Sulfotransferase molecular biology: cDNAs and genes

RICHARD M. WEINSHILBOUM,[1] DIANE M. OTTERNESS, IBRAHIM A. AKSOY, THOMAS C. WOOD, CHENGTAO HER, AND REBECCA B. RAFTOGIANIS

Department of Pharmacology, Mayo Medical School/Mayo Clinic/Mayo Foundation, Rochester, Minnesota 55905, USA

As has occurred in most fields, rapid progress has been made in our understanding of the molecular structure of the cDNAs and genes for the cytosolic sulfotransferases. More than 30 ST cDNAs have been isolated from several vertebrate species as well as from plants. These enzymes share significant sequence homology at both the level of the cDNAs as well as in their intron/exon structure, suggesting that the cytosolic STs have evolved from a common ancestral gene. The continued characterization of both the ST cDNAs and the structural genes that encode them will provide insights into the possible mechanisms of regulation of expression of the different forms as well as the basis for the allelic differences observed in the expression of the STs. We can expect that the number of ST cDNAs will increase rapidly, especially as new forms are cloned from different species. This review describes our understanding of some of the relationships between the STs at a molecular level at this time in a rapidly growing field. Such a review provides an important summation of the field and a basis with which to evaluate future new ST cDNAs and genes that are certain to be characterized.

—Charles N. Falany, Coordinating Editor

ABSTRACT **Sulfotransferase (ST) enzymes catalyze the sulfate conjugation of many hormones, neurotransmitters, drugs, and xenobiotic compounds. These reactions result in enhanced renal excretion of the sulfate-conjugated reaction products, but they can also lead to the formation of "bioactivated" metabolites. ST enzymes are members of an emerging gene superfamily that presently includes phenol ST (PST), hydroxysteroid ST (HSST), and, in plants, flavonol ST (FST) "families," members of which share at least 45% amino acid sequence identity. These families can be further subdivided into "subfamilies" that are at least 60% identical in amino acid sequence. For example, the PST family includes both PST and estrogen ST (EST) subfamilies. Amino acid sequence motifs exist within ST enzymes that are conserved throughout phylogeny. These signature sequences may be involved in the binding of 3′-phosphoadenosine-5′-phosphosulfate, the cosubstrate for the sulfonation reaction. There are presently five known human cytosolic ST enzymes: an EST, an HSST, and three PSTs. cDNAs and genes for all of these enzymes have been cloned, and chromosomal localizations have been reported for all five genes. Genes for these human enzymes, as well as those of other mammalian cytosolic ST enzymes that have been cloned, show a high degree of structural homology, with conservation of the locations of most intron/exon splice junctions. Human ST enzyme expression varies among individuals. Functionally significant genetic polymorphisms for ST enzymes in humans have been reported, and other molecular genetic mechanisms that might be involved in the regulation of the expression of these enzymes are being explored. Knowledge of the molecular biology of cytosolic ST enzymes, when placed within a context provided by decades of biochemical research, promises to significantly enhance our understanding of the regulation of the sulfate conjugation of hormones, neurotransmitters, and drugs.—Weinshilboum, R. M., Otterness, D. M., Aksoy, I. A., Wood, T. C., Her, C., Raftogianis, R. B. Sulfotransferase molecular biology: cDNAs and genes. *FASEB J.* 11, 3–14 (1997)**

Key Words: sulfation · sulfonation · sulfate conjugation · gene structure · chromosomal localization · molecular regulation

SULFATE CONJUGATION IS AN important pathway in the biotransformation of many hormones, neurotransmitters, drugs, and xenobiotic compounds (1–4). Flavonols and other endogenous compounds in plants may also undergo sulfate conjugation (5). The cytosolic sulfotransferase (ST)[2]

[1] Correspondence.

[2] Abbreviations: ST, sulfotransferase; DHEA, dehydroepiandrosterone; PAPS, 3′-phosphoadenosine-5′-phosphosulfate; TS, thermostable; TL, thermolabile; FSTs, flavonol STs; SMP, senescence marker protein; EST, estrogen ST; UTR, untranslated region.

0892-6638/97/0011-0003/$01.50 © FASEB

enzymes that catalyze these reactions are members of an emerging gene superfamily. Although the reactions catalyzed by these enzymes have historically been referred to as "sulfation," they are more accurately described chemically as sulfonation reactions. Sulfonation increases the water solubility of most compounds, and therefore their renal excretion, but it can also result in bioactivation to form active metabolites of drugs or carcinogens from procarcinogens (1–4). Sulfate conjugation was first described as a pathway in biotransformation by Baumann in 1876 (6). However, only within the past two decades has the biochemistry of ST enzymes been well characterized, and only within the past few years, with a steadily accelerating pace, has application of the techniques of molecular biology begun to make it possible to determine definitively the number of these enzymes, their substrate specificities, regulation, and evolutionary relationships.

ST enzymes catalyze a so-called "phase II" pathway in drug and xenobiotic metabolism—a conjugation reaction. ST isoforms, like those of other drug-metabolizing enzymes, display significant overlap in their substrate specificities, resulting in confusion when biochemical techniques have been used to study these enzymes—confusion compounded by the existence of several different "streams" of ST enzyme research. For example, pharmacologic studies have tended to address the role of these enzymes in the biotransformation of drugs and xenobiotic compounds, whereas endocrinologists have focused attention on the sulfonation of hormones, especially steroid hormones. These fields of research have tended to develop in parallel, with the use of different names for what have eventually proved to be identical enzymes. That situation is not unique, but rather mirrors the historical development of research involving other drug-metabolizing enzyme systems such as the cytochromes P450 and the UDP-glucuronosyltransferases. Within the past decade, and particularly within the past 5 years, cDNAs and genes for many cytosolic ST enzymes have been cloned and characterized. The results of those studies, when evaluated within a context provided by decades of biochemical work, have enhanced our ability to classify these enzymes and to study their function and regulation. To place our rapidly evolving understanding of the molecular biology of cytosolic ST enzymes within that context, it will be necessary to review briefly the biochemical background against which molecular experiments have been performed. The reader is referred to other articles in this *FASEB J.* series in which the biochemistry and function of ST enzymes are described in greater detail.

ST BIOCHEMISTRY: AN OVERVIEW

The enzymes that catalyze the sulfonation of drugs, hormones, and neurotransmitters in animals, like those in plants that catalyze similar reactions with flavonols as substrates, are cytosolic in subcellular localization (1–5). Although there are also membrane-bound ST enzymes that catalyze the sulfonation of polysaccharides and tyrosyl residues within proteins (7, 8), the subsequent discussion will focus entirely on cytosolic STs—enzymes that are predominantly dimers, both homo- and heterodimers, with monomer M_r values that vary from 30 to 36 kDa (3, 4, 9). However, in plants and in some mammals, these enzymes can exist as catalytically active monomers (5, 10). In the past, ST enzymes were classified on the basis of their substrate specificities or the general nature of the reactions that they catalyzed (2, 4). The hydroxyl group is the most common target for sulfonation, and included among classes of compounds that undergo O-sulfonation are "simple" planar phenols such as 4-nitrophenol, phenolic monoamine neurotransmitters such as dopamine, estrogens and other phenolic steroid hormones, nonphenolic hydroxysteroids such as dehydroepiandrosterone (DHEA), and, in plants, flavonols such as quercetin (4). The cosubstrate for these reactions is 3′-phosphoadenosine-5′-phosphosulfate (PAPS) (11). All cytosolic ST enzymes are included within the International Union of Biochemistry nomenclature category of "sulfotransferases," EC 2.8.2 (12). However, since these enzymes display significant overlap in their substrate specificities, and since virtually every laboratory studying them has adopted its own system of trivial nomenclature, it is often difficult to know exactly which enzyme is being described in any given research communication. Ultimately, to dispel such uncertainty would require determination of the primary amino acid sequence and the application of a unique name to each gene product. As a result of the cloning of cytosolic ST enzyme cDNAs and genes, that may soon become possible.

