΔOD/(min x μg x 0.021) (28).

(−)-Benzo[a]pyrene-7,8-trans-7,8-dihydriodil oxidation assay

Reactions to measure BlαP-7,8-diol metabolism were carried out as described (10). Briefly, an appropriate volume of microsomes to contain 5 μmol of CYP1B1 was resuspended in a final volume of 198 μl incubation buffer (100 mM NaPO4, 5 mM MgCl2, 1.4 mM NADPH final concentrations, pH 7.5). Each allele was assayed in triplicate at 0.0, 0.625, 1.25, 2.5, 5 and 10 μM BlαP-7,8-diol. All reactions were carried out at low light levels in silanized glass test tubes. Reactions were pre-incubated at 37°C for 2 min, 2 μl of the appropriate 100× substrate stock in methanol was added; tubes were vortexed briefly and returned to 37°C for 15 min. Reactions were stopped by adding 1 vol ice-cold acetone and then extracted twice with 2 vol of ethyl acetate, dried under nitrogen and resuspended in an appropriate volume of HPLC grade methanol (150–300 μl) to maintain relatively consistent metabolite concentrations. These were stored at −20°C until use. We confirmed that the reaction conditions were in the linear range with respect to time and protein concentration (data not shown).

HPLC analysis was performed to separate the products as described with some modifications (10). Injections were performed either manually or using a Waters WISP712 autosampler. The 60 min run on a Bondclone C-18 column (Phenomenex, Torrance, CA) contained a 30 min isotropic phase of 50:50 methanol:H2O followed by a 10 min gradient up to 100% methanol to elute the substrate, at constant flow rate of 1 ml/min. UV (Waters Tunable Absorbance Detector 486) and fluorescence (Ranin Dynamax FL-2, Varian, Walnut Creek, CA) signals were recorded and analyzed using Shimadzu Class-VP Chromatography software (Columbia, MD). Standard curves using authentic standards demonstrated that the quantification is linear between 2.5 and 25 pmol (data not shown). Standards at four concentrations were run every few weeks and slopes varied <10% (CV). Michaelis–Menten kinetics were fit to data using Prism software (GraphPad, San Diego, CA). Any experiment where the calculated KM was lower than the lowest substrate concentration used was discarded as having insufficient data.

**Genotype, allele and haplotype frequencies**

Two novel PCR–RFLP assays were used to detect the four CYP1B1 polymorphisms: one assay for Arg48Gly and Ala119Ser, another for Val432Leu and Asn453Ser. For the Arg48Gly and Ala119Ser polymorphisms, the two PCR primers used were (forward) 5′-GCC CTA ACC TCC TAC TCA GGC AAG-3′ and (reverse) 5′-CGG CAG CCG AAA CAC AC-3′, which generate a 1223 bp fragment. The PCR products were sequenced on an ABI 3370 genetic analyzer. All samples were genotyped for the Arg48Gly and Ala119Ser polymorphisms: one assay for Arg48Gly and Ala119Ser, another for Val432Leu and Asn453Ser, which generate a 283 bp fragment. The PCR components were the same as previous, except 1.5 mM MgCl2 was used. The PCR cycling conditions used were 94°C for 4 min; 33 cycles of 94°C for 15 s, 55°C for 30 s, 72°C for 30 s; and 72°C for 4 min. PCR products were incubated for 3 h at 37°C with 2.5 U each BsrI and MwoI. The Val452 allele contains a polymorphic BsrI site; the Leu432 allele does not. The Asn453Ser allele contains a polymorphic MwoI site; the Ser453 allele does not.

For both assays, PCR was performed on a GeneAmp® 9600 thermocycler (Applied Biosystems, Foster City, CA), and products were resolved on a 3.25% Metaphor® agarose gel (FMC BioProducts, Rockland, ME) stained with ethidium bromide and visualized using an Eagle Eye II®still video system (Stratagene, La Jolla, CA).

