British Journal of Pharmacology (2003) 139, 1373–1382
www.nature.com/bjp

Human estrogen sulfotransferase (SULT1E1) pharmacogenomics: gene resequencing and functional genomics

[1]Araba A. Adjei, [1]Bianca A. Thomae, [1]Janel L. Prondzinski, [2]Bruce W. Eckloff, [2]Eric D. Wieben & *,[1]Richard M. Weinshilboum

[1]Department of Molecular Pharmacology & Experimental Therapeutics, Mayo Medical School-Mayo Clinic-Mayo Foundation, Rochester, MN 55905, U.S.A. and [2]Department of Biochemistry and Molecular Biology, Mayo Medical School-Mayo Clinic-Mayo Foundation, Rochester, MN 55905, U.S.A.

1 Estrogens are used as drugs and estrogen exposure is a risk factor for hormone-dependent diseases such as breast cancer. Sulfate conjugation is an important pathway for estrogen metabolism. The sulfotransferase (SULT) enzyme SULT1E1 has the lowest K_m values for estrogens and catecholestrogens of the 10 known human SULT isoforms.

2 We previously cloned and characterized the human SULT1E1 cDNA and gene as steps toward pharmacogenetic studies. In the present experiments, we set out to determine whether common, functionally significant genetic polymorphisms might exist for ***SULT1E1***. As a first step, we 'resequenced' the eight ***SULT1E1*** exons and exon–intron splice junctions as well as portions of the 5′-flanking region using DNA from 60 African-American and 60 Caucasian-American subjects.

3 In all, 23 polymorphisms, 22 single nucleotide polymorphisms (SNPs) and one insertion deletion were observed. There were three nonsynonymous coding SNPs (cSNPs) that altered the following encoded amino acids: Asp22Tyr, Ala32Val and Pro253His. Among these, 12 pairs of SNPs were tightly linked. In addition, 12 unambiguous ***SULT1E1*** haplotypes were identified, including six that were common to both populations studied.

4 Transient expression in COS-1 cells of constructs containing the three nonsynonymous cSNPs showed significant decreases in SULT1E1 activity for the Tyr22 and Val32 allozymes, with corresponding decreases in levels of immunoreactive protein. There were no changes in levels of either activity or immunoreactive protein for the His253 allozyme. Apparent K_m values of the Val32 allozyme for the two cosubstrates for the reaction, 17β-estradiol and 3′-phosphoadenosine 5′-phosphosulfate, were not significantly different from those of the wild-type enzyme, but there was a two- to three-fold increase in K_m values for the His253 allozyme and a greater than five-fold increase for the Tyr22 allozyme.

5 These observations raise the possibility that genetically determined variation in SULT1E1-catalyzed estrogen sulfation might contribute to the pathophysiology of estrogen-dependent diseases as well as variation in the biotransformation of exogenously administered estrogens.
British Journal of Pharmacology (2003) **139,** 1373–1382. doi:10.1038/sj.bjp.0705369

Keywords: Estrogen sulfation; sulfotransferase; SULT1E1; pharmacogenomics; genetic polymorphisms; gene resequencing; functional genomics

Abbreviations: cSNPs, coding single nucleotide polymorphisms; DMEM, Dulbecco's modified Eagle's medium; E1, estrone; E2, 17β-estradiol; K_m, Michaelis constant; PAPS, 3′-phosphoadenosine 5′-phosphosulfate; SULT, sulfotransferase; V_{max}, maximum velocity

Introduction

Estrogens are used as therapeutic agents and exposure to these steroids is a risk factor for the development of hormone-dependent disease (Colditz, 1996; Henderson & Bernstein, 1996). A major pathway for estrogen metabolism in humans is sulfate conjugation catalyzed by sulfotransferase (SULT) enzymes (Hobkirk, 1993; Aksoy *et al*., 1994; Falany *et al*., 1995). Steroid sulfation can potentially contribute to the control of levels of biologically active hormone in target tissues; it might also form a hormone 'reservoir' after the hydrolysis of sulfate conjugates by sulfatases (Pasqualini & Chetrite, 1996), and it enhances the water solubility–and thus the renal excretion–of steroids (Hobkirk, 1993; Strott, 1996). In all, 10 SULT isoforms are known to be expressed in human tissue. These enzymes catalyze the transfer of a sulfuryl group from 3′-phosphoadenosine 5′-phosphosulfate (PAPS) to an acceptor substrate. SULT1E1 catalyzes the sulfate conjugation of estrone (E1), 17β-estradiol (E2), catecholestrogens and 2-

*Author for correspondence;
E-mail: weinshilboum.richard@mayo.edu
Supported by NIH Grants RO1 GM35720 (RMW), RO1 GM28157 (RMW), UO1 GM61388 (RMW and EDW) and PO1 CA82267 (RMW), and by DOD Grant DAMD 17-99-1-9281(AAA). The DNA resequencing data for ***SULT1E1*** described in this manuscript have been deposited in the NIH-sponsored Pharmacogenetics Research Network database PharmGKB with accession number PKB_N00000019.
Advance online publication: 14 July 2003

methoxyestradiol with the lowest K_m values–in the nM range–of the 10 known human SULT isoforms (Falany *et al*., 1995; Zhang *et al*., 1998; Adjei *et al*., 2001; Adjei & Weinshilboum, 2002). As a result, SULT1E1 is able to catalyze the sulfation of physiological concentrations of estrogens (Pasqualini & Chetrite, 1996). SULT1E1 is expressed in many human tissues including liver and jejunum (Aksoy *et al*., 1994; Her *et al*., 1996), mammary epithelium cells (Falany & Falany, 1996; Qian *et al*., 1998), endometrium (Falany *et al*., 1998) and testis (Song *et al*., 1995; 1997; Qian & Song, 1999).

We previously cloned and characterized the human SULT1E1 cDNA and gene (Aksoy *et al*., 1994; Her *et al*., 1995) as steps toward studies of possible inherited variation in estrogen sulfation. The human SULT1E1 gene is approximately 20 kb in length, consists of eight exons and maps to chromosome 4q13 (Her *et al*., 1995). We also reported large individual variations in levels of SULT1E1 immunoreactive protein in the small intestine (Her *et al*., 1996). Large individual variations were also observed in human liver samples (Song *et al*., 1998). In the present experiments, we have applied a genotype-to-phenotype strategy to determine whether ***SULT1E1***, like the genes encoding other human SULT isoforms (Raftogianis *et al*., 1997; 1999; Freimuth *et al*., 2001; Thomae *et al*., 2002), might contain common, functionally significant genetic polymorphisms. Specifically, we resequenced ***SULT1E1*** using DNA from 60 African-American and 60 Caucasian-American subjects and observed 23 polymorphisms, including three nonsynonymous cSNPs–one in DNA from an African-American and two in DNA from Caucasian-American subjects. We then created expression constructs for these nonsynonymous cSNPs and performed functional genomic studies of recombinant allozymes encoded by these variant alleles. Those results showed decreased levels of basal activity and decreased levels of immunoreactive protein for two of the three variant allozymes, as compared with the wild-type enzyme. Substrate kinetic studies also showed differences in apparent K_m and V_{max}/K_m ratios for two of the variant allozymes. These results raise the possibility of inherited variation in the SULT1E1-catalyzed sulfate conjugation of estrogens and structurally related steroid compounds.

Methods

DNA samples

DNA from 60 African-American and 60 Caucasian-American subjects was obtained from the Coriell Cell Repository (Camden, NJ, U.S.A.). Specifically, 60 samples each from the two 100-item sample sets, HD100AA and HD100CAU, were used to perform these experiments. All of the DNA samples had been anonymized by the National Institutes of Health prior to their deposit in the Coriell Cell Repository, and all subjects had provided written consent for the use of their DNA for experimental purposes. The present studies were reviewed and approved by the Mayo Clinic Institutional Review Board.

