

DMD #79221

Posttranscriptional regulation of UGT2B10 hepatic expression and activity by alternative splicing

Adrien Labriet, Eric P. Allain, Michèle Rouleau, Yannick Audet-Delage, Lyne Villeneuve and Chantal Guillemette

Pharmacogenomics Laboratory, Centre Hospitalier Universitaire (CHU) de Québec Research Center – Université Laval, Québec, QC, Canada and Faculty of Pharmacy, Université Laval, Québec, Canada

Canada Research Chair in Pharmacogenomics (CG)

DMD #79221

Short title: Regulation of UGT2B10 by alternative splicing

Corresponding Author:

Chantal Guillemette, Ph.D.

Pharmacogenomics Laboratory

CHU de Québec Research Center – Univesité Laval

2705 Boul. Laurier, R4701.5

Québec, Canada, G1V 4G2

Tel. (418) 654-2296

E-mail: Chantal.Guillemette@crchudequebec.ulaval.ca

Abstract: 248 words

Introduction: 722

Material and Methods: 1811

Results: 1566

Discussion: 1686

Body text: 5785

References: 60

NUMBER OF:

Tables: 0

Figures: 7

Supplemental Tables: 1

Supplemental Figures: 6

DMD #79221

Non-standard abbreviations:

AS: Alternative splicing

ER: Endoplasmic reticulum

MS-MRM: Mass spectrometry – multiple reaction monitoring

RT-PCR: Reverse transcription - PCR

RNA-seq: RNA sequencing

TPM: transcripts per million

UDP-GlcA: UDP-glucuronic acid

UGT: UDP-glucuronosyltransferases

DMD #79221

Abstract

The detoxification enzyme UDP-glucuronosyltransferase UGT2B10 is specialized in the N-linked glucuronidation of many drugs and xenobiotics. Preferred substrates possess tertiary aliphatic amines and heterocyclic amines such as tobacco carcinogens and several anti-depressants and anti-psychotics. We hypothesized that alternative splicing (AS) constitutes a mean to regulate steady state levels of UGT2B10 and enzyme activity. We established the transcriptome of *UGT2B10* in normal and tumoral tissues of multiple individuals. Highest expression was in the liver, where ten AS transcripts represented 50% of the *UGT2B10* transcriptome in 50 normal livers and 44 hepatocellular carcinomas. One abundant class of transcripts involves a novel exonic sequence and leads to two alternative (alt.) variants with novel in-frame C-termini of 10 or 65 amino acids. Their hepatic expression was highly variable among individuals, correlated with canonical transcript levels, and was 3.5 fold higher in tumors. Evidence for their translation in liver tissues was acquired by mass spectrometry. In cell models, they co-localized with the enzyme and influenced the conjugation of amitriptyline and levomedetomidine by repressing or activating the enzyme (40-70%; $P < 0.01$), in a cell context-specific manner. A high turnover rate for the alt. proteins, regulated by the proteasome, was observed in contrast to the more stable UGT2B10 enzyme. Moreover, a drug-induced remodelling of *UGT2B10* splicing was demonstrated in the HepaRG hepatic cell model, which favored alt. variants expression over the canonical transcript. Our findings support a significant contribution of AS in the regulation of UGT2B10 expression in the liver with an impact on enzyme activity.

DMD #79221

Introduction

N-linked glucuronidation is an important inactivation route for amine-containing drugs and xenobiotics (Kaivosaari et al., 2011; Kato et al., 2013). Two of the 19 UDP-glucuronosyltransferases (UGTs), UGT2B10 and UGT1A4, are the main drivers of N-linked glucuronidation (Chen et al., 2008a; Kerdpin et al., 2009; Kato et al., 2013). Long described as an orphan UGT, the discovery that UGT2B10 is crucial for the detoxification of tobacco carcinogens has raised much attention and rationalized the importance to characterize this unique UGT (Chen et al., 2007; Kaivosaari et al., 2007; Chen et al., 2008b; Berg et al., 2010; Murphy et al., 2014). UGT2B10 is one of the main liver UGT enzymes, based on mRNA (quantitative PCR and deep RNA-sequencing data) and protein levels (quantitative mass spectrometry-based proteomics data) (Court et al., 2012; Fallon et al., 2013; Margaillan et al., 2015; Tourancheau et al., 2017). Expression has also been reported in the breast, testis, gallbladder, tongue and tonsils although at much lower levels than hepatic expression (Haakensen et al., 2010; Jones and Lazarus, 2014).