**Results**

**Genotype, allele and haplotype frequencies**

The public health importance of any inherited functional difference in CYP1B1 activity is dependent on both the magnitude of the difference and the allele frequency in the population. Therefore, we genotyped individuals of Asian, African and European ancestry in order to estimate allele and haplotype frequencies in these groups (Tables I and II). We found all loci were in Hardy–Weinberg equilibrium within all three populations for all four loci. Because haplotype determination is complicated by incomplete genotype data, we only used individuals with complete data at all loci for that analysis: 98 Asians, 98 African-Americans and 143 European-Americans. While direct haplotype determination is possible for individuals heterozygous at most at one locus, estimation must be used for those heterozygous at two or more loci. We estimated haplotypes using the Estimated Haplotypes (EH) algorithm (29).

| Table I. Genotype and allele frequencies of CYP1B1 polymorphisms |
|-------------------------|-------------------------|-------------------------|-------------------------|
|                         | Asian ancestry (n = 120)  | African ancestry (n = 111) | European ancestry (n = 181) |
|                         | (+)/b                   | (+)/−                  | (+)/−                  | (+)/−                  |
| Arg48Gly                | 67% (76)                | 29% (33)               | 4% (5)                 | 25% (26)               | 48% (49)               | 26% (27)               | 50% (58)               |
| Val432Leu               | 2% (2)                  | 14% (14)               | 84% (87)               | 96% (102)              | 4% (4)                 | 0% (0)                 | 0.02                   |
| Asn453Ser               | 99% (102)               | 1% (1)                 | 0% (0)                 | 96% (102)              | 4% (4)                 | 0% (0)                 | 1% (1)                 |

*Population sample sizes represent the total individuals genotyped. The numbers of readable genotypes are listed in the table.

**Mutations decrease activity of CYP1B1 against B[a]P-7,8-diol**

For the Val432Leu and Asn453Ser polymorphisms, the two PCR primers used were (forward) 5′-GCC TGT CAC TAT TCC TCA TGC G-3′ and (reverse) 5′-GTG AGC CAG GAT GGA GAT GGA A-3′, which generate a 283 bp fragment. The PCR components were the same as previous, except 1.5 mM MgCl2 was used. The PCR cycling conditions used were 94°C for 4 min; 33 cycles of 94°C for 15 s, 57°C for 30 s, 72°C for 30 s; and 72°C for 4 min. PCR products were incubated for 3 h at 37°C with 2.5 U each BsrI and MwoI. The Val452 allele contains a polymorphic BsrI site; the Leu432 allele does not. The Asn453Ser allele contains a polymorphic MwoI site; the Ser453 allele does not.

For both assays, PCR was performed on a GeneAmp® 9600 thermocycler (Applied Biosystems, Foster City, CA), and products were resolved on a 3.25% Metaphor® agarose gel (FMC BioProducts, Rockland, ME) stained with ethidium bromide and visualized using an Eagle Eye II®still video system (Stratagene, La Jolla, CA).

**Results**

**Genotype, allele and haplotype frequencies**

The public health importance of any inherited functional difference in CYP1B1 activity is dependent on both the magnitude of the difference and the allele frequency in the population. Therefore, we genotyped individuals of Asian, African and European ancestry in order to estimate allele and haplotype frequencies in these groups (Tables I and II). We found all loci were in Hardy–Weinberg equilibrium for each population. Allele frequencies differed widely among population groups. For example, the Gly48 allele occurred at a frequency of 0.19 in Asians, 0.5 in African-Americans and 0.33 in European-Americans, while the Leu432 allele occurred at frequencies of 0.91, 0.29 and 0.58, respectively. Asn453Ser was the most rare polymorphism among those of Asian and African descent (0.005 and 0.02, respectively) but occurred more commonly among those of European descent (0.16).
Using the individuals for whom we had complete genotype information at each of the four polymorphic sites, we prepared estimates of haplotype frequency in the population samples (Table II). Applying the EH algorithm, RAVN, RALN, GSLN, GSVN and RALS haplotypes were predicted to occur among all three populations, although at different frequencies. The first three haplotypes were common among all groups. The RALN haplotype was estimated to account for 72.5% of alleles among Asians, but only 7% and 9% in African- and European-Americans, respectively. RAVN was the most common haplotype observed in the latter two populations (41 and 42%). Of note, among Asians and European-Americans the sequences at amino acid positions 48 and 119 were closely linked; the Arg48 occurred with Ala119 and Gly48 occurred with Ser119. The GAVN haplotype was uniquely observed among those of African descent, presumably a recombinant haplotype between RAVN and GSVN. The data indicate that all Ser453 alleles occur on Leu432 alleles, suggesting a more recent origin for the Ser453 polymorphism.