It is important to review briefly our current understanding of the biochemistry of cytosolic ST enzymes in humans because they have been studied extensively and because of the important role of sulfonation in drug, hormone, and neurotransmitter metabolism (1–4). Human tissues are presently known to express at least five ST enzymes. These five enzymes differ with regard to their substrate specificities, inhibitor sensitivities, thermal stabilities, and regulation (i.e., their levels of activity are regulated independently). Three are PSTs, one is an EST, and one is an HSST. Biochemical experiments conducted before ST cDNAs and genes had been cloned unequivocally identified only three human cytosolic ST activities (1, 3, 4). Those biochemical studies demonstrated the existence of one form of PST that was thermostable (TS), catalyzed the sulfate conjugation of micromolar concentrations of 4-nitrophenol and other simple phenols, and was sensitive to the ST inhibitor 2,6-dichloro-4-nitrophenol. A separate form of PST expressed in human tissues was thermolabile (TL), preferentially catalyzed the sulfonation of micromolar concentrations of dopamine and other phenolic monoamines, and was relatively resistant to 2,6-dichloro-4-nitrophenol inhibition. These two enzymes have been referred to as either the TS and TL or, on the basis of their substrate specificities, as the phenol-preferring (P) and monoamine-preferring (M) forms of PST, respectively (1, 4). The third biochemically defined cytosolic

ST in humans was an HSST that catalyzed the sulfonation of cholesterol, bile acids, and 3-β-hydroxysteroids such as DHEA (13, 14). That enzyme has been referred to most frequently as DHEA ST (13). Because all three of these ST enzymes have been reported to be capable of catalyzing the sulfonation of estrogens (13, 15, 16), the question of whether a separate and distinct EST existed in human tissues remained open until a cDNA for a human EST was cloned and expressed (17). Subsequently, cDNA and gene cloning experiments demonstrated the existence of two separate human genes that encode two separate TS PST cDNAs, both located on the short arm of human chromosome 16 in close proximity to the gene for TL PST (18–20). Therefore, the number of known cytosolic ST enzymes in humans has increased in a short span of time from three to five, and now includes an EST, DHEA ST, and three PSTs. Similar developments have occurred with other species. The subsequent discussion will outline ways in which the cloning of cDNAs for cytosolic ST enzymes has begun to make it possible to classify these enzymes on the basis of their primary amino acid sequence, facilitate the study of their function, and make it possible to clone and structurally characterize ST genes. Finally, ways in which this molecular information is leading us back to functional biology will be described briefly.

ST cDNAs: CLONING AND CLASSIFICATION

cDNA cloning was the first step taken in studies of the molecular biology of ST enzymes. Thirty eukaryotic ST cDNAs had been reported at the time of this review. When evaluating reported cDNAs, we arbitrarily assumed that two sequences represented alleles at a single genetic locus rather than the products of different ST genes if their encoded amino acid sequences were more than 97% identical. On the basis of encoded amino acid sequences, these 30 cDNAs could be classified as members of at least three large "families," one of which included at least two "subfamilies" (**Fig. 1**) (4, 18, 20). Members of each family of enzymes showed at least 45% amino acid sequence identity, whereas members of subfamilies were 60% or more identical in amino acid sequence. Although these percentages were selected arbitrarily, they are very similar to those that have proved useful in classifying other gene superfamilies such as the cytochromes P450 (21). Two of these ST enzyme families are expressed in mammals, a PST and an HSST family, whereas a family of flavonol STs (FSTs) is expressed in plants (Fig. 1). There are at least two subfamilies within the PST family: the "PSTs," enzymes that preferentially catalyze the sulfate conjugation of simple planar phenols, and the "ESTs," which catalyze the sulfonation of estrogens with very low K_m values. For purposes of this review, the species from which each cDNA was cloned is indicated by a small letter or letters in which "r" represents rat, "m" mouse, "h" human, "b" bovine, "gp" guinea pig, "mf" *Macaca fascicularis*, "fc" *Flaveria chloraefolia*, "fb" *Flaveria bidentis*, and "at" *Arabidopsis thaliana*. The names of the enzymes themselves, once again with a few exceptions, are abbreviated PST, EST, HSST, or FST. Exceptions include a rat liver enzyme within the PST family that is designated AAFST for "acetylaminofluorene" ST (22), a rat liver enzyme designated 1B1ST that has been reported to catalyze the sulfonation of dopamine, 4-nitrophenol, and thyroid hormones (23), and a rat liver enzyme designated SMP for "senescence marker protein," a name that will be explained subsequently (24). The rAAFST, r1B1ST, and atST cDNAs are the most striking exceptions to the family and subfamily classification outlined in the dendrogram

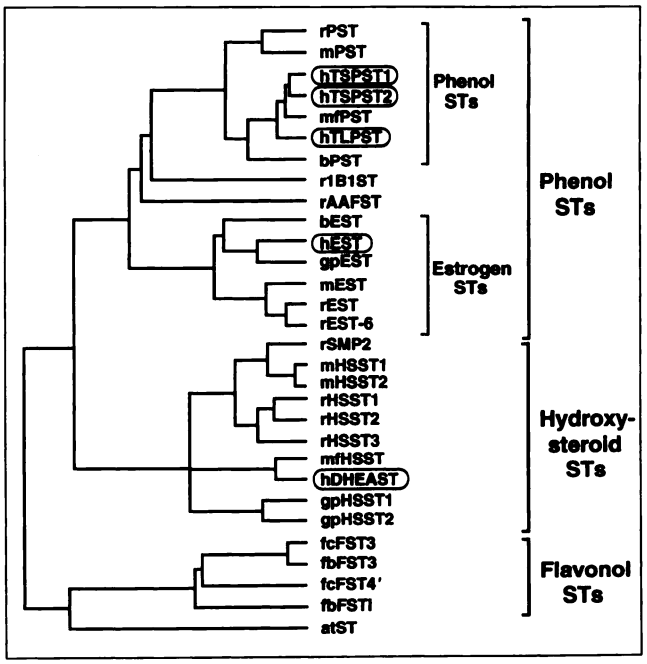

Figure 1. ST enzyme classification on the basis of primary amino acid sequence. The dendrogram depicts the degree of amino acid sequence identity among cytosolic ST enzymes. It was constructed by use of the PILEUP program (61). This classification scheme is a modification of that proposed by Weinshilboum and Otterness (4), Otterness et al. (39), Her et al. (18), and Raftogianis et al. (20) in which "families" of enzymes, phenol STs (PSTs), hydroxysteroid STs (HSSTs), and flavonol STs (FSTs), include cDNAs that encode proteins with 45% or greater sequence identity, whereas "subfamilies" have 60% or greater amino acid sequence identity. Functional designations of families are shown at the extreme right of the figure, with subfamily designations listed to the left of family names. Human enzymes are "circled." Abbreviations for ST enzymes include species designations, with "r" representing rat, "m" mouse, "h" human, "b" bovine, "gp" guinea pig, "mf" *Macaca fascicularis*, "fc" *Flaveria chloraefolia*, "fb" *Flaveria bidentis*, and "at" *Arabidopsis thaliana*. The names of individual enzymes are abbreviated PST, HSST, FST, EST for estrogen ST, AAFST for N-hydroxyl-2-acetylaminofluorene ST, and 1B1ST for a rat liver ST with multiple substrates. TL represents "thermolabile", TS "thermostable" and SMP "senescence marker protein." Sources of sequence data include the following: rPST (62), mPST (63), hTSPST1 (64), hTSPST2 (65), mfPST (66), hTLPST (28), bPST (60), r1B1ST (67), rAAFST (22), bEST (29), hEST (17), gpEST (68), mEST (69), rEST (70), rEST-6 (71), rSMP2 (24), mHSST1 (72), mHSST2 (73), rHSST1 (74), rHSST2 (75), rHSST3 (76), mfHSST (77), hDHEAST (30), gpHSST1 (78), gpHSST2 (79), fcFST3 (32), fbFST3 (80), fcFST4' (32), fbFST1 (81), and atST (25).

TABLE 1. *Cytosolic ST cDNA abbreviations and GenBank accession numbers*[a]

cDNA	cDNA Designation(s) by Reporting Authors	Accession #	References
PST FAMILY			
rPST	PST-1	X52883, S42994	(62)
	Mx-STb	L19998	(82)
	AST IV	X68640	(83)
	Tyrosine-ester ST	-----	(84)
	PST-1	-----	(85)
mPST	mST_{p1}	L02331	(63)
hTSPST1	P-PST-1	L19999	(64)
	HAST1	L10819	(26)
	HAST2	L19955, U09031	(27)
	P-PST	X84654	(86)
	ST1A3	X78283	(65)
	H-PST	U26309	(87)
hTSPST2	ST1A2	X78282	(65)
	HAST4V	U28169	(88)
	HAST4	U28170	(88)
mfPST	monPST-1	D85514	(66)
hTLPST	HAST3	L19956	(27)
	TL PST	U08032	(28)
	hEST	L25275	(16)
	m-PST	X84653	(86)
	hM-PST	-----	(89)
bPST	PST	U35253	(60)
rlB1ST	ST1B1	----	(67)
	rT/DST	U38419	(23)
rAAFST	ST1C1, HAST-1	L22339	(22)
bEST	OST	X56395, M54942, M38672	(29)
hEST	hEST	U08098	(17)
	hEST-1	S77383	(90)
gpEST	EST	S45979, U09552	(68)
mEST	mEST	S78182	(69)
rEST	EST_r	M86758	(70)
	rEST-3	S76489	(71)
	Ste1	U50204	(91)
	Ste2	U50205	(91)
rEST-6	rEST-6	S76490	(71)
HSST FAMILY			
rSMP2	SMP-2	J02643	(24)
mHSST1	mST_{a1}	L02335	(72)
mHSST2	mST_{a2}	L27121	(73)
rHSST1	ST-20	M31363	(74)
	ST-21a	D14987	(76)
	ST-21b	D14988	(76)
rHSST2	ST-40	M33329	(75)
	ST-41	X63410	(92)
rHSST3	ST-60	D14989	(76)
mfHSST	monHST-1	D85521	(77)
hDHEAST	DHEA ST	U08024 (A) U08025 (G)	(30)
	hST_a	S43859, S43861, L02337	(93)
	DHEA-ST8	X70222 S53620 L20000	(94)
	HST-hfa	X84816	(95)
gpHSST1	gpHST	U06871	(78)
gpHSST2	gpHST2	U35115	(79)
	gpPreg-ST	U55944	(96)
FST FAMILY			
fcFST3	pFST3	M84135	(32)
fbFST3	pBFST3	U10275	(80)
fcFST4'	pFST4'	M84136	(32)
fbFST!	pBFSTX	U10277	(81)
atST	RaRo47	Z46823	(25)

[a] The left-hand column lists cDNA designations in the same order as depicted in Fig. 1. Abbreviations are defined in the legend to Fig. 1. cDNA designations used by the authors reporting the sequences, as well as GenBank accession numbers for each sequence and references are listed to the right. The table demonstrates the difficulty of knowing exactly which ST enzyme is being described in any given research communication.

shown in Fig. 1. On the basis of amino acid sequence identity of encoded proteins, both of the rat enzymes fall within the PST family but neither fits well within either the PST or EST subfamilies, a situation similar to that of the plant atST, with sequence identity to the FSTs that varies only from 41 to 43% (25). The two rat enzymes may be members of additional subfamilies within the PST family, and the *Arabidopsis* ST may be the first member of a separate family of plant ST enzymes.

