Isolation and Characterization of the UGT2B28 cDNA Encoding a Novel Human Steroid Conjugating UDP-Glucuronosyltransferase†

Éric Lévesque,‡ David Turgeon,‡ Jean-Sébastien Carrier,‡ Valérie Montminy, Martin Beaulieu, and Alain Bélanger*

Oncology and Molecular Endocrinology Research Center, CHUL Research Center, Laval University, Quebec City, Quebec, Canada G1V 4G2

Received November 13, 2000; Revised Manuscript Received January 16, 2001

ABSTRACT: UDP-glucuronosyltransferase (UGT) enzymes belonging to the UGT2B subfamily catalyze the transfer of glucuronic acid to a large number of endogenous compounds, particularly steroids, to facilitate their elimination from target cells. A novel human UGT2B cDNA of 1666 bp was isolated and encodes a 529-amino acid protein named UGT2B28 type I. Glucuronidation assays demonstrated that UGT2B28 type I catalyzes the conjugation of endogenous and exogenous compounds. The tissue distribution of UGT2B28 revealed the expression of the type I transcript in the liver, breast, and LNCaP cells. Two other UGT2B cDNAs were isolated, and sequence analysis led to the identification of two truncated UGT2B28 species. UGT2B28 type II differs from type I by a deletion of 308 bp in the cofactor binding domain, whereas UGT2B28 type III lacks 351 bp in the putative substrate binding domain. All UGT2B28 isoforms are encoded by a single *UGT2B28* gene which has a genomic organization similar to that of the other *UGT2B* genes characterized thus far. Although no substrates could be identified for the shorter isoforms, the three subtypes were shown to be located in the endoplasmic reticulum and the perinuclear membrane, demonstrating that the missing domains are not required for the subcellular localization of these UGT2B proteins. However, all the domains remain necessary for observing glucuronidation activity. The expression of UGT2B28 type I in the breast and liver suggests a role of this enzyme in the androgen and estrogen metabolism in these tissues.

The UDP-glucuronosyltransferase (UGT)[1] enzymes are endoplasmic reticulum (ER) membrane-bound proteins which catalyze the biotransformation of hydrophobic molecules, including steroid hormones, into glucuronide derivatives (*1*). The addition of glucuronic acid from UDP-glucuronic acid (UDPGA) has been proposed to cause a steric hindrance of the molecule, abolishing the affinity for its specific receptor and rendering the steroid molecule more polar, inactive, and easier to eliminate into the circulation (*2*). This concept is supported by high circulating levels of androsterone-glucuronide (ADT-G) and 3α-diol-G, which were postulated to reflect androgen metabolism in human extrahepatic tissues (*3*−*7*). It has been thought that the androgen metabolites were devoid of biological functions. However, recent demonstrations suggest that 3α-diol has an important physiological role in parturition in mice, and also modulates the activity of an orphan receptor in the mouse (*8*). Thus, 3α-hydroxysteroid dehydrogenase (3α-HSD), which was previously considered to contribute to androgen catabolism, is also involved in the activation of steroids. Glucuronidation, which was shown to occur in extrahepatic tissues, thus takes on major importance in converting androgen metabolites (*7*, *9*, *10*).

The enzymes responsible for the glucuronidation reaction have been divided into two families, UGT1 and UGT2, on the basis of amino acid sequence homology (*11*). The UGT1 gene locus is localized on chromosome 2q37 and is composed of at least 16 exons. Four exons in the 3′ region of the locus encode the identical carboxy-terminal domain of each UGT1 protein (*12*). These four common exons are coupled with one of the 12 different versions of exon 1 by exon sharing, leading to the formation of 12 different UGT1 transcripts (*12*). Promoter elements are found upstream of each exon 1, providing several modes of regulation for the expression of nine proteins and three pseudogenes (*12*). UGT1 enzymes demonstrated conjugation activity principally on bilirubin and exogenous compounds, but recently, substrates such as steroid hormones have been identified (*13*−*18*).

The enzymes of the UGT2 subfamily are further regrouped into two subfamilies (*11*). The UGT2A family members were isolated from rat (*19*), bovine (*20*), and human (*21*) olfactive epithelium, where they may act as terminators of odorant signals (*19*). In humans, six cDNA clones encoding UGT2B4 (*22*, *23*), UGT2B7 (*24*, *25*), UGT2B10 (*26*), UGT2B11 (*27*), UGT2B15 (*28*, *29*), and UGT2B17 (*30*) have been characterized. These widely expressed UGT2B enzymes conjugate

† This work was supported by the Medical Research Council (MRC) of Canada, the Fonds de la Recherche en Santé du Québec (FRSQ), and Endorecherche. É.L. is a holder of a scholarship from the Medical Research Council (MRC) of Canada. D.T. is a holder of a scholarship from the Fonds de la Recherche en Santé du Québec (FRSQ).

* To whom the correspondence should be addressed. Telephone: (418)654-2296. Fax: (418)654-2761. E-mail: alain.belanger@crchul.ulaval.ca.

‡ These authors contributed equally to this work.

[1] Abbreviations: UGT, UDP-glucuronosyltransferase; DHT, dihydrotestosterone; 3α-diol, androstane-3α,17β-diol; ADT, androsterone; E_3, estriol; E_2, estradiol; E_1, estrone; G, glucuronide; LCA, litocholic acid; HDCA, hyodeoxycholic acid; HSD, hydroxysteroid dehydrogenase; UDPGA, UDP-glucuronic acid; DAPI, diamidino-2-phenylindole; PDI, protein disulfide isomerase.

10.1021/bi002607y CCC: $20.00
Published on Web 03/09/2001

several classes of steroid molecules showing overlapping but distinct substrate specificities (*14*). UGT2B4 demonstrates glucuronidation activity toward 3α-diol, estriol (E_3), and 4-hydroxyestrone (*22*, *31*). UGT2B7, UGT2B15, and UGT2B17 catalyze the conjugation at the 17*β*-hydroxy position of DHT, testosterone (Testo), and 3α-diol, whereas UGT2B7 and UGT2B17 also glucuronidate steroids at position 3α-hydroxy as shown by conjugation of ADT (*28*−*31*). To date, the stable expression in HK293 cells of the two human isoforms, UGT2B10 and UGT2B11, revealed no transferase activity toward several endogenous and exogenous compounds, and therefore, they are considered orphan UGT enzymes (*26*, *27*). The characterization of PAC clones containing more than 750 kb of genomic DNA demonstrated the presence of several *UGT2B* genes and pseudogenes within a cluster on chromosome 4q13-q21.1 (*32*). Characterization of the genes encoding human UGT2B4 (*33*), UGT2B7 (*34*, *35*), UGT2B15 (*32*), and UGT2B17 (*36*), in addition to rat UGT2B1 (*37*) and UGT2B2 (*38*), demonstrated a genomic structure composed of six exons, which are, in opposition of the *UGT1* genes, all exclusive to a specific *UGT2B* gene.

Due to the high degree of homology of UGT2B isoforms and the variable abundance of each transcript, it is difficult by conventional screening techniques to isolate a specific human UGT2B cDNA. However, on the basis of our results, it was obvious that novel steroid UGT2B enzymes remain unidentified since steroids such as cortisol, aldosterone, and 5α-reduced progesterone metabolites are conjugated in human tissues. Furthermore, the presence of a 3α-specific UGT transferase in both rat (UGT2B2) (*39*) and monkey (UGT2B18) (*40*), combined with a high human plasma level of ADT-G that cannot be fully explained by UGT2B7 (*41*) and UGT2B17 (*30*), suggests that other human UGT enzymes remained uncharacterized.

Here, we report the characterization of a novel cDNA and its corresponding gene, which encodes the UGT2B28 type I protein. UGT2B28 type I conjugates 5*β*-androstane 3α,17*β*-diol, estradiol, ADT, Testo, 4-methylumbelliferone, and bile acids. UGT2B28 type I is expressed in the liver and mammary gland, suggesting a role of this enzyme in steroid metabolism in these tissues. In addition, two other UGT2B cDNAs were isolated, UGT2B28 type II and III which lack several amino acids in the putative UDPGA and substrate binding domains. Although no biological functions were attributed to these shorter isoforms, the findings presented here suggest for the first time that a single UGT2B gene encodes highly related proteins with a variable number of structural motifs.

MATERIALS AND METHODS

Materials. UDP-glucuronic acid and all aglycons were obtained from Sigma (St. Louis, MO) and ICN Pharmaceutical Inc. (Montreal, PQ). Radioinert steroids were purchased from Steraloids Inc. (Wilton, NH). UDP-[^{14}C]glucuronic acid (285 mCi/mmol), [α-^{32}P]dCTP (3000 Ci/mmol), and [^{35}S]-methionine were obtained from Dupont NEN (Boston, MA). Lipofectin and Geneticin (G418) were obtained from Gibco BRL (Burlington, ON), and Blasticidin was obtained from Invitrogen (Carlsbad, CA). Protein assay reagents were obtained from Bio-Rad (Richmond, CA). Restriction enzymes and other molecular biology reagents were from Pharmacia LKB Biotechnology Inc. (Milwaukee, WI), Stratagene (La Jolla, CA), Gibco BRL, and Boehringer Mannheim (Indianapolis, IN). *Pfu* DNA polymerase was purchased from Stratagene (La Jolla, CA). HK293, ZR-75-1, MCF-7, T47-D, and LNCaP cells were obtained from the American Type Culture Collection (Rockville, MD). Total RNA from human liver, prostate, adrenal gland, testis, mammary gland, kidney, stomach, small intestine, skeletal muscle, spleen, ovary, and lung was purchased from Clontech (Palo Alto, CA).

cDNA Isolation of Human UGT2B28 Type III. Affinity-purified LNCaP cell mRNA was used to construct a cDNA library in the *λ*ZAP express vector, which was screened as previously described (*30*). After approximately 10^6 recombinants had been screened, 30 positive clones were isolated from the LNCaP cell library. The UGT2B28 type III cDNA clone was isolated from the LNCaP cell cDNA library and was sequenced in both directions using specific UGT oligonucleotides (*42*).

Cloning of Human UGT2B28 Type I and Type II. UGT2B28 type I and type II were obtained by PCR amplification of an LNCaP cell cDNA library using 100 pmol of the sense primer (5′-ACCAGGATGGCTCTGAAGTG-GACTTCAGTTCTTCTGCTGATACATCTCG-3′) and 100 pmol of the antisense primer (5′-TTGACATCAGTTTAT-TCCAGCAAGA-3′). The PCR conditions were as follows: 94 °C for 30 s, 60 °C for 30 s, and 72 °C for 2 min over the course of 35 cycles, using *Pfu* polymerase. PCR products were separated by electrophoresis on a 1.0% agarose gel, and two fragments of 1358 and 1666 bp were purified. Purified amplicons were shotgun cloned into the *pBluescript* SK vector digested with *Eco*RV and sequenced in both directions. PCRs (using proofreading *Pfu* polymerase) were performed with oligos mentioned above on recombinant UGT2B28 plasmids, and then, the PCR products were gel purified and were subcloned in an *Eco*RV pcDNA6 HisB vector (Invitrogen). The recombinant cDNAs were both sequenced as previously described (*42*).