SULT1E1 *resequencing*

To make it possible to amplify all ***SULT1E1*** exons and splice junctions as well as a portion of the 5′-flanking region of the gene, eight separate PCR amplifications were performed with each of the 120 DNA samples studied. M13 'tags' were added to the 5′-ends of each primer to make it possible to use dye-primer DNA-sequencing chemistry to facilitate the identification of heterozygous bases (Chadwick *et al*., 1996). Primers were designed to hybridize within introns, within the 5′-flanking region or within the 3′-untranslated region of the terminal exon at locations selected to avoid repetitive sequence. Sequences of all primers used in these and subsequent experiments are listed in Table 1. Amplification reactions were performed with AmpliTaq Gold DNA polymerase (Perkin-Elmer, Foster City, CA, U.S.A.) with a 'hot start' to help ensure amplification specificity. The 50 μl reaction mixture contained 2.5 U of DNA polymerase, 5 μl of a 10-fold diluted DNA sample (160–190 ng DNA), 12.5 pmol of each primer (7 pmol for exon 7), 0.05 mM dNTP (Boehringer Mannheim, Indianapolis, IN, U.S.A.) and 5 μl of $10 \times$ PCR buffer containing 15 mM $MgCl_2$ (Perkin-Elmer). PCR cycling parameters involved a 12 min 'hot start' at 94°C, followed by 35 cycles (40 cycles for exon 7) of 94°C for 30 s, 55°C (68°C for the exon 7) for 30 s and 45 s at 72°C– with a final 10 min extension at 72°C. All reactions were performed in a Perkin-Elmer model 9700 thermal cycler. Amplicons were sequenced on both strands in the Mayo Molecular Core Facility with an ABI 377 DNA sequencer using BigDye™ (Perkin-Elmer) dye primer-sequencing chemistry. To exclude PCR-induced artefacts, independent amplifications were performed for those samples in which a SNP was observed only once or for any sample with an ambiguous chromatogram. Chromatograms were analyzed using the PolyPhred 3.0 (Nickerson *et al*., 1997) and Consed 8.0 (Gordon *et al*., 1998) programs from the University of Washington. The University of Wisconsin GCG software package, Version 10, was also used to analyze nucleotide sequence. GenBank accession numbers for the ***SULT1E1*** reference sequences used in these experiments were U20514–U20521 (Her *et al*., 1995).

Expression constructs and transient expression

Site-directed mutagenesis was performed using the QuickChange kit (Stratagene, La Jolla, CA, U.S.A.) to create DNA constructs for alleles containing the three ***SULT1E1*** nonsynonymous cSNPs observed during the resequencing experiments. Primer sequences used to create those constructs are listed in Table 1. Sequences of the constructs were confirmed by sequencing both strands after cloning them into the eukaryotic expression vector pCR3.1 (Invitrogen, Carlsbad, CA, U.S.A.). Expression constructs for the SULT1E1 wild-type and variant cDNAs, as well as 'empty' pCR3.1 that lacked an insert, were then transfected into COS-1 cells in serum-free Dulbecco's modified Eagles medium (DMEM) (BioWhittaker, Walkersville, MD, U.S.A.) using the TransFast reagent (Promega, Madison, WI, U.S.A.), as suggested by the manufacturer, at a charge ratio of 1 : 1. Construct DNA (7 μg) was cotransfected with 7 μg of pSV-β-galactosidase DNA (Promega) as a control to make it possible to correct for transfection efficiency. After 48 h, the COS-1 cells were harvested in 5 mM potassium phosphate buffer, pH 7.5, and were homogenized with a Polytron homogenizer (Brinkmann Instruments, Westbury, NY, U.S.A.). The homogenates were centrifuged at $100{,}000 \times g$ for 1 h, and supernatant preparations were stored at −80°C.

Table 1 Human ***SULT1E1*** gene resequencing and site-directed mutagenesis primers

Primer name	*Primer location*	*Primer sequence*
Gene resequencing		
UF(−289)	5′-FR	GCAGGATATTTCTACATCTCCATGAATGAACATGACT
I1R(147)	Intron 1	GCTTCACATCATTAATTAACTAAAGTATCAAATCAAGACTTTGGC
I1F(−170)	Intron 1	CTCTCTAGTTACCCAAACTATTTGATATGCAACTTTGC
I2R(145)	Intron 2	GAGCTACCTTTTCTATGTCCATATCCAAACTACCG
I2F(−151)	Intron 2	ATAGAAAATATTTCCTGAGTCTGTGGCTATTCAGACACC
I3R(122)	Intron 3	GCTGTCTTATGTAGAAGACCTGATACTAATTGCCATTC
I3F(−196)	Intron 3	TAGGCATGCAATGCATAATAATTACACCATGGGGAATG
I4R(185)	Intron 4	TGGCAAAAGACAGAGTTGGAATTAAAATATAGACTCTCTGAC
I4F(−183)	Intron 4	AAACCACTGTCACCTCAGGTTATTGAAGATGTCTT
I5R(134)	Intron 5	ATGCTTGCTCTTAAACCTCCAGGCCCCTTTAGA
I5F(−171)	Intron 5	CATGCTTTGCCTCTCTTGCTGGAGAGAACCT
I6R(167)	Intron 6	GCTTCAAATCTATGCTAAAGTATCTGTATTATTTTGGTCCTTTCC
I6F(−160)	Intron 6	CACAGCTTTTATAAAATTCCCCCAATTAGATTTCTCATTAGAAATC
I7R(132)	Intron 7	TCAAATATGAAAGACTGCTGAAGAAAACTTAAGCTGGGTT
I7F(−170)	Intron 7	CATCTTTGTAAGCCCCCAAAAGTATATCATTAAAGGTATAC
DR1022	3′-FR	AGTTAAACAAAAATTTAAAAAGAAAATGTCAACATAATCCATGA
Site-directed mutagenesis		
F(−12)	Exon 1	CAGTGTACCACAATGAATTCTG
R890	Exon 7	CCTTCTTAGATCTCAGTTCGAA
F48	Exon 2	GATTCTAATGTATAAA***t***ATTTTGTCAAATATTG
R80	Exon 2	CAATATTTGACAAAAT***a***TTTATACATTAGAATC
F81	Exon 2	GGATAATGTGGAAG***t***GTTCCAGGCAAGAC
R109	Exon 2	GTCTTGCCTGGAAC***a***CTTCCACATTATCC
F743	Exon 7	CCAGAAATTGTCGC***a***CTTCATGAGAAAGG
R772	Exon 7	CCTTTCTCATGAAG***t***GCGACAATTTCTGG

A universal M13 forward sequence (5′-TGTAAAACGACGGCCAGT-3′) was added to the 5′-ends of all forward primers, and a universal M13 reverse sequence (5′-CAGGAAACAGCTATGACC-3′) was added to the 5′-ends of all reverse primers. Bold, italicized lower case letters in the primers are the mutated bases used to perform the site-directed mutagenesis. 'F' represents forward; 'R' reverse; 'U' upstream; 'I' intron; 'D' downstream and 'FR' flanking region. The numbering scheme for primers located in exons and the 5′-FR is based on the cDNA sequence, with the 'A' at the translation initiation codon designated as (+1). Positions 5′ and 3′ to that location were assigned negative or positive numbers, respectively. Intron-based primers were numbered on the basis of nucleotide distance from splice junctions, with (+1) as the first nucleotide at the 5′-end, and (−1) as the first nucleotide at the 3′-end of the intron.

Enzyme assays, substrate kinetic and thermal stability studies

Recombinant wildtype and variant allozymes were assayed for SULT1E1 enzyme activity using the method of Foldes & Meek (1973) as modified by Aksoy *et al*. (1994). Specifically, 50 nM E2 served as the sulfate acceptor substrate and 0.4 μM ^{35}S-PAPS as the sulfate donor cosubstrate. Enzyme reactions were incubated for 20 min and included 8 mM dithiothreitol, 1.25 mM $MgCl_2$ and 10 mM potassium phosphate buffer, pH 6.5, in a total volume of 160 μl. Blanks were samples that did not contain E2. One unit of enzyme activity represented the formation of 1 nmol of E2 sulfate per hour of incubation at 37°C. Levels of enzyme activity were corrected for transfection efficiency by measuring β-galactosidase activity using the β-galactosidase Assay System (Promega), as suggested by the manufacturer. Standard curves were constructed that ranged from 1 to 6 mU β-galactosidase activity, and all samples fell within this range. Protein concentrations were determined using the method by Bradford (1976), and protein concentrations for COS-1 cell cytosol preparations ranged from 0.94 to 1.13 $\mu g\,\mu l^{-1}$.