**Microsome characterization**

Specific protein concentrations in microsomes containing each of the 16 presumed normal alleles averaged 24 pmol/mg (range 15–32 pmol/mg) as determined by western blot (data not shown). This is consistent with earlier experience with the RAVN construct (26). Among the CYP1B1 mutants, all single amino acid substitutions resulted in expressed protein at levels similar to the normal alleles, except Gly365Trp, which was only found at 3 pmol/mg (data not shown). A 10 bp deletion construct that causes a frame-shift and premature stop did not produce any detectable protein. It is not known whether this is because the truncated protein is not expressed, is unstable, or has lost epitopes for the polyclonal antibody. Cytochrome c reductase activity levels averaged 134 ± 35 (SD) pmol/µg/min (range 86–220 pmol/µg/min; data not shown). The variance was <15% for measurements taken on a single day, and there was similar good agreement between repeated measures of an individual sample. There was no observable relationship between allele sequence and protein concentration or reductase level.

**B[a]P-7,8-diol metabolism**

All 16 alleles with common polymorphisms produced the four expected tetrol products, (±)benzo[a]pyrene-r-7,t-8,c-9,c-10-tetrahydrotetrol (RTCC), (±)benzo[a]pyrene-r-7,t-8,c-9,t-10-tetrahydrotetrol (RTCT), (±)benzo[a]pyrene-r-7,t-8,t-9,c-10-tetrahydrotetrol (RTTC) and (±)benzo[a]pyrene-r-7,t-8,t-9,t-10-tetrahydrotetrol (RTTT), which were identified by co-migration with authentic standards (Figure 2). Vector

### Table II. Estimated haplotype frequencies (%) for CYP1B1

<table>
<thead>
<tr>
<th>Allelic variant</th>
<th>Asian ancestry % (n)ᵇ</th>
<th>African ancestry % (n)ᵇ</th>
<th>European ancestry % (n)ᵇ</th>
</tr>
</thead>
<tbody>
<tr>
<td>RAVN</td>
<td>8.0% (16)</td>
<td>41.0% (80)</td>
<td>42.0% (123)</td>
</tr>
<tr>
<td>RAVS</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>RALN</td>
<td>72.5% (142)</td>
<td>7.0% (14)</td>
<td>9.0% (27)</td>
</tr>
<tr>
<td>RALS</td>
<td>0.5% (1)</td>
<td>2.0% (4)</td>
<td>15.0% (44)</td>
</tr>
<tr>
<td>GSVN</td>
<td>3.0% (6)</td>
<td>28.0% (55)</td>
<td>1.0% (3)</td>
</tr>
<tr>
<td>GSVS</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>GSLN</td>
<td>16.0% (31)</td>
<td>18.0% (35)</td>
<td>33.0% (97)</td>
</tr>
<tr>
<td>GSLS</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>RSVN</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>RSVS</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>RSLN</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>RSLS</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>GAVN</td>
<td>–</td>
<td>4.0% (4)</td>
<td>–</td>
</tr>
<tr>
<td>GAVS</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>GALN</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>GALS</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
</tbody>
</table>

ᵇHaplotypes estimated using EH algorithm (29).

ᵇᵇPercent and number (n) for each population, see Table I.

ᵇᶜDash (–) indicates haplotype not predicted from observed genotypes.

---

**Fig. 2.** Chromatogram of B[a]P-7,8-diol metabolites produced by CYP1B1. Separation of B[a]P-7,8,9,10-tetrols produced in the metabolism of 10 μM B[a]P-7,8-diol by CYP1B1. Tetrols are separated in 35 min by 50:50 methanol:water on a C18-reverse phase HPLC column, and detected by fluorescence with 340 excitation, 420 emission. Pure methanol is required to elute B[a]P-7,8-diol. Peaks were identified by co-migration with authentic standards (insert), which were used to generate standard curves for quantification of products. 1, RTTC; 2, RTCT; 3, RTTT; 4, RTCC; 5, B[a]P-7,8-diol; *, unknown.
control microsomes did not produce any metabolic products from B[a]P-7,8-diol (data not shown). Mutant constructs had such low activity that only the one or two most common tetrols were quantifiable. There was very little secondary metabolism observed at the substrate concentrations used, although CYP1B1 is able to further oxidize the tetrols, producing uncharacterized products that migrate between 4 and 10 min (data not shown). Three unidentified peaks were observed in samples with active CYP1B1, and probably represent triol products (5,10).