Isolation of the UGT2B28 Gene. A full-length UGT2B28 cDNA probe was used to screen a RCPI-3 PAC library (BACPAC Resources Center at CHORI, Oakland, CA) as previously described (*35*). One positive genomic PAC clone was isolated from this screen (PAC clone 23123) and was obtained from the BACPAC Resources Center. The 23123 PAC clone was digested with *Eco*RI, *Hin*dIII, or *Eco*RV, and fragments were shotgun cloned into the *Bluescript* SK$^+$ vector (Stratagene) and transformed into *Escherichia coli* DH5α cells. Recombinant colonies were screened successively with probes corresponding to the different exons of UGT2B4, 2B7, 2B15, and 2B17, which were obtained by PCR amplification. Probe labeling and hybridization conditions are identical to those used for Northern blot analysis. Sequencing of the genomic clones was performed with the dideoxynucleotide chain termination method (*42*). The size of the introns was determined by PCR amplification using PAC clone 23123 as a template and oligonucleotides specific to the introns of the *UGT2B28* gene (Table 1). PCR was preformed using *Tth* polymerase mix (Clontech). Amplifications were performed for 30 cycles at 95 °C for 15 s, 59 °C for 15 s, and 68 °C for 1−10 min depending on the intron length. The sizes of the PCR fragments were determined by agarose gel electrophoresis.

Table 1: Primers Used for the Determination of the Intron Sizes of the UGT2B28 Gene[a]

amplicon	primer
sense	
intron 1	5′-AAGAATTTGTTTAATCGGGAACTTGAAGA-3′
exon 2/intron 2/ exon 3	5′-ATCCACTTTTTCTTTTCTTTATTCCTGTCA-3′
intron 3	5′-TGTGGAAAACTACTGAAAGAGGCTGTTAAG-3′
intron 3/exon 4/ intron 4	5′-TGTGGAAAACTACTGAAAGAGGCTGTTAAG-3′
exon 4/intron 4	5′-TACAGTTCTAACATTCTATAATTTTT-3′
antisense	
intron 1	5′-TACATAACTTTCTGAAAGGGGGTTAGAATT-3′
exon 2/intron 2/ exon 3	5′-TCTTTCAGTAGTTTTCCACACCAGTAAGGC-3′
intron 3	5′-AATTATAGAATGTTAGAACTGTAAAAAGGG-3′
intron 3/exon 4/ intron 4	5′-CGTACTTGTTTAGGAGATGTAATTGAAGT-3′
exon 4/intron 4	5′-CGTACTTGTTTAGGAGATGTAATTGAAGT-3′

[a] The sequences of the oligonucleotides used to estimate the size of the introns (indicated on the left) are shown. The sense primers represent the 5′ boundaries, and the antisense primers represent the 3′ boundaries of the PCR products.

In Vitro Transcription/Translation Assay. The pcDNA6 HisB vectors containing the UGT2B28 types I−III cDNAs (Invitrogen) were transcribed using T7 RNA polymerase and translated in the presence of [^{35}S]methionine, using the TNT-coupled rabbit reticulocyte lysate system from Promega Corp. (Madison, WI). The reaction products were separated on 10% SDS−PAGE and exposed on Hyperfilm-MP for 3 h.

Stable Expression of UGT2B28 Types I−III. HK293 cells were grown in Dulbecco's modified Eagle's medium containing 4.5 g/L glucose, 10 mM HEPES, 110 μg/mL sodium pyruvate, 100 IU/mL penicillin, 100 μg/mL streptomycin, and 10% fetal bovine serum in a humidified incubator, with an atmosphere of 5% CO_2, at 37 °C. Twenty micrograms each of pcDNA6 HisB-UGT2B28 types I−III were used to transfect HK293 cells using Lipofectin, according to the manufacturer's instructions (Gibco BRL). Forty-eight hours post-transfection, stable transfectants were selected in medium containing 10 μg/mL Blasticidin. After colony selection, nine monoclonal cell lines stably expressing human UGT2B28 type I, II, or III were isolated.

Northern Blot Analysis. Total RNA was isolated from human embryonic kidney cells (HK293) untransfected and stably expressing UGT2B28 types I−III, according to the Tri reagent acid phenol protocol as specified by the supplier (Molecular Research Center, Inc., Cincinnati, OH). Twenty micrograms of total RNA was electrophoresed on a 1.0% agarose gel and transferred to a GeneScreen Plus membrane (Boston, MA). A pool of UGT2B cDNA probes encoding UGT2B4, UGT2B7, and UGT2B15 was radiolabeled with the randomly primed DNA labeling kit (Boehringer Mannheim) in the presence of [α-^{32}P]dCTP. The Northern blot was prehybridized in 40% formamide, 5× Denhardt's, 5× SSPE, 50 mM Tris-HCl (pH 8.0), 1.0% SDS, and 100 μg/mL salmon sperm DNA for 4 h at 42 °C. Hybridization was performed in the same solution for 16 h at 42 °C with 1.0×10^6 cpm/mL cDNA probe. The blot was washed twice in 0.5× SSC and 0.1% SDS at 42 °C for 10 min and exposed for 36 h at −80 °C to XAR Hyperfilm with an intensifying screen (Kodak Corp., Rochester, NY).

Preparation of Microsomal Fractions. Microsomes were prepared by differential centrifugation. HK293 cells stably expressing UGT2B28 types I−III were homogenized in 4 mM K_2HPO_4 (pH 7.0), 20% glycerol, 1 mM EDTA, and 0.5 mM dithiothreitol, and the cellular homogenate was centrifuged at 12000*g* for 20 min to remove nuclei, unbroken cells, and mitochondria. The pellet was discarded, and the supernatant was centrifuged at 105000*g* for 60 min to obtain the microsomal pellet, which was resuspended in homogenization buffer and stored at −20 °C.

Western Blot Analysis. To ascertain the expression of the human UGT2B28 proteins, 20 μg of microsomal protein from HK293 cells and HK293 cells stably expressing UGT2B28 types I−III were separated by 10% SDS−PAGE. The gel was transferred onto a nitrocellulose membrane and probed with EL-93 anti-UGT2B17 antisera (1:2000 dilution) as reported previously (*28*). An anti-rabbit IgG horse antibody conjugated with peroxidase (Amersham) was used as the second antibody, and the resulting immunocomplexes were visualized using the enhanced chemiluminescence kit (Renaissance, Québec City, PQ) following the manufacturer's instructions and exposed to Hyperfilm-MP for 1 h (Kodak Corp.).

Glucuronidation Assay Using Microsome Preparations. Enzymatic assays were performed using 12 μM [^{14}C]UDP-glucuronic acid (UDPGA), 94 μM unlabeled UDPGA, 200 μM aglycone, and 20 μg of proteins from microsome preparations in 50 mM Tris-HCl (pH 7.5), 10 mM $MgCl_2$, 100 μg/mL phosphatidylcholine, and 8.5 mM saccharolactone in a final volume of 100 μL. Assays were performed for 16 h at 37 °C and were terminated by adding 100 μL of methanol. Samples were centrifuged at 14 000 rpm for 1 min in an Eppendorf microcentrifuge to remove the precipitated proteins. One hundred microliters of the aqueous phase was applied onto TLC plates (0.25 mm thick silica gel, Whatman, Maidstone, U.K.) and chromatographed in a solvent of toluene, methanol, and acid acetic (7:3:1) as previously reported (*22*, *40*). The TLC plates were exposed for 24 h, and the extent of glucuronidation was assessed by Phosphorimager (Molecular Dynamics, Sunnyvale, CA).

Quantitation of Free and Glucuronidated Steroids in Breast Cyst Fluid. Seventy-five women underwent a needle aspiration of a breast cyst. After aspiration, the gross cyst disease fluid was stored at −20 °C until it was assayed and all breast cyst samples were analyzed simultaneously to avoid interassay variation. Samples were analyzed following a procedure originally designed for the quantitation of plasma steroids (*43*−*45*). Assay of the unconjugated steroids was performed by extraction and column chromatography, while that for steroid glucuronide involved an extraction followed by treatment with β-glucuronidase and column chromatography (*6*).

All RIA results are shown as the mean ± standard error of the mean of triplicate determinations on individual samples. Radioimmunoassays were calculated according to the method of Rodbart and Lewald (*46*, *47*).

Tissue Distribution of the UGT2B28 Type I−III Transcripts. The tissue distribution of human UGT2B28 was determined using a specific RT-PCR technique. The reverse transcription reaction was performed using 10 μg of total human RNA. The reaction was carried out using 200 units of Superscript II (Gibco BRL) in 125 μM dNTPs, 10 mM Tris-HCl (pH 8.3), 50 mM KCl, 1.5 mM $MgCl_2$, 32.4 units of RNA Guard, and 100 pmol of the antisense primer (5′-

TTGACATCAGTTTATTCCAGCAAGA-3′) in a total volume of 20 μL for 90 min at 42 °C. After inactivation of the reverse transcriptase for 5 min at 100 °C, 2 μL of the reverse transcription product was used as a template in a PCR mixture containing 10 mM Tris-HCl (pH 8.3), 50 mM KCl, 1.5 mM $MgCl_2$, 0.1 mM dNTP, and 1.25 units of *Pfu* DNA polymerase in a total volume of 100 μL. The reaction was carried out using 100 pmol of the specific sense primer (5′-AGTTCTTCTGCTGATACATCTCG-3′) and antisense primer (5′-AGTGGTACTGGAACCAGGTGAGGTGAC-3′). The PCR was performed for 35 cycles (1 min at 94 °C, 1 min at 55 °C, and 1 min at 72 °C), after which one-fifth of the PCR product was electrophoresed on an ethidium bromide-stained 1.0% agarose gel, and the DNA fragments that were obtained were visualized under UV light. All room-temperature reactions were controlled by using specific oligonucleotides for glyceraldehyde phosphate dehydrogenase (GAPDH). The identity of all PCR products was verified by direct sequencing (*22*, *28*, *30*, *40*, *48*).