UGT2B10 displays a preference for tertiary aliphatic amines and heterocyclic amines. These structures are found in several clinically used drugs such as antihistamines, anti-psychotics and anti-depressants including several of the tricyclic class such as imipramine and amitriptyline (Kaivosaari et al., 2008; Kaivosaari et al., 2011; Kato et al., 2013; Kazmi et al., 2015; Pattanawongsa et al., 2016) whereas endogenous substrates have yet to be identified. Although UGT1A4 substrate preference significantly overlaps with that of UGT2B10, the latter presents a greater affinity and clearance for many

DMD #79221

tertiary cyclic amines at therapeutic concentration (Kato et al., 2013). One structural determinant of the specificity towards amine substrates may be the residues Pro40 of UGT1A4 and Leu34 of UGT2B10 located in their substrate binding domain, a position that is otherwise a strictly conserved histidine residue (His40, coordinates of UGT1A1) in all human UGTs (Kerdpin et al., 2009).

Genetic studies in smokers have established a direct link between single nucleotide polymorphisms (SNPs) at the *UGT2B10* gene locus (4q13.2), UGT2B10 activity and metabolism of nicotine and cotinine (Chen et al., 2007; Chen et al., 2012; Murphy et al., 2014; Patel et al., 2015; Ware et al., 2016; Murphy, 2017). Two relatively frequent polymorphisms are associated with significant decreased nicotine and cotinine glucuronidation in humans. One variant creates the missense Asp67Tyr coding variant (rs61750900) that abolishes UGT2B10 glucuronidation activity (Chen et al., 2007), while the other variant (rs116294140/rs2942857) alters the splice acceptor site between intron 2 and exon 3, and is thought to create an unstable mRNA (Murphy et al., 2014; Fowler et al., 2015). Consistently, individuals homozygous for either of these SNPs have very low to undetected nicotine and cotinine glucuronides in urine, further indicating that UGT2B10 is the main UGT responsible for glucuronidation of nicotine and cotinine (Chen et al., 2007; Murphy et al., 2014). It also supports the main role of this detoxification pathway in tobacco carcinogenesis. These genetic variations display an important ethnic bias. Whereas the Asp67Tyr variation is most frequent in Caucasians (nearly 10%) (Chen et al., 2007; Murphy et al., 2014), the splice variant rs2942857 prevails in approximately 35% of African Americans and in up to 50% of subjects of African origins (NCBI; Murphy et al., 2014). Of interest, the splice variant was also

DMD #79221

reported to affect the metabolism of the antipsychotic preclinical drug RO5263397 whose main clearance route is N-glucuronidation (Fowler et al., 2015). It is thus likely that these SNPs also affect the glucuronidation of prescribed drugs conjugated by UGT2B10 such as amitriptyline (Kaivosari et al., 2008; Kato et al., 2013).

Recent studies by our group revealed that alternative splicing (AS) largely expands the UGT transcriptome (Tourancheau et al., 2016; Tourancheau et al., 2017). In turn, alternative (alt.) isoforms modulate the activity of UGT enzymes, suggesting that AS programs may contribute to interindividual variability in xenobiotics metabolism and cancer susceptibility (Bellemare et al., 2010b; Menard et al., 2013; Rouleau et al., 2014; Rouleau et al., 2016). Given the clinical importance of UGT2B10, by virtue of its detoxification functions towards tobacco by-products, anti-depressors and anti-psychotics, we hypothesized that AS represents a mean to regulate steady-state levels of UGT2B10 mRNA and enzyme activity. The goal of this study was to investigate in more details the expression of UGT2B10 and evaluate the influence of alternative isoforms on the metabolic functions of UGT2B10.

DMD #79221

Materials and Methods

Analysis of UGT2B10 mRNA expression

To carefully analyze *UGT2B10* expression and AS patterns, raw RNA sequencing (RNA-seq) data were downloaded from public databases and realigned to the recently established UGT transcriptome (Tourancheau et al., 2016; GTEx (<http://www.gtexportal.org/home/>) and TCGA (<https://gdc.cancer.gov/>)). RNA-seq datasets were obtained from dbGaP at <http://www.ncbi.nlm.nih.gov/gap> through dbGaP accession number phs000424.v6.p1; project ID 13346. The GTEx normal liver data (downloaded on June 7, 2017) were from 50 healthy individuals [29 males; 18 females; 3 unknown; median age: 55 years old (range 21-69 years old), mostly Caucasians]. The TCGA cancer liver data (downloaded on August 22, 2017), all hepatocellular carcinoma, were from 44 individuals [29 males, 15 females; median age: 62 years old (range 18-82 years old), 16 Caucasians, 22 Asians, and 5 African-Americans]. GTEx RNA-seq data of bladder (n=6), breast (n=51), colon (n=49), lung (n=48), kidney (n=35) prostate (n=50), and skin (n=53) were similarly obtained. HepaRG RNA-seq data (GSE 71446, accessed March 9, 2017; (Li et al., 2015)) were downloaded from the NCBI Gene expression omnibus database. For each individual RNA-seq data, quality of FASTQ files was assessed with FastQC before quality trimming with Trimmomatic v0.36 (Bolger et al., 2014). UGT transcript quantification was done with trimmed reads from each sample and a custom UGT transcript annotation (Tourancheau et al., 2016) using Kallisto v0.43 (Bray et al., 2016). Data (counts) was then upper-quantile normalized and further adjusted using housekeeping genes as described (de Jonge et al., 2007) using the EDASeq and RUVSeq packages for R version 3.2.2. Differential expression of UGT