Apparent K_m values were also determined with both cosubstrates for each allozyme studied. Since SULTs can display profound substrate inhibition (Weinshilboum & Otterness, 1994), apparent K_m and V_{max} values for E2 were determined by using a two-step process. Initial experiments used 10-fold serial dilutions of E2 that ranged from 10^{-3} to 10^{-9} M. Maximal activity for all allozymes was found at E2 concentrations near 100 nM. A second set of experiments was then performed in the presence of 0.4 μM PAPS using eight concentrations of E2 that varied from 3.1 to 400 nM. For the determination of apparent K_m values for PAPS, seven concentrations of PAPS that ranged from 2.7 to 150 nM were assayed in the presence of 50 nM E2.

Enzyme thermal stability was measured as described by Reiter *et al*. (1982; 1983). Specifically, diluted COS-1 cell supernatant preparations for each of the recombinant SULT1E1 allozymes were incubated in a water bath for 15 min at temperatures that ranged from 28 to 46°C. The samples were placed on ice immediately after incubation. Aliquots of the same supernatant preparations were also kept on ice as controls. Enzyme activity was then measured in both heated and unheated samples. Blank values were determined for each temperature studied.

Western blot analysis

Quantitative Western blot analyses were performed with recombinant SULT1E1 allozymes. The level of immunoreactive protein was measured using a rabbit polyclonal antibody directed against SULT1E1 amino acids 1–13, with the addition of cysteine at the carboxy terminus. This antibody and its characteristics have been described in detail elsewhere (Her *et al*., 1996). The quantity of COS-1 cell cytosol for each allozyme loaded on 12.5% acrylamide gels was adjusted to achieve equal β-galactosidase activity, thus correcting for transfection efficiency. Proteins were separated by SDS–

PAGE prior to transfer to nitrocellulose membranes (Scleicher and Schuell, Keene, NH), and SULT1E1 was detected by use of the ECL Western Blotting System (Amersham Pharmacia, Piscataway, NJ, U.S.A.). An AMBIS Radioanalytic Imaging System, Quant Probe Version 4.31 (Ambis, Inc., San Diego, CA, U.S.A.) was used to quantitate immunoreactive protein, and the data were expressed as a percentage of the wild-type SULT1E1 band intensity on that gel.

Data analysis

Average levels of recombinant allozyme activity and immunoreactive protein were compared by ANOVA performed with the StatView program, version 4.5 (Abacus concepts, Berkeley, CA, U.S.A.). Apparent K_m values were calculated with the method of Wilkinson (1961) using a computer program written by Cleland (1963). Points that deviated from linearity on double inverse plots as a result of substrate inhibition were not used in these analyses. T_{50} values, that is, values for temperatures resulting in 50% thermal inactivation, were calculated using the GraphPad Prism computer program (GraphPad Software, Inc., San Diego, CA, U.S.A.). Average apparent K_m and T_{50} values were compared among allozymes by the use of an unpaired Student's *t*-test. Linkage among ***SULT1E1*** polymorphisms was determined by testing all possible pairwise combinations of SNPs and calculating D' values–a method for reporting linkage data that is independent of allele frequency (Hartl & Clark, 1997; Hedrick, 2000). Haplotype analysis was performed by use of the E–M algorithm (Excoffier & Slatkin, 1995; Long *et al*., 1995).

Results

The series of experiments described in this study began with the resequencing of ***SULT1E1*** using 120 DNA samples obtained from the Coriell Cell Repository (Camden, NJ, U.S.A.). Of these samples, 60 were from African-American and 60 from Caucasian-American subjects. All ***SULT1E1*** exons were resequenced, including splice junctions, as well as approximately 180 bp of the 5′-flanking region of the gene–an area that contains the 'TATA box'. In all, 23 polymorphisms were observed, including three nonsynonymous cSNPs. Linkage relationships and haplotypes were determined, and expression constructs were created for the wild-type and three variant allozymes encoded by alleles with the nonsynonymous cSNPs to make it possible to perform functional genomic experiments. After expression in COS-1 cells, levels of enzyme activity, levels of immunoreactive protein, substrate kinetic properties and thermal stability were determined for all four SULT1E1 recombinant allozymes. Although differences in substrate kinetics were observed, the major functional effects of changes in amino-acid sequence as a result of the cSNPs were related to the level of immunoreactive protein. These results will now make it possible to test the hypothesis that ***SULT1E1*** genetic polymorphisms might contribute to individual variation in estrogen biotransformation and effect.

SULT1E1 *resequencing and polymorphisms*

The first step in this series of experiments involved resequencing ***SULT1E1***. A total of approximately 730,000 bp of DNA sequence was analyzed for the 120 DNA samples studied. When this sequence was compared to the consensus ***SULT1E1*** gene sequence (GenBank accession numbers U20514–U20521), 23 polymorphisms, including one insertion–deletion and 22 SNPs, were observed in the 240 alleles resequenced (Figure 1 and Table 2). In all, 18 polymorphisms were present in DNA samples from African-American, and 13 in DNA from Caucasian-American subjects. Of the SNPs, 10 in the African-American subjects were not observed in the Caucasian-American subjects, and five of those present in the Caucasian-American samples were absent in DNA from African-American subjects. In total, 13 of the 18 polymorphisms in DNA from African-American subjects and eight of the 13 in DNA from Caucasian-American subjects had frequencies greater than 1%, and, as a result, would be considered 'common' in those populations. SNPs seen in only a single DNA sample were always verified by performing a separate, independent PCR amplification, followed by sequencing.

Three nonsynonymous cSNPs were observed, each of which was present in only a single heterozygous sample. One (G64 T) was present only in DNA from an African-American subject, while the other two, C95 T and C758A, were observed in DNA samples from Caucasian-American subjects. These three cSNPs resulted in changes in encoded amino acids Asp22Tyr, Ala32Val and Pro258His, respectively. We verified the presence of these three cSNPs both by performing independent PCR amplifications, followed by DNA sequencing, and by performing RFLP assays with the same Coriell DNA samples as template for the initial PCR, followed by digestion with *Hyp*188III (Asp22Tyr), *Mwo*I (Ala32Val) or *Dra*III (Pro253-His), respectively.

We also compared the SNPs that we observed during our resequencing studies with those present in publicly accessible databases. A total of 16 human ***SULT1E1*** variant sequences had been deposited in the SNP database (www.ncbi.nlm.nih.

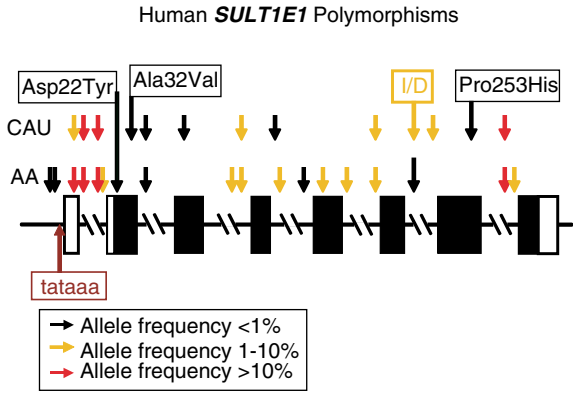

Figure 1 Human ***SULT1E1*** genetic polymorphisms. The figure shows a schematic representation of the human SULT1E1 gene with the locations of polymorphisms indicated by arrows. Black rectangles represent the open reading frame and white rectangles represent portions of exons that encode untranslated region (UTR) sequence. 'AA' represents data obtained with DNA from African-American subjects and 'CA' represents data obtained using DNA from Caucasian-American subjects. Changes in encoded amino acids resulting from the presence of nonsynonymous cSNPs are also indicated. Colors of arrows indicate allele frequencies in the two populations studied. Red arrows represent frequencies of greater than 10%; yellow arrows represent frequencies from 1 to 10% and black arrows represent polymorphisms with frequencies of less than 1%.