Immunofluorescence Visualization. HeLa cells (20000) were plated in culture slides (Biocoat Culture Slides, Becton Dickinson, Flanklin Lakes, NJ) and incubated for 16 h. Transfection was then performed with Exgen 500 (MBI Fermentas, Flamborough, ON) according to the manufacturer's instructions. Twenty hours post-transfection, cells were washed three times with cold PBS, and then fixed for 15 min with Formalin (Sigma Chemical Co.). Cells were washed five times with PBS before permeabilization with PBS containing 0.5% Triton X-100 (Sigma Chemical Corp.) for 15 min at room temperature. After three washes with PBS, the cells were incubated for 1 h in blocking buffer [6% dry milk and 3% bovine serum albumin in Tris-buffered saline (TBS)] to minimize unspecific binding of antibodies. Primary antibodies rabbit anti-*c*-myc (Upstate Biotechnology, Lake Placid, NY) and mouse anti-PDI (Stressgen Biotechnology Corp., Victoria, BC) were added in binding buffer (blocking buffer diluted 1:10 with TBS) with a 1:100 dilution. Slides were incubated overnight at 4 °C and then washed twice with PBS containing 0.2% Tween 20 and twice with PBS. Secondary antibodies (Alexa Fluor 594 goat anti-rabbit and Alexa Fluor 488 goat anti-mouse from Molecular Probes Inc., Eugene, OR) were added simultaneously at a 1:1000 dilution in binding buffer, and slides were incubated for 1 h at room temperature in the dark. Washes were carried out as described for primary antibodies. Diamidino-2-phenylindole (DAPI) counterstain was achieved by incubating the slides for 10 min in the dark at room temperature with 0.1 μM dilactate DAPI (Molecular Probes Inc.). Cells were intensively washed with water, allowed to dry, and mounted with a Prolong Anti-Fade kit (Molecular Probes Inc.). For visualization, a Leica DMR-B epifluorescence microscope (Leica Inc., Deerfield, IL) was used in combination with standard photography. Typically, exposure times varied between 0.5 and 15 s with Kodak Gold 100 ASA film.

RESULTS

Isolation of the UGT2B28 cDNA. To obtain novel human extrahepatic steroid metabolizing UGT enzymes, we screened an LNCaP cell cDNA library. Due to the high degree of homology between the UGT2B cDNA clones isolated to date, the library was screened using a pool of radiolabeled cDNA probes synthesized from the human UGT2B4, UGT2B7, and UGT2B15 cDNAs. After approximately 10^6 nonamplified recombinants had been screened, 30 positive clones were isolated. Among these positives clones we identified cDNAs encoding UGT2B4(D^{458}) (*22*, *49*), UGT2B4-(E^{458}) (*22*), UGT2B10 (*26*), UGT2B11 (*27*), UGT2B15(Y^{85}) (*28*), UGT2B15(D^{85}) (*28*, *29*), UGT2B17 (*30*), and a novel 1.5 kb UGT2B cDNA. Sequence analysis revealed that this cDNA is 1500 bp in length and contains an open reading frame of 1236 bases flanked by a 5′-untranslated region of 2 bp and a 3′-untranslated region of 262 bp. Comparison of the nucleic acid sequence with other known UGT2B enzymes highlighted the absence of 351 bp within the putative substrate binding domain (Figure 1). To confirm the expression of this isoform in LNCaP cells and human tissues, we performed RT-PCR using primers specific for this transcript. In addition to the anticipated 1097 bp amplicon, we amplified two other longer cDNA fragments of 1140 and 1448 bp. To obtain the corresponding full-length cDNAs, we performed PCR with specific primers (with one primer in the highly variable 3′-untranslated region) (Figure 1) on an LNCaP cell cDNA library which led to amplification products of 1666 and 1358 bp. The two fragments were cloned into the pcDNA6 HisB plasmid and sequenced with specific primers.

Sequence analysis of the two longer cDNA species revealed a coding sequence identical with the previous initial 1.5 kb cDNA, except that they both possess the missing 351 bp not found in the putative exon 1 of the shortest cDNA. Furthermore, the nucleotide region of the 1140 bp amplicon was identical to the other two cDNAs with the exception of nucleotides 1003−1310 (from the ATG) that were missing in the region encompassing the putative UDPGA binding site (Figure 1). The sequences of the full-length cDNA clones, including the 5′- and 3′-flanking regions, are 1851, 1543, and 1500 pb in length, and they were named UGT2B28 type I (GenBank entry AF177272), UGT2B28 type II (GenBank entry AF177273), and UGT2B28 type III (GenBank entry AF177274), respectively, according to their respective length (Figure 2).

The 1851 bp UGT2B28 type I cDNA contains an open reading frame of 1587 bp encoding a characteristic 529-amino acid UGT protein (Figure 2). Two putative polyadenylation signals with the sequences AATAAA and AATAAT are present at nucleotides 1809 and 1821, respectively, with a poly(A+) tail starting at position 1830 (Figure 1). The protein primary structure, deduced from the nucleotide sequence (Figure 1), contains an ER targeting signal, a membrane anchoring region, and substrate and cofactor binding domains. The deletion of 308 nucleotides found in the UGT2B28 type II cDNA led to the presence of a premature stop codon. The protein encoded by this cDNA is thus composed of 335 amino acids (Figure 1). The protein contains the ER leader sequence and the substrate binding domain but is without the UDPGA binding domain and the transmembrane region. At the opposite of the latter isoform, the 351 bp deletion of the UGT2B28 type III cDNA does not alter the open reading frame of the protein, leaving the conserved UDPGA binding domain intact in the C-terminal region. The 412-amino acid protein contains a characteristic signal peptide directing the protein to the ER and a hydrophobic transmembrane region at the C-terminal end of the protein (Figure 2). Sequence analysis of the three cDNAs

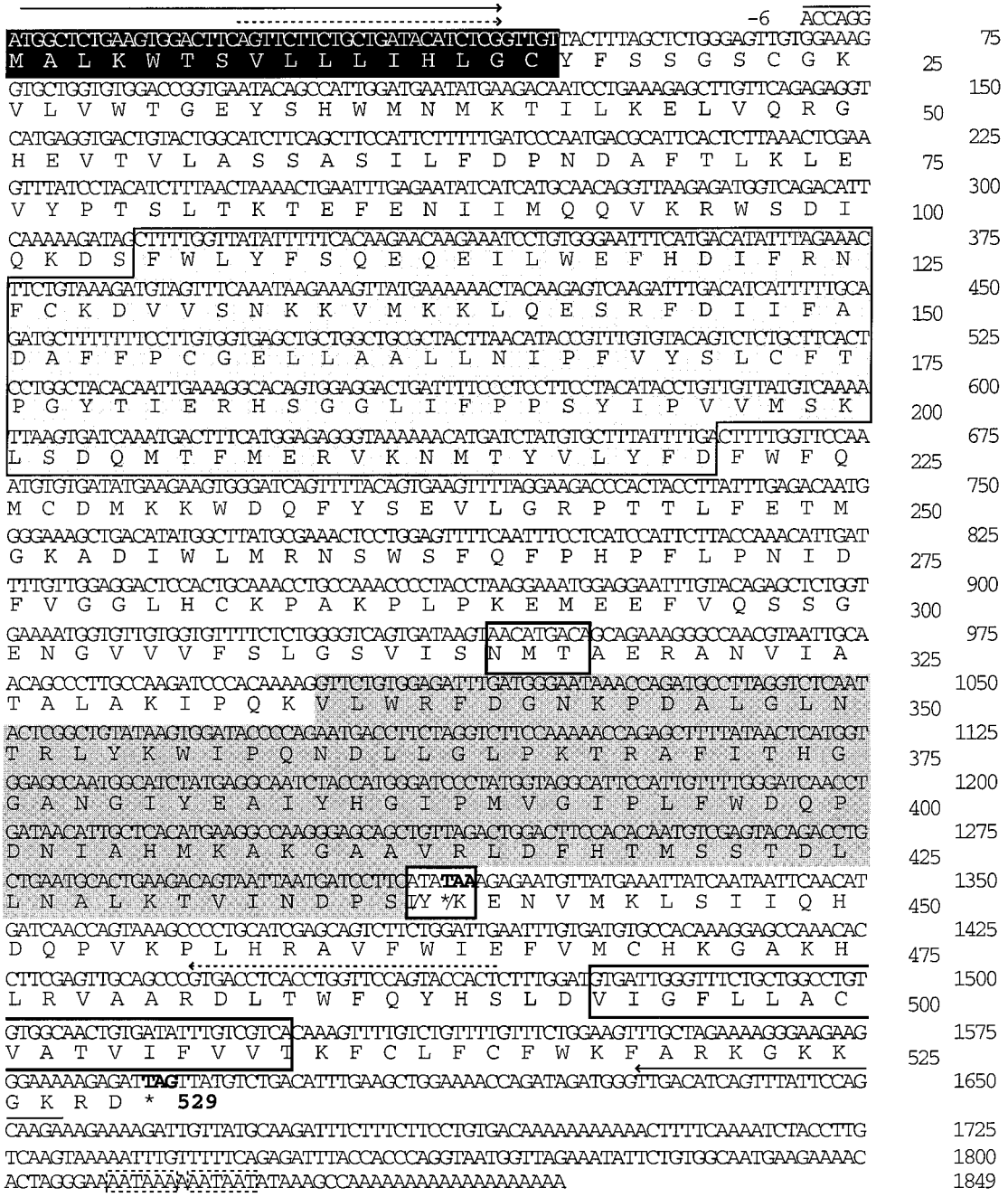

FIGURE 1: Nucleotide and amino acid sequences of UGT2B28. The putative membrane-insertion signal peptide is denoted by a black box, and the anchoring domain is denoted by the light gray box. The boxed residues (NMT) identify the potential glycosylation site. The sequence identity from residue 105 to 221 (shaded box) is absent in UGT2B28 type III (GenBank entry AF177274). UGT2B28 type II (GenBank entry AF177273) does not contain the sequence from residues 335 to 437 and possesses a premature stop codon at nucleotide 1006, which is denoted with an asterisk. Continuous arrows show primers used to amplify full-length UGT2B28 type I (GenBank entry AF177272), whereas the dashed arrows represent primers used for RT-PCRs.

revealed that all UGT2B28 subtypes contain only one potential asparagine-linked glycosylation site [NX(S/T)] present at residues 315−317 of UGT2B28 types I and II and 198−200 of UGT2B28 type III (Figure 2).

Amino acid sequence alignment with other UGT2B enzymes demonstrated that UGT2B10 and UGT2B11 are the isoforms which are the most homologous to UGT2B28 type I (Figure 3 and Table 2). UGT2B10 and UGT2B11 are two orphan UGTs and are 89 and 95% identical to UGT2B28 type I, respectively. The amino-terminal region from amino acids 1−290 of UGT2B28 type I is 86 and 94% identical to human UGT2B10 and UGT2B11, respectively, whereas the carboxy-terminal region from residues 291−529 of these UGT2B isoforms was more than 92% identical (Table 2). The UGT2B28 type I protein is 76−85% identical to the other human UGT isoforms (Table 2 and Figure 3B).