Relationships among the amino acid sequences of proteins encoded by these 30 cDNAs are depicted graphically in the dendrogram shown in Fig. 1, and synonyms for the descriptive designations used in Fig. 1 are listed in **Table 1** in the same order that the enzymes are listed in the dendrogram. The multiple names for the same enzyme listed in Table 1 demonstrate why this field of research has become a scientific "tower of Babel" and why it will be so important to develop a single, universally accepted system of nomenclature for ST enzymes. For example, one group of investigators refers to all human PST isoforms as "HAST" for "human aryl ST" (26, 27), whereas another group uses exactly the same designation for a rat enzyme as an abbreviation for "hydroxylacetylaminofluorene ST" (22)—rAAFST in Fig. 1 and Table 1. The rat HAST falls

structurally into what is probably a separate subfamily from that of the human HAST enzymes. An equally confusing situation exists for the PST sequence named h-TLPST in Fig. 1, also referred to as EST in the list of synonyms in Table 1. The story of this cDNA also provides a striking example of the relative separation of pharmacologic and endocrinologic ST research. More than a decade of biochemical pharmacologic research resulted in the characterization of the TL or M form of PST as an enzyme with a high affinity for monoamine neurotransmitters such as dopamine (1, 3, 4). That enzyme had been purified and studied not only with regard to substrate specificity, but also with respect to many other properties such as sensitivity to 2,6-dichloro-4-nitrophenol inhibition and thermal stability. Several different laboratories cloned and characterized the cDNA for TL PST virtually simultaneously (16, 27, 28). One of those cDNAs was cloned from a human placental library by an endocrine research group that referred to it as an estrogen ST (EST) cDNA because the encoded protein was capable of catalyzing estrogen sulfonation (16). However, that protein fell within the PST subfamily on the basis of amino acid sequence and the biochemical properties of the expressed enzyme were clearly those of TL PST (28). This cDNA has continued to be referred to in the pharmacologic literature as one that encodes the TL or M form of PST, whereas in the endocrinologic literature it is referred to as an EST cDNA. That situation is understandable on the basis of the separation that occurs all too often between different areas of scientific investigation, a separation that our expanding knowledge of ST molecular biology now offers some hope of bridging. Although functional designations have been retained in the names temporarily assigned to ST families and subfamilies shown in Fig. 1, the classification scheme itself is based entirely on amino acid sequence, not on function. Therefore, the scheme depicted in the figure should be thought of as only a "way station" along a path that could eventually lead to a nomenclature that is entirely "function neutral," a nomenclature such as that already applied successfully to the cytochromes P450 (21). However, since such a common nomenclature has yet to be developed and accepted, in this discussion we will continue to refer to these enzymes as PSTs, ESTs, HSSTs, or FSTs.

The first ST cDNA cloned was the rat liver SMP2 cDNA reported by Chatterjee et al. (24). That cDNA was not cloned, however, as a result of an investigation of enzymes that catalyze sulfate conjugation; rather, those investigators were studying a rat liver protein that displayed age- and gender-dependent variation in expression, a so-called "senescence marker protein." Although SMP2 has never been expressed and studied as an ST enzyme, its amino acid sequence clearly places it within the HSST family (Fig. 1). The first ST cDNA cloned as a direct result of studies of sulfate conjugation was a bovine placental EST cDNA cloned after purification of the bovine placental enzyme, determination of its partial amino acid sequence, and the design of degenerate oligonucleotides that could be used to screen a cDNA library (29). As with many protein families, the initial round of ST cDNA cloning required purification of the enzyme of interest, followed by utilization of the purified enzyme to obtain antibodies that could be used to screen cDNA expression libraries or to determine partial amino acid sequence that could be used, as it was for the bovine placental EST, to screen a cDNA library or to design primers for use with the polymerase chain reaction (PCR). After the first three ST cDNAs (one EST, one PST, and one HSST) had been cloned, it was noted that these enzymes contained several areas of very high sequence homology. As discussed subsequently, it has been speculated these ST "signature sequences" might participate in the binding of the cofactor for the reaction, PAPS (4). These highly conserved sequences also proved useful in a practical sense to make it possible to design oligonucleotide primers for the PCR that could be used to amplify and isolate other ST cDNAs (30). Once a large enough number of those cDNAs had been cloned, the process was accelerated by taking advantage of the high sequence homology within families and subfamilies of these enzymes to seek additional isoforms within a species or to clone cross-species orthologues. That process will undoubtedly be accelerated further by the development of panels of cDNA "expressed sequence tags" such as those available from the I.M.A.G.E. consortium (31). The final outcome at the time of this review was that thirty cDNAs from six mammalian and three plant species had been reported (Fig. 1, Table 1). Isolation of ST cDNAs has made it possible to express them and, thus, to study the biochemical properties of the proteins they encode. The results of those experiments will be described in detail in other reviews in this *FASEB J.* series. Cloning of ST cDNAs also revealed several areas of high amino acid sequence homology among these enzymes.

ST SIGNATURE SEQUENCES

Alignment of the amino acid sequences of ST enzymes revealed at least four areas of sequence that were highly conserved throughout phylogeny (32). Two of those regions were particularly striking, one located relatively near the amino terminus (region I), and another near the carboxyl terminus of the enzyme (region IV) (4, 32). Recognition of these highly conserved sequences initially served a practical function in facilitating the cloning of ST cDNAs (30), but because PAPS is a cosubstrate for the reactions catalyzed by all of these enzymes, it was also speculated that these areas of high sequence homology might be involved in PAPS binding (4). A review of ST enzyme biochemistry and molecular biology written when only 11 cDNAs had been cloned identified the amino terminus region I sequence as YPKSGTxW, in which "x" represented any amino acid and the region IV sequence near the carboxyl terminus was RKGxxGDWKNxFT (4). As shown in **Table 2**, those early observations have held up surprisingly well during a period when the number of cDNAs has nearly

TABLE 2. *Cytosolic ST "signature sequences"* [a]

ENZYME	REGION I SEQUENCE	POSITION	REGION IV SEQUENCE	POSITION
rPST	TYPKSGTTW	41-49	RKGTTGDWKNTFT	253-265
mPST	TYPKSGTNW	48-56	RKGTIGDWKNTFT	260-272
hTSPST1	TYPKSGTTW	45-53	RKGMAGDWKTTFT	257-269
hTSPST2	TYPKSGTTW	45-53	RKGMAGDWKTTFT	257-269
mfPST	TYPKSGTTW	45-53	RKGMTGDWKTTFT	257-269
hTLPST	TYPKSGTTW	45-53	RKGMAGDWKTTFT	257-269
bPST	TYPKSGTTW	45-53	RKGITGDWKSTFT	257-269
r1B1ST	TYPKSGTTW	45-53	RKGVVGDWKNYFT	258-270
rAAPST	TYAPAGTTW	53-61	RKGMPGDWKNYFT	266-278
bEST	TYPKSGTTW	45-53	RKGDVGDWKNHFT	257-269
hEST	TYPKSGTTW	44-52	RKGITGDWKNHFT	256-268
gpEST	AYPKSGTTW	46-54	RKGISGDWKNHFT	258-270
mEST	TYPKSGTTW	45-53	RKGIIGDWKNHFP	257-269
rEST	TYPKSGSTW	45-53	RKGIVGDWKNHFP	257-269
rEST-6	TYPKSGSTW	45-53	RKGIVGDWRNHFP	257-269
rSMP2	TYPKSGTNW	41-49	RKGTTGDWKNHFT	244-256
mHSST1	TYPKSGTNW	41-49	RKGTIGDWKNHFT	247-259
mHSST2	TYPKSGTNW	41-49	RKGTTGDWKNHFT	247-259
rHSST1	AYPKSGTNW	40-48	RKGTVGDWKNHFT	246-258
rHSST2	TYPKSGTNW	40-48	RNGTTGDWKNHFT	246-258
rHSST3	TYPKSGTNW	41-49	RKGTTGDWKNHFT	247-259
mfHSST	TYPKSGTNW	41-49	RKGISGDWKNHLT	247-259
hDHEAST	TYPKSGTNW	41-49	RKGVSGDWKNHFT	247-259
gpHSST1	TYPKSGTNW	41-49	RKGTVGDWKNHFT	247-259
gpHSST2	TYPKSGTNW	41-49	RKGVVGDWKHHFT	247-259
fcPST3	SYPKSGTTW	56-64	RKGKDGDWKNYFT	276-288
fbPST3	SYPKSGTTW	56-64	RKGKDGDWKNYFT	277-289
fcPST4'	SYPKSGTTW	66-74	RKAKDGDWKNYFT	285-297
fbPST4	SAPKTGTTW	56-64	RKGQDGDWKNYFT	274-286
atST	TNPKSGTTW	72-80	RKGEIGGWRDSFE	290-302
CONSENSUS	TYPKSGTxW		RKGxxGDWKxxFT	

