0026-895X/97/020212-09$3.00/0

MOLECULAR PHARMACOLOGY, **52:**212–220 (1997).

Differential Expression of the *UGT1A* Locus in Human Liver, Biliary, and Gastric Tissue: Identification of *UGT1A7* and *UGT1A10* Transcripts in Extrahepatic Tissue

CHRISTIAN P. STRASSBURG, KARL OLDHAFER, MICHAEL P. MANNS, and ROBERT H. TUKEY

University of California, San Diego, Department of Pharmacology, Cancer Center, La Jolla, California 92093-0636 (C.P.S., R.H.T.), and Departments of Abdominal and Transplant Surgery (K.O.) and Gastroenterology and Hepatology (M.P.M.), Medizinische Hochschule Hannover, 30625 Hannover, Germany

Received March 19, 1997; Accepted April 21, 1997

SUMMARY

Family 1 UDP-glucuronosyltransferases (UGTs) (UGT1A) are encoded by a locus that predicts the existence of at least nine individual proteins. The different proteins are generated by exon-sharing, which results in the production of a family of proteins that contain identical, 245-amino acid, carboxyl-terminal domains and an amino-terminal region of approximately 280 amino acids. The diversity of the *UGT1A* locus suggests the existence of complex regulation, most likely designed to account for the variable and specific glucuronidation requirements. However, the tissue-specific and extrahepatic regulation of the complete *UGT1A* locus has not been defined to date. In this study, quantitative duplex reverse transcription-polymerase chain reaction was used to analyze UGT1A RNA expression in 16 hepatic, four biliary, and two gastric human tissue specimens. UGT1A3 and UGT1A6 were found to be expressed in the three tissues, whereas UGT1A5 and UGT1A8 were not expressed. Hepatocellular and biliary tissue expressed UGT1A1 and UGT1A4 but hepatocellular tissue uniquely expressed UGT1A9, whereas biliary tissue expressed UGT1A10. In contrast to hepatocellular tissue, gastric tissue expressed UGT1A7 in addition to UGT1A10. The expression of UGT1A9 in hepatic tissue, UGT1A7 in gastric tissue, and UGT1A10 in biliary and gastric tissue provides evidence for the selective regulation of the *UGT1A* locus in hepatic and extrahepatic tissues. The newly identified UGT1A7 and UGT1A10 transcripts were cloned and found to be 95.86% identical. Sequence analysis confirmed two proteins with divergent amino termini of 285 residues and identical carboxyl termini of 245 residues. This study provides evidence for hepatic and extrahepatic regulation of the human *UGT1A* locus and identifies two novel extrahepatic transcripts of the UGT1A family.

The UGTs are a family of proteins localized to the endoplasmic reticulum. These proteins conjugate, through glucuronidation, a host of compounds, including drugs, endogenous metabolites such as steroids, bilirubin, and bile acids, and the catabolic end products of dietary ingredients. Glucuronide metabolites are rendered more hydrophilic and can therefore be eliminated by excretion in bile or urine (1). To date, >30 individual transferase gene products have been identified from different species and classified into two families, termed UGT1 and UGT2. The selection and identification of these two families are based on their predicted amino acid sequence similarities (2, 3). The UGT2 proteins most likely evolved by gene duplication events, with the structural genes being individually encoded on chromosome 4 (4). In contrast to members of the UGT2 family, UGT1A gene products encode proteins that have identical carboxyl-terminal domains of 245 amino acids (5, 6) but display considerable divergence in the amino-terminal 280 amino acids. The UGT1A proteins are encoded on chromosome 2 (7, 8), where at least 12 divergent first exons[1] located at the 5′ end of the

A portion of this work was presented at the 7th International Workshop on Glucuronidation and the Glucuronosyltransferases, University of Iowa, Iowa City, IA, May 19–22, 1996. This work was supported in part by Deutsche Forschungsgemeinschaft Grant Str493/2–1 (C.P.S.) and United States Public Health Service Grant GM49135 (R.H.T.).

[1] A nomenclature system for the UGT1A proteins is based on recommendations made at the 7th International Workshop on Glucuronidation and the Glucuronosyltransferases, University of Iowa, Iowa City, IA, May 19–22, 1996. The identification of the locus has been described, with the flanking exon 1 sequences assigned in alphabetical order. For example, the first six common functional exon sequences were originally designated UGT1A, UGT1BP, UGT1C, UGT1D, UGT1E, and UGT1F. The nomenclature that is recommended identifies the proteins as UGT1A1, UGT1A3, UGT1A4, UGT1A5, and UGT1A6, respectively (9). *UGT1BP* does not contain an open reading frame and is felt to be a pseudogene. Additional exon sequences have recently been identified and deposited in GenBank (Ida Owens, National Institutes of Health, personal communication); those proteins would in turn be classified based on their exon locations within the locus. In this manuscript, we classify the respective RNAs in the same way as in the recommended classification of the proteins. To date, 12 flanking exon 1 sequences have been identified, nine of which could encode functional proteins. These RNAs and the proteins that

ABBREVIATIONS: UGT, UDP-glucuronosyltransferase; DRT, duplex reverse transcription; PCR, polymerase chain reaction; bp, base pair(s); dNTP, deoxynucleoside triphosphate; RT, reverse transcription.

locus are selectively spliced with four constant exons (exon 2–5) to form unique RNA transcripts (9). Nucleic acid sequence analysis indicates that UGT1A2 (9), UGT1A11, and UGT1A12 (GenBank sequences U39951 and U35992, respectively, for the latter two) encode pseudogenes (10). Each first exon is flanked by polymerase II recognition sequences, suggesting that each gene may be selectively regulated.

Corresponding cDNAs from the *UGT1A* locus that encode UGT1A1 (5), UGT1A3 (11), UGT1A4 (5, 12), UGT1A6 (13), and UGT1A9 (6) have been identified from human liver cDNA libraries, because the liver is generally regarded as the primary site of human glucuronidation. The expression of these cDNAs has demonstrated that the UGTs display a broad range of substrate specificity, which indicates that each UGT has evolved to broaden the ability of this gene family to recognize the many different, structurally unique agents that are targeted for glucuronidation. It appears that a complex mode of regulation is required for control of this closely related family of proteins. However, individual regulation of the unique UGT1A transcripts has not been conclusively demonstrated to date.

There is approximately 93–94% nucleic acid sequence homology between the first exons that encode UGT1A3 to UGT1A5, as well as between the first exons that encode UGT1A7 to UGT1A10. The first exon sequences that encode UGT1A7–1A10 have diverged from UGT1A3–1A5 and display approximately 50% identity, whereas the first exon that encodes UGT1A1 displays the highest degree of similarity to UGT1A3–1A5, being 58% similar to UGT1A4. The exon encoding UGT1A6 is most similar to those of UGT1A7 to UGT1A10, sharing 54% nucleic acid sequence identity with UGT1A9. As a result of the high degree of nucleic acid sequence identity between the first exon sequences that encode the UGT1A3–1A5 and UGT1A7–1A10 sequences, identification of unique gene expression requires the use of methods that can distinguish the minor differences in nucleic acid sequences. An example of this strategy has been the use of oligonucleotides in hybridization experiments, and this has led to the observation that UGT1A3 and UGT1A4 but not UGT1A5 are expressed in human liver (9). However, the precise hepatic and extrahepatic tissue distribution and expression of the entire *UGT1A* locus in human tissues have not been examined. In particular, the question of whether UGT1A transcripts that are not detectable in liver tissue are regulated in other human tissues has not been addressed.