Isolation of the UGT2B28 Gene. To isolate the human *UGT2B28* gene, a human genomic PAC library was screened with a full-length UGT2B28 cDNA probe and the PAC clone 23123 was isolated. Restriction enzyme digestion of this PAC clone at the *Not*I sites, which flank the insert in the cloning vector, yielded a DNA insert of approximately 100 kb (not shown). Screening of the PAC 23123 *Eco*RI, *Hin*dIII, and *Eco*RV fragments cloned into the *pBluescript* SK$^+$ vector

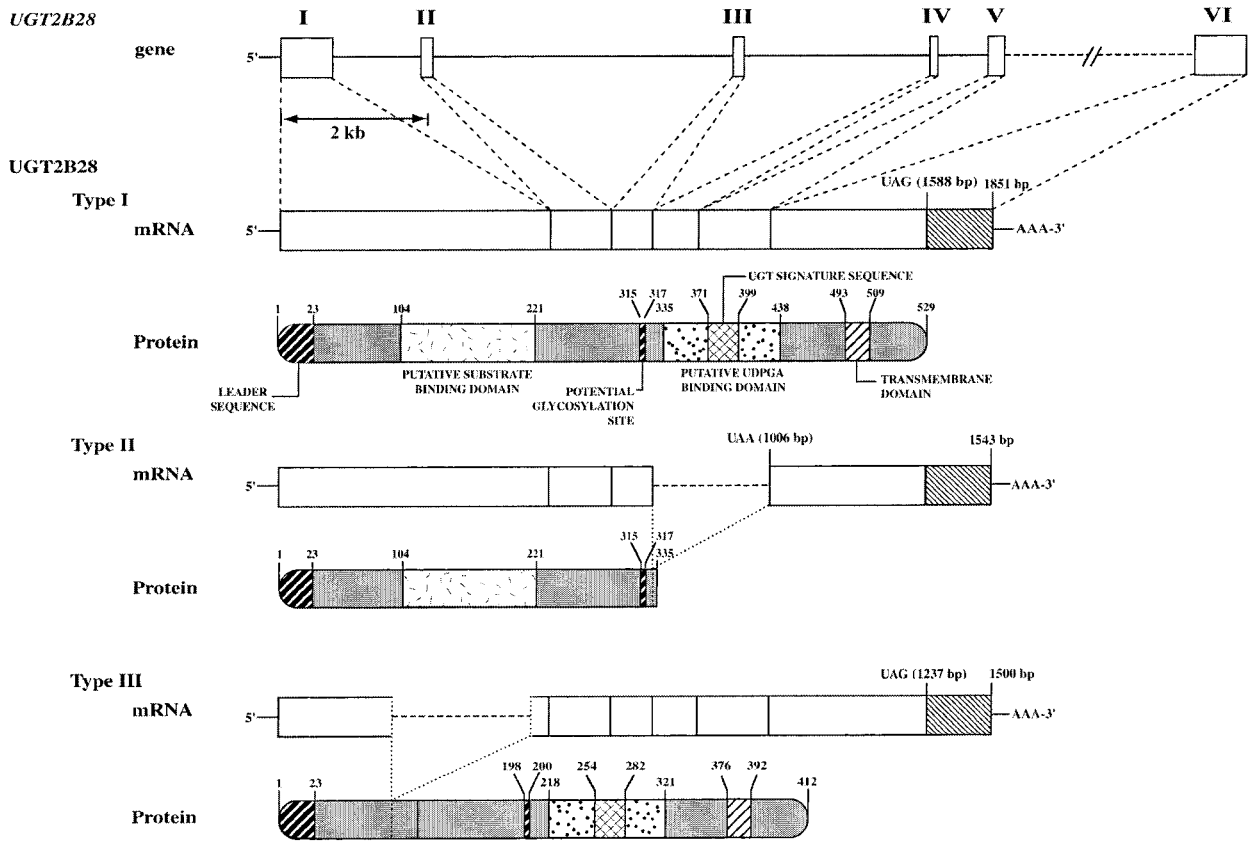

FIGURE 2: Schematic representation of the human *UGT2B28* gene and the encoded mRNA species and proteins. The dashed line indicates that this region of the gene was not present in the PAC 23123 clone and was deduced from the common UGT2B genomic organization and the full-length UGT2B28 cDNA. The protein's putative structural motifs are as indicated.

led to the identification of exons 1−5 of the *UGT2B28* gene. Exons 1−5 are 721, 149, 132, 88, and 220 bases long, respectively, which is identical to all characterized UGT2B genes (Figure 2). PCR amplification was used to determine the appropriate size of the introns, which vary from 700 bp to 4.5 kb (Figure 2). Exon 6 of the *UGT2B28* gene was not present in this PAC clone, and due to the high degree of homology between the versions of exon 6 from all human UGT2B genes (more than 95%), we were not able to isolate this exon in several PAC clones containing UGT2B genes. However, using identical screening conditions, the entire *UGT2B11* gene was isolated in the PAC 23123 clone, indicating that the two highly homologous genes, *UGT2B11* and *UGT2B28*, are clustered together on chromosome 4.

Expression of UGT2B28 Type I−III Proteins. To determine the ability of each UGT2B28 cDNA to encode the predicted proteins, pcDNA6 HisB vectors containing each cDNA were transcribed by T7 RNA polymerase, and translated with rabbit reticulocyte lysate (Figure 4A). The radiolabeled peptides that were produced were 52, 35, and 42 kDa as expected for UGT2B28 types I−III, respectively. Stable expression of these three cDNAs (containing the ORF regions) into HK293 cells demonstrated the expression of transcripts of the appropriate size in each cell line (Figure 4B). As shown in Figure 4B, the HK293-UGT2B28 type I−III stable cell lines expressed 1.6, 1.4, and 1.3 kb messenger RNA, respectively. To ascertain the level of expression of UGT2B28 type I−III proteins in these stably transfected HK293 cells, Western blot analysis was performed on a microsomal preparation using a polyclonal antibody specific to the UGT2B enzymes (EL-93). The three UGT2B28 proteins detected had apparent molecular masses of 52, 35, and 42 kDa as already observed in the in vitro transcription/translation assay (Figure 4C). UGT proteins were absent in the untransfected HK293 cell microsome preparation as previously described (Figure 4C) (*22*, *28*).

Transferase Activity of UGT2B28 Enzymes. To determine the substrate specificity of the UGT2B28 enzymes, microsomal preparations from cells stably expressing UGT2B28 types I−III were incubated for 16 h (UGT2B28 type I is stable for 16 h under these conditions; not shown) in the presence of different substrates (200 μM), 12 and 94 μM radiolabeled and unlabeled UDPGA. Under these assay conditions, cofactor availability may slightly reduce the velocity of the reaction; however, we and others have used this concentration to obtain the highest signal-to-background ratio, and it accurately indicates the transferase activity (*30*, *50*). UGT2B28 type I protein glucuronidated eugenol, 5β-androstane 3α,17β-diol, etiocholanolone, 5α-androstane 3α,-17β-diol (3α-diol), 4-methylumbelliferone, 1-naphthol, estradiol (E_2), ADT, testosterone, hyodeoxycholic acid (HDCA), and lithocholic acid (LCA) (Table 2 and Figure 5). Eugenol, a phenolic compound which is glucuronidated by the majority of UGT2B isoenzymes, was the best substrate for glucuronidation by UGT2B28 type I with a velocity of 131 pmol h^{-1} (mg of protein)$^{-1}$. Steroid molecules highly conjugated by UGT2B28 type I included 5β-androstane 3α,17β-diol [88 pmol h^{-1} (mg of protein)$^{-1}$], ADT [38 pmol h^{-1} (mg of protein)$^{-1}$], and E_2 [46 pmol h^{-1} (mg of protein)$^{-1}$] (Figure 5). A nontoxic bile acid, HDCA, was also significantly glucuronidated at a rate of 34 pmol h^{-1} (mg of protein)$^{-1}$ by HK293 microsomal fractions expressing UGT2B28 type