Table 2 Human ***SULT1E1*** genetic polymorphisms

Location	*Nucleotide*	*Sequence change*	*Amino-acid change*	*Frequency of variant allele*	
				African-American subjects	*Caucasian-American subjects*
5′-Flanking	−232	G→A		0.008	0.000
5′-Flanking	−190	C→G		0.008	0.000
Exon 1	−64	G→A		0.200	0.067
Intron 1	I1(69)	A→G		0.225	0.492
Intron 1	I1(−73)	G→C		0.300	0.383
Intron 1	I1(−20)	A→T		0.092	0.000
Exon 2	64	G→T	Asp22Tyr	0.008	0.000
Exon 2	95	C→T	Ala32Val	0.000	0.008
Intron 2	I2(22)	T→C		0.008	0.008
Exon 3	237	T→C		0.000	0.008
Intron 3	I3(−137)	T→G		0.017	0.008
Intron 3	I3(−80)	A→G		0.017	0.000
Intron 4	I4(69)	A→T		0.000	0.033
Intron 4	I4(139)	A→T		0.017	0.008
Intron 4	I4(−23)	A→G		0.008	0.000
Exon 5	459	C→T		0.017	0.000
Intron 5	I5(55)	C→T		0.017	0.000
Intron 5	I5(−10)	C→G		0.083	0.017
Intron 6	I6(55)	I/D		0.017	0.033
Intron 6	I6(−39)	T→C		0.000	0.017
Exon 7	758	C→A	Pro253His	0.000	0.008
Intron 7	I7(−121)	C→T		0.333	0.108
Intron 7	I7(−63)	T→G		0.017	0.000

Locations, sequence alterations and frequencies of polymorphisms observed in the two populations studied are listed. The numbering scheme is that described in the legend for Table 1. Polymorphisms within exons are boxed. I/D refers to an insertion–deletion event in which a G in the most common (wild type) allele was deleted.

gov/SNP). Of these, 15 SNPs were located in introns and one was in the 3′-flanking region. Only three of the SNPs, I1(69), I5(−10) and I7(−121), were observed during our resequencing studies. The remaining 13 publicly available SNPs were located in areas of genomic sequence not studied in our experiments. Iida *et al*. (2001) resequenced ***SULT1E1*** using 48 DNA samples from Japanese subjects–a different ethnic group than those included in the present study. They observed 27 SNPs, but only four of the SNPs, those at nucleotide −64 in exon one, I1(69) and I1(−73) in intron one and I3(−137) in intron three were observed in our studies. These investigators did not report allele frequencies for their population (Iida *et al*., 2001). The remaining SNPs reported by Iida *et al*. (2001), like those deposited in dbSNP, were located in intronic areas that we did not resequence. Therefore, of the 23 ***SULT1E1*** polymorphisms that we observed, only six–and none of the nonsynonymous cSNPs–were publicly available.

SULT1E1 *linkage and haplotype analysis*

We next performed linkage analysis for the various polymorphisms that we observed in ***SULT1E1***. To do that, pairwise combinations of all 23 polymorphisms were analyzed, and D' values were calculated (Hartl & Clark, 1997; Hedrick, 2000). D' values can range from $+1.0$ when two polymorphisms are maximally positively associated to -1.0, when two polymorphisms never occur together (Hartl & Clark, 1997; Hedrick, 2000). The results of that analysis (Table 3) showed that if a P-value of $\leqslant 0.01$ was selected as the 'cutoff' (Table 3), nine and six pairs of polymorphisms were positively linked in the African-American and Caucasian-American populations, respectively, D' values were not calculated for the three nonsynonymous cSNPs because of their low allele frequencies.

Table 3 ***SULT1E1*** linkage analysis

Nucleotide positions of polymorphism pairs		D′ *value*	
		African-American	*Caucasian-American*
−232	I3(−80)	1	–
−232	459	1	–
−232	I6(55)	1	–
−232	I7(−63)	1	–
−64	I1(−73)	−1	–
−64	I7(−121)	0.93	0.85
I1(69)	I5(−10)	1	1
I1(69)	I1(−73)	–	0.65
I1(−73)	I7(−121)	−1	–
I3(−80)	I6(55)	1	1
I3(−80)	I6(−39)	–	1
I4(−23)	I5(55)	1	–
459	I7(−63)	1	–
I6(55)	I6(−39)	–	1

Pairwise analyses were performed for ***SULT1E1*** polymorphisms. The numbering for the locations of polymorphisms is described in the legend for Table 1. Only polymorphism pairs that had P-values $\leqslant 0.01$ have been listed.

It is becoming clear that haplotype may prove to be important for understanding functional variation as a result of genetic polymorphisms (Drysdale *et al*., 2000). Therefore, we also analyzed ***SULT1E1*** haplotypes (Table 4). In all, 12 unambiguous haplotypes, that is, those deduced entirely on the basis of genotype, were observed in the 120 DNA samples resequenced, including 10 in African-American and eight in Caucasian-American subjects (Table 4). These unequivocal haplotypes accounted for 89 and 92% of all samples from the two groups, respectively. Haplotype 'assignments' were then

Table 4 Human ***SULT1E1*** haplotype analysis

	Nucleotide position		Exon 1	Intron 1			Exon 2		Intron 4	Intron 5		Exon 7	Intron 7
	Frequency		-64	69	-73	20	64	95	69	55	-10	758	-121
Allele designation	African-A merican	Caucasian-American											
*1A	0.250	0.071	G	A	C	A	G	C	A	C	C	C	C
*1B	0.190	0.040	A	A	G	A	G	C	A	C	C	C	T
*1C	0.106	0.326	G	A	G	A	G	C	A	C	C	C	C
*1D	0.103	0.045	G	A	G	A	G	C	A	C	C	C	T
*1E	0.067	0.102	G	G	G	A	G	C	A	C	G	C	C
*1F	0.064	—	G	A	G	T	G	C	A	C	C	C	C
*1G	0.050	0.276	G	G	C	A	G	C	A	C	C	C	C
*1H	0.040	—	G	G	G	A	G	C	A	C	C	C	T
*1I	0.010	—	A	A	G	A	G	C	A	C	C	C	C
*1J	0.008	—	G	A	G	A	G	C	A	T	C	C	C
*1K	—	0.045	G	G	G	A	G	C	A	C	C	C	C
*1L	—	0.009	G	A	G	A	G	C	T	C	C	C	C
*2	0.008	—	G	A	G	A	T	C	A	C	C	C	C
*3	—	0.008	G	A	G	A	G	T	A	C	C	C	C
*4	—	0.008	G	A	G	A	G	C	A	C	C	A	C

The 12 unambiguous haplotypes–those labeled ***1***–observed in the 120 DNA samples resequenced are listed. Sequence within each of the alleles that differed from the ***SULT1E1*** consensus sequence (***1C***) has been shown as white type against a black background. Initial designations of haplotypes were made on the basis of the encoded amino-acid sequence, with the wild-type sequence being designated ***1***. 'Letter' designations were then added based on descending allele frequencies, starting with haplotypes present in both ethnic groups and then making assignments based on haplotypes observed in samples from African-American subjects. Although we were unable to determine their haplotypes unequivocally, we have also listed the Tyr22 variant as *2, the Val32 variant as *3 and the His253 as *4.

made on the basis of the encoded amino-acid sequence of the allozyme, with wildtype sequences being designated ***1***. We also assigned letter designations to haplotypes based on descending allele frequencies, starting with the African-American subjects, that is, ***1A*** was more frequent than ***1B***, and both were more frequent than ***1C*** (Table 4). We assigned ***2*** to the Tyr22 variant, ***3*** to the Val32 variant and ***4*** to the His253 variant even though unambiguous haplotypes could not be assigned for these samples. Finally, we were able, using the E–M algorithm (Excoffier & Slatkin, 1995; Long *et al*., 1995), to infer a total of 12 and 10 additional haplotypes for DNA samples from African-American and Caucasian-American subjects, respectively. Data with regard to these additional inferred haplotypes can be obtained from the authors upon request.