In the present study, DRT-PCR has been used to measure transcripts of all known members of the *UGT1A* locus in human hepatic, biliary, and gastric tissue. These data provide evidence for a differential and complex regulation of the locus and demonstrate that extrahepatic *UGT1A* regulation plays an important role in the selective expression of UGT1A transcripts.

Materials and Methods

Tissue procurement. Liver tissue samples were collected from patients undergoing hemihepatectomy or liver transplantation (n = 16) and stomach tissue samples were collected from patients undergoing gastrectomy (n = 2) at the University of Hannover Medical Center (Hannover, Germany). Healthy tissue that displayed no visible signs of necrosis was obtained from the resection specimen and immediately snap-frozen in liquid nitrogen. Biliary epithelium was obtained by dissection of the mucosa of healthy gallbladders (n = 4) and immediately snap-frozen in liquid nitrogen. All samples were continuously stored at −80° until analysis.

RNA isolation. Approximately 200 mg of frozen tissue were pulverized in liquid nitrogen, and the frozen tissue powder was immediately lysed in acidic phenol/guanidinium isothiocyanate solution (TriPure; Boehringer Mannheim, Mannheim, Germany). Before lysis, samples were not allowed to thaw at any point of the procedure. The mixture was homogenized by vortex-mixing for 1 min, and the RNA was extracted following the protocol of Chomczynski (14). Briefly, the cell mixture of 1.2 ml containing acidic phenol/guanidinium isothiocyanate was separated into aqueous and organic phases by the addition of 0.2 ml of chloroform and subsequent centrifugation at 12,000 × g. The aqueous phase was collected and the RNA was precipitated by the addition of 1 volume of 2-propanol and centrifugation in an Eppendorf microcentrifuge. Pellets were washed twice with 75% ethanol and resuspended in diethylpyrocarbonate-treated water. RNA was quantified and purity was monitored by spectrophotometry at 260 and 280 nm. To monitor for possible DNA contamination, PCR amplification of β-actin cDNA was performed using primers that span the exon 4/intron 5/exon 5 junction. Genomic DNA template leads to a 312-bp product and the cDNA template generates a 202-bp product, which was observed for all isolated RNA samples. RNA samples were stored at −80° until analysis.

Oligonucleotide primers for DRT-PCR. Seventeen primers for the amplification of the human UGT RNA transcripts as well as β-actin sequences were generated, by using automated phosphoamidite chemistry, at the University of California, San Diego, Cancer Center Molecular Biology Core Facility. The UGT primers were either specific for regions in each of the unique exon 1 sequences or specific for sequences that were common to all of the UGT1A RNAs. The primers used in this study are listed below, with the numbers for each used throughout the manuscript. All of the sequences have been deposited in GenBank, and their base positions are indicated based on the sequence listed by the accession number, as follows: 1) β-actin sense primer (accession number M10278, bases 942–962), 5′-ggcggcaccaccatgtaccct-3′; 2) β-actin antisense primer (accession number M10278, bases 1123–1143), 5′-aggggccggactcgtcatact-3; 3) common UGT-1 sense primer (accession number M57899, bases 1042–1062), 5′- tcgaatcttgcgaacaacacg-3′; 4) common UGT-1 antisense primer (accession number M57899, bases 1529–1509), 5′-atgaaggccactgtcagcacg-3′; 5) common exon 5 antisense primer (accession number M57899, bases 1056–1036), 5′-gttcgcaagattcgatggtcg-3′; 6) common exon 1 antisense primer (accession number M57899, bases 842–822), 5′-ccaatgaagaccatgttgggc-3′; 7) exon 1A1 sense primer (accession number M57899, bases 412–432), 5′-aacaaggagctcatggcctcc-3′; 8) exon 1A3 sense primer (accession number M84127, bases 477–501), 5′-tgttgaacaatatgtctttggtcta-3′; 9) exon 1A4 sense primer (accession number M57951, bases 287–307), 5′-gaaggaatttgatcgcgttac-3′; 10) exon 1A5 sense primer (accession number M84129, bases 285–303), 5′-ggtggtggtcctcaccctg-3′; 11) exon 1A6 sense primer (accession number M84130, bases 686–706), 5′-tcctggctgagtatttgggcc-3′; 12) exon 1A7 sense primer (accession number U39570, bases 443–461), 5′-tgccgatgctcgctggacg-3′; 13) exon 1A8 sense primer (accession number U42604, bases 761–781), 5′-ggtcttcgccaggggaatagg-3′; 14) exon 1A9 sense primer (accession number S55985, bases 676–699), 5′-ggaggaacatttattatgccaccg-3′; 15) exon 1A10 sense primer (accession number U39550, bases 822–844), 5′-cctctttcctatgtccccaatga-3′; 16) exons 1A3–5 common sense primer (accession number M57951, bases 667–687), 5′-acatgctctacccctctggccc-3′; 17) exons 1A7–10 common sense primer (accession number S55985, bases 578–598), 5′-cagtgccctgctcctctttcc-3′.

A schematic diagram of the position of each primer relative to the coding region of the RNA and the expected RT-PCR product sizes is shown in Fig. 1. Primers for human β-actin (primers 1 and 2) have been previously published (15). Antisense primers 4, 5, and 6 anneal

they encode are designated UGT1A1 through UGT1A10, corresponding to their respective locations in the *UGT1A* locus.

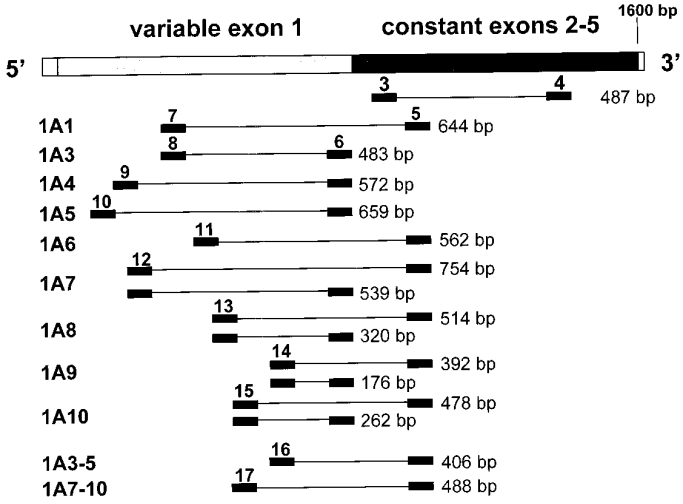

Fig. 1. Structure of family 1 UGTs and localization of PCR primers for all nonpseudogene isoforms. Isoform-specific sense primers are located within the variable first exons and lead to transcript sizes of different sizes. The two antisense primers are located either within an area common to all first exons or within the 5′ portion of the common 3′ sequence (exons 2–5). Note that one primer pair is located entirely within the constant domain and therefore could lead to the amplification of all UGT1A transcripts.

to all of the UGT1A transcripts. Sense primer 3 and antisense primer 4 generate a 487-bp transcript that is representative of all of the UGT1A transcripts. Primers 16 and 17 anneal to conserved sequences of UGT1A3, UGT1A4, and UGT1A5 and UGT1A7, UGT1A8, UGT1A9, and UGT1A10, respectively. Primers 7–15 (Fig. 1) anneal to unique sequences located within the divergent first exons of the UGT RNA. A combination of 5′ exon 1 sense primers and primer 6 allows for the amplification of genomic DNA and cDNA sequences, whereas the combination of the 5′ exon 1 sense primers and primer 5 allows for amplification of only cDNA sequences. Nonspecific primer binding was excluded by computerized databank search analysis using the GenBank Blastn software. Primers were designed with melting temperature values ranging from 51° to 63°.