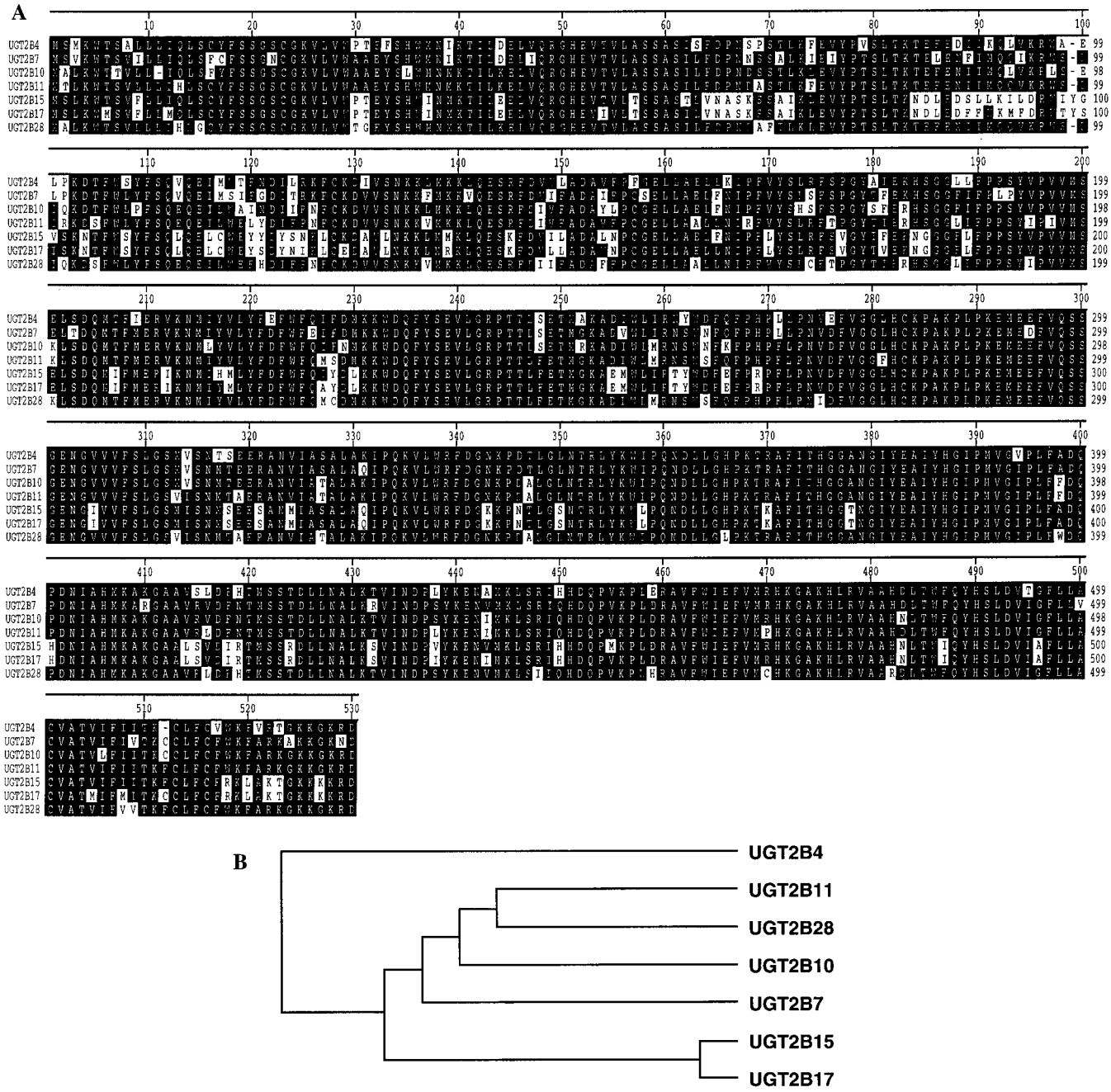

FIGURE 3: (A) Amino acid sequence of UGT2B28 type I and alignment with the sequence of other UGT2B isoenzymes. The amino acid sequences were deduced from the nucleotide sequences. (B) Dendrogram of homology between human UGT2B enzymes.

Table 2: Homology between the Deduced Nucleic Acid and Amino Acid Sequences of UGT2B28 Type I and Other UGT2B Isoenzymes[a]

		UGT2B28 type I amino acid identity (%)			
	steroid substrates	nucleic acid	amino-terminal domain (residues 1−290)	carboxy-terminal domain (residues 291−530)	overall
UGT2B4	4-OHE$_1$, E$_3$, 3α-diol	89	80	88	84
UGT2B7	3α-diol, DHT, T, ADT, 4-OHE$_1$, E$_3$	89	80	91	85
UGT2B10	ND	94	86	92	89
UGT2B11	ND	97	94	95	95
UGT2B15	3α-diol, DHT, T	84	70	82	76
UGT2B17	DHT, ADT, 3α-diol, T	84	70	81	76

[a] The nucleic acid identity (%) is as indicated. The sequence identities of the amino-terminal domain from residue 1 to 290, the carboxy-terminal domain from residue 291 to 529, and the entire protein are as indicated. The steroid specificity of each enzyme is described. Abbreviations: ADT, androsterone; E$_3$, estriol; E$_2$, estradiol; 4-OHE$_1$, 1,3,5,10-estratrien-3,4-diol-17-one; 3α-diol, 5α-androstane-3α,17β-diol; DHT, dihydrotestosterone; T, testosterone; THE, tetrahydrocortisone. ND (not detectable) indicates that no substrate has been identified.

I. On the other hand, UGT2B28 types II and III have not demonstrated glucuronidation activity toward the 35 endogenous and exogenous compounds that were tested (data not shown). Glucuronidation of these compounds was not

Table 3: Substrates Tested Which Are Not Glucuronidated by UGT2B28 Type I[a]

C19 steroids
dihydrotestosterone (DHT)
epiandrosterone
dehydroepiandrosterone (DHEA)
androst-5-ene-3β,17β-diol (Δ5-diol)
5α-androstane-3β,17β-diol (3β-diol)
C18 steroids
estrogens
estrone
estriol
catecholestrogens
1,3,5,10-estratriene-2,3-diol-17-one
1,3,5,10-estratriene-3,4-diol-17-one
1,3,5,10-estratriene-2,3,17β-triol
1,3,5,10-estratriene-3,4,17β-triol
C21 steroids
pregnenolone
17-OH-pregnenolone
5α-pregnane-3α-ol-20-one
5α-pregnane-3α,20α-diol
aldosterone
cortisol
11-deoxycortisol (11-DOC)
5α-tetrahydrocortisone
5β-tetrahydrocortisone
bile acids
chenodeoxycholic acid
cholic acid
exogenous compound
p,p′-biphenyl

[a] Determination of transferase activity was performed in the presence of 12 μM [^{14}C]UDP-glucuronic acid and 94 μM unlabeled UDP-glucuronic acid to optimize sensitivity of the assay.

detected using control microsomal extracts from untransfected HK293 cells as previously reported (*30*).

Breast Cyst Fluid Concentrations of Conjugated and Unconjugated Steroids. Figure 5B shows the concentrations in cyst fluid of unconjugated and conjugated C18 and C19 steroids, namely, 3α-diol, ADT, E_1, and E_2. The ADT levels at 6.2 nmol/L were higher than those of 3α-diol, E_2, and E_1 (Figure 5B). On the other hand, ADT-G, the major steroid conjugate found in breast cyst fluid, was found at a concentration of 125 nmol/L that exceeds by more than 10- and 100-fold that of 3α-diol-G, E_1-G, and E_2-G (*47*).

Tissue Distribution of the UGT2B28 Transcripts. In the past few years, we have shown that RNase protection analysis was not sufficiently sensitive for detection of UGT2B mRNA species in extrahepatic cells, and therefore, here we used a specific RT-PCR approach to determine the tissue distribution of UGT2B28 transcripts in human tissues. Under these conditions, the three UGT2B28 subtypes were detected in liver, mammary gland, and LNCaP cells. In addition, the 1097 bp amplicon corresponding to the UGT2B28 type III transcript was detected in the kidney, testis, prostate, BPH, placenta, lung, stomach, small intestine, spleen, and skeletal muscle, and in breast cancer cell lines such as MCF-7, T47-D, and ZR-75-1 (Figure 6). The three transcripts were not detected in the adrenal, ovary, or HK293 cells. The PCR conditions used here were optimized for all UGT2B28 transcripts since only one set of primers was designed for all reactions as shown in Figure 1. Therefore, knowing that identical oligonucleotides were used for all reactions, one would expect to obtain all UGT2B28 transcripts expressed

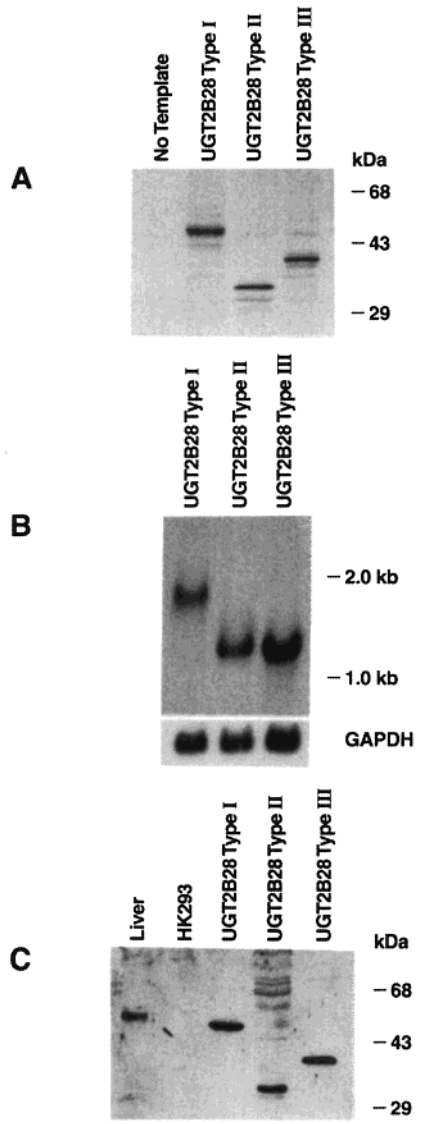

FIGURE 4: (A) In vitro transcription/translation mechanism of UGT2B28. The pcDNA6 HisB vector containing UGT2B28 type I−III cDNA was transcribed using T7 RNA polymerase and translated in the presence of [^{35}S]methionine, using TNT-coupled rabbit reticulocyte lysate. Reaction products were separated via 10% SDS−PAGE and exposed on Hyperfilm. (B) Northern blot analysis of UGT2B transcripts isolated from stably transfected HK293-UGT2B28 cells. Twenty micrograms of total RNA from HK293 cells stably expressing UGT2B28 types I−III were electrophoresed on a 1% agarose gel, transferred onto a GeneScreen membrane, and hybridized with a pool of UGT2B cDNA probes. (C) Western blot analysis using an anti-UGT2B antibody. Twenty micrograms of microsomal proteins from HK293 cells stably expressing UGT2B28 was separated by 10% SDS−PAGE, transferred onto a nitrocellulose membrane, and probed with the polyclonal EL-93 antibody.

in a specific human tissue, which is exactly the case in the liver and LNCaP cells and the mammary gland (Figure 6). The amplification of the shortest UGT2B28 transcripts in each tissue is also indicative of the specificity of the PCRs since no other UGT2B isoforms have been shown to possess shorter mRNA species. Under these conditions, we were not able to demonstrate the presence of other UGT2B28 transcripts. Finally, all PCR products were subcloned, and the identity of each DNA fragment was confirmed by direct sequencing (*48*). RT-PCR amplification of the GAPDH transcript was used as a positive control for the RT-PCRs