DMD #79221

isoforms was assessed using the edgeR package for R. Normalized counts were then converted to counts per million (CPM) or transcripts per million (TPM) using transcript length. Reverse Transcription-PCR (RT-PCR) analysis of *UGT2B10* transcripts was as described (Tourancheau et al., 2016), using primer sequences provided in Table S1. The Basic Local Alignment Search Tool (BLAST) of the National Center for Biotechnology Information (NCBI, <https://blast.ncbi.nlm.nih.gov/Blast.cgi>) served to search sequence similarity between the novel UGT2B10 sequences and other genes in humans and other species. The Protein BLAST (blastp) suite was used to search "Non-redundant protein sequences" and "Reference proteins" whereas Translated nucleotide (tblastn) suite was used to search the "Nucleotide collection" and "Reference RNA sequences" with the unique amino acid sequences of alt. proteins.

Expression vectors and human cell models (HEK293 and HepG2)

To study alt. transcripts and proteins, expression vectors were produced from the *UGT2B10_v1* pcDNA6 construct (Beaulieu et al., 1998) using the Q5 Site-Directed Mutagenesis kit (New England Biolabs Ltd., Whitby, ON, Canada). Sequences of purified primers used for mutagenesis are provided in **Supplemental Table 1**. *UGT2B10_v1* coding sequence was also cloned in the pcDNA3.1 vector to produce co-expression cell models (below). All constructs were verified by Sanger sequencing. For the expression of alt. proteins tagged with V5-his (for immunoprecipitations studies), the stop codon of each coding sequence cloned in the pcDNA6 vector was removed using the Q5 Site-Directed Mutagenesis kit to permit in-frame V5-his expression.

DMD #79221

HEK293 and HepG2 cell lines were obtained from the American Type Culture Collection (Manassas, VA, USA) and grown as described (Levesque et al., 1997; Menard et al., 2013). HEK293 cells (UGT negative; 2×10^6 cells in a 10 cm-plate) were transfected with 1 μ g of each construct using Effectene (Qiagen, Toronto, ON, Canada), and HepG2 cells (UGT positive; 1×10^7 cells) were transfected with 20 μ g of each construct by electroporation with the Neon Transfection System (Invitrogen, ThermoFisher Scientific, Ottawa, ON, Canada) as per manufacturers' instructions. HEK293 cells stably expressing the UGT2B10 cDNAs were established by supplementing cell culture media with blasticidin (Wisent, St-Bruno, QC, Canada; 10 μ g/ml). Clones were selected based on UGT2B10 expression detected by immunoblotting with antibodies specified below. HEK293 cells co-expressing UGT2B10_v1 (encoding the canonical enzyme) and alt. UGT2B10 (with novel sequences in C-termini) were established by subsequent transfection of HEK293 clones expressing alt. sequences with the UGT2B10_v1-pcDNA3.1 construct and selection with G418 (Invitrogen, 1 mg/ml). HepG2 cells, which express UGT2B10_v1 endogenously, were transfected with the constructs expressing alt. variants and clone selection was with blasticidin. Control HEK293 and HepG2 cells were produced by transfection with the parental vector pcDNA6 and selection as above.

Analysis of protein expression

Antibodies. The rabbit polyclonal anti-UGT2B10 #1845 produced in-house against GST-UGT2B11 (aa 60-140) by Dr Alain Bélanger's group (Chouinard et al., 2006), was used for the immunodetection of UGT2B10 in HepG2 (1:5000) and HEK293 (1:10 000) cell models. Our analysis revealed that this antibody detects UGT2B10, UGT2B11 and less efficiently UGT2B28 (**Supplemental Figure 1**). The monoclonal anti-UGT2B10 antibody

DMD #79221

(Abcam ab57685) was used for immunoprecipitation in human liver. Cell compartment-specific antibodies used for immunofluorescence (IF) were anti-58K Golgi protein (1:100; ab27043, Abcam Inc., Toronto, ON, Canada), anti-protein disulfide isomerase (PDI: 1:100; ab2792, Abcam) for the ER, and anti-lamin (1:200; sc-376248, Santa Cruz Biotechnology, Dallas, TX, USA). DNA was stained with DRAQ5 (1:2000; ThermoFisher Scientific). Anti-calnexin was from Enzo Life Sciences (Farmingdale, NY, USA; ADI-SPA-860; 1:5000).