Northern blot analysis. Northern blot analysis was performed with denaturing agarose gels containing 7% formaldehyde (16). Briefly, 15 μg of total hepatic or biliary RNA were electrophoresed in a 1% agarose gel, containing 40 mM 3-(*N*-morpholino)propanesulfonic acid, pH 7.0, 10 mM sodium acetate, 1 mM EDTA, and 7% formaldehyde, before transfer to nitrocellulose membranes (Immobilon-NC; Millipore, Bedford, MA). A cDNA restriction fragment from bases 1042 to 1467 was isolated from the UGT1A1 cDNA (5) (kindly provided by Dr. Ida Owens, National Institutes of Health) constant region by digestion with *Taq*I/*Kpn*I, and the fragment was gel-purified (Qiagen, Hilden, Germany). The cDNA was then labeled with ^{32}P by nick translation, and approximately 10^8 cpm/ml was used in hybridization. Filters were washed with 0.2× standard saline citrate/0.1% sodium dodecyl sulfate at room temperature and 0.1× standard saline citrate/0.1% sodium dodecyl sulfate at 42° (1× standard saline citrate = 15 mM sodium citrate, pH 7.0, containing 150 mM NaCl). Dried filters were exposed to X-ray film at −80°.

Southern blot analysis. PCR products were separated by electrophoresis in an 1% agarose gel. The gel was incubated for 20 min in a buffer containing 0.5 M NaOH and 1.5 M NaCl and was then neutralized for 45 min in a buffer containing 1.5 M NaCl and 1 M Tris·HCl, pH 7.4, after which the DNA was transferred to nitrocellulose filters. A 721-bp UGT1A1 *Xho*I/*Eco*RI fragment of exon 1 was excised from cDNA and column-purified. For detection of UGT1A6, a 790-bp *Eco*RI/*Pvu*II fragment of exon 1 was excised from the complete UGT1A6 cDNA (13) (kindly provided by Dr. Brian Burchell, Ninewells Hospital and Medical School, Dundee, Scotland). Fragments were labeled by nick-translation. Hybridization of the probes to Southern blots was carried out as described for Northern blot analysis.

RT. Three micrograms of total RNA were denatured in the presence of 0.5 μg of oligo(dT) primer at 70° for 10 min and quick-chilled on ice. In a volume of 19 μl, containing 20 mM Tris·HCl, pH 8.4, 50 mM KCl, 2.5 mM $MgCl_2$, 10 mM dithiothreitol, and 0.5 mM levels of each dNTP, the RNA was incubated at 42° for 5 min before the addition of 200 units of reverse transcriptase (SuperscriptII; Gibco BRL, Gaithersburg, MD). The final volume of 20 μl was incubated at 42° for 50 min and at 70° for 15 min and chilled on ice.

DRT-PCR for UGT1A1, UGT1A6, UGT1A3–1A5, and UGT1A7–1A10. The cDNA mixture, in a volume of 96 μl, contained 3 mM $MgCl_2$, 50 mM KCl, 20 mM Tris·HCl, pH 8.4, 0.2 mM levels of each dNTP, 2 μM primer 1, and 2 μM primer 2. Five units of VENT (exo$^-$) DNA polymerase (NEB, Beverly, MA) were incubated at 94° for 3 min, followed by six cycles of 94° for 1 min, 59° for 1 min, and then 72° for 1 min. After six cycles, human β-actin primers in 4 μl were added to achieve a final concentration of 0.4 μM, bringing the final reaction volume to 100 μl. Cycling was continued to a total of 32 cycles, followed by a 7-min incubation at 72°.

DRT-PCR for UGT1A3, UGT1A4, and UGT1A5. Buffer conditions were identical to those described for UGT1A1, UGT1A6, UGT1A3–1A5, and UGT1A7–1A10. The reaction mixture was incubated at 94° for 3 min, followed by six cycles of 94° for 1 min, 56° for 1 min, and then 72° for 1 min. Human β-actin primers were added to a concentration of 0.4 mM, and cycling was continued for a total of 32 cycles, followed by 7 min of incubation at 72°. PCR with a human genomic DNA template was performed using 1 μg of human genomic lymphocytic DNA and the aforementioned conditions for 35 cycles without the addition of β-actin primers.

DRT-PCR for UGT1A7, UGT1A8, UGT1A9, and UGT1A10. Each reaction was performed in a volume of 96 μl, containing the cDNA mixture and 2.38 mM $MgCl_2$, 50 mM KCl, 20 mM Tris·HCl, pH 8.4, 0.2 mM levels of each dNTP, 2 μM primer 1, and 2 μM primer 2. Five units of *Taq* DNA polymerase (Promega, Madison, WI) were added and the reactions were initially incubated at 94° for 3 min, followed by six cycles of 94° for 1 min, 64° for 1 min, and then 72° for 1 min. Human β-actin primers were added to a concentration of 0.4 mM and a final reaction volume of 100 μl. Cycling was continued for a total of 32 cycles, followed by a 7-min incubation at 72°. PCR with a human genomic DNA template was performed using 1 μg of human genomic lymphocytic DNA and the aforementioned conditions for 35 cycles without the addition of β-actin primers. Each DRT-PCR was performed in duplicate in a Perkin Elmer 480 thermocycler. Control reactions without template and without DNA polymerase were included throughout the analysis.

All primer pairs used are summarized in Fig. 1. All PCR products (UGT1A1, UGT1A3, UGT1A4, UGT1A6, UGT1A7, UGT1A9, and UGT1A10) were purified and analyzed by dideoxy sequencing to confirm the expected individual exon 1 sequences.

Quantification of PCR products. Ten microliters of each PCR were resolved by electrophoresis in a 2% agarose gel containing 1 μg/ml ethidium bromide. Gels were photographed using Polaroid type 665 positive/negative film (Polaroid, Cambridge, MA). Negatives were used for laser quantification in an LKB 2222–020 UltroScan XL densitometer (Pharmacia LKB, Bromma, Sweden). The area beneath the graph of each peak was determined, and means were calculated from duplicate samples. Peaks for human β-actin were used as internal standards for each RNA sample and were used to determine the relative UGT band intensity, according to the following formula: (mean peak area for UGT/mean peak area for β-actin) × 100 = relative arbitrary units. Data were analyzed and expressed with Student's *t* test, using Microsoft Excel software.

Cloning and analysis of the UGT1A10 transcript. Total biliary RNA was used to synthesize cDNA as described above. The UGT1A10 transcript was cloned by PCR amplification. The reaction volume for the PCR was 100 μl of 20 mM Tris·HCl, pH 8.4, and

contained 5 μl of the cDNA mixture, 2.75 mM $MgCl_2$, 50 mM KCl, 0.2 mM levels of each dNTP, 0.4 mM levels of the sense primer 5′-tcgctctagatgatggctccgc-3′, 0.4 mM levels of the antisense primer 5′-attctcgagccacctctcaatgg-3′, and 5 units of VENT (exo$^-$) DNA polymerase. Cycling was carried out for 35 rounds of 94° for 1 min followed by 58° for 1 min and then 72° for 2 min, with a final incubation of 7 min at 72°. The sense primer was designed to incorporate an *Xba*I restriction endonuclease cleavage site 5′ to the initiation methionine ATG codon, whereas the antisense primer contained an *Xho*I site 3′ to the stop TGA codon. The amplified product of approximately 1600 bp was digested with *Xba*I and *Xho*I, and the purified insert was cloned into the *Xba*I/*Xho*I sites of the KS+ pBluescript vector (Stratagene, La Jolla, CA). Plasmids transformed into *Escherichia coli* JM109 cells were screened for inserts through α-complementation. The sequence of the insert was determined by dideoxy sequence analysis.