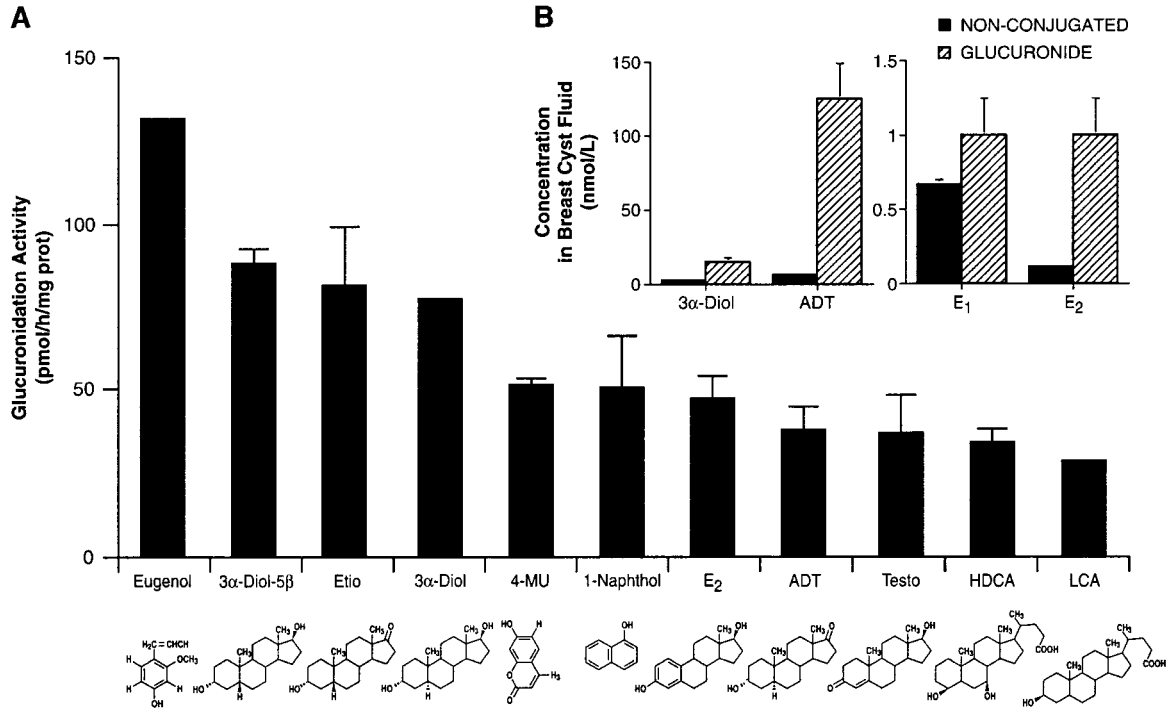

FIGURE 5: (A) Rates of glucuronidation catalyzed by the UGT2B28 type I isoenzyme toward endogenous and exogenous compounds. Initial screening for reactive substrates was performed in the presence of 12 μM [^{14}C]UDP-glucuronic acid and 94 μM unlabeled UDP-glucuronic acid to optimize the sensitivity of the assay. The subsequent quantitation of glucuronide formation was optimized to measure the glucuronidation activity of UGT2B28 type I for the various substrates and was performed in the presence of 12 μM [^{14}C]UDP-glucuronic acid and 94 μM unlabeled UDP-glucuronic acid as previously reported (*22*, *28*, *30*). Activity values are the means ± standard deviation of three independent experiments. (B) Breast cyst fluid concentrations (nanomoles per liter) of free and glucuronide derivatives of 3α-diol, ADT, E_2, and E_1 obtained from 75 women who underwent needle aspiration of breast cysts (*47*).

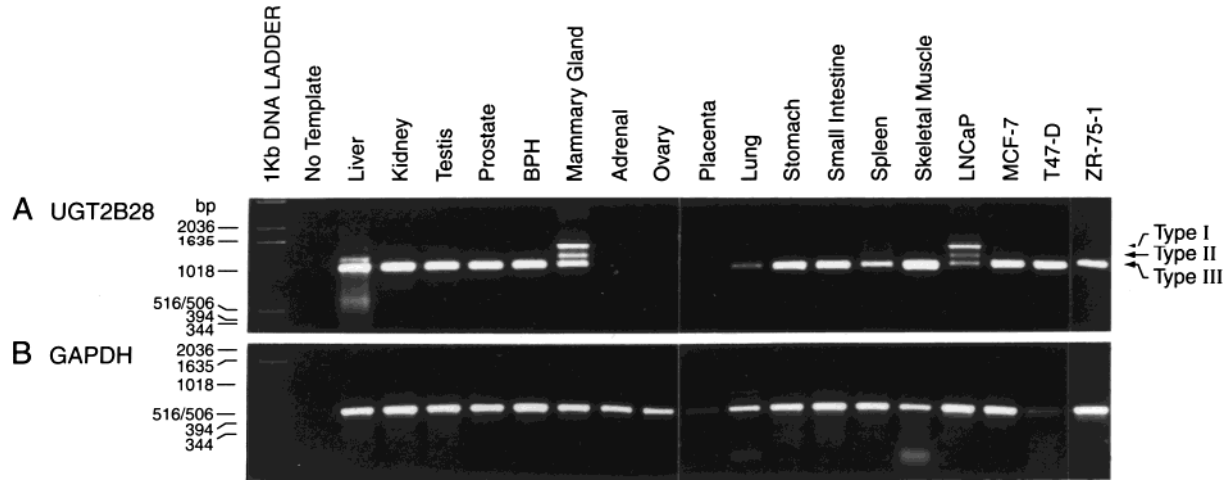

FIGURE 6: Tissue distribution of UGT2B28 transcripts. Total RNA isolated from human liver, kidney, testis, prostate, benign prostate hyperplasia (BPH), mammary gland, adrenal, ovary, placenta, lung, stomach, small intestine, spleen, skeletal muscle, LNCaP, MCF-7, T47-D, ZR-75-1, and UGT2B28 types I−III were analyzed by specific RT-PCR analysis as described in Materials and Methods. One-fifth of each RT-PCR product was separated on a 1.0% agarose ethidium bromide-stained gel. The integrity of the RNA samples and the polymerase chain reactions were demonstrated by production of the 545 bp product corresponding to GAPDH.

and yielded products of appropriate size in each tissue sample (Figure 6).

Subcellular Localization of UGT2B28 Proteins. Immunofluorescence analysis demonstrated the expression of UGT2B28 types I−III in the perinuclear structure as well as in the endoplasmic reticulum (Figure 7A,D,G). Protein disulfide isomerase (PDI) located in the endoplasmic reticulum colocalized with UGT2B28 types I−III in both organelles (Figure 7B,E,H). Coloration with diamidino-2-phenylindole (DAPI) was restricted to the nucleus (Figure 7C,F,I). Approximately one-tenth of the total cells were transfected with UGT2B28 type I (Figure 7A−C), type II (Figure 7D−F), and type III (Figure 7G−I).

DISCUSSION

In this study, we report the isolation of a novel cDNA and its corresponding gene, which encode a novel UDP-glucuronosyltransferase protein, UGT2B28. The UGT2B28 type I cDNA encodes a full-length UGT2B enzyme with all the structural domains that are essential for UGT activity. Indeed, this perinuclear and ER-located enzyme possesses a conventional leader sequence targeting the protein in the

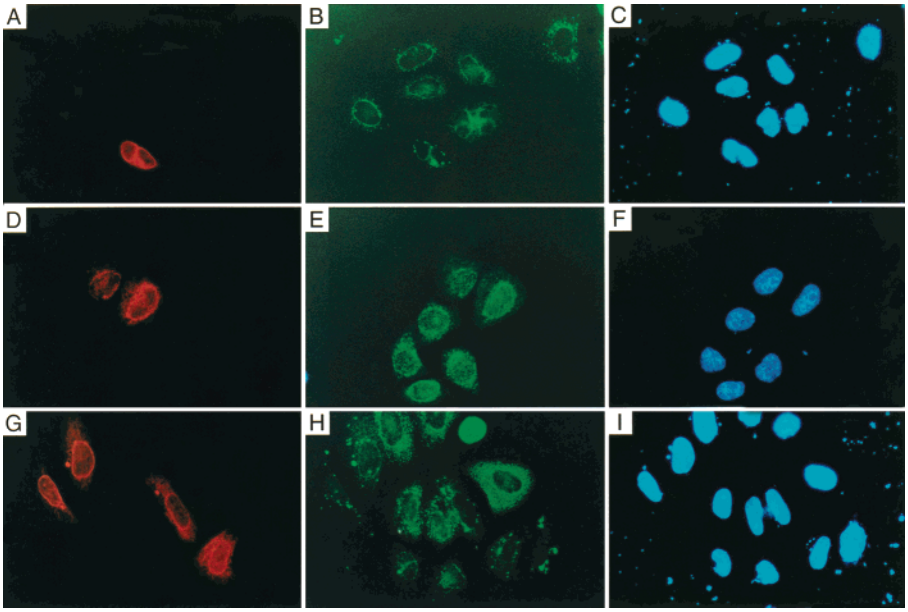

FIGURE 7: Immunofluorescence localization of UGT2B28 types I−III. HeLa cells were transfected with respective UGT2B28 isoenzymes, and cells were fixed and then treated with antibodies: (A−C) UGT2B28 type I, (D−F) UGT2B28 type II, and (G−I) UGT2B28 type III. Panels A, D, and G show anti-*c*-myc fusion tag fluorescence and thus represent the UGT2B28 protein; panels B, E, and H corresponded to anti-PDI fluorescence, and nuclei were stained with DAPI in panels C, F, and I.

membrane of these organelles, an ER-anchoring signal at the carboxyl end of the protein, a substrate binding domain in the less conserved amino half of the enzyme, and a characteristic UDPGA binding site in the highly conserved C-terminal region. As found among other members of the UGT2B subfamily, the 529-amino acid polypeptide encoded by UGT2B28 type I is less homologous to the other UGT2B enzymes within the amino-terminal region (residues 1−290) than the carboxy-terminal region (residues 291−529/530) (Figure 3A and Table 2) (*14*). Sequence alignment with other UGT2B proteins demonstrates the high degree of identity between UGT2B10 and UGT2B11, two orphan UGT2B proteins (*26*, *27*) (Figure 3A and Table 2). Stable expression of UGT2B28 type I in HK293 cells demonstrated the expression of a 1.6 kb transcript which encodes a 52 kDa transferase capable of conjugating bile acids and steroid molecules such as androgens and estrogens. Eleven of the 35 substrates that were tested were found to be conjugated by UGT2B28 type I, including ADT, Testo, 3α-diol, 5*β*-androstane-3α,17*β*-diol, E_2, and bile acids such as HDCA and lithocholic acid. Conjugation by UGT2B28 type I occurs at both the 3α and 17*β* positions of the steroid molecule, as shown by its ability to glucuronidate ADT and Testo. In addition to C19 steroids, UGT2B28 type I conjugates C18 steroids, as demonstrated by glucuronidation of E_2. However, UGT2B28 type I was unable to catalyze the conjugation of C21 steroids.

Tissue distribution of UGT2B28 transcripts was assessed by specific RT-PCR. The expression of UGT2B28 type I was observed in the liver and mammary gland. The human liver is known to be a major tissue involved in the formation of 5*β*-androstane 3α,17*β*-diol due to the high level of expression of the enzyme 5*β*-reductase (*51*−*53*). The highest UGT2B28 transferase activity was obtained with 5*β*-androstane 3α,17*β*-diol as the endogenous substrate, and therefore, the role of this enzyme in hepatocytes can be related to the irreversible inactivation and elimination of this C19 metabolite. To date, only UGT2B7, in addition to UGT2B28, has been shown to conjugate 5*β*-androstane 3α,17*β*-diol. UGT2B28 type I also glucuronidates bile acids, and it may contribute to HDCA and LCA glucuronidation in the liver, in addition to UGT2B4, UGT2B7, UGT1A3, and UGT1A8 (*22*, *54*, *55*). The extrahepatic expression of the UGT2B28 type I transcript was restricted to the mammary gland, which is an estrogen- and androgen-dependent tissue (*56*). Figure 5B shows that the breast cyst fluid contains significant amounts of androgen and estrogen glucuronides which is consistent with the expression and activity of UGT2B28 in the breast tissue (*57*). We observed the expression of the active UGT2B28 type I protein within the ER and perinuclear membranes. The physiological significance of the expression in the perinuclear structure is likely to be related to the regulation of steroid hormones action. By conjugating steroid molecules before their entrance into the nucleus, UGT enzymes could modulate the availability of steroid molecules for their receptors and therefore influence the hormonal response. In addition to the glucuronidation of C18 and C19 steroids by UGT2B28, the presence of this enzyme in breast tissue suggests that this enzyme may be involved in the inactivation of androgen and estrogen in this tissue (Figure 8).