[a] The highly conserved region I and region IV amino acid sequences found in cytosolic ST enzymes are shown. Each protein is listed in the same order as depicted in the dendrogram in Fig. 1. Abbreviations are defined in the legend to Fig. 1. "Position" refers to amino acid number within the protein. Consensus sequences for regions I and IV are listed at the bottom. "x" indicates any amino acid. The black columns represent amino acids with greater than 93% identity of the amino acid at that position while the black columns with an asterisk have greater than 93% identity within groups of amino acids having similar properties as defined by Dayhoff et al. (99). Nonidentical residues in any column are "boxed" and shown black on white. Arrows indicate invariant amino acids.

tripled. It has been pointed out that the region IV sequence is similar to a motif termed the glycine-rich phosphate binding loop, the so-called "P-loop," present in some ATP and GTP binding proteins (33). That portion of region IV that includes the sequence GxxGxxK might be a homologue for the glycine-rich region, followed by a conserved lysine, found in some P-loop motifs (33). Site-directed mutagenesis experiments performed with guinea pig EST showed that substitution of alanines for G and K in this region resulted in loss of enzyme activity and loss of the ability of ^{35}S-PAPS to function as a photoaffinity ligand for the enzyme (33). Site-directed mutagenesis studies performed with the protein encoded by the plant fcFST3 cDNA (Fig. 1), suggested that the invariant lysine 59 present within region I (Table 2) might participate in the stabilization of an intermediate formed during the sulfonation reaction (34). Substitution for the invariant initial arginine present in region IV (Table 2) of fcFST3 gave results suggesting that this residue might be involved in PAPS binding (34). Both sets of site-directed mutagenesis data focused attention on the possible involvement of region IV in PAPS binding.

In summary, ST cDNA cloning has made it possible to determine the primary amino acid sequences of a large number of these enzymes; to express them in a variety of expression systems and, thus, to characterize the biochemical properties of the encoded proteins; to begin to develop a rational classification for cytosolic ST enzymes; to understand potential evolutionary relationships among these enzymes; and to gain insight into the function of ST amino acid sequence motifs that are highly conserved throughout phylogeny. cDNA isolation also made it possible to clone and structurally characterize the genes for these enzymes.

ST GENE STRUCTURES AND CHROMOSOMAL LOCALIZATIONS

Cloning of cDNAs for cytosolic ST enzymes was an important step required to clone and characterize the genes for these enzymes. Although fewer ST genes than cDNAs have been characterized, genes for all five of the known human cytosolic ST enzymes have been cloned and structurally characterized (18–20, 35–39). The structures of those genes, as well as a rat PST gene (40), a guinea pig EST gene (41), and two partial rat HSST gene sequences, SMP2A and SMP2B (42), are depicted graphically in **Fig. 2**. The structural homology is striking. All of the PST and EST genes have seven exons, listed as exons II through VIII in **Fig. 3**, that encode cDNA open reading frame sequence; locations of intron/exon splice junctions are highly conserved, both within and among species. The reason that the initial exons for the human PST genes have been labeled IA, IB, and IC will be discussed subsequently. The only HSST gene that has been completely characterized structurally at this time, the human DHEA ST gene, *STD* (39), also shows significant structural similarity within the 3′-portion of the gene with the final four PST and EST exons. Although the rat HSST gene sequences that have been reported are incomplete, their second exons are identical in length with that of the human DHEA ST gene (42) (Fig. 3). If the genes for the PST and HSST families shown in Fig. 1 evolved from a common ancestor, fusion of the second and third exons encoding the open reading frame may have occurred in the evolutionary line leading to HSSTs or, conversely, that exon was divided in the line leading to the PST family of enzymes. The HUGO Nomenclature Committee has assigned the gene symbols *STP* and *STM* to the human TS and TL PST genes, and the symbols *STE* and *STD* to the human EST and DHEA ST genes, respectively. Analogous symbols have been applied to mouse ST genes (**Table 3**). However, as can be seen in Table 3, confusion similar to that which exists in the cDNA nomenclature has already crept into the gene nomenclature, since the *STP2* gene in humans does not encode a protein homologous to that encoded by the gene designated *Stp2* in the mouse—i.e., these are not orthologous genes.

The chromosomal localizations of ST genes in humans and mice are also listed in Table 3. The human EST gene, *STE*, has been localized to the long arm of human chromosome 4 (38) in an area syntenic with the location of an apparently orthologous gene in the mouse (43). *STD*, the gene for the only known human HSST, DHEA ST, has been localized to the long arm of human chromosome 19 (44) in an area syntenic with the location of an HSST gene found on mouse chromosome 7 (43). Finally, a mouse PST gene, *Stp*, is located on mouse chromosome 7 in an area syntenic with the short arm of human chromosome 16 (45, 46), the location of a gene complex that includes all three known human PST genes, *STP1*, *STP2*, and *STM* (18, 47, 48). *STP1* in humans is located approximately 45 kb 5′ to the location of *STP2*, and the two genes are aligned "head to tail" (20). Those two genes, in turn, are located approximately 100 kb telomeric to the gene for the TL or M form of PST, *STM* (49). It has been speculated that these three PST genes on the short arm of human chromosome 16 originated as a result of gene duplication events or gene duplication plus recombination (20).

In summary, at this relatively early stage in the study of ST genes, the picture that emerges is one of a high degree of structural homology, both within and across species boundaries. However, many questions remain to be answered. For example, until structural characterization of all ST genes within a species has been completed, it will remain unclear whether cDNAs with very similar sequences represent the products of different genes or of different alleles at a single locus. The question of the existence of alleles also raises the issue of molecular genetic mechanisms involved in functionally significant regulation of the expression or properties of ST enzymes.

ST PHARMACOGENETICS

Knowledge of the molecular biology of ST enzymes has begun to make it possible to study mechanisms involved

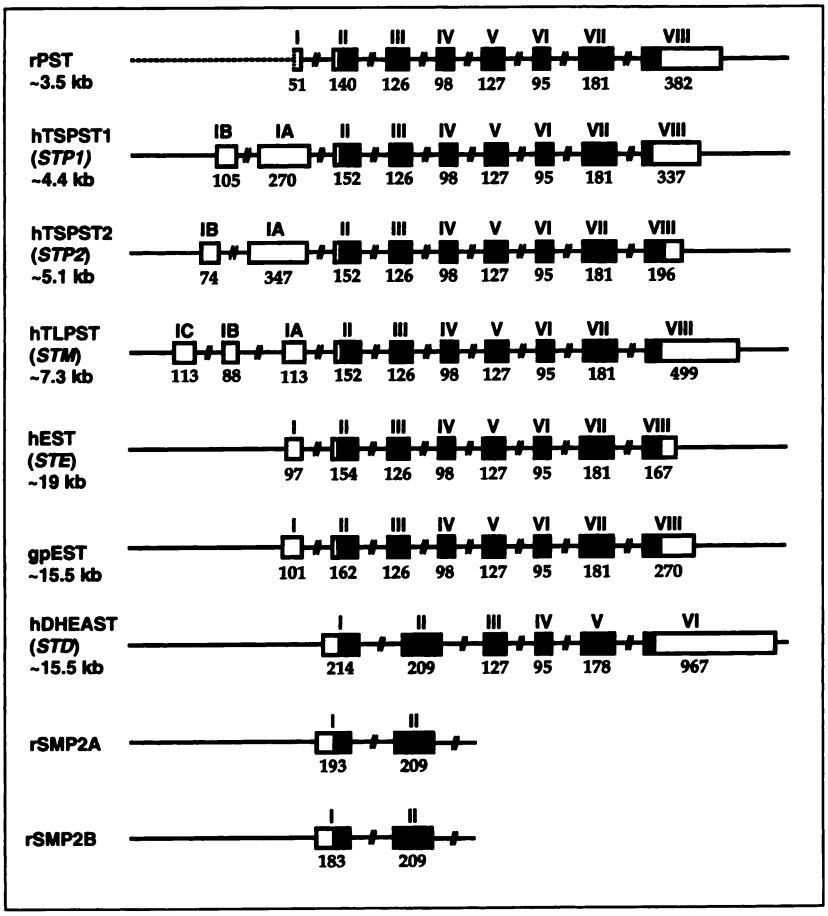

Figure 2. ST gene structures. The structures of genes for cytosolic ST enzymes that have been described are shown in the same order that the enzymes appear in the dendrogram in Fig. 1. Abbreviations for enzyme names are defined in the legend for Fig. 1. The partially characterized rSMP2A and rSMP2B genes are listed at the bottom of the figure, since complete structures for those genes have not been reported (42). Black rectangles represent coding regions, and open rectangles represent noncoding portions of exons. Numbers beneath exons represent their lengths in bp; numbers beneath gene symbols represent approximate gene lengths, in kb, from the initial through the final exon.