Recombinant allozyme enzyme activity and Western blot analysis

We next addressed the possible functional consequences of the ***SULT1E1*** nonsynonymous cSNPs. Expression constructs for the three ***SULT1E1*** nonsynonymous cSNPS (ORF nucleotides 64, 95 and 758), as well as a construct with the wildtype sequence were expressed in COS-1 cells. A mammalian expression system was used in these experiments to help ensure appropriate post-translational modification as well as the presence of the mammalian protein degradation machinery. Many examples of genetic changes in a single amino acid that result in significant decreases in levels of immunoreactive protein–due in some cases to alterations in proteasome-mediated protein degradation–have been reported (Szumlanski *et al*., 1996; Preuss *et al*., 1998; Tai *et al*., 1999; Freimuth *et al*., 2001; Thomae *et al*., 2002). Two of the three recombinant SULT1E1 allozymes, Tyr22 and Val32, displayed significantly decreased levels of enzyme activity when compared with that of the 'wild-type' enzyme, while the level of

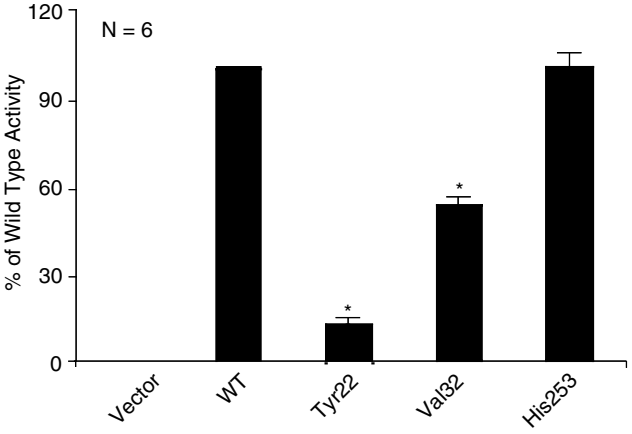

Figure 2 Recombinant human SULT1E1 allozyme enzyme activity. Average levels of enzyme activity are shown for each of the recombinant SULT1E1 allozymes assayed with E2 as the sulfate acceptor substrate. All values have been corrected for transfection efficiency. Each bar represents the average of six independent transfections (mean ± s.e.m.). *Indicates $P<0.001$ when compared to the wild-type allozyme.

activity of the His253 allozyme was similar to that of the wild-type enzyme (Figure 2). All of these data were corrected for transfection efficiency. The Asp22Tyr polymorphism resulted in a reduction in enzyme activity of approximately 90%, while the Ala32Val polymorphism resulted in a 42% reduction (Figure 2). These observations raised the question of mechanism(s) responsible for differences in basal levels of enzyme activity.

One possible mechanism responsible for decreased level of enzyme activity would involve a decrease in the quantity of enzyme protein. Therefore, recombinant cytosol preparations from six independent transfections for each allozyme were subjected to quantitative Western blot analysis. The sequence of the peptide used to generate the rabbit polyclonal antibody

used to perform these Western blot analyses did not include any of the amino acids changed in the variant SULT1E1 allozymes (Her *et al*., 1996). Average levels of immunoreactive SULT1E1 protein, corrected for transfection efficiency, paralleled relative levels of enzyme activity for all the three of the variant allozymes (Figures 2 and 3). A representative Western blot with two different exposure times, 10 and 60 s, is shown in Figure 3b. Immunoreactive protein for the Tyr22 allozyme could only be detected after the longer exposure time (Figure 3b). These observations represent additional evidence that the alteration of only a single amino acid as a result of a genetic polymorphism can significantly alter the level of enzyme protein.

Recombinant allozyme substrate kinetics and thermal stability

Enzyme kinetic properties could also be affected by changes in amino-acid sequence. Therefore, we determined apparent K_m values for each of the recombinant SULT1E1 allozymes with both E2 and PAPS as substrates. Since SULTs can display profound substrate inhibition by the sulfate acceptor cosubstrate (Weinshilboum & Otterness, 1994), these experiments were conducted in two stages. During the first series of studies, a wide range of substrate concentrations was tested. In the second set of experiments, based on the results of the initial studies, the concentration range was narrowed to values appropriate for the determination of apparent K_m values, but no data points that displayed substrate inhibition were used to calculate these values. There were statistically significant differences among SULT1E1 allozymes in apparent $K_m$values for both E2 and PAPS (Table 5). As a result, the decreased level of basal enzyme activity for the Tyr22 allozyme could be attributed both to alterations in level of enzyme protein and to alterations in substrate kinetics–but predominantly to decreased protein (Figures 2, 3 and Table 5). However, substrate kinetics for the Val32 allozyme were only slightly different from those the wild-type enzyme. Therefore, the decrease in level of enzyme activity for that allozyme appeared to result primarily from the decreased level of immunoreactive protein. Although there were no differences in levels of either enzyme activity or immunoreactive protein for the His253 allozyme, that allozyme did display significant increases in both K_m and V_{max} values when compared with the wild-type enzyme–differences that may have 'offset' each other in their effect on level of enzyme activity. The apparent K_m values of wild-type SULT1E1 for E2 and PAPS that we determined fell within the range of values that have been reported previously by other authors (Zhang *et al*., 1998; Fujita *et al*., 1999; Hempel *et al*., 2000; Otake *et al*., 2000; Pedersen *et al*., 2002). That was encouraging since our studies, unlike those for several previous reports, used cytosol preparations rather than purified SULT1E1 to determine kinetic constants.

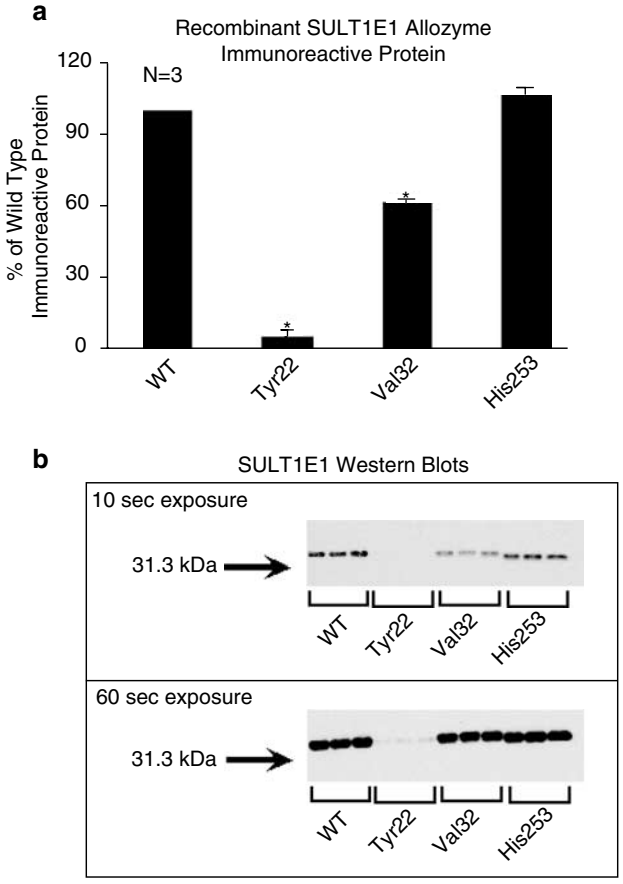

Figure 3 Recombinant human SULT1E1 allozyme Western blot analysis. (a) Average levels of immunoreactive SULT1E1 protein are shown for each of the recombinant allozymes, expressed as a percentage of the wild-type protein. Each bar represents the average of three independent transfections (mean ± s.e.m.). *Indicates $P<0.0006$ when compared to the wild-type level of immunoreactive protein. (b) SULT1E1 Western blots with 10 and 60 s exposures. Three samples for each of the allozymes were loaded on the gel on the basis of β-galactosidase activity to correct for variations in transfection efficiency. The Tyr22 allozyme protein was present (see 60 s exposure), but in greatly reduced quantity.

Finally, studies of variant allozymes for other enzymes have shown that decreased levels of enzyme protein as a result of genetic polymorphisms have often shown an association with decreased enzyme thermal stability (Scanlon *et al*., 1979; Kang *et al*., 1991; Raftogianis *et al*., 1997; Preuss *et al*., 1998; Shield & Weinshilboum, 2002). Therefore, we assayed the thermal stability of SULT1E1 allozymes. When T_{50} values, the temperature at which the enzyme was 50% inactivated, were determined, the Tyr22 and Val32 allozymes both showed reduced T_{50} values when compared with the wild-type enzyme–although the decrease for the Val32 allozyme was not statistically significant (Table 5).