Of the two epoxides formed by oxidation of B[a]P-7,8-diol, the trans-diol-epoxide (DE2, yielding RTTC and RTTT) is thought to be more reactive with DNA than the cis-diol-epoxide (DE1, yielding RTCC and RTCT). Therefore, we first examined whether any of the polymorphisms change the reaction stereospecificity by calculating the ratio of trans- and cis-epoxide formation at high substrate concentration. The ratio of DE2:DE1 ranged from 1.5 to 2.4, and none were significantly different from RAVN (Kruskal–Wallis, Dunn’s post test for multiple comparisons, data not shown). This is in good agreement with a previously reported ratio of 2.4 for RAVN (10). All further analysis therefore compared enzymes by the production of the highly favored tetrol, RTTC, derived from the biologically significant DE2.

Five missense mutations and one 10 bp deletion associated previously with congenital glaucoma (14,16) were assayed for their ability to metabolize B[a]P-7,8-diol at a saturating substrate concentration of 10 μM. Each reaction included 10 pmol enzyme. Since the deletion construct could not be detected on a western blot, we used 350 μg of total protein, twice the average total protein in other experiments. All mutant alleles except the deletion construct had low levels of metabolic activity, ranging from 0.03–0.13 pmol RTTC/min/pmol P450, which represents 3–12% of normal activity for the RAVN clone (Figure 3). The deletion mutant generated no detectable product.

All enzymes examined followed Michaelis–Menten kinetics, as shown by representative curves for three common haplotypes (Figure 4). Initial estimates of kinetic parameters for each of the 16 presumed normal alleles indicated a potential 6-fold range of $K_M$ values, from 0.76 to 4.67 μM, a 3-fold range in $k_{cat}$, from 0.44 to 1.35 pmol RTTC/min/pmol CYP1B1, and a 5-fold range in $k_{cat}/K_M$, ranging between 0.24 and 1.25 (Table III). Both Michaelis–Menten curve fitting and Eadie–Hofstee plots yielded similar estimates of the kinetic parameters (data not shown). There was no pattern of any consistent effect on the various parameters with a particular polymorphism. We therefore chose to repeat measurements for those haplotypes with the highest and lowest values for each constant in order to be able to assess the differences statistically: RAVN and RSVN had the highest and lowest $K_M$ values, respectively, while RAVS and RALS represent...
Table III. Estimate of kinetic parameters for 16 CYP1B1 alleles

<table>
<thead>
<tr>
<th>Allelic variant</th>
<th>$K_M$ (µM)</th>
<th>$k_{cat}$ (min$^{-1}$)</th>
<th>$k_{cat}/K_M$</th>
</tr>
</thead>
<tbody>
<tr>
<td>RAVN</td>
<td>4.67 ± 1.13</td>
<td>1.14 ± 0.13</td>
<td>0.24</td>
</tr>
<tr>
<td>RAVS</td>
<td>1.08 ± 0.26</td>
<td>1.35 ± 0.08</td>
<td>1.25</td>
</tr>
<tr>
<td>RALN</td>
<td>1.84 ± 0.25</td>
<td>1.02 ± 0.04</td>
<td>0.55</td>
</tr>
<tr>
<td>RALS</td>
<td>0.76 ± 0.20</td>
<td>0.44 ± 0.03</td>
<td>0.58</td>
</tr>
<tr>
<td>GSVN</td>
<td>2.90 ± 0.52</td>
<td>0.72 ± 0.05</td>
<td>0.26</td>
</tr>
<tr>
<td>GSVS</td>
<td>2.56 ± 0.43</td>
<td>1.00 ± 0.06</td>
<td>0.39</td>
</tr>
<tr>
<td>GSLN</td>
<td>1.71 ± 0.49</td>
<td>0.64 ± 0.06</td>
<td>0.37</td>
</tr>
<tr>
<td>GSLS</td>
<td>2.08 ± 0.75</td>
<td>0.84 ± 0.12</td>
<td>0.40</td>
</tr>
<tr>
<td>RSVN</td>
<td>0.76 ± 0.20</td>
<td>0.94 ± 0.05</td>
<td>1.23</td>
</tr>
<tr>
<td>RSVS</td>
<td>1.04 ± 0.31</td>
<td>0.66 ± 0.05</td>
<td>0.64</td>
</tr>
<tr>
<td>RSLN</td>
<td>1.73 ± 0.24</td>
<td>0.60 ± 0.03</td>
<td>0.40</td>
</tr>
<tr>
<td>RSLS</td>
<td>1.94 ± 0.40</td>
<td>0.78 ± 0.06</td>
<td>0.40</td>
</tr>
<tr>
<td>GAVN</td>
<td>0.86 ± 0.25</td>
<td>0.68 ± 0.04</td>
<td>0.79</td>
</tr>
<tr>
<td>GAVS</td>
<td>1.25 ± 0.27</td>
<td>0.53 ± 0.03</td>
<td>0.42</td>
</tr>
<tr>
<td>GALN</td>
<td>1.29 ± 0.32</td>
<td>0.77 ± 0.06</td>
<td>0.59</td>
</tr>
<tr>
<td>GALS</td>
<td>1.32 ± 0.32</td>
<td>0.56 ± 0.04</td>
<td>0.42</td>
</tr>
</tbody>
</table>