in the molecular genetic regulation of these enzymes. Many drug and xenobiotic metabolizing enzymes show large inherited variations in their regulation and/or properties. "Pharmacogenetics" has taught us to expect that genetic polymorphisms and other genetic mechanisms may play important roles in controlling variation among individuals in the expression or properties of enzymes involved in the biotransformation of drugs or other chemicals and, thus, in individual differences in the toxicity or therapeutic efficacy of these agents (50, 51). Biochemical studies have already demonstrated the existence of genetic polymorphisms that regulate PST activities in human tissues (52–54). Those experiments, performed before the cloning of any ST cDNA or gene, were made possible by the observation that both TS PST and TL PST (the enzymes encoded by the genes *STP1* and *STM*) were expressed in an easily accessible human tissue: the blood platelet (55), thus making population and family studies possible. Genetic polymorphisms were shown to be involved in the regulation of TS PST levels of activity and thermal stability (53) not only in the platelet, but also in the human brain, small intestine, and liver (54). A genetic polymorphism regulating platelet TL PST activity was also described on the basis of the results of family studies (52). The presence of large differences among individuals in the expression of DHEA ST activity in human liver and small intestine (56, 57), and of immunoreactive EST in the human small intestine (57), raise the possibility that genetic polymorphisms might also exist for those enzymes. Several variant nucleotide sequences that alter encoded amino acids have already been described for TS PST1 cDNAs (20), as have restriction fragment length polymorphisms (58). It should now be possible to test the hypothesis that these DNA nucleotide sequence variants for *STP1* might cosegregate with functional differences in levels of TS PST1 enzyme activity or thermal stability.

ST MOLECULAR REGULATION

Included among other molecular genetic mechanisms that might be involved in the regulation of ST enzyme expression are alternative splicing of RNA transcripts or the existence of alternative sites of transcription initiation. Data in support of such mechanisms have already been obtained for the human PST genes. It was noted previously that the

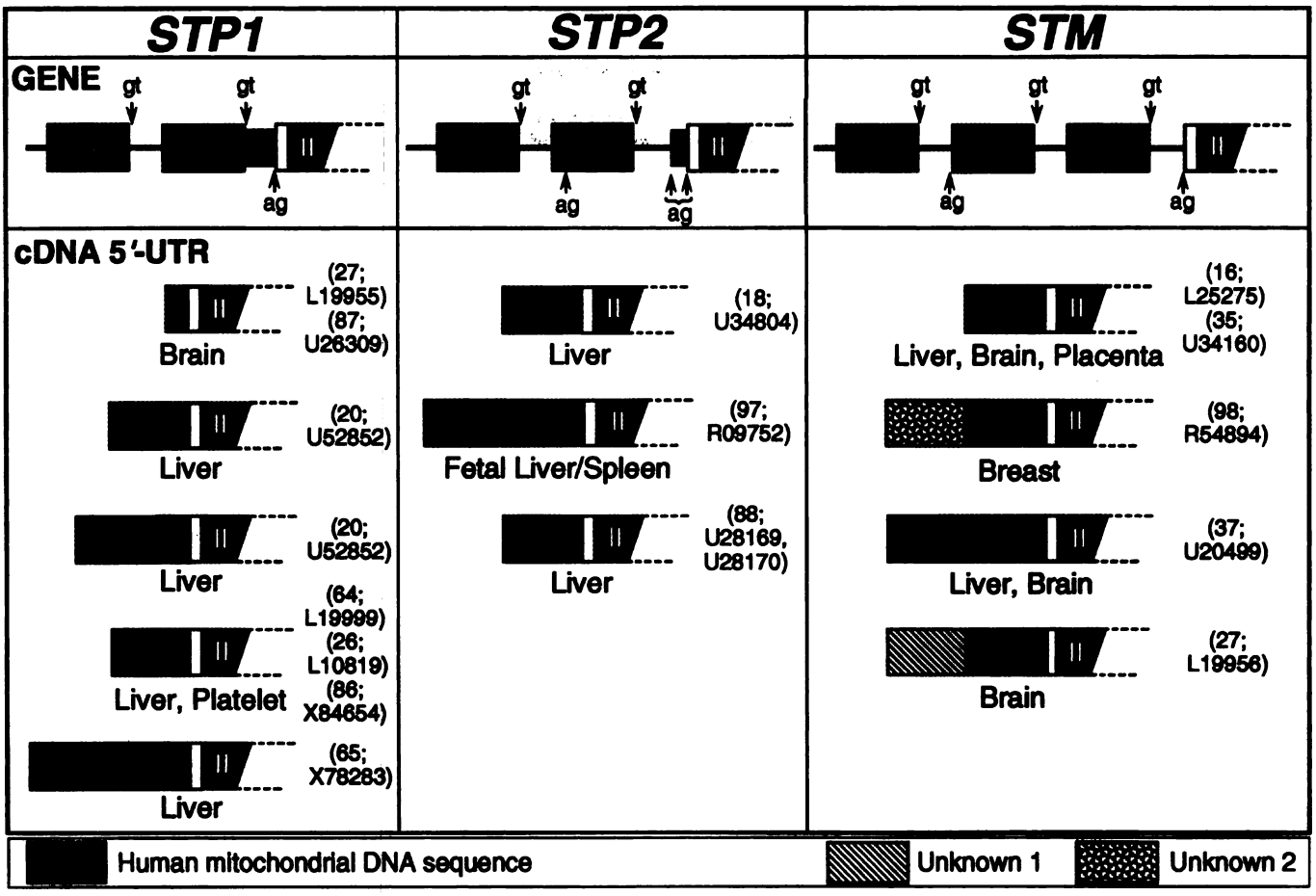

Figure 3. Human PST cDNA 5′-UTR structures. At the top, the figure depicts schematically 5′-exons of the three known cytosolic human PST genes, *STP1*, *STP2*, and *STM*. Exons are numbered as shown in Fig. 2. 5′-UTR sequences that have been reported for cDNAs encoded by each gene are shown below the appropriate gene structure to demonstrate variable 5′-exon usage as a result of alternative transcription initiation and/or alternative splicing. The figure includes references as well as GenBank accession numbers for each cDNA 5′-UTR structure depicted.

noncoding 5′-exons for these genes were labeled IA, IB, and IC in Fig. 2. In the course of cloning human PST cDNAs, application of 5′-rapid amplification of cDNA ends (5′-RACE) demonstrated that several different 5′-untranslated region (UTR) sequences could be present with any given cDNA open reading frame. Furthermore, when the genes had been cloned, these different 5′-UTR sequences were represented within the 5′-flanking regions of the genes at the locations shown schematically in Fig. 3. Figure 3 illustrates the fact that, for each of the three human PST genes, *STP1*, *STP2*, and *STM*, a variety of 5′-UTR sequences can be generated by alternative transcription initiation, alternative splicing, or both. The possible functional implications for tissue-specific expression or regulation during growth and development of this repertoire of 5′-UTR sequences represents one area for future investigation. Finally, there is already evidence that the expression of ST mRNA can be regulated by hormones or during the course of growth and development (59, 60). Obviously, studies of molecular mechanisms involved in the regulation of ST enzyme expression are still in their infancy, but there is no doubt that those data will increase rapidly in the future.

CONCLUSIONS

Knowledge of ST molecular biology has increased in an exponential fashion during the past decade. Information obtained by cloning and characterizing cDNAs and genes for these cytosolic enzymes has begun to make it possible to classify them, to express their cDNAs, to study the biochemical properties of the expressed proteins, and to begin to study the molecular genetic regulation of these enzymes. All of this information has been developed within a context provided by decades of convergent research on ST enzymes and sulfonation from many different fields. Sulfate conjugation has been studied by pharmacologists because of an interest in the biotransformation of drugs and xenobiotic compounds; by neuroscientists because of the role of sulfate conjugation in the metabolism of neurotransmitters such as the catecholamines; by endocrinologists because of the role of these enzymes in the biotransformation of hormones, particularly steroid hormones; and by gastroenterologists because of the importance of sulfonation in the metabolism of bile acids (1–4). Each of these lines of investigation has contributed significantly to our understanding of this family of enzymes. One of the many ad-

TABLE 3. *Cytosolic ST gene chromosomal localizations*[a]

ST GENE CHROMOSOMAL LOCALIZATIONS						
Human Gene	Human cDNA	Human Gene Chromosomal Localization	Mouse Gene	Mouse cDNA	Mouse Gene Chromosomal Localization	Syntenic Human Chromosomal Region
STP1	TSPST1	16p12.1-p11.2	*Stp*	PST	7 (distal)	16p12.1-p11.2
STP2	TSPST2	16p12.1-p11.2	—	—	—	—
STM	TLPST	16p11.2	—	—	—	—
—	—	—	*Stp2*	AAFST	17	19p13 or 6p
STE	EST	4q13.1	*Ste*	EST	5 (medial)	4q
STD	DHEAST	19q13.3	*Std*	HSST	7 (proximal)	19q13

[a] The table lists each gene for which a chromosomal localization has been reported, as well as the HUGO Nomenclature Committee approved symbol for the human gene. Human cDNA names are defined in the legend to Fig. 1. Comparable genes in the mouse with their chromosomal localizations, as well as areas of human chromosomes syntenic with the mouse chromosomal localizations, are also listed. The mouse "Stp2" gene was localized with the rat AAFST cDNA (22, 46) and is not orthologous to any presently described human ST cDNA or gene.

vantages of our rapidly expanding understanding of the molecular biology of ST enzymes is that the molecular data are helping to integrate work on this important group of enzymes across scientific disciplines. Ultimately, molecular knowledge will return the focus to biological function and will serve to enhance understanding of the role of sulfonation in the therapeutic efficacy and toxicity of drugs, in individual differences in the biotransformation of hormones and neurotransmitters, and in disease processes such as carcinogenesis.