SULT1E1 polymorphisms and crystal structures: The X-ray crystal structure of mouse SULT1E1 has been solved in the presence of 3′-phosphoadenosine 5′-phosphate at a resolution of 2.5 Å (Kakuta *et al*., 1997). Recently, Pedersen *et al*. (2002) also solved the crystal structure of human SULT1E1 and reported that the substrate-binding pocket of human SULT1E1 is very similar to that of the mouse enzyme. Figure 4 depicts the human SULT1E1 structure determined by X-ray crystallography, with locations of the polymorphisms that we observed indicated in relation to known substrate-binding regions of the enzyme. The Asp22Tyr polymorphism that we observed was located 14 residues downstream from the conserved SULT 'Region I' sequence motif (Weinshilboum & Otterness, 1994; Varin *et al*., 1995; Weinshilboum *et al*., 1997), the E2 substrate-binding region. This residue was located at the entrance to the substrate-binding pocket. Therefore, this

Table 5 Human SULT1E1 allozyme substrate kinetics and thermal stability

SULT1E1 allozyme	*E2 as varied substrate*			*PAPS as varied substrate*			*Thermal stability*
	Apparent K_m *(nM)*	V_{max}	$V_{max}/K_m \times (100)$	*Apparent* K_m *(nM)*	V_{max}	$V_{max}/K_m \times (100)$	T_{50} *(°C)*
Wild type	30.0±5.0[a]	26.8±4.4	89.3	56.0±2.9[a]	27.2±0.7[a]	48.5	38.5±0.92
Tyr22	220±30[a]	4.7±0.3[b]	2.2	240±9.2[a]	1.9±0.1[a]	0.8	35.0±0.32[c]
Val32	44.0±5.0[a]	22.8±6.6	51.7	65.0±4.2[a]	5.7±0.1[a]	8.8	36.7±1.2
His253	97.0±7.0[a]	58.6±2.5[b]	60.4	180±7.5[a]	34.8±0.2[a]	19.3	38.4±0.46

Values are expressed as mean±s.e.m. ($n=3$). V_{max} is expressed as nmol/h/β-galactosidase units. [a]Indicates that all values differed significantly from each other ($P \leqslant 0.04$); [b]indicates that values differed significantly from each other ($P \leqslant 0.01$) except for the wild-type and Val32 allozymes; and [c]indicates that the value differed significantly from the wildtype and His253 allozymes ($P \leqslant 0.003$).

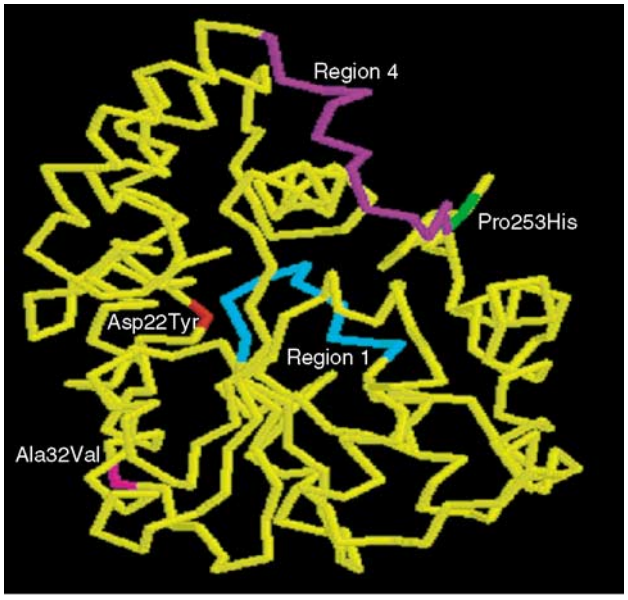

Figure 4 Human SULT1E1 crystal structure showing polymorphism locations. The X-ray crystal structure of the human SULT1E1 monomer is shown with the highly conserved 'Region 1' (blue) and 'Region 4' (purple) sequences highlighted as well as the locations of the three genetic polymorphisms observed in the course of the present study.

change in amino acid from one with a polar to one with an acidic side chain could potentially influence access of the substrate to the active site–a possible explanation for the increase in apparent K_m value which we observed (Table 5). In contrast, the Ala32Val polymorphism involved two residues with hydrophobic side chains. This alteration in amino acid resulted in both decreased enzyme activity and decreased immunoreactive protein (Figures 2 and 3), but without a substantial effect on substrate kinetics. On the basis of the X-ray crystal structure, this amino acid was located on the surface the protein outside the putative substrate-binding site. The third change in amino acid, Pro253His, was located three amino acids upstream of the 'Region IV' conserved SULT sequence motif, the PAPS-binding site (Weinshilboum & Otterness, 1994; Varin *et al*., 1995; Weinshilboum *et al*., 1997). This final polymorphism resulted in a two- to three-fold increase in K_m and V_{max} values for E2, while the K_m for PAPS increased approximately three-fold, but with only a slight increase in V_{max} (Table 5). Although this slight increase in V_{max} was statistically significant, its *in vivo* functional significance remains to be determined.

Discussion

Estrogens are used as drugs–especially in postmenopausal hormone replacement therapy (Barratt-Connor & Grady, 1998). In addition, epidemiologic studies have shown repeatedly that cumulative lifetime exposure to estrogen is an important risk factor for breast cancer (Colditz, 1997). Estrogen inactivation in humans occurs, in part, as a result of sulfate conjugation catalyzed by SULTs, with SULT1E1 having the lowest K_m values of known human SULT isoforms for these steroid hormones (Hobkirk, 1993; Aksoy *et al*., 1994; Falany *et al*., 1995; Strott, 1996; Adjei & Weinshilboum, 2002). SULT1E1 also catalyzes the sulfate conjugation of catecholestrogens, 2-methoxyestradiol (an experimental antineoplastic agent) and other structurally related steroids (Aksoy *et al*., 1994; Falany *et al*., 1995; Adjei *et al*., 2001; Adjei & Weinshilboum, 2002). Many human SULTs are genetically polymorphic, and those polymorphisms can be functionally significant (Raftogianis *et al*., 1997; 1999; Freimuth *et al*., 2001; Thomae *et al*., 2002)–raising the possibility that SULT1E1 might also display genetically mediated variation in function.

Therefore, we set out to determine whether ***SULT1E1*** might have genetic polymorphisms, whether those polymorphisms might differ among ethnic groups and whether they might have functional implications. We began by resequencing the eight exons, intron sequence flanking those exons and a portion of the 5′-flanking region of the gene using DNA samples from African-American and Caucasian-American subjects. In all, 23 polymorphisms, including three nonsynonymous cSNPs, were observed in the 120 DNA samples (240 alleles) resequenced. There were more SNPs in African-American DNA samples than in the samples from Caucasian-American subjects (Table 2, Figure 1). This observation is compatible with previous reports of greater DNA sequence diversity in African populations than in other ethnic groups (Kaessmann *et al*., 1999; Thomae *et al*., 2002; Tishkoff & Williams, 2002). Several of the ***SULT1E1*** polymorphisms were tightly linked (Table 3), and a series of unambiguous haplotypes were observed in both populations (Table 4). Expression constructs were then created for the three nonsynonymous ***SULT1E1*** cSNPs observed, and those constructs were used to transfect COS-1 cells transiently. The His253 and wild-type allozymes were similar with respect to levels of basal activity and immunoreactive protein (Figures 2 and 3). However, there were decreases in basal levels of enzyme activity for variant allozymes Tyr22 and Val32, decreases that were associated with decreased levels of immunoreactive protein (Figures 2 and 3). Furthermore, the thermal stability of the Tyr22 and Val32 was decreased

(Table 4). These decreases in thermal stability may be associated with alterations in protein folding (Wang & Moult, 2001)–a hypothesis that will have to be pursued in the course of future studies. Finally, substrate kinetic studies showed that apparent K_m values for E2 and PAPS also varied among allozymes (Table 5).