Primary structures of UGT2B28 types II and III are identical to UGT2B28 type I, except that these cDNAs are shorter due to the absence of nucleotides within the amino (residues 1−290) and carboxyl (residues 291−529) portion of the protein. Nucleic acid alignment of UGT2B28 type II cDNA with the *UGT2B28* gene demonstrated that the 308 bp deletion from nucleotides 1003−1310 relative to the adenine of the initiation codon corresponds to exons 4 and 5. Unadequate maturation of the UGT2B28 primary transcripts, such as exon skipping, could be responsible for the expression of the truncated UGT2B28 type II mRNA. The deletion of this portion of the protein does not preserve the open reading frame; thus, a premature stop codon is found at nucleotide 1006. Therefore, the protein encoded, composed of only 335 amino acids, does not contain the putative

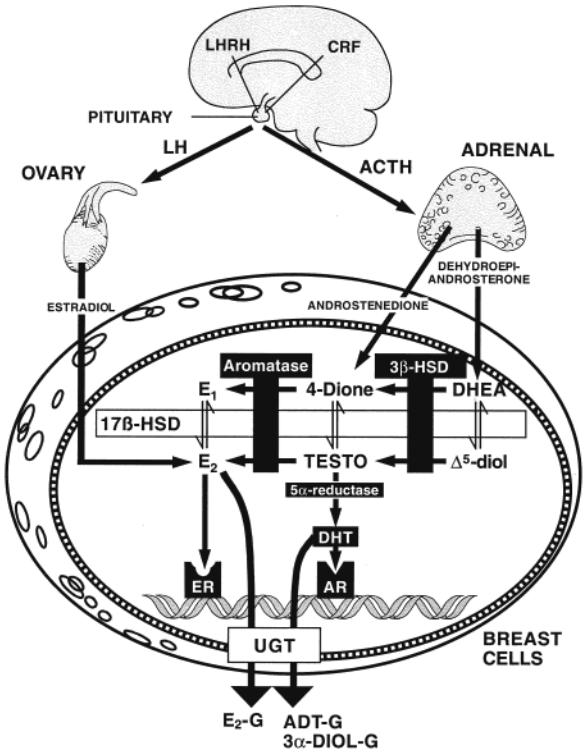

FIGURE 8: Schematic representation of the putative physiological role of UGT2B28 type I in breast cells. Steroid molecules such as DHT and estradiol are conjugated and inactivated by UGT2B28 and then excreted in the circulation. These conjugated molecules are finally eliminated in bile or urine.

UDPGA binding domain, the UGT signature sequence, or the transmembrane domain. Comparison of the shortest cDNA, UGT2B28 type III, with the *UGT2B28* gene sequence demonstrates the absence of 351 bp within the region corresponding to exon 1 of this gene. This deletion of nucleotides 312−662 does not change the open reading frame, and the predicted peptide is composed of 412 amino acids. However, UGT2B28 type III lacks the putative substrate binding domain, but contains all the other UGT2B characteristic domains.

Stable expression of the three UGT2B28 isoforms in the HK293 cell line demonstrated the expression of transcripts of 1.4 and 1.3 kb corresponding to the UGT2B28 type II and III coding regions, respectively. The expression of UGT2B mRNA species shorter than 1.6 kb was previously observed in monkey testis (*58*). However, in both human and monkey, the isolation and characterization of shorter UGT2B transcripts have not been reported. In vitro transcription/translation experiments and Western blot analysis performed with microsomal proteins obtained from the stably transfected HK293 cells clearly demonstrate that both UGT2B28 transcripts encoded microsomal proteins of 35 (type II) and 42 kDa (type III). Interestingly, despite the absence of the transmembrane domain and the dilysine motif of UGT2B28 type II, immunofluorescence localization demonstrates that this protein is localized in the ER and perinuclear membranes as found for UGT2B28 type I. These results are in accordance with previous observations indicating that the carboxy-terminal transmembrane region as well as the dilysine motif are not essential components for the localization of UGT2B proteins in the ER (*59*−*61*). Indeed, the localization of the UGT2B28 type II protein can be attributed to the signal peptide and amino acids 141−240. It has been demonstrated that both domains could be implicated in the targeting of the protein within the ER, and that the latter stretch of amino acids would certainly be responsible for the retention of the protein within this organelle. Although the residues implicated in the protein−ER association are not fully characterized, it has been suggested that residues 141−240 may form a buried α-helix which interacts with the lipid bilayer via hydrophobic interactions (*60*, *62*). These results, obtained with the first naturally occurring UGT protein devoid of the consensus ER retention signal, suggest that amino-terminal residues might be involved in maintaining UGT proteins in the ER. However, it was demonstrated by progressive truncation of the carboxyl end of the chimeric UGT protein that this region could be implicated in the correct folding of UGT2B enzymes within the membrane. Thus, it was not surprising to obtain no conjugation activity with UGT2B28 type II.

Integration of UGT2B28 type III does not present such interrogations since the translated peptide contains all the classical domains required for the expression of ER proteins. The 412-amino acid protein contains the leader sequence, the membrane anchoring domain, and the dilysine motif, and thus, it was not surprising to observe the expression of this peptide by immunofluorescence within the ER membrane. Interestingly, residues 105−221 of a characteristic UGT2B enzyme are absent in UGT2B28 type III. This 117-amino acid deletion partially disrupts the 100-amino acid stretch of the internal targeting-retention motif described by Ouzzine et al. (*62*). These results, obtained with both UGT2B28 types II and III, suggest that the localization of the UGT proteins implicates several domains, but that these domains are not required simultaneously for the endogenous expression of UGT proteins within the ER membrane. As found for UGT2B28 type II, UGT2B28 type III has not demonstrated conjugation activity toward the substrates that have been tested. The substrate binding domain, which is proposed to reside in the amino-terminal domain (residues 1−291), is restrained in the UGT2B28 type III isoenzyme; thus, the interaction between the aglycone and the protein is restricted to a smaller region which probably adopts a different conformational structure. Therefore, these data suggest that all the different domains of UGT2B proteins are not required simultaneously for the localization of these enzymes in the ER; however, they all remain necessary for the correct folding of the protein and for transferase activity. Soon, the use of photoaffinity probes, which are analogues of UDPGA and substrates, will certainly provide important information about the folding and the interactions of the different domains of the UGT2B28 proteins.

UGT2B28 type II is coexpressed with the type I protein in the liver, mammary gland, and LNCaP cell line, whereas UGT2B28 type III was observed in the vast majority of human peripheral tissues. The role of types II and III remains unclear, since these UGT2B28 subtypes are not active on the substrates used in this study. However, Western blot analyses have demonstrated that UGT2B28 type III was capable of homodimerization (not shown), and therefore, it is possible that this isoenzyme may form dimers with other UGT proteins, as previously observed with some UGT1 and UGT2 isoforms (*63*, *64*). Furthermore, given a recent

hypothesis by Iyanagi (*63*), it is also conceivable that these shorter isoforms might be involved in cofactor transport. Western blot analyses were performed on different monkey tissues with the EL-93 antibody (*58*, *65*). Interestingly, we identified shorter UGT2B isoforms in some tissues, with molecular masses similar to those of UGT2B28 types II and III (*58*). Although it does not demonstrate the expression of the shorter UGT2B28 isoforms in human tissues, it suggests that UGT proteins, with lower molecular masses, are expressed in extrahepatic tissues.

Previous studies have shown that all human UGT2B genes are composed of six exons with a very similar intron/exon organization. The coexpression of UGT2B28 types I and II and the dramatic differences in the expression of UGT2B28 type III suggest a tissue-specific processing of these mRNA species. The exact mechanism by which this specific expression pattern occurs is unclear, but it may arise from an unconventional primary transcript maturation implicating exon skipping or alternative splicing. This explanation is highly probable in the case of the UGT2B28 type II transcript which lacks exons 4 and 5 of the corresponding *UGT2B28* gene. Therefore, a specific processing of the primary transcript, skipping the region corresponding to exons 4 and 5, would lead to the expression of the UGT2B28 type II cDNA. In the case of UGT2B28 type III, the absence of 351 nucleotides in the first part of the cDNA, which is usually encoded by a single exon, raised the possibility that this gene presents structural features different from the actual common UGT2B organization. However, the characterization of the *UGT2B28* gene revealed a genomic organization similar to those of the *UGT2B4* (*33*), *UGT2B7* (*34*, *35*), *UGT2B15* (*32*), and *UGT2B17* (*36*) genes. The presence of cryptic splicing sites inside exon 1 could constitute a mechanism by which an alternative splicing process could lead to the expression of two different mRNAs. A similar splicing mechanism is used by the UGT1 family to produce at least 12 different mRNAs by splicing the 12 specific versions of exon 1 to common exons 2−5 (*12*), demonstrating the importance of post-transcriptional events in the expression of UGT enzymes.

In summary, we report here the isolation and characterization of a novel human UGT2B steroid metabolizing enzyme. We show that the *UGT2B28* gene encodes three different mRNA species which lead to the formation of novel UGT proteins with different structural domains. Our results also indicate that UGT2B28 transcripts present a differential expression pattern in human tissues. The UGT2B28 type I enzyme was shown to be active on C19 and C18 steroids, and its expression in liver and breast tissue further supports the possibility that UGT2B28 type I is a relevant enzyme involved in steroid inactivation. Immunofluorescence studies have also shown that despite the deletion of significant regions of UGT2B28, all three subtypes of the UGT2B28 protein have the same ER and perinuclear subcellular localization.

ACKNOWLEDGMENT

We thank Dr. Pei Min Rong and Pierre-Michel Simard for technical assistance, Olivier Barbier for reticulocyte lysates, and Thierry Sabourin for helping in the specific RT-PCR tissue distribution experiment. We thank Dr. Chantal Guillemette and Dr. Dean W. Hum for their critical reading of the manuscript.