We thank Luanne Wussow for her assistance with the preparation of this manuscript and its graphical material. Supported in part by NIH grants RO1 GM 28157 and RO1 GM 35720 and by a supplement to RO1 GM 35720 supported by the Office of Research on Women's Health.

REFERENCES

1. Weinshilboum, R. M. (1986) Phenol sulfotransferase in humans: Properties, regulation, and function. *Federation Proc.* **45**, 2223–2228
2. Mulder, G. J., and Jakoby, W. B. (1990) Sulfation. In *Conjugation Reactions in Drug Metabolism* (Mulder, G. J., ed) pp. 107–161, Taylor & Francis Ltd., New York
3. Falany, C. N. (1991) Molecular enzymology of human liver cytosolic sulfotransferases. *Trends Pharmacol. Sci.* **12**, 255–259
4. Weinshilboum, R. M., and Otterness, D. M. (1994) Sulfotransferase enzymes. In *Conjugation-Deconjugation Reactions in Drug Metabolism and Toxicity* (Kauffman, F. C., ed) Chapt. 22, "Handbook of Experimental Pharmacology" series, Vol. 112, pp. 45–78, Springer-Verlag, Berlin Heidelberg
5. Varin, L., and Ibrahim, R. K. (1989) Partial purification and characterization of three flavonol-specific sulfotransferases from *Flaveria chloraefolia*. *Plant Physiol.* **90**, 977–981
6. Baumann, E. (1876) Ueber Sulfosaüren im Harn. *Berlin Dtsch. Chem. Ges.* 54–58
7. Huttner, W. B. (1982) Sulphation of tyrosine residues—a widespread modification of proteins. *Nature (London)* **299**, 273–276
8. Hashimoto, Y., Orellana, A., Gil, G., and Hirschberg, C. B. (1992) Molecular cloning and expression of rat liver N-heparan sulfate sulfotransferase. *J. Biol. Chem.* **267**, 15744–15750
9. Kiehlbauch, C. C., Lam, Y. F., and Ringer, D. P. (1995) Homodimeric and heterodimeric aryl sulfotransferases catalyze the sulfuric acid esterification of N-hydroxyl-2-acetylaminofluorene. *J. Biol. Chem.* **270**, 18941–18947
10. Takikawa, H., Stolz, A., and Kaplowitz, N. (1986) Purification of a 32.5 kDa monomeric sulfotransferase from rat liver with activity for bile acids and phenolic steroids. *FEBS Lett.* **207**, 193–197
11. Farooqui, A. A. (1980) 3′-Phosphoadenosine-5′-phosphosulfate metabolism in mammalian tissues. *Int. J. Biochem.* **12**, 529–536
12. Nomenclature Committee of the International Union of Biochemistry on the Nomenclature and Classification of Enzyme-Catalyzed Reactions. (1992) *Enzyme Nomenclature 1992*, pp. 299–303, Academic Press, New York
13. Falany, C. N., Vazquez, M. E., and Kalb, J. M. (1989) Purification and characterization of human liver dehydroepiandrosterone sulphotransferase. *Biochem. J.* **260**, 641–646
14. Aksoy, I. A., Otterness, D. M., and Weinshilboum, R. M. (1993) Cholesterol sulfation in human liver: catalysis by dehydroepiandrosterone sulfotransferase. *Drug Metab. Dispos.* **21**, 268–276
15. Hernández, J. S., Watson, R. W. G., Wood, T. C., and Weinshilboum, R. M. (1992) Sulfation of estrone and 17β-estradiol in human liver: catalysis by thermostable phenol sulfotransferase and by dehydroepiandrosterone sulfotransferase. *Drug Metab. Dispos.* **20**, 413–422
16. Bernier, F., Lopez-Solache, I., Labrie, F., and Luu-The, V. (1994) Cloning and expression of cDNA encoding human placental estrogen sulfotransferase. *Mol. Cell. Endocrinol.* **99**, R11–R15
17. Aksoy, I. A., Wood, T. C., and Weinshilboum, R. M. (1994) Human liver estrogen sulfotransferase: cDNA cloning, expression and biochemical characterization. *Biochem. Biophys. Res. Commun.* **200**, 1621–1629
18. Her, C., Raftogianis, R., and Weinshilboum, R. M. (1996) Human phenol sulfotransferase *STP2* gene: molecular cloning, structural characterization and chromosomal localization. *Genomics* **33**, 409–420
19. Bernier, F., Soucy, P., and Luu-The, V. (1996) Human phenol sulfotransferase gene contains two alternative promoters: structure and expression of the gene. *DNA Cell Biol.* **15**, 367–375
20. Raftogianis, R., Her, C., and Weinshilboum, R. M. (1996) Human phenol sulfotransferase pharmacogenetics: *STP1* gene cloning and structural characterization. *Pharmacogenetics* **6**, 473–487
21. Nebert, D. W., Nelson, D. R., Coon, M. J., Estabrook, R. W., Feyereisen, R., Fujii-Kuriyama, Y., Gonzalez, F. J., Guengerich, F. P., Gunsalus, I. C., Johnson, E. F., Loper, J. C., Sato, R., Waterman, M. R., and Waxman, D. J. (1991) The P450 superfamily: update on new sequences, gene mapping and recommended nomenclature. *DNA Cell Biol.* **10**, 1–14