aMicrosomes were pooled from at least two preparations for each allele.
bParameters reported are the best estimate and the standard deviation of that estimate from a single experiment at five concentrations performed in triplicate using pooled microsomes.

The ratio of best fit estimates of each parameter, as shown in columns 2 and 3.

Table IV. Replicated kinetic parameters for six CYP1B1 alleles

<table>
<thead>
<tr>
<th>Allelic variant</th>
<th>$K_M$ (µM)</th>
<th>$k_{cat}$ (min$^{-1}$)</th>
<th>$k_{cat}/K_M$</th>
</tr>
</thead>
<tbody>
<tr>
<td>RAVN</td>
<td>2.79 ± 0.80</td>
<td>0.85 ± 0.17</td>
<td>0.33 ± 0.12</td>
</tr>
<tr>
<td>RAVS</td>
<td>1.38 ± 0.25</td>
<td>0.88 ± 0.31</td>
<td>0.68 ± 0.35</td>
</tr>
<tr>
<td>RSVN</td>
<td>1.71 ± 0.61</td>
<td>0.72 ± 0.24</td>
<td>0.50 ± 0.36</td>
</tr>
<tr>
<td>RALS</td>
<td>1.32 ± 0.41</td>
<td>0.48 ± 0.16</td>
<td>0.40 ± 0.18</td>
</tr>
<tr>
<td>RALN</td>
<td>1.79 ± 0.04</td>
<td>0.76 ± 0.40</td>
<td>0.42 ± 0.22</td>
</tr>
<tr>
<td>GSVN</td>
<td>2.79 ± 0.75</td>
<td>0.63 ± 0.10</td>
<td>0.24 ± 0.06</td>
</tr>
</tbody>
</table>

Each allelic variant was assayed in five to nine experiments, using two to three independent microsomal preparations.

$P < 0.05$, $^aP < 0.01$ compared with RAVN, by Kruskal–Wallis and Dunn’s post test for multiple comparisons.

The highest and lowest $k_{cat}$ values. This set included the alleles, RAVN and RAVS, representing the extremes of $k_{cat}/K_M$. We also included in this analysis two of the most common haplotypes, RALN and GSVN.

Average values for at least three independent experiments with at least two independent microsomal preparations are shown in Table IV. In this analysis, both RALS and RAVS appeared to have $K_M$ values that are significantly lower than RAVN, which has the highest, when analyzed using Kruskal–Wallis and Dunn’s post-test for multiple comparisons. $K_M$ values were found to be 1.38 ± 0.25, 1.32 ± 0.41 and 2.79 ± 0.8 for RAVS, RALN and RAVN, respectively. There are no significant differences in either $k_{cat}$ or catalytic efficiency ($k_{cat}/K_M$). This suggests that these enzymes all perform with similar efficiency when the concentration of substrate is very small compared with $K_M$.