Cloning and analysis of the UGT1A7 transcript. Total gastric RNA was used to synthesize cDNA as described above. The UGT1A7 transcript was amplified by PCR in a volume of 100 μl, under the same conditions as for the cloning of UGT1A10 cDNA. The sense primer 5′-ctgggatccagttctctgatggctcgtgc-3′ incorporated a *Bam*HI restriction endonuclease site upstream of the ATG start codon; the antisense primer 5′-cacaagcttctcaatgggtcttggatttgtgg-3′ was designed to include a *Hin*dIII restriction endonuclease site downstream of the TGA stop codon. The product of approximately 1600 bp was digested with *Bam*HI and *Hin*dIII and directionally cloned into a *Bam*HI- and *Hin*dIII-digested KS+ pBluescript vector. The sequence was determined by dideoxy sequence analysis, using specific UGT1A7 oligonucleotide primers.

Results

Overall expression of the *UGT1A* locus in hepatic and biliary tissue. Northern blot analysis using a conserved UGT1A 3′ portion of the UGT1A1 cDNA demonstrated that RNA transcripts of approximately 2.6 kilobases were generated in both hepatic and biliary tissue. The overall level of expression was greater in hepatic than biliary tissue. A similar result was evident when DRT-PCR was performed using primers 3 and 4, which recognized the conserved region of the expressed UGT1A RNAs (Fig. 2). Although there were significant differences in the overall expression levels of the *UGT1A* locus in hepatic and biliary tissue, experiments were designed to selectively identify the transcripts that were expressed in these tissues.

In the first series of experiments, a DRT-PCR assay was developed that could be used to quantitate the different UGT1A transcripts in hepatic and biliary tissue. Because multiple RNA samples were to be evaluated for expression of the different UGTs, kinetic profiles for the DRT-PCR assays were first determined, to ensure linearity in the synthesis of both the UGT1A and β-actin transcripts. UGT1A quantitative DRT-PCR with primers 3 and 4 as well as actin primers 1 and 2 resulted in the expected bands of 487 bp and 202 bp. When the PCR was terminated at 23, 25, 27, 29, 31, and 35 cycles and band intensities were quantified by laser densitometry, amplification of both cDNAs was found to be linear between 23 and 35 cycles (Fig. 3). Based upon this result, subsequent analysis was therefore conducted at 32 cycles.

Specific analysis of UGT1A1 and UGT1A6. The nucleic acid sequence in exon 1 of UGT1A1 and UGT1A6 is divergent from the exon 1 sequence of the other UGT1A RNAs. To ensure that the UGT1A1 and UGT1A6 primers would specifically amplify these RNA transcripts, DRT-PCR analysis was conducted using the UGT1A1 and UGT1A6 primers, as well as conserved primers that amplified UGT1A3 to UGT1A5 (primers 16 and 5) and UGT1A7 to UGT1A10 (primers 17 and 5) transcripts. After resolution of the PCR products in 2% agarose, Southern blot analysis was conducted using labeled cDNA fragments that encoded the 5′ divergent region of UGT1A1 and UGT1A6. Specific amplifi-

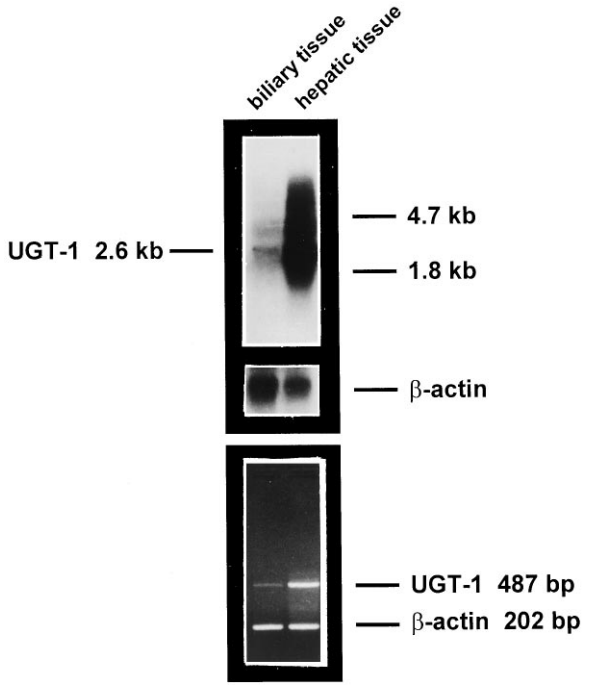

Fig. 2. Northern blot (*top*) and DRT-PCR (*bottom*) analysis of UGT1A expression in human biliary (*left lane*) and hepatic (*right lane*) tissue. Hybridization with a ^{32}P-labeled *Taq*I/*Kpn*I cDNA fragment of the UGT1A1 constant region identified a 2.6-kilobase band that was expressed in both hepatic and biliary tissue. This finding was reproduced by DRT-PCR coamplification of the constant UGT1A domain (487 bp) in the presence of human β-actin (202 bp).

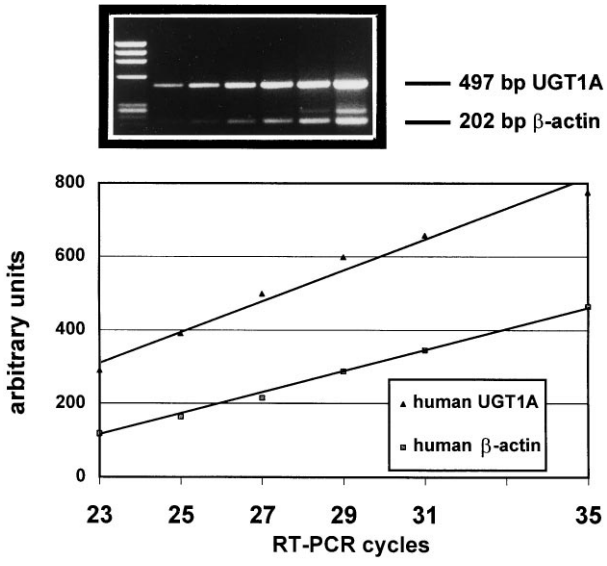

Fig. 3. Kinetics of DRT-PCR amplification of UGT1A transcripts and human β-actin. *Top*, coamplification was stopped after 23, 25, 27, 29, 31, or 35 cycles, and the products were separated in a 2% agarose gel containing ethidium bromide. *Bottom,* the laser densitometric quantification of the bands from the agarose gel is shown. Amplification of UGT and β-actin was linear between cycles 23 and 35.

cation of UGT1A1 and UGT1A6 resulted in transcript sizes of 644 bp and 562 bp, respectively, whereas transcripts specific for UGT1A3–1A5 and UGT1A7–1A10 were generated that were 488 and 406 bp, respectively (Fig. 4). Southern blot analysis probed with an exon 1-specific UGT1A1 cDNA annealed only to the 644-bp UGT1A1 transcript, demonstrating that UGT1A3–1A5 or UGT1A7–1A10 gene products could not be recognized by UGT1A1 exon 1 sequences. In a similar experiment, hybridization with the UGT1A6 exon 1 cDNA annealed to only UGT1A6 transcripts. These experiments demonstrated that the UGT1A1 and UGT1A6 primers could be used to selectively identify the specific expression of these gene products in liver.