22. Nagata, K., Ozawa, S., Miyata, M., Shimada, M., Gong, D.-W., Yamazoe, Y., and Kato, R. (1993) Isolation and expression of a cDNA encoding a male-specific rat sulfotransferase that catalyzes activation of *N*-hydroxy-2-acetylaminofluorene. *J. Biol. Chem.* **268**, 24720–24725
23. Sakakibara, Y., Takami, Y., Zwieb, C., Nakayama, T., Suiko, M., Nakajima, H., and Liu, M.-C. (1995) Purification, characterization, and molecular cloning of a novel rat liver dopa/tyrosine sulfotransferase. *J. Biol. Chem.* **270**, 30470–30478
24. Chatterjee, B., Majumdar, D., Ozbilen, O., Murty, C. V. R., and Roy, A. K. (1987) Molecular cloning and characterization of cDNA for androgen-repressible rat liver protein, SMP-2. *J. Biol. Chem.* **262**, 822–825
25. Lacomme, C., and Roby, D. (1996) Molecular cloning of a sulfotransferase in *Arabidopsis thaliana* and regulation during development and in response to infection with pathogenic bacteria. *Plant Mol. Biol.* **30**, 995–1008
26. Zhu, X., Veronese, M. E., Sansom, L. N., and McManus, M. E. (1993) Molecular characterisation of a human aryl sulfotransferase cDNA. *Biochem. Biophys. Res. Commun.* **192**, 671–676
27. Zhu, X., Veronese, M. E., Bernard, C. C. A., Sansom, L. N., and McManus, M. E. (1993) Identification of two human brain aryl sulfotransferase cDNAs. *Biochem. Biophys. Res. Commun.* **195**, 120–127
28. Wood, T. C., Aksoy, I. A., Aksoy, S., and Weinshilboum, R. M. (1994) Human liver thermolabile phenol sulfotransferase: cDNA cloning, expression and characterization. *Biochem. Biophys. Res. Commun.* **198**, 1119–1127
29. Nash, A. R., Glenn, W. K., Moore, S. S., Kerr, J., Thompson, A. R., and Thompson, E. O. P. (1988) Oestrogen sulfotransferase: molecular cloning and sequencing of cDNA for bovine placental enzyme. *Australian J. Biol. Sci.* **41**, 507–516
30. Otterness, D. M., Wieben, E. D., Wood, T. C., Watson, R. W. G., Madden, B. J., McCormick, D. J., and Weinshilboum, R. M. (1992) Human liver dehydroepiandrosterone sulfotransferase: molecular cloning and expression of cDNA. *Mol. Pharmacol.* **41**, 865–872
31. Lennon, G., Auffray, C., Polymeropoulos, M., and Soares, M. B. (1996) The I.M.A.G.E. Consortium: an integrated molecular analysis of genomes and their expression. *Genomics* **33**, 151–152
32. Varin, L., DeLuca, V., Ibrahim, R. K., and Brisson, N. (1992) Molecular characterization of two plant flavonol sulfotransferases. *Proc. Natl. Acad. Sci. USA* **89**, 1286–1290
33. Komatsu, K., Driscoll, W. J., Koh, Y., and Strott, C. A. (1994) A P-loop related motif (GxxGxxK) highly conserved in sulfotransferases is required for binding the activated sulfate donor. *Biochem. Biophys. Res. Commun.* **204**, 1178–1185
34. Marsolais, F., and Varin, L. (1995) Identification of amino acid residues critical to catalysis and cosubstrate binding in the flavonol 3-sulfotransferase. *J. Biol. Chem.* **270**, 30458–30463
35. Bernier, F., Leblanc, G., Labrie, F., and Luu-The, V. (1994) Structure of human estrogen and aryl sulfotransferase gene: two mRNA species issued from a single gene. *J. Biol. Chem.* **269**, 28200–28205
36. Dooley, T. P., Probst, P., Monroe, P. B., Mole, S. E., Liu, Z., and Doggett, N. A. (1994) Genetic organization and DNA sequence of the human catecholamine-sulfating phenol sulfotransferase gene (STM). *Biochem. Biophys. Res. Commun.* **205**, 1325–1332
37. Aksoy, I. A., and Weinshilboum, R. M. (1995) Human thermolabile phenol sulfotransferase gene (STM): molecular cloning and structural characterization. *Biochem. Biophys. Res. Commun.* **208**, 786–795
38. Her, C., Aksoy, I. A., Kimura, S., Brandriff, B. F., Wasmuth, J. J., and Weinshilboum, R. M. (1995) Human estrogen sulfotransferase gene (STE): cloning, structure and chromosomal localization. *Genomics* **29**, 16–23
39. Otterness, D. M., Her, C., Aksoy, S., Kimura, S., Wieben, E. D., and Weinshilboum, R. M. (1995) Human dehydroepiandrosterone sulfotransferase gene: molecular cloning and structural characterization. *DNA Cell Biol.* **14**, 331–341
40. Khan, A. S., Taylor, B. R., Chung, K., Etheridge, J., Gonzales, R., and Ringer, D. P. (1993) Genomic structure of rat liver aryl sulfotransferase IV-encoding gene. *Gene* **137**, 321–326
41. Chiba, H., Komatsu, K., Lee, Y. C., Timizuka, T., and Strott, C. A. (1995) The 3′-terminal exon of the family of steroid and phenol sulfotransferase genes is spliced at the N-terminal glycine of the universally conserved GXXGXXK motif that forms the sulfonate donor binding site. *Proc. Natl. Acad. Sci. USA* **92**, 8176–8179
42. Song, C.-S., Kim, J. M., Roy, A. K., and Chatterjee, B. (1990) Structure and regulation of the senescence marker protein 2 gene promoter. *Biochemistry* **29**, 542–551
43. Leiter, E. H., and Chapman, H. D. (1994) Obesity-induced diabetes (Diabesity) in C57BL/KsJ mice produces aberrant *trans*-regulation of sex steroid sulfotransferase genes. *J. Clin. Invest.* **93**, 2007–2013
44. Otterness, D. M., Mohrenweiser, H. W., Brandriff, B. F., and Weinshilboum, R. M. (1995) Dehydroepiandrosterone sulfotransferase gene (STD): localization to human chromosome 19q13.3. Cytogenet. *Cell Genet.* **70**, 45–47
45. Dooley, T. P., Obermoeller, R. D., Leiter, E. H., Chapman, H. D., Falany, C. N., Deng, Z., and Siciliano, M. J. (1993) Mapping of the phenol sulfotransferase gene (STP) to human chromosome 16p12.1-p11.2 and to mouse chromosome 7. *Genomics* **18**, 440–443
46. Khan, A. S., Taylor, B. R., Filie, J. D., Ringer, D. P., and Kozak, C. A. (1995) Rat phenol-preferring sulfotransferase genes (*Stp* and *Stp2*): localization to mouse chromosomes 7 and 17. *Genomics* **26**, 417–419
47. Aksoy, I. A., Callen, D. F., Apostolou, S., Her, C., and Weinshilboum, R. M. (1994) Thermolabile phenol sulfotransferase (STM): localization to human chromosome 16p11.2. *Genomics* **23**, 275–277
48. Dooley, T. P., Mitchison, H. M., Munroe, P. B., Probst, P., Neal, M., Siciliano, M. J., Deng, Z., Doggett, N. A., Callen, D. F., Gardiner, R. M., and Mole, S. E. (1994) Mapping of two phenol sulfotransferase genes, *STP* and *STM*, to 16p: candidate genes for Batten disease. *Biochem. Biophys. Res. Commun.* **205**, 482–489
49. Järvelä, I. E., Mitchison, H. M., O'Rawe, A. M., Munroe, P. B., Taschner, P. E. M., de Vos, N., Lerner, T. J., D'Arigo, K. L., Callen, D. F., Thompson, A. D., Knight, M., Marrone, B. L., Mundt, M. O., Meincke, L., Breuning, M. H., Gardiner, R. M., Doggett, N. A., and Mole, S. E. (1995) YAC and cosmid contigs spanning the Batten disease (*CLN3*) region at 16p12.1-p11.2. *Genomics* **29**, 478–489
50. Kalow, W. (1992) *Pharmacogenetics of Drug Metabolism*. Pergamon Press, New York
51. Price Evans, D. A. (1993) *Genetic Factors in Drug Therapy: Clinical and Molecular Pharmacogenetics*, Cambridge University Press
52. Price, R. A., Cox, N. J., Spielman, R. S., Van Loon, J., Maidak, B. L., and Weinshilboum, R. M. (1988) Inheritance of human platelet thermolabile phenol sulfotransferase (TL PST) activity. *Genet. Epidemiol.* **5**, 1–15
53. Price, R. A., Spielman, R. S., Lucena, A. L., Van Loon, J. A., Maidak, B. L., and Weinshilboum, R. M. (1989) Genetic polymorphism for human platelet thermostable phenol sulfotransferase (TS PST) activity. *Genetics* **122**, 905–914
54. Weinshilboum, R. (1990) Sulfotransferase pharmacogenetics. *Pharmacol. Ther.* **45**, 93–107
55. Hart, R. F., Renskers, K. J., Nelson, E. B., and Roth, J. A. (1979) Localization and characterization of phenol sulfotransferase in human platelets. *Life Sci.* **24**, 125–130
56. Aksoy, I. A., Sochorová, V., and Weinshilboum, R. (1993) Human liver dehydroepiandrosterone sulfotransferase: nature and extent of individual variation. *Clin. Pharmacol. Ther.* **54**, 498–506
57. Her, C., Szumlanski, C. L., Aksoy, I. A., and Weinshilboum, R. M. (1996) Human jejunal estrogen sulfotransferase and dehydroepiandrosterine sulfotransferase: immunochemical characterization of individual variation. *Drug Metab. Dispos.* **24**, 1382–1335
58. Henkel, R. D., Galindo, L. V., and Dooley, T. P. (1994) An *Eco*RI RFLP in human phenol sulfotransferase genes. *Human Mol. Genet.* **3**, 2264
59. Runge-Morris, M., and Wilusz, J. (1991) Age and gender-related gene expression of hydroxysteroid sulfotransferase-a in rat liver. *Biochem. Biophys. Res. Commun.* **175**, 1051–1056
60. Schauss, S. J., Henry, T., Palmatier, R., Halvorson, L., Dannenbring, R., and Beckmann, J. D. (1995) Characterization of bovine tracheobronchial phenol sulphotransferase cDNA and detection of mRNA regulation by cortisol. *Biochem. J.* **311**, 1–9
61. Feng, D.-F., and Doolittle, R. F. (1987) Progressive sequence alignment as a prerequisite to correct phylogenetic trees. *J. Mol. Evol.* **25**, 351–360
62. Ozawa, S., Nagata, K., Gong, D., Yamazoe, Y., and Kato, R. (1990) Nucleotide sequence of a full-length cDNA (PST-1) for aryl sulfotransferase from rat liver. *Nucleic Acids Res.* **18**, 4001
63. Kong, A.-N. T., Ma, M., Tao, D., and Yang, L. (1993) Molecular cloning of cDNA encoding the phenol/aryl form of sulfotransferase (mST_{p1}) from mouse liver. *Biochim. Biophys. Acta* **1171**, 315–318
64. Wilborn, T. W., Comer, K. A., Dooley, T. P., Reardon, I. M., Heinrikson, R. L., and Falany, C. N. (1993) Sequence analysis and expression of the cDNA for the phenol sulfating form of human liver phenol sulfotransferase. *Mol. Pharmacol.* **43**, 70–77
65. Ozawa, S. H., Nagata, K., Shimada, M., Ueda, M., Tsuzuki, T., Yamazoe, Y., and Kato, R. (1995) Primary structures and properties of two related forms of aryl sulfotransferase in human liver. *Pharmacogenetics* **5**, S135–S140
66. Ogura, K., Satsukawa, M., Okuda, H., and Watabe, T. (1996) cDNA cloning and bacterial expression of monkey liver phenol sulfotransferase. GenBank Accession number D85514
67. Yamazoe, Y., Nagata, K., Ozawa, S., and Kato, R. (1994) Structural similarity and diversity of sulfotransferases. *Chem. Biol. Interact.* **92**, 107–117
68. Oeda, T., Lee, Y. C., Driscoll, W. J., Chen, H.-C., and Strott, C. A. (1992) Molecular cloning and expression of a full-length complementary DNA encoding the guinea pig adrenocortical estrogen sulfotransferase. *Mol. Endocrinol.* **6**, 1216–1226
69. Song, W.-C., Moore, R., McLachlan, J. A., and 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
70. Demyan, W. F., Song, C. S., Kim, D. S., Her, S., Gallwitz, W., Rao, T. R., Siomczynska, M., Chatterjee, B., and Roy, A. K. (1992) Estrogen sulfotransferase of the rat liver: complementary DNA cloning and age- and sex-specific regulation of messenger RNA. *Mol. Endocrinol.* **6**, 589–597
71. Falany, J. L., Krasnykh, V., Mikheeva, G., and Falany, C. N. (1995) Isolation and expression of an isoform of rat estrogen sulfotransferase. *J. Steroid Biochem. Mol. Biol.* **52**, 35–44