In order to test the reliability of the differences predicted by the kinetic parameters, we compared the reaction rates for RAVS and RALS with RAVN at one of two substrate concentrations. Varying reaction time provides a robust measure of rate with which to compare enzymes. For enzymes with the same $k_{cat}$, a 2-fold lower $K_M$ means the higher affinity enzyme will generate product at twice the rate at a substrate concentration equal to the lower $K_M$; so the slope of the curve product versus time for RAVS should be approximately twice that of RAVN at 1 µM B[α]P, as would the slope for RALS if the apparent 2-fold difference in $k_{cat}$ is truly significant. If RALS has a 2-fold lower $k_{cat}$ as well, the difference in rate at low concentration will be much less apparent, but at saturating concentrations, the 2-fold difference in $k_{cat}$ would be clearly observed. The rates of RTTC production by each of the Ser453 alleles were not different when compared with RAVN at either 1 or 10 µM B[α]P-7,8-diol (Figure 5), meaning we were unable to confirm that RALS and RAVS have 2-fold lower $K_M$ compared with RAVN. RALS showed a trend towards a lower $k_{cat}$ compared with RAVN (Figure 5D), but much smaller than the 2-fold difference suggested in the earlier experiments, and this difference could not be established with statistical significance.

Discussion

Several presumed low activity alleles of CYP1B1 have been identified as co-segregating with one form of primary congenital glaucoma in family studies of the inherited disease. We have demonstrated that six of these mutations severely reduce enzyme function. Similar results have been reported for two of the mutations, G61E and R469W, in the metabolism of steroids (30). Yeast microsomes containing enzymes with any of five missense mutations exhibit between 3 and 12% of normal RAVN activity for the activation of B[α]P-7,8-diol, while a deletion mutation shows no activity and cannot be detected on western blot by a polyclonal antibody, suggesting that the truncated protein is unstable. Our results suggest an explanation for the observation, made in a study of congenital glaucoma among the Amish, that individuals homozygous for a substitution mutation had a less severe phenotype than those homozygous for a truncating mutation (19).

In order to test individual or interactive effects of four commonly occurring polymorphisms (Arg48Gly, Ala19Ser, Val432Leu and Asn453Ser), we systematically created all 16 possible recombinant alleles and expressed them in yeast. In contrast to the glaucoma-associated mutations, these polymorphisms have modest effects on CYP1B1 kinetics. There were no differences that could be demonstrated in either $k_{cat}$ or catalytic efficiency ($k_{cat}/K_M$) for the activation of B[α]P-7,8-diol. Repeated kinetic measurements revealed a 2-fold lower $K_M$ in two haplotypes, RAVS and RALS, compared with the wild-type RAVN for this substrate. This difference, $K_M = 1.3$ µM versus $K_M = 2.8$ µM, is a small but statistically significant difference ($P < 0.05$). However, comparing metabolic rates at single substrate concentrations, the alleles did not appear to differ.

The biological relevance of the kinetic differences observed needs to be evaluated in three contexts: (i) typical exposure levels to the carcinogen in the general population; (ii) the range of likely enzyme induction; and (iii) the frequency of the higher risk alleles in the population. Dietary exposures to carcinogens such as B[α]P are in the nanograms/day range, which means that in non-smokers, cellular concentrations are likely to be quite low compared with $K_M$. As at low substrate concentrations kinetics are determined by the catalytic efficiency ($k_{cat}/K_M$), which is the first order rate constant for the reaction when substrate is much lower than $K_M$, our results suggest that there will not be differences in susceptibility correlated with the genotypes investigated for non-smokers. In smokers, elevated exposures have been shown to lead to...
enzyme induction, particularly in the lung (31). This induction leads to mRNA expression levels that vary by 40-fold, which, if the protein is mostly active, would create a greater variation in total enzyme activity than that due to the small, observed differences in $K_M$.

Small differences in individual risk due to variation in metabolism may have a large impact on population-based risk estimates if the risk factor is at high frequency in the population. Allele frequency estimates in the population groups examined in this study are consistent with those reported in the literature (20,22,23,32). The RALN, RAVN, GSLN and GSVN haplotypes (all functional) together account for the majority of alleles across all population groups tested. Alleles containing the Ser453 variant (RAVS and RALS, the possible high activity alleles) are uncommon, with RAVS being undetectable in our small populations. The RALS allele occurs between 0.5% in Asians and 15% in European-Americans. It would appear that risk associated with this allele would be most easily detected in those of European descent. One study has looked for an increased risk of breast cancer due to the Ser453 variant among 328 Caucasians and found no association (23). Statistical power calculations ($\alpha < 0.05$ and $\beta > 80$) suggest that a study of over 600 individuals would be needed in order to detect a 2-fold increase in risk among homozygous RALS individuals with reasonable probability.