In summary, we have detected and functionally characterized genetic variation in the human ***SULT1E1*** gene. Although the three nonsynonymous cSNPs that we observed occurred relatively infrequently and, therefore, might represent mutations rather than polymorphisms (Table 2), functional genomic studies showed that two of them resulted in significant decreases in both levels of enzyme activity and quantity of enzyme protein (Figures 2 and 3). In addition, allozymes encoded by these variant alleles also displayed alterations in substrate kinetics (Table 5). The *in vivo* functional implications of these observations remain to be determined, as do the possible functional consequences of the other polymorphisms that we observed. The recent observation that disruption of the SULT1E1 gene in mice results in male infertility (Qian *et al*., 2001) raises the question of whether a similar phenomenon might occur in humans. For example, an individual homozygous for the Tyr22 allozyme might be at risk for this phenotype–although that possibility is purely speculative in the absence of the identification and study of such subjects. Estrogen exposure is also a major risk factor for hormone-dependent neoplasia such as breast cancer (Henderson & Bernstein, 1996). Therefore, future studies might focus on these variants as risk factors for both reproductive disorders and hormone-dependent cancer. The present study also represents an example of the way in which a genotype-to-phenotype strategy can be used to discover genetic polymorphisms and test their functional significance. Finally, our results raise the possibility that ethnic-specific genetic variation in ***SULT1E1***–a gene encoding an estrogen-metabolizing enzyme–may contribute to individual differences in the biotransformation of this steroid hormone. If that proves to be the case *in vivo*, these 'pharmacogenetic' variations could contribute to individual differences in lifetime estrogen exposure as well as variation in the metabolism and, thus, the effect of exogenously administered estrogens.

We thank Mrs Luanne Wussow for her assistance with the preparation of this manuscript and Dr Daniel Schaid for performing the linkage and haplotype analyses.

References

ADJEI, A.A. & WEINSHILBOUM, R.M. (2002). Catecholestrogen sulfation: possible role in carcinogenesis. *Biochem. Biophys. Res. Commun.,* **292,** 402–408.

ADJEI, A.A., WOOD, T.C. & WEINSHILBOUM, R.M. (2001). 2-Methoxyestradiol (2-ME_2) sulfation: possible metabolic pathway. *Clin. Pharm. Ther.,* **69,** 75.

AKSOY, I.A., WOOD, T.C. & WEINSHILBOUM, R.M. (1994). Human liver estrogen sulfotransferase: cDNA cloning, expression and biochemical characterization. *Biochem. Biophys. Res. Commun.,* **200,** 1621–1629.

BARRATT-CONNOR, E. & GRADY, D. (1998). Hormone replacement therapy, heart disease and other considerations. *Ann. Rev. Public Health,* **19,** 55–72.

BRADFORD, M.M. (1976). A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principles of protein–dye binding. *Anal. Biochem.,* **72,** 248–254.

CHADWICK, R.B., CONRAD, M.P., MCGINNIS, M.D., JOHNSTON-DOW, L., SPURGEON, S.L. & KRONICK, M.N. (1996). Heterozygote and mutation detection by direct automated fluorescent DNA sequencing using a mutant *Taq* DNA polymerase. *BioTechniques,* **20,** 676–683.

CLELAND, W.W. (1963). Computer programmes for processing enzyme kinetic data. *Nature,* **198,** 463–465.

COLDITZ, G.A. (1997). A biomathematical model of breast cancer incidence: the contribution of reproductive factors to variation in breast cancer incidence. In: *General Motors Cancer Research Foundation. Accomplishments in Cancer Research 1996*. ed. Fortner, J.G. & Shaap, P.A., pp. 116–121. Philadelphia: Lippincott-Raven.

DRYSDALE, C.M., MCGRAW, D.W., STACK, C.B., STEPHENS, J.C., JUDSON, R.S., NANDABALAN, K., ARNOLD, K., RUANO, G. & LIGGETT, S.B. (2000). Complex promoter and coding region β_2-adrenergic receptor haplotypes alter receptor expression and predict *in vivo* responsiveness. *Proc. Natl. Acad. Sci. U.S.A.,* **97,** 10483–10488.

EXCOFFIER, L. & SLATKIN, M. (1995). Maximum-likelihood estimation of molecular haplotype frequencies in a diploid population. *Mol. Biol. Evol.,* **12,** 921–927.

FALANY, C.N., KRASNYKH, V. & FALANY, J.L. (1995). Bacterial expression and characterization of a cDNA for human liver estrogen sulfotransferase. *J. Steroid Biochem. Mol. Biol.,* **52,** 529–539.

FALANY, J.L., AZZIZ, R. & FALANY, C.N. (1998). Identification and characterization of cytosolic sulfotransferases in normal human endometrium. *Chem. Biol. Interact.,* **109,** 329–339.

FALANY, J.L. & FALANY, C.N. (1996). Regulation of estrogen sulfotransferase in human endometrial adenocarcinoma cells by progesterone. *Endocrinology,* **137,** 1395–1401.

FOLDES, A. & MEEK, J.L. (1973). Rat brain phenolsulfotransferase–partial purification and some properties. *Biochim. Biophys. Acta,* **327,** 365–374.

FREIMUTH, R.R., ECKLOFF, B., WIEBEN, E.D. & WEINSHILBOUM, R.M. (2001). Human sulfotransferase SULT1C1 pharmacogenetics: gene resequencing and functional genomic studies. *Pharmacogenetics,* **11,** 747–756.

FUJITA, K., NAGATA, K., YAMAZAKI, T., WATANABE, E., SHIMADA, M. & YAMAZOE, Y. (1999). Enzymatic characterization of human cytosolic sulfotransferases; identification of ST1B2 as a thyroid hormone sulfotransferase. *Biol. Pharm. Bull.,* **22,** 446–452.

GORDON, D., ABAJIAN, C. & GREEN, P. (1998). *Consed*: a graphical tool for sequence finishing. *Genome Res.,* **8,** 195–202.

HARTL, D.L. & CLARK, A.G. (1997). *Principles of Population Genetics*. 3rd edn. pp. 96–106.. Sunderland, MA: Sinauer Associates, Inc.

HEDRICK, P.W. (2000). *Genetics of Populations*. 2nd edn. pp. 396–405.. Sudbury MA: Jones and Bartlett.

HEMPEL, N., BARNETT, L., BOLTON-GROB, R.M., LIYOU,, N.E. & MCMANUS,, M.E. (2000). Site-directed mutagenesis of the substrate-binding cleft of human estrogen sulfotransferase. *Biochem. Biophys. Res. Commun.,* **276,** 224–230.

HENDERSON, B.E. & BERNSTEIN, L. (1996). Endogenous and exogenous hormone factors. In: *Diseases of the Breast*. ed. Harris, J.R., Lippman, M.E., Morrow, M. and Hellman, S., pp. 185–200. Philadelphia: Lippincott-Raven.

HER, C., AKSOY, I.A., KIMURA, S., BRANDRIFF, B.F., WASMUTH, J.J. & WEINSHILBOUM, R.M. (1995). Human estrogen sulfotransferase gene (STE): cloning, structure and chromosomal localization. *Genomics,* **29,** 16–23.

HER, C., SZUMLANSKI, C., AKSOY, I.A. & WEINSHILBOUM, R.M. (1996). Human jejunal estrogen sulfotransferase and dehydroepiandrosterone sulfotransferase: immunochemical characterization of individual variation. *Drug Met. Dispos.,* **24,** 1328–1335.

HOBKIRK, R. (1993). Steroid sulfation. *Trends Endocrinol. Metab.,* **4,** 69–74.

IIDA, A., SEKINE, A., SAITO, S., KITAMURA, Y., KITAMOTO, T., OSAWA, S., MISHIMA, C. & NAKAMARA, Y. (2001). Catalog of 320 single nucleotide polymorphisms (SNPs) in 20 quinone oxidoreductase and sulfotransferase genes. *J. Hum. Genet.,* **46,** 251–259.

KAESSMANN, H., HEISSIG, F., VON HAESELER, A. & PÄÄBO, S. (1999). DNA sequence variation in a non-coding region of low recombination on the human X chromosome. *Nat. Genet.,* **22,** 78–81.

KAKUTA, Y., PEDERSEN, L.G., CARTER, C.W., NEGISHI, M. & PEDERSEN, L.C. (1997). Crystal structure of estrogen sulphotransferase. *Nat. Struct. Biol.,* **4,** 904–908.

KANG, S.S., WONG, P.W., BOCK, H.G., HORWITZ, A. & GRIX, A. (1991). Intermediate hyperhomocysteinemia resulting from compound heterozygosity of methylenetetrahydrofolate reductase mutations. *Am. J. Hum. Genet.,* **48,** 546–551.