72. Kong, A.-N. T., Tao, D., Ma, M., and Yang, L. (1993) Molecular cloning of the alcohol/hydroxysteroid form (mST_{a1}) of sulfotransferase from mouse liver. *Pharmaceut. Res.* **10**, 627–630
73. Kong, A-N. T., and Fei, P. (1994) Molecular cloning of three sulfotransferase cDNAs from mouse liver. *Chem. Biol. Interact.* **92**, 161–168
74. Ogura, K., Kajita, J., Narihata, H., Watabe, T., Ozawa, S., Nagata, K., Yamazoe, Y., and Kato, R. (1989) Cloning and sequence analysis of a rat liver cDNA encoding hydroxysteroid sulfotransferase. *Biochem. Biophys. Res. Commun.* **165**, 168–174
75. Ogura, K., Kajita, J., Narihata, H., Watabe, T., Ozawa, S., Nagata, K., Yamazoe, Y., and Kato, R. (1990) cDNA cloning of the hydroxysteroid sulfotransferase STa sharing a strong homology in amino acid sequence with the senescence marker protein SMP-2 in rat livers. *Biochem. Biophys. Res. Commun.* **166**, 1494–1500
76. Watabe, T., Ogura, K., Satsukawa, M., Okuda, H., and Hiratsuka, A. (1994) Molecular cloning and functions of rat liver hydroxysteroid sulfotransferases catalysing covalent binding of carcinogenic polycyclic arylmethanols in DNA. *Chem. Biol. Interact.* **92**, 87–105
77. Ogura, K., Satsukawa, M., Kato, K., Okuda, H., and Watabe, T. (1996) Molecular cloning of monkey liver hydroxysteroid sulfotransferase. GenBank Accession number D85521
78. Lee, Y. C., Park, C.-S., and Strott, C. A. (1994) Molecular cloning of a chiral-specific 3α-hydroxysteroid sulfotransferase. *J. Biol. Chem.* **269**, 15838–15845
79. Luu, N. X., Driscoll, W. J., Martin, B. M., and Strott, C. A. (1995) Molecular cloning and expression of a guinea pig 3-hydroxysteroid sulfotransferase distinct from chiral-specific 3α-hydroxysteroid sulfotransferase. *Biochem. Biophys. Res. Commun.* **217**, 1078–1086
80. Ananvoranich, S., Varin, L., Gulick, P., and Ibrahim, R. (1994) Cloning and regulation of flavonol 3-sulfotransferase in cell suspension cultures of *Flaveria bidentis*. *Plant Physiol.* **106**, 485–491
81. Ananvoranich, S., Gulick, P., and Ibrahim, R. K. (1995) Flavonol sulfotransferase-like cDNA clone from *Flaveria bidentis*. *Plant Physiol.* **107**, 1019–1020
82. Hirshey, S. J., Dooley, T. P., Reardon, I. M., Henrikson, R. L., and Falany, C. N. (1992) Sequence analysis, in vitro translation, and expression of the cDNA for rat liver minoxidil sulfotransferase. *Mol. Pharmacol.* **42**, 257–264
83. Yerokun, T., Etheredge, J. L., Norton, T. R., Carter, H. A., Chung, K. H., Birckbichler, P. J., and Ringer, D. P. (1992) Characterization of a complementary DNA for rat liver aryl sulfotransferase IV and use in evaluating the hepatic gene transcript levels of rat at various stages of 2-acetylaminofluorene-induced hepatocarcinogenesis. *Cancer Res.* **52**, 4779–4786
84. Chen, X., Yang, Y.-S., Zheng, Y., Martin, B. M., Duffel, M. W., and Jakoby, W. B. (1992) Tyrosine-ester sulfotransferase from rat liver: bacterial expression and identification. *Protein Expression Purif.* **3**, 421–426
85. Cruickshank, D., Sansom, L. N., Veronese, M. E., Mojaarabi, B., McManus, M. E., and Zhu, X. (1993) cDNA expression studies of rat liver aryl sulphotransferase. *Biochem. Biophys. Res. Commun.* **191**, 295–301
86. Jones, A. L., Hagen, M., Coughtrie, M. W. H., Roberts, R. C., and Glatt, H. (1995) Human platelet phenolsulfotransferase: cDNA cloning, stable expression in V79 cells and identification of a novel allelic variant of the phenol-sulfating form. *Biochem. Biophys. Res. Commun.* **208**, 855–862
87. Hwang, S. R., Kohn, A. B., and Hook, V. Y. H. (1995) Molecular cloning of an isoform of phenol sulfotransferase from human brain hippocampus. *Biochem. Biophys. Res. Commun.* **207**, 701–707
88. Zhu, X., Veronese, M. E., Iocco, P., and McManus, M. E. (1996) cDNA cloning and expression of a new form of human aryl sulfotransferase. *Int. J. Biochem. Cell. Biol.* **28**, 565–571
89. Ganguly, T. C., Krasnykh, V., and Falany, C. N. (1995) Bacterial expression and kinetic characterization of the human monoamine-sulfating form of phenol sulfotransferase. *Drug Metab. Dispos.* **23**, 945–950
90. Falany, C. N., Krasnykh, V., and Falany, J. L. (1995) Bacterial expression and characterization of a cDNA for human liver estrogen sulfotransferase. *J. Steroid Biochem. Mol. Biol.* **52**, 529–539
91. Rikke, B. A., and Roy, A. K. (1996) Structural relationships among members of the mammalian sulfotransferase gene family. *Biochim. Biophys. Acta* **1307**, 331–338
92. Ogura, K., Satsukawa, M., Okuda, H., Hiratsuka, A., and Watabe, T. (1994) Major hydroxysteroid sulfotransferase STe in rat liver cytosol may consist of two microheterogeneous subunits. *Chem. Biol. Interact.* **92**, 129–144
93. Kong, A.-N. T., Yang, L., Ma, M., Tao, D., and Bjornsson, T. D. (1992) Molecular cloning of the alcohol/hydroxysteroid form (hST_a) of sulfotransferase from human liver. *Biochem. Biophys. Res. Commun.* **187**, 448–454
94. Comer, K. A., Falany, J. L., and Falany, C. N. (1993) Cloning and expression of human liver dehydroepiandrosterone sulfotransferase. *Biochem. J.* **289**, 233–240
95. Forbes, K. J., Hagen, M., Glatt, H., Hume, R., and Coughtrie, M. W. (1995) Human fetal adrenal hydroxysteroid sulphotransferase: cDNA cloning, stable expression in V79 cells and functional characterisation of the expressed enzyme. *Mol. Cell Endocrinol.* **112**, 53–60
96. Dufort, I., Tremblay, Y., Bélanger, A., Labrie, F., and Luu-The, V. (1996) Isolation and characterization of a stereospecific 3 -hydroxysteroid sulfotransferase (pregnenolone sulfotransferase) cDNA. *DNA and Cell Biol.* **15**, 481–487
97. I.M.A.G.E. Consortium CloneID 128106. Lennon, G., Auffray, C., Polymeropoulos, M., and Soares, M. B. (1996) The I.M.A.G.E. Consortium: an integrated molecular analysis of genomes and their expression. *Genomics* **33**, 151–152. GenBank Accession number R09752
98. I.M.A.G.E. Consortium CloneID 154523. Lennon, G., Auffray, C., Polymeropoulos, M., and Soares, M. B. (1996) The I.M.A.G.E. Consortium: an integrated molecular analysis of genomes and their expression. *Genomics* **33**, 151–152. GenBank Accession number R54894
99. Dayhoff, M. O., Barker, W. C., and Hunt, L. T. (1983) Establishing homologies in protein sequences. *Methods Enzymol.* **19**, 524–539