The locations of the rare mutations examined within CYP1B1 are consistent with our results that the mutations associated with congenital glaucoma have a large impact on function. Those point mutations associated with congenital glaucoma are spread throughout the protein but are generally well-conserved residues (14). Gly61Glu is in the hinge region and Arg469Trp is in the heme-binding region, both highly conserved segments of the protein, although the latter amino acid is not conserved outside of the CYP1 family. Pro437Leu is situated on the boundary of the so-called meander region and is highly conserved, while Gly365Trp is on the boundary of helix J and is conserved across several human CYP families and within the CYP1 family across species. Only Asp374Asn is in a non-conserved region between helices J and K. However, this substitution exchanges an acidic residue with a neutral one, and could have important consequences for proper protein folding, which could in turn disrupt the active site or substrate binding. In contrast, the common SNPs examined here are well outside conserved regions with the exception of Asn453Ser, which is found within the meander region.

Fig. 5. Rate determinations for RAVS, RALS and RAVN. Rates at low and high substrate concentration for RAVS (triangles, A and B) or RALS (inverted triangles, C and D) compared with wild-type RAVN (squares). Experiments were performed in duplicate at 0, 10, 15 and 20 min with B[a]P-7,8-diol concentrations of 1 $\mu$M (A and C), where a difference in slope would reflect a difference in $K_M$, or 10 $\mu$M (B and D), where a difference in slope would reflect a difference in $k_{cat}$. None of the slopes were significantly different from wild-type. Symbols specify mean, error bars indicate range where range is greater than the size of the symbol; open symbols indicate only a single data point was available.
Although Asn453Ser does not represent a conserved residue in the CYP1 family broadly, it is conserved in the CYP1B1 sequence of human, mouse and rat.

Only one other group has compared four CYP1B1 alleles for the metabolism of B[a]P-7,8-diol, and they also do not report differences, although the experiment is done in Escherichia coli and much higher overall KM values were obtained than reported here (33). This discrepancy in KM measurements between E.coli and yeast experiments is consistent with that found for estradiol metabolism, and may be due to inadequate association between CYP1B1 and reductase in the E.coli plasma membranes (26,34–36). A second paper that compared alleles in the metabolism of B[a]P cannot be evaluated as the assay was not linear with time for the interval used, so kinetic parameters cannot be accurately derived (37).

CYP1B1 can metabolize a broad range of substrates, many of which are carcinogenic (12). CYP1B1 is also the most catalytically efficient 17β-estradiol 4-hydroxylase described to date (26). This metabolite is believed to be important in estrogen carcinogenesis (38). A thorough study of estradiol metabolism in yeast needs to be undertaken, as it is not known whether kinetic analysis can be generalized across substrates. Groups investigating estradiol metabolism so far in E.coli have not found differences in kcat, and have divergent findings for KM. Reports include no differences, a higher KM value for RAVN compared with other alleles, and higher KM values for RALN and RALS compared with others (33,34,36,37).

Logical pursuit of genetic variation as a risk factor for disease rests upon the biological plausibility that the variation could affect the disease process. For a metabolic enzyme hypothesized to affect cancer risk, the critical variation is in the contribution it makes to the availability of an ultimate carcinogen. Here we have taken a rigorous biochemical approach to investigate nine single amino acid substitutions and one deletion in CYP1B1, encompassing both disease-associated mutations and all possible haplotypes from four common polymorphisms, in order to better understand the range of normal and abnormal function of this enzyme in the production of B[a]P diol epoxides. We have begun the investigation of amino acid substitutions with B[a]P-7,8-diol because it is a well characterized and ubiquitously carcinogenic. CYP1B1, like most other CYP enzymes, is able to activate a wide range of substrates, and a complete characterization of how genetic polymorphisms might affect cancer risk will require a broader investigation into the metabolism of other substrates. As we build an understanding of how gene-environment interactions can be generalized and predicted, a systematic approach to examining the functional consequences of polymorphisms could ultimately allow us to predict what types of genetic changes in a family of enzymes would have important effects on particular substrates.

Acknowledgements

Thank you to Dr Randy Bryant for his advice and feedback. This research was supported by the National Institutes of Health grants ES08148 and ES03819.