Analysis of UGT1A3 to UGT1A5 and UGT1A7 to UGT1A10. To examine the specificity of the primers generated toward these UGT1A transcripts, PCR analysis was performed with each pair of primers using human lymphocyte DNA in parallel with RT-PCR. As shown in Fig. 5, *UGT1A3*, *UGT1A4*, and *UGT1A5* genomic sequences were amplified using lymphocyte DNA. When the same primers were used with liver cDNA, only UGT1A3 and UGT1A4 transcripts were identified. A similar result was observed when these primers were coamplified with β-actin primers. Because the UGT1A5 primers amplified only genomic DNA, this result demonstrates that these primers are specific for UGT1A5 sequences. In a separate experiment, the 572-bp UGT1A4 RT-PCR product was gel-purified and used as the template in a PCR that included both the UGT1A4 and UGT1A3 primers. In this experiment, only the 572-bp UGT1A4 transcript was reamplified (data not shown), demonstrating that the UGT1A4 and UGT1A3 primers could also be used to selectively identify these gene products.

In a similar series of experiments, amplification of *UGT1A7*, *UGT1A8*, *UGT1A9*, and *UGT1A10*, using genomic DNA, resulted in the banding patterns that were predicted from the location of the primer sequences in each exon, as shown in Fig. 5B. DRT-PCR analysis with these same primers demonstrated that only UGT1A9 was expressed in human liver. The detection of UGT1A9 transcripts was also demonstrated when the cDNA was coamplified with human β-actin primers (Fig. 5B, *lanes 7* and *11*). When PCR products

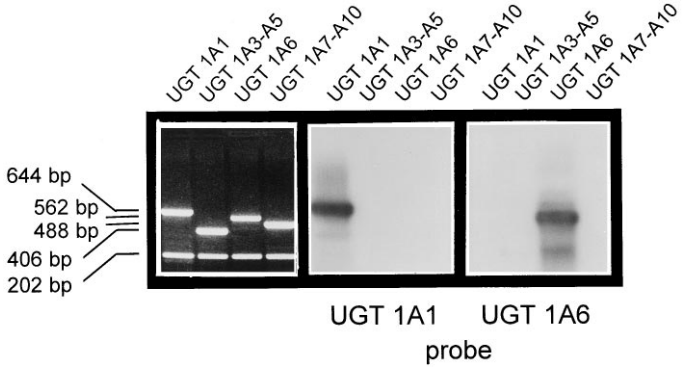

Fig. 4. Southern blot analysis. DRT-PCR amplification products of UGT1A1, of sequences common to the UGT1A3–1A5 cluster, of UGT1A6, and of sequences common to the UGT1A7–1A10 cluster were blotted and probed with ^{32}P-labeled exon 1 fragments of UGT1A1 (*Xho*I/*Eco*RI) and UGT1A6 (*Eco*RI/*Pvu*II). *Left*, DRT-PCR products of expected sizes separated in a 2% ethidium bromide-stained gel. *Middle* and *right,* hybridization patterns of UGT1A1 and UGT1A6 cDNA labels with the respective DRT-PCR amplicons.

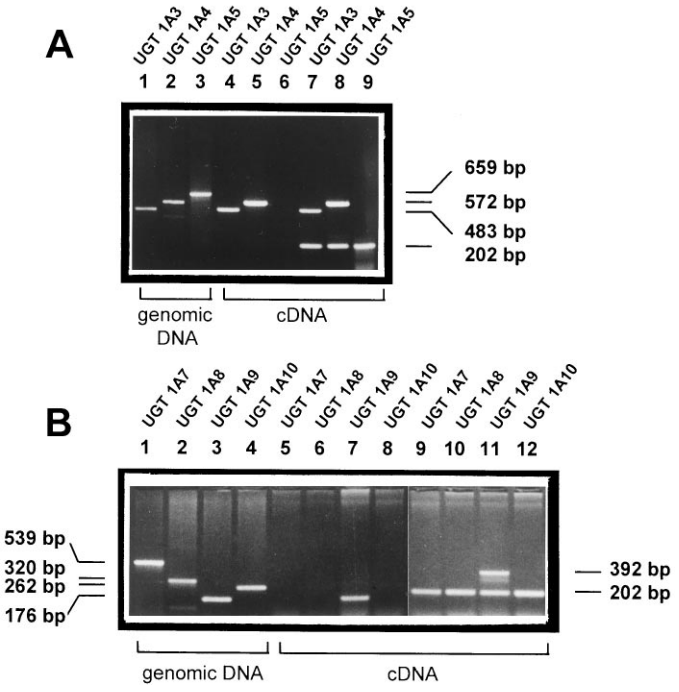

Fig. 5. A, Specificity of UGT1A primers. PCR was performed with human lymphocytic genomic DNA, which resulted in transcripts of 483 bp, 572 bp, and 659 bp (*lanes 1-3*). RT-PCR (*lanes 4-6*) and DRT-PCR coamplification with β-actin (*lanes 7-9*) using human hepatic cDNA template demonstrated transcripts of UGT1A3 and UGT1A4 but not UGT1A5. B, Specificity and function of the UGT1A7, UGT1A8, UGT1A9, and UGT1A10 primers. PCR was performed with human lymphocytic genomic DNA, which resulted in transcripts of the expected sizes of 539 bp, 320 bp, 176 bp, and 262 bp (*lanes 1-4*). RT-PCR using human hepatic cDNA template demonstrated amplification of only the UGT1A9 176-bp band (*lanes 5-8*). DRT-PCR was performed using the same sense primers as in *lanes 1-8* and an antisense primer located 179 bp further downstream within the UGT1A constant domain, for ease of band discrimination from human β-actin (202 bp, *lanes 9-12).*

generated from genomic DNA with the UGT1A7, UGT1A8, or UGT1A10 primers were purified and used as templates, no product was generated by PCR when UGT1A9 primers were used (data not shown). Confirmation that the amplified DRT-PCR product encoded the respective gene products was further established by direct dideoxy sequence analysis. Together, these experiments demonstrate that the selection of primers to identify gene products encoded by the *UGT1A* locus could be used in DRT-PCR to selectively amplify transcripts expressed in human tissues.

Differential regulation of UGT1A4, UGT1A7, UGT1A9, and UGT1A10 in human hepatic, biliary, and gastric tissue. DRT-PCR to detect UGT1A1 and UGT1A3–1A10 transcripts was performed using the 16 hepatic, four biliary, and two gastric tissue RNA samples. An example of these results is shown in Fig. 6. UGT1A1 was expressed in none of the two gastric samples, three of the four biliary samples, and all 16 hepatocellular samples. RNA transcripts encoding UGT1A3, UGT1A4, and UGT1A6 were expressed in all hepatocellular and biliary tissue samples, whereas UGT1A5, UGT1A7, and UGT1A8 were absent. In contrast, both gastric tissue samples expressed UGT1A3 and UGT1A6 but not UGT1A4. Gastric tissue additionally expressed UGT1A7, a transcript found in neither of the other samples