REFERENCES

1. Dutton, G. J. (1980) *Glucuronidation of drugs and other compounds*, CRC Press, Boca Raton, FL.
2. Bélanger, A., Hum, D. W., Beaulieu, M., Lévesque, É., Guillemette, C., Tchernof, A., Bélanger, G., Turgeon, D., and Dubois, S. (1998) *J. Steroid Biochem. Mol. Biol. 65*, 301−310.
3. Mauvais-Jarvis, P., Charransol, G., and Bobas-Masson, F. B. (1973) *J. Clin. Endocrinol. Metab. 36*, 452−459.
4. Morimoto, I., Edmiston, A., Hawks, D., and Horton, R. (1981) *J. Clin. Endocrinol. Metab. 52*, 772−778.
5. Moghissi, E., Ablan, F., and Horton, R. (1984) *J. Clin. Endocrinol. Metab. 59*, 417−421.
6. Brochu, M., and Bélanger, A. (1987) *J. Clin. Endocrinol. Metab. 64*, 1283−1287.
7. Bélanger, A., Brochu, M., Lacoste, D., Noel, C., Labrie, F., Dupont, A., Cusan, L., Caron, S., and Couture, J. (1991) *J. Steroid Biochem. Mol. Biol. 40*, 593−598.
8. Mahendroo, M. S., Cala, K. M., and Russell, D. W. (1996) *Mol. Endocrinol. 10*, 380−392.
9. Roy, A. K. (1992) *Proc. Soc. Exp. Biol. Med. 3*, 265−272.
10. Song, W. C., and Melner, M. H. (2000) *Endocrinology 141*, 1587−1589.
11. Mackenzie, P. I., Owens, I. S., Burchell, B., Bock, K. W., Bairoch, A., Belanger, A., Fournel Gigleux, S., Green, M., Hum, D. W., Iyanagi, T., Lancet, D., Louisot, P., Magdalou, J., Chowdhury, J. R., Ritter, J. K., Schachter, H., Tephly, T. R., Tipton, K. F., and Nebert, D. W. (1997) *Pharmacogenetics 7*, 255−269.
12. Owens, I. S., and Ritter, J. K. (1995) *Prog. Nucleic Acids Res. 51*, 306−338.
13. Ebner, T., and Burchell, B. (1993) *Drug Metab. Dispos. 21*, 50−55.
14. Hum, D. W., Bélanger, A., Lévesque, É., Barbier, O., Beaulieu, M., Albert, C., Vallée, M., Guillemette, C., Tchernoff, A., Turgeon, D., and Dubois, S. (1999) *J. Steroid Biochem. Mol. Biol. 69*, 413−423.
15. Senafi, S. B., Clarke, D. J., and Burchell, B. (1994) *Biochem. J. 303*, 233−240.
16. King, C. D., Green, M. D., Rios, G. R., Coffman, B. L., Owens, I. S., Bishop, W. P., and Tephly, T. R. (1996) *Arch. Biochem. Biophys. 332*, 92−100.
17. Green, M. D., and Tephly, T. R. (1996) *Drug Metab. Dispos. 24*, 356−363.
18. Mojarrabi, B., Butler, R., and Mackenzie, P. I. (1996) *Biochem. Biophys. Res. Commun. 225*, 785−790.
19. Lazard, D., Zupko, K., Poria, Y., Nef, P., Lazarovits, J., Horn, S., Khen, M., and Lancet, D. (1991) *Nature 349*, 790−793.
20. Lazard, D., Tal, N., Rubinstein, M., Khen, M., Lancet, D., and Zupko, K. (1990) *Biochemistry 29*, 7433−7440.
21. Jedlitschky, G., Cassidy, A. J., Sales, M., Pratt, N., and Burchell, B. (1999) *Biochem. J. 340*, 837−843.
22. Lévesque, É., Beaulieu, M., Hum, D. W., and Bélanger, A. (1999) *Pharmacogenetics 9*, 207−216.
23. Jackson, M. R., McCarthy, L. R., Harding, D., Wilson, S., Coughtrie, M. W., and Burchell, B. (1987) *Biochem. J. 242*, 581−588.
24. Ritter, J. K., Sheen, Y. Y., and Owens, I. S. (1990) *J. Biol. Chem. 265*, 7900−7906.
25. Jin, C., Miners, J. O., Lillywhite, K. J., and Mackenzie, P. I. (1993) *J. Pharmacol. Exp. Ther. 264*, 475−479.
26. Jin, C. J., Miners, J. O., Lillywhite, K. J., and Mackenzie, P. I. (1993) *Biochem. Biophys. Res. Commun. 194*, 496−503.
27. Beaulieu, M., Lévesque, E., Hum, D. W., and Bélanger, A. (1998) *Biochem. Biophys. Res. Commun. 248*, 44−50.
28. Lévesque, É., Beaulieu, M., Green, M. D., Tephly, T. R., Bélanger, A., and Hum, D. W. (1997) *Pharmacogenetics 7*, 317−325.
29. Chen, F., Ritter, J. K., Wang, M. G., Mcbride, O. W., Lubet, R. A., and Owens, I. S. (1993) *Biochemistry 32*, 10648−10657.

30. Beaulieu, M., Lévesque, É., Hum, D. W., and Bélanger, A. (1996) *J. Biol. Chem. 271*, 22855−22862.
31. Turgeon, D., Carrier, J.-S., Lévesque, É., Hum, D. W., and Bélanger, A. (2001) *Endocrinology 142*, 778−787.
32. Turgeon, D., Carrier, J.-S., Lévesque, É., Beatty, B. G., Bélanger, A., and Hum, D. W. (2000) *J. Mol. Biol. 295*, 489−504.
33. Monaghan, G., Burchell, B., and Boxer, M. (1997) *Mamm. Genome 8*, 692−694.
34. Ishii, Y., Hansen, A. J., and Mackenzie, P. I. (2000) *Mol. Pharmacol. 57*, 940−947.
35. Carrier, J.-S., Turgeon, D., Journeault, K., Hum, D. W., and Bélanger, A. (2000) *Biochem. Biophys. Res. Commun. 272*, 616−621.
36. Beaulieu, M., Lévesque, É., Tchernof, A., Beatty, B., Bélanger, A., and Hum, D. W. (1997) *DNA Cell Biol. 16*, 1143−1155.
37. Mackenzie, P. I., and Rodbourn, L. (1990) *J. Biol. Chem. 265*, 11328−11332.
38. Haque, S. J., Peterson, D. D., Nebert, D. W., and Mackenzie, P. I. (1991) *DNA Cell Biol. 10*, 515−524.
39. Mackenzie, P. I. (1986) *J. Biol. Chem. 261*, 14112−12117.
40. Beaulieu, M., Levesque, E., Barbier, O., Turgeon, D., Belanger, G., Hum, D. W., and Belanger, A. (1998) *J. Mol. Biol. 275*, 785−794.
41. Coffman, B. L., King, C. D., Rios, G. R., and Tephly, T. R. (1998) *Drug Metab. Dispos. 26*, 73−77.
42. Sanger, F., Nicklen, S., and Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. U.S.A. 74*, 5463−5467.
43. Brochu, M., and Bélanger, A. (1987) *J. Clin. Endocrinol. Metab. 60*, 294−298.
44. Bélanger, A., Brochu, M., and Cliche, J. (1984) *J. Steroid Biochem. 24*, 1069−1072.
45. Bélanger, A., Brochu, M., and Cliche, J. (1986) *J. Clin. Endocrinol. Metab. 62*, 812−815.
46. Rodbard, D., and Lewald, J. E. (1970) in *2nd Karolina Symposium on Research Methods in Reproductive Endocrinology*, pp 79−103, Bogtryklereit Forum, Copenhagen, Denmark.
47. Bélanger, A., and Labrie, F. (1990) in *Biochemistry of breast cyst fluid. Correlation with breast cancer risk* (Angeli, B., Chasalow and Dogliotti, Eds.) pp 93−100, Annals of the New York Academy of Science, New York.
48. Rhéaume, E., Sirois, I., Labrie, F., and Simard, J. (1991) *Nucleic Acids Res. 19*, 6060.
49. Fournel-Gigleux, S., Jackson, M. R., Wooster, R., and Burchell, B. (1989) *FEBS Lett. 243*, 119−122.
50. Green, M., Oturu, E. M., and Tephly, T. R. (1994) *Drug Metab. Dispos. 22*, 799−805.
51. Okuda, A., and Okuda, K. (1984) *J. Biol. Chem. 259*, 7519−7524.
52. Sugimoto, Y., Yoshida, M., and Tamaoki, B. (1990) *J. Steroid Biochem. Mol. Biol. 37*, 717−724.
53. Iyer, R. B., Binstock, J. M., Schwartz, I. S., Gordon, G. G., Weinstein, B. I., and Southren, A. L. (1990) *Steroids 55*, 495−500.
54. Cheng, Z., Radominska-Pandya, A., and Tephly, T. R. (1999) *Drug Metab. Dispos. 27*, 1165−1170.
55. Gall, W. E., Zawada, G., Mojarrabi, B., Tephly, T. R., Green, M. D., Coffman, B. L., Mackenzie, P. I., and Radominska-Pandya, A. (1999) *J. Steroid Biochem. Mol. Biol. 70*, 101−108.
56. Gingras, S., Cote, S., and Simard, J. (2000) *Mol. Endocrinol. 14*, 229−240.
57. Bélanger, A., Caron, S., Labrie, F., Naldoni, C., Dogliotti, L., and Angeli, A. (1990) *Eur. J. Cancer 26*, 277−281.
58. Bélanger, G., Beaulieu, M., Lévesque, E., Hum, D. W., and Bélanger, A. (1997) *DNA Cell Biol. 16*, 1195−1205.
59. Meech, R., Yogalingam, G., and Mackenzie, P. I. (1996) *DNA Cell Biol. 15*, 489−494.
60. Ouzzine, M., Magdalou, J., Burchell, B., and Fournel-Gigleux, S. (1999) *FEBS Lett. 454*, 187−191.
61. Meech, R., and Mackenzie, P. I. (1998) *Arch. Biochem. Biophys. 356*, 77−85.
62. Ouzzine, M., Magdalou, J., Burchell, B., and Fournel-Gigleux, S. (1999) *J. Biol. Chem. 274*, 31401−31409.
63. Ikushiro, S., Emi, Y., and Iyanagi, T. (1997) *Biochemistry 36*, 7154−7161.
64. Meech, R., and Mackenzie, P. I. (1997) *J. Biol. Chem. 272*, 26913−26917.
65. Albert, C., Barbier, O., Vallee, M., Beaudry, G., Belanger, A., and Hum, D. W. (2000) *Endocrinology 141*, 2472−2480.

BI002607Y