References


5. Yang,S.K., McCourt,D.W., Roller,P.P. and Gelboin,H.V. (1976) Enzymatic conversion of benzo(a)pyrene leading predominantly to the diol-epoxide r-7,8-dihydrodiol-3,9,10-oxo-7,8,9,10-tetrahydrobenzo(a)pyrene through a single enantiomer of r-7,8-dihydrodiol-7,8-dihydrobenzo(a)pyrene. Proc. Natl Acad. Sci. USA, 73, 2594–2598.


20. Watanabe, J., Shimada, T., Gillam, E.M., Ikuta, T., Suemasu, K., Higashi, Y.,
    polymorphism with incidence to breast and lung cancer. *Pharmacogenetics*,
    10, 25–33.
    Gao, Y.T. (2000) Genetic polymorphism of cytochrome P450 1B1 and risk
22. Fritsche, E., Bruning, T., Jonkmanns, C., Ko, Y., Bolt, H.M. and Abel, J.
    (1999) Detection of cytochrome P450 1B1 Bfr I polymorphism: genotype
    distribution in healthy German individuals and in patients with colorectal
    of cytochrome P450 1B1 (CYP1B1) polymorphism with steroid receptor
    status in breast cancer [published erratum appears in *Cancer Res*. 1999
24. Ko, Y., Abel, J., Harth, V. et al. (2001) Association of CYP1B1 codon 432
    mutant allele in head and neck squamous cell cancer is reflected by
25. Goodman, M.T., McDiffic, K., Kolonel, L.N., Terada, K., Donlon, T.A.,
    Wilkens, L.R., Guo, C. and Le Marchand, L. (2001) Case-control study of
    ovarian cancer and polymorphisms in genes involved in catecholestrogen
    formation and metabolism. *Cancer Epidemiol. Biomarkers Prev.*, 10,
    209–216.
    (1996) 17β-Estradiol hydroxylation catalyzed by human cytochrome p450
27. Walker, N.J., Crofts, F.G., Li, Y., Lax, S.F., Hayes, C.L., Strickland, P.T.,
    cytochrome P450 1B1 (CYP1B1) protein in the livers of TCDD-treated
    rats: detection using polyclonal antibodies raised to histidine-tagged fusion
    proteins produced and purified from bacteria. *Carcinogenesis*, 19,
    395–402.
    Hayes, A.W. (ed.), *Principles and Methods of Toxicology*, 2nd Edn.
29. Terwilliger, J.D. and Ott, J. (1994) Linkage disequilibrium between alleles
    at marker loci. In *Handbook of Human Genetic Linkage*. Johns Hopkins
    University Press, Baltimore, MD, pp. 188–198.
    two mutations of human CYP1B1, G61E and R469W, on stability and
    endogenous steroid substrate metabolism. *Pharmacogenetics*, 11,
    793–801.
    chemical carcinogenesis. In *Environmental Health Sciences*. Johns
    Hopkins School of Public Health, Baltimore, p. 171.
    frequency distribution of genetic polymorphisms in the CYP1A1 and
    CYP1B1 genes in Japanese and Caucasian populations. *Xenobiotica*, 30,
    Polymorphisms in P450 CYP1B1 affect the conversion of estradiol to the
    potentially carcinogenic metabolite 4-hydroxyestradiol. *Pharmacogenetics*,
    10, 343–353.
    Cytochrome P450 1B1 (CYP1B1) pharmacogenetics: association of
    polymorphisms with functional differences in estrogen hydroxylation
35. McLellan, R.A., Oscarson, M., Hidebrand, M., Leidvik, B., Jonsson, E.,
    Otter, C. and Ingelman-Sundberg, M. (2000) Characterization and func-
    tional analysis of two common human cytochrome P450 1B1 variants.
36. Shimada, T., Watanabe, J., Kawajiri, K., Sutter, T.R., Guengerich, F.P.,
    1607–1613.
37. Shimada, T., Watanabe, J., Inoue, K., Guengerich, F.P. and Gillam, E.M.
    (2001) Specificity of 17beta-oestradiol and benzo(a)pyrene oxidation by
    polymorphic human cytochrome P450B1 variants substituted at residues

Received January 20, 2003; revised April 23, 2003; accepted May 4, 2003