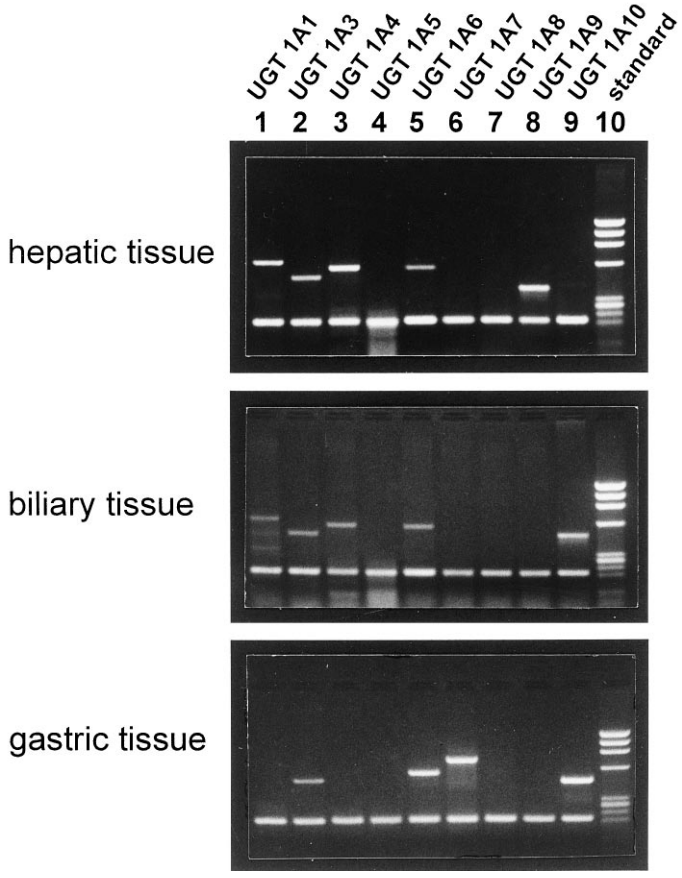

Fig. 6. Tissue-specific hepatic and extrahepatic expression of the human *UGT1A* locus. DRT-PCR analysis of all predicted UGT1A transcripts in human hepatic, biliary, and gastric tissues was performed. *Top*, ethidium bromide-stained 2% agarose gel demonstrating the typical UGT1A repertoire found in all 16 human hepatic tissue samples. UGT1A1, UGT1A3, UGT1A4, UGT1A6, and UGT1A9 are expressed; UGT1A5, UGT1A7, UGT1A8, and UGT1A10 are not. *Middle* and *bottom*, typical UGT1A repertoire of human biliary and gastric tissue samples, respectively. Note the exclusive expression of UGT1A9 in hepatic tissue and of UGT1A7 in gastric tissue. UGT1A10 transcripts are specifically localized to the nonhepatocellular tissue types.

investigated. Whereas UGT1A9 was expressed in all hepatocellular samples, there was no expression of this gene product in the four biliary samples or the two gastric samples. However, UGT1A10 was expressed in the four biliary tissues and the two gastric tissues but not observed in any of the 16 hepatocellular samples. UGT1A4, UGT1A7, UGT1A9, and UGT1A10 appear to be selectively regulated in human hepatic, biliary, and gastric tissue.

The quantitative DRT-PCR also allowed a comparison of UGT1A gene expression in hepatocellular and biliary tissues. A statistical comparison of the two gastric samples was not performed because of the small sample size. Quantification was accomplished by measuring the levels of each transcript relative to that of β-actin. The statistical analysis of *UGT1A* gene expression is presented in Fig. 7. The overall expression of *UGT1A* in hepatic tissue was significantly more abundant than in biliary tissue, but this was not uniformly observed for all UGT1A transcripts. UGT1A3, UGT1A4, and UGT1A6 were all expressed less abundantly in biliary than in hepatic tissue; however, the expression levels of UGT1A1 were not significantly different. Moreover, the relative differences in RNA expression between hepatic and biliary tissues were similar for UGT1A3 and UGT1A4, but the expression levels of UGT1A6 between the two tissues differed from those observed for UGT1A3 and UGT1A4. This indicates that these genes, although expressed in the same tissues, are potentially regulated differentially. The overall UGT1A expression levels in gastric tissue appear to be similar to those found in hepatic tissue, as is evident in the sample in Fig. 6.

Characterization of the novel biliary tissue-derived human UGT1A10 and gastric tissue-derived human UGT1A7 transcripts. A DRT-PCR was performed using primers that were designed to amplify the open reading frame of the UGT1A7 and UGT1A10 transcripts, as outlined in Materials and Methods. A 1613-bp UGT1A7 cDNA and a 1611-bp UGT1A10 cDNA were amplified, cloned, and characterized by DNA sequence analysis. The nucleic acid sequences of UGT1A7 and UGT1A10 were 95.86% identical, with all differences being located within the first 847 residues of the cDNA. The 746 nucleic acids at the 3′end of the cDNA were identical in the two clones. Translation of both sequences revealed an open reading frame of 1593 nucleic acids. The UGT1A7 and UGT1A10 proteins differ with respect to 53 amino acid residues in the amino-terminal portion, as shown in Fig. 8. As observed for all previously cloned UGT1A proteins, UGT1A7 and UGT1A10 share an identical carboxyl terminus of 245 amino acids, representing exons 2–5, and a variable amino terminus of 285 amino acids, representing the variable first exon. These data suggest that the identified clones represent two novel human UGT1A transcripts.

The UGT1A7 and UGT1A10 proteins display the most striking differences between amino acids 3 and 24. In this region none of the residues appear to be conserved (Fig. 8). In most instances, approximately 24 amino-terminal amino acids encode a signal sequence. Interestingly, analysis of the amino-terminal region for signal sequence residues (17) indicates that UGT1A10 does not contain a hydrophobic leader sequence, which could suggest that the amino-terminal region of UGT1A10 may not be removed as predicted for the other UGTs. In contrast, UGT1A7 exhibits a normal leader sequence region with a predicted cleavage site after amino acid 24 (Fig. 8, *box*).

Discussion

UGTs play an important role in human metabolism. Homeostasis depends upon the availability of functional detoxification systems capable of eliminating lipophilic compounds such as endogenous metabolic intermediates, steroid hormones, drugs, and xenobiotic substances. The human UGT1A cDNAs encoding UGT1A1, UGT1A3, UGT1A4, UGT1A6, and UGT1A9 have been cloned and the expressed proteins studied for substrate specificities. These UGTs have been shown to display a wide variety of substrate selectivities, including the ability to conjugate small and large planar phenols, steroids such as estriol and estrone, tertiary and quarternary amines, benz(α)pyrenes, 2-acetylaminofluorene metabolites, and the heme synthetic byproduct bilirubin (11, 15, 18, 19). The diverse nature of substrate specificity elicited by the UGT1A family of proteins may have evolved to meet complex requirements of glucuronidation in various organs and during exposure to an array of different compounds. This