LONG, J.C., WILLIAMS, R.C. & URBANEK, M. (1995). An E–M algorithm and testing strategy for multiple-locus haplotypes. *Am. J. Hum. Genet.,* **56,** 799–810.

NICKERSON, D.A., TOBE, V.O. & TAYLOR, S.L. (1997). PolyPhred: automating the detection and genotyping of single nucleotide substitutions using fluorescence-based resequencing. *Nucleic Acids Res.,* **25,** 2745–2751.

OTAKE, Y., NOLAN, A.L., WALLE, U.K. & WALLE, T. (2000). Quercetin and resveratrol potently reduce estrogen sulfotransferase activity in normal human mammary epithelial cells. *J. Steroid Biochem. Mol. Biol.,* **73,** 265–270.

PASQUALINI, J.R. & CHETRITE, G. (1996). Activity, regulation and expression of sulfatase, sulfotransferase, and 17β-hydroxysteroid dehydrogenase in breast cancer. In: *Hormone-dependent Cancer*. ed. Pasqualini, J.R. & Katzenellenbogen, B.S., Chapter 2, pp. 25–80. New York: Marcel-Dekker, Inc.

PEDERSEN, L.C., PETROTCHENKO, E., SHEVTSOV, S. & NEGISHI, M. (2002). Crystal structure of the human estrogen sulfotransferase–PAPS complex: evidence for catalytic role of Ser137 in the sulfuryl transfer reaction. *J. Biol. Chem.,* **277,** 17928–17932.

PREUSS, C.V., WOOD, T.C., SZUMLANSKI, C.L., RAFTOGIANIS, R.B., OTTERNESS, D.M., GIRARD, B., SCOTT, M.C. & WEINSHILBOUM, R.M. (1998). Human histamine *N*-methyltransferase pharmacogenetics: common genetic polymorphisms that alter activity. *Mol. Pharmacol.,* **53,** 708–717.

QIAN, Y., DENG, C. & SONG, W.C. (1998). Expression of estrogen sulfotransferase in MCF-7 cells by cDNA transfection suppresses the estrogen response: potential role of the enzyme in regulating estrogen-dependent growth of breast epithelial cells. *J. Pharmacol. Exp. Ther.,* **286,** 555–560.

QIAN, Y.M. & SONG, W.C. (1999). Regulation of estrogen sulfotransferase expression in Leydig cells by cyclic adenosine 3′,5′-monophosphate and androgen. *Endocrinology,* **140,** 1048–1053.

QIAN, Y.M., SUN, X.J., TONG, M.H., LI, X.P., RICHA, J. & SONG, W.C. (2001). Targeted disruption of the mouse estrogen sulfotransferase gene reveals a role of estrogen metabolism in intracrine and paracrine estrogen regulation. *Endocrinology,* **142,** 5342–5350.

RAFTOGIANIS, R.B., WOOD, T.C., OTTERNESS, D.M., VAN LOON, J.A. & WEINSHILBOUM, R.M. (1997). Phenol sulfotransferase pharmacogenetics in humans: association of common ***SULT1A1*** alleles with TS PST phenotype. *Biochem. Biophys. Res. Commun.,* **239,** 298–304.

RAFTOGIANIS, R.B., WOOD, T.C. & WEINSHILBOUM, R.M. (1999). Human phenol sulfotransferases ***SULT1A2*** and ***SULT1A1***: genetic polymorphisms, allozyme properties and human liver genotype–phenotype correlations. *Biochem. Pharmacol.,* **58,** 605–610.

REITER, C., MWALUKO, G., DUNNETTE, J., VAN LOON, J. & WEINSHILBOUM, R. (1983). Thermolabile and thermostable human platelet phenol sulfotransferase: substrate specificity and physical separation. *Naunyn-Schmiedeberg's Archv. Pharmacol,* **324,** 140–147.

REITER, C., MWALUKO, G. & WEINSHILBOUM, R. (1982). Human platelet phenol sulfotransferase (PST): substrate specificity of thermolabile and thermostable forms. *Fed. Proc.,* **41,** 1478.

SCANLON, P.D., RAYMOND, F.A. & WEINSHILBOUM, R.M. (1979). Catechol-*O*-methyltransferase: thermolabile enzyme in erythrocytes of subjects homozygous for the allele for low activity. *Science,* **203,** 63–65.

SHIELD, A.J. & WEINSHILBOUM, R.M. (2002). Human catechol *O*-methyltransferase (COMT) polymorphisms: functional genomic studies. *Drug Met. Rev.,* **34,** 153.

SONG, W.-C., MOORE, R., MCLACHLAN, J.A. & NEGISHI, M. (1995). Molecular characterization of a testis-specific estrogen sulfotransferase and aberrant liver expression in obese and diabetogenic C57BL/KsJ-db/db mice. *Endocrinology,* **136,** 2477–2484.

SONG, W.-C., QIAN, Y. & LI, A.P. (1998). Estrogen sulfotransferase expression in the human liver: marked interindividual variation and lack of gender specificity. *J. Pharmacol. Exp. Ther.,* **284,** 1197–1202.

SONG, W.-C., QIAN, Y., SUN, X. & NEGISHI, M. (1997). Cellular localization and regulation of expression of testicular estrogen sulfotransferase. *Endocrinology,* **138,** 5006–5012.

STROTT, C.A. (1996). Steroid sulfotransferases. *Endocr. Rev.,* **17,** 670–697.

SZUMLANSKI, C., OTTERNESS, D., HER, C., LEE, D., BRANDRIFF, B., KELSELL, D., SPURR, N., LENNARD, L., WIEBEN, E. & WEINSHILBOUM, R. (1996). Thiopurine methyltransferase pharmacogenetics: human gene cloning and characterization of a common polymorphism. *DNA Cell Biol.,* **15,** 17–30.

TAI, H.-L., FESSING, M.Y., BONTEN, E.J., YANISHEVSKY, Y., D'AZZO, A., KRYNETSKI, E.Y. & EVANS, W.E. (1999). Enhanced proteasomal degradation of mutant human thiopurine *S*-methyltransferase (TPMT) in mammalian cells: mechanism for TPMT protein deficiency inherited by *TPMT*2, TPMT*3A, TPMT*3B* or *TPMT*3C*. *Pharmacogenetics,* **9,** 641–650.

THOMAE, B.A., ECKLOFF, B.W., FREIMUTH, R.R., WIEBEN, E.D. & WEINSHILBOUM, R.M. (2002). Human sulfotransferase SULT2A1 pharmacogenetics: genotype to phenotype studies. *Pharmacogenomics J.,* **2,** 48–56.

TISHKOFF, S.A. & WILLIAMS, S.M. (2002). Genetic analysis of African populations: human evolution and complex disease. *Nat. Rev. Genet.,* **3,** 611–621.

VARIN, L., DELUCA, V., IBRAHIM, R.K. & BRISSON, N. (1995). Molecular characterization of two plant flavonol sulfotransferases. *Proc. Natl. Acad. Sci. U.S.A.,* **89,** 1286–1290.

WANG, Z. & MOULT, J. (2001). SNPs, protein structure and disease. *Hum. Mutat.,* **17,** 263–270.

WEINSHILBOUM, R. & OTTERNESS, D. (1994). Sulfotransferase enzymes. In: *Conjugation–Deconjugation Reactions in Drug Metabolism and Toxicity*. ed. Kaufmann, F.C., Chapter 2, 'Handbook of Experimental Pharmacology' series, Vol. 112, pp. 45–78. Berlin, Heidelberg: Springer-Verlag.

WEINSHILBOUM, R.M., OTTERNESS, D.M., AKSOY, I.A., WOOD, T.C., HER, C. & RAFTOGIANIS, R.B. (1997). Sulfotransferase molecular biology: cDNAs and genes. *FASEB J.,* **11,** 3–14.

WILKINSON, G.N. (1961). Statistical estimations in enzyme kinetics. *Biochem. J.,* **80,** 324–332.

ZHANG, H., VARLAMOVA, O., VARGAS, P.M., FALANY, C.N. & LEYH, T.X. (1998). Sulfuryl transfer: the catalytic mechanism of human estrogen sulfotransferase. *J. Biol. Chem.,* **273,** 10888–10892.

(Received February 5, 2003
Revised May 1, 2003
Accepted May 13, 2003)