Mass spectrometry – multiple reaction monitoring (MS-MRM). Detection of peptides unique to alt. UGT2B10 was as described (Rouleau et al., 2016) with minor modifications. Briefly, human liver S9 fraction (8 mg protein; Xenotech LLC, Lenexa, KS, USA) was lysed for 45 min on ice in a total volume of 4 ml Lysis Buffer containing 0.05 M Tris-HCl pH 7.4, 0.15 M NaCl, 1% (w/v) Igepal CA-630 (Sigma-Aldrich), 1 mM dithiothreitol, and Complete protease inhibitor cocktail (Roche, Laval, QC, Canada). Lysates were centrifuged for 15 min at 13,000 g, and UGT2B10 was immunoprecipitated with 10 µg of the monoclonal anti-UGT2B10 (Abcam ab57685) for 1 h at 4°C on an orbital shaker. Protein complexes were captured by an overnight incubation at 4°C with Protein G-coated magnetic beads (200 µl Dynabeads, ThermoFisher Scientific). Beads were washed in lysis buffer and with 50 mM ammonium bicarbonate and stored at –20°C until analysis. Tryptic digests of UGT2B10 were prepared and analyzed by MS-coupled multiple reaction monitoring on a 6500 QTRAP hybrid triple quadrupole/linear ion trap mass spectrometer (Sciex, Concord, ON, Canada) as described (Rouleau et al., 2016). Briefly, MS analyses were conducted with an ionspray voltage of 2500 V in positive ion mode. Peptides were desalted on a 200 µm × 6 mm chip trap ChromXP C18

DMD #79221

column, 3 μm (Eksigent, Sciex), at 2 $\mu\text{l}/\text{min}$ solvent A (0.1% formic acid). Peptides were then eluted at a flow rate of 1 $\mu\text{l}/\text{min}$ with a 30-min linear gradient from 5 to 40% solvent B (acetonitrile with 0.1% formic acid) and a 10-min linear gradient from 40 to 95% solvent B. MRM-MS analyses were performed using the four most intense transitions for each of the target peptides for the light and heavy forms. The UGT2B10 signature peptides were detected in tryptic digests of the immunoprecipitated UGT2B10 samples, and peptide identity was confirmed by co-injection of isotopically labeled [$^{13}\text{C}_6$, $^{15}\text{N}_2$]Lys and [$^{13}\text{C}_6$, $^{15}\text{N}_4$]Arg synthetic peptides (Pierce Protein Biotechnology, ThermoFisher Scientific).

Glucuronidation assays – For enzymatic assays in intact cells (*in situ* assays), two cell models were employed (HEK293 and HepG2 cells). Cells were seeded in 24-well plates (HEK293: 8×10^4 cells/well, HepG2: 2.25×10^5 cells/well). Assays were initiated 72 h after seeding by replacing the culture medium with fresh medium (1 ml/well) containing a UGT2B10 substrate (amitriptyline, 7.5 μM and 150 μM ; levomedetomidine, 7.5 μM and 75 μM). All substrates were obtained from Sigma-Aldrich. Cells were incubated for 4 h, media were then collected and stored at -20°C until glucuronide (G) quantification by high-performance liquid chromatography tandem mass spectrometry. Assays were replicated in at least two independent experiments in triplicates. Separation of amitriptyline and levomedetomidine was performed onto an ACE Phenyl column 3 μM packing material, 100×4.6 mm (Canadian Life Science, ON, Canada). Isocratic condition with 70% methanol/30% water/3 mM ammonium formate with a flow rate of 0.9 ml/min was used to elute amitriptyline-G. A linear gradient was used to elute levomedetomidine-G, with 5% methanol/95% water/1 mM ammonium formate as initial

DMD #79221

conditions followed by a 10 min-linear gradient to 90% methanol/10% water/ 1 mM ammonium formate. The glucuronides were quantified by tandem mass spectrometry (API 6500; Biosystems-Sciex, Concord, ON, Canada). The following mass ion transitions (m/z) were used: 377.1 \rightarrow 201.1 for levomedetomidine-G and 454.2 \rightarrow 191.1 for amitriptyline-G. Glucuronidation activity (area/hour/mg protein/UGT level) was normalized for the expression of the UGT2B10 enzyme in each cell model, determined by densitometry scanning of band intensity on immunoblots with antibody #1845. Cycloheximide glucuronidation assays were conducted using microsomes isolated from human liver (Xenotech LLC), HepG2-pcDNA6