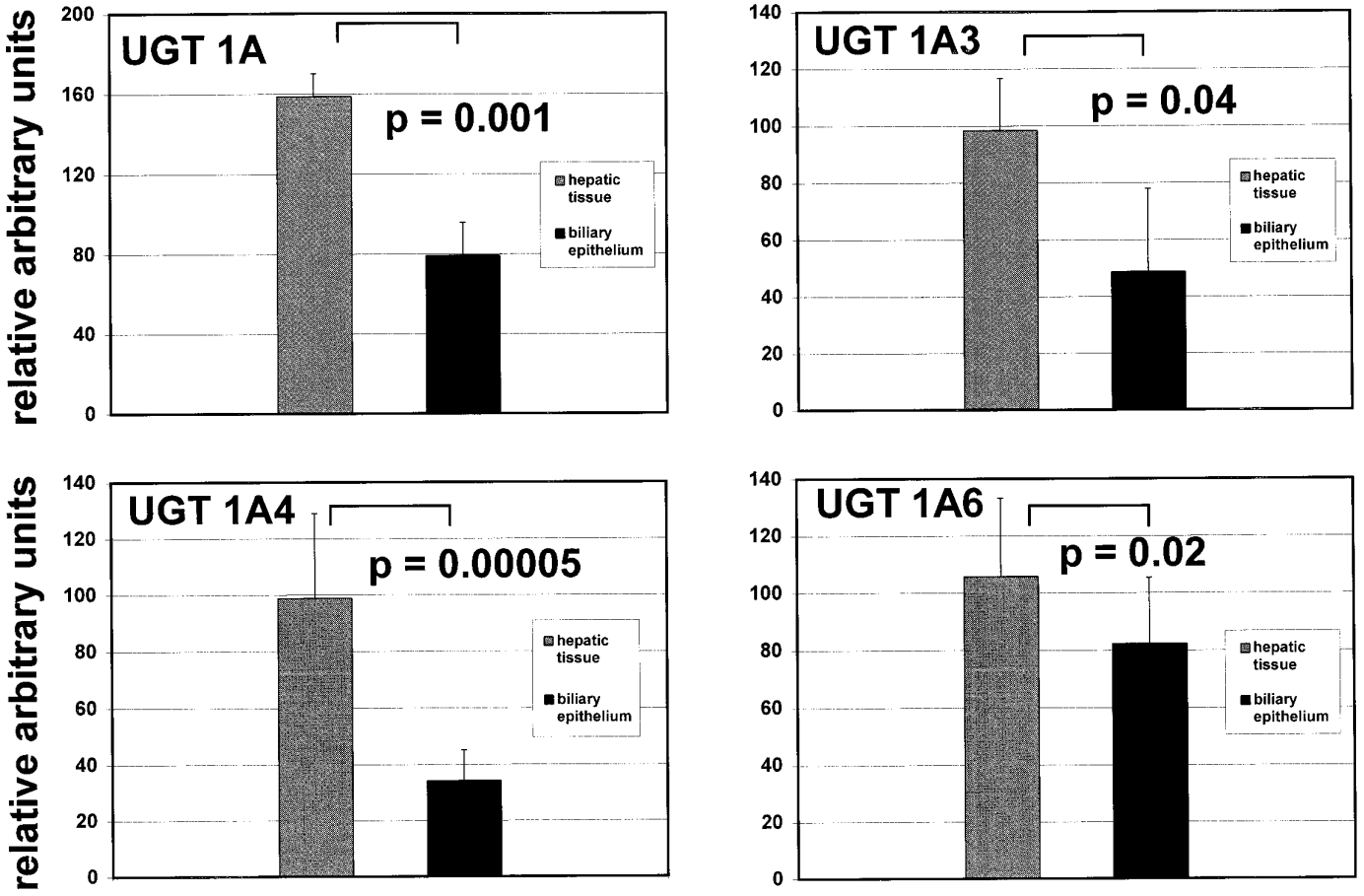

Fig. 7. Quantitative differences in the expression of the *UGT1A* locus in human hepatic and biliary tissue. DRT-PCR demonstrates a significantly lower expression of UGT1A in biliary tissue ($p = 0.001$) (compare with Fig. 2). Analysis of different RNA transcripts shows that UGT1A3 ($p = 0.04$), UGT1A4 ($p = 0.00005$), and UGT1A6 ($p = 0.02$) significantly contribute to this finding. No significant differences were observed for UGT1A1 ($p = 0.09$).

would require a mode of regulation of individual UGT1A transcripts in the liver and in extrahepatic tissues. Although DNA sequence analysis of the 5′ flanking regions of each exon 1 sequence has identified polymerase II regulatory elements, which indicates that each *UGT1A* gene could be regulated independently, evidence for their differential expression in human tissues is not available. The data presented in this study now provide evidence for differential expression of the *UGT1A* locus in human hepatic and extrahepatic tissues.

To demonstrate the precise repertoire of UGT1A transcripts in tissues, methods that guarantee specificity, despite high levels of sequence identity between target transcripts, had to be used. Sequence identity between the first exons of the cluster encoding *UGT1A3*, *UGT1A4*, and *UGT1A5* and the cluster encoding *UGT1A7*, *UGT1A8*, *UGT1A9*, and *UGT1A10* is approximately 93%. Alignments of *UGT1A* exon 1 sequences identified divergent areas for specific PCR primer annealing that differed in as little as two residues between related *UGT1A* first exons. With these primers it was possible to specifically amplify *UGT1A* first exon sequences from genomic DNA and cDNA templates. A universally applicable nonradioactive DRT-PCR assay was developed that enabled the specific and quantitative detection of all individual UGT1A transcripts that were expressed in either hepatocellular, biliary, or gastric tissues.

DRT-PCR analysis of 16 human hepatocellular, four biliary, and two gastric tissue samples demonstrated patterns of expression that were unique for each tissue. The expression of UGT1A1, UGT1A3, UGT1A4, and UGT1A6 was observed in both hepatocellular and biliary tissues, although the overall expression levels were lower in biliary tissue. Gastric tissue was characterized by the expression of UGT1A3 and UGT1A6 but not UGT1A1 and UGT1A4. Transcripts encoding UGT1A5 and UGT1A8 were absent in all three tissue types. Significantly, the hepatocellular expression of the *UGT1A* locus was characterized by the unique expression of UGT1A9. This particular form was not expressed in either biliary or gastric tissue. In contrast, gastric epithelium was characterized by the unique expression of UGT1A7, which was not detected in hepatocellular or biliary tissue. Additionally, the expression of UGT1A10 was specifically localized to both of the nonhepatocellular tissues. This is the first documentation of tissue-specific differential expression of the *UGT1A* locus in human hepatic and extrahepatic epithelium. It is also the first example of the expression of UGT1A7 and UGT1A10 transcripts in humans.

These results are in agreement with previous findings that document the expression of UGT1A3 and UGT1A4 but not UGT1A5 in liver tissue (9). The UGT cDNAs encoding UGT1A1, UGT1A3, UGT1A4, UGT1A6, and UGT1A9 (6, 9, 11, 13) have been cloned. Retrospectively, the isolation and cloning of these UGT1A cDNAs from human liver RNA are

Fig. 8. Amino acid sequence alignment of the novel extrahepatic UGT1A7 and UGT1A10 transcripts. Shown are the 530-amino acid sequences of biliary tissue-derived UGT1A10 and gastric tissue-derived UGT1A7. For UGT1A10, only divergent residues are listed. All 53 amino acid differences are restricted to the first 285 amino acids, representing the variable UGT1A first exon. The carboxyl-terminal 245 residues representing exons 2–5 are identical. *Box*, projected signal peptide region, which is characterized by a high degree of divergence. In contrast to UGT1A10, this region is hydrophobic in UGT1A7 and predicts a signal peptide with a cleavage site after residue 24. Genbank accession numbers: UGT1A7, U89507; UGT1A10, U89508.

not surprising, because these forms represent the extent of the specific UGT1A repertoire of hepatocytes. UGT1A transcripts expressed in hepatocellular and biliary epithelium, both of which function in close anatomical and functional proximity of the liver, are similar except for the switching of UGT1A9 and UGT1A10 regulation. This may indicate regulatory requirements in these tissues that are more similar than those of gastric epithelium. In gastric epithelium, UGT1A1 and UGT1A4 transcripts, the proteins of which are, for example, involved in bilirubin glucuronidation, are absent. In addition, UGT1A7 is regulated as an additional extrahepatic UGT1A transcript. These findings suggest a complex mode of UGT1A regulation in various human organs.

The use of RT-PCR to detect the expression of UGT1A RNA also provided an opportunity to examine the levels of expression in hepatocellular and biliary tissues. Biliary tissue displayed significantly lower levels of expression of UGT1A3, UGT1A4, and UGT1A6, compared with liver. However, the expression levels of UGT1A1 varied significantly in biliary tissue, compared with hepatocellular tissue. Coupled with the observation that UGT1A9 and UGT1A10 are differentially regulated in these tissues, the variation in the levels of UGT1A1 expression could reflect differential glucuronidation requirements of bile ducts and gallbladder versus hepatic tissue and represents additional evidence that the *UGT1A* locus is regulated differentially and in a tissue-specific fashion.

The newly identified human extrahepatic UGT1A7 and UGT1A10 transcripts were cloned and characterized. Both cDNAs exhibit an open reading frame of 1593 bp. The resulting 530-amino acid proteins conform to other cloned UGT1A gene products. UGT1A7 and UGT1A10 encode a 285-residue variable amino-terminal portion representing exon 1 and a carboxyl-terminal portion of 245 amino acids, which is encoded by exons 2–5. Both proteins belong to the UGT1A7 to UGT1A10 cluster, as evident from the significant sequence similarities. UGT1A7 and UGT1A10 exhibit an overall identity of 95.86%, with all 53 amino acid differences being located within the first 283 residues of the protein. Analysis of the predicted amino acid sequence of UGT1A10 also indicates considerable divergence in the amino-terminal portion of the protein that is believed to encode the signal leader sequence. This hydrophobic domain, which is present on most of the UGTs, encodes the nascent polypeptide that interacts with signal recognition particles (20), a process that stops translation and targets the ribosome-signal recognition particle-polypeptide complex to the endoplasmic reticulum. As translation resumes and the protein is translocated through the membrane, the leader sequence is removed. Based upon the predicted amino acid sequence, this appears to be inherent to all of the known UGTs (21). This property of the UGT leader sequence has also been demonstrated experimentally (22). In comparison with UGT1A7, however, the amino-terminal portion of UGT1A10 does not possess a characteristic leader sequence when analyzed by the method of Von Heijne (17) (Fig. 8). Because transport of proteins through the membrane does not always require typical signal sequences, as dis-

played by the ATP-binding cassette transport proteins (23), integration of UGT1A10 into the membrane may proceed by a mechanism independent of that of the other UGTs. This may be significant in light of the nonhepatocellular expression of UGT1A10.

In summary, the results provide convincing evidence for complex and tissue-specific regulation underlying human *UGT1A* control. The identification of two novel UGT1A transcripts not expressed in hepatocellular epithelium underscores the extrahepatic significance of the UGT1A family.

References

1. Dutton, G. J. *Glucuronidation of Drugs and Other Compounds.* CRC Press, Boca Raton, FL (1980).
2. Tukey, R. H., and E. F. Johnson. Molecular aspects of the drug-metabolizing enzymes, in *Principles of Drug Action* (W. Pratt and P. Taylor, eds.), Ed. 3. Churchill Livingstone, New York, 423–468 (1990).
3. Burchell, B., D. W. Nebert, D. R. Nelson, K. W. Bock, T. Iyanagi, P. M. Jansen, D. Lancet, G. J. Mulder, J. R. Chowdhury, G. Siest, T. R. Tephly, and P. I. Mackenzie. The UDP-glucuronosyltransferase gene superfamily: suggested nomenclature based on evolutionary divergence. *DNA Cell Biol.* **10:** 487–494 (1991).
4. Monaghan, G., D. J. Clarke, S. Povey, C. G. See, M. Boxer, and B. Burchell. Isolation of a human YAC contig encompassing a cluster of *UGT2* genes and its regional localization to chromosome 4q13. *Genomics* **23:**496–499 (1994).
5. Ritter, J. K., J. M. Crawford, and I. S. Owens. Cloning of two human liver bilirubin UDP-glucuronosyltransferase cDNAs with expression in COS-1 cells. *J. Biol. Chem.* **266:**1043–1047 (1991).
6. Wooster, R., L. Sutherland, T. Ebner, D. Clarke, O. Da Cruz e Silva, and B. Burchell. Cloning and stable expression of a new member of the human liver phenol/bilirubin:UDP-glucuronosyltransferase cDNA family. *Biochem. J.* **278:**465–469 (1991).
7. Moghrabi, N., L. Sutherland, R. Wooster, S. Povey, M. Boxer, and B. Burchell. Chromosomal assignment of human phenol and bilirubin UDP-glucuronosyltransferase genes (*UGT1A*-subfamily). *Ann. Hum. Genet.* **56:**81–91 (1992).
8. Harding, D., S. J. Jeremiah, S. Povey, and B. Burchell. Chromosomal mapping of a human phenol UDP-glucuronosyltransferase, *GNT1. Ann. Hum. Genet.* **54:**17–21 (1990).
9. Ritter, J. K., F. Chen, Y. Y. Sheen, H. M. Tran, S. Kimura, M. T. Yeatman, and I. S. Owens. A novel complex locus *UGT1* encodes human bilirubin, phenol and other UDP-glucuronosyltransferase isozymes with identical carboxyl termini. *J. Biol. Chem.* **267:**3257–3261 (1992).
10. Owens, I. S., and J. K. Ritter. Gene structure at the human *UGT1* locus creates diversity in isozyme structure, substrate specificity, and regulation. *Prog. Nucleic Acids Res.* **51:**306–338 (1995).
11. Mojarrabi, B., R. Butler, and P. I. Mackenzie. cDNA cloning and characterization of the human UDP-glucuronosyltransferase, UGT1A3. *Biochem. Biophys. Res. Commun.* **225:**785–790 (1996).
12. Green, M. D., and T. R. Tephly. Glucuronidation of amines and hydroxylated xenobiotics and endobiotics catalyzed by expressed human UGT1.4 protein. *Drug Metab. Dispos.* **24:**356–363 (1996).
13. Harding, D., S. Fournel-Gigleux, M. R. Jackson, and B. Burchell. Cloning and substrate specificity of a human phenol UDP-glucuronosyltransferase expressed in COS-7 cells. *Proc. Natl. Acad. Sci. USA* **85:** 8381–8385 (1988).
14. Chomczynski, P. A reagent to the single-step simultaneous isolation of RNA, DNA and proteins from cell and tissue samples. *BioTechniques* **15:**532–537 (1993).
15. Wei, Q., X. Xu, L. Cheng, R. J. Legerski, and F. Ali-Osman. Simultaneous amplification of four DNA repair genes and *beta* actin in human lymphocytes by multiplex reverse transcriptase-PCR. *Cancer Res.* **55:**5025–5029 (1995).
16. Tukey, R. H., N. T. Nguyen, M. Green, and T. R. Tephly. Cloning and characterization of rabbit liver UDP-glucuronosyltransferase cDNAs: developmental and inducible expression of the 4-hydroxybiphenyl *UGT2B13* gene. *J. Biol. Chem.* **268:**15260–15266 (1993).
17. Von Heijne, G. A new method for predicting signal sequence cleavage sites. *Nucleic Acids Res.* **14:**4683–4690 (1986).
18. Ebner, T., R. P. Remmel, and B. Burchell. Human bilirubin UDP-glucuronosyltransferase catalyzes the glucuronidation of ethinylestradiol. *Mol. Pharmacol.* **43:**649–654 (1993).
19. Green, M. D., W. P. Bishop, and T. R. Tephly. Expressed human UGT1.4 protein catalyzes the formation of quaternary ammonium-linked glucuronides. *Drug Metab. Dispos.* **23:**299–302 (1995).
20. Rapoport, T. A. Protein transport across the endoplasmic reticulum membrane: facts, models, mysteries. *FASEB J.* **5:**2792–2798 (1991).
21. Von Heijne, G. Signal sequences: the limits of variation. *J. Mol. Biol.* **184:**99–105 (1985).
22. Mackenzie, P. I., and I. S. Owens. Cleavage of nascent UDP-glucuronosyltransferase from rat liver by dog pancreatic microsomes. *Biochem. Biophys. Res. Commun.* **122:**1441–1449 (1984).
23. Kuchler, K., and J. Thorner. Secretion of peptides and proteins lacking hydrophobic signal sequences: the role of adenosine triphosphate-driven membrane translocators. *Endocr. Rev.* **13:**499–514 (1992).

Send reprint requests to: Robert H. Tukey, Ph.D., University of California, San Diego, Department of Pharmacology, La Jolla, CA 92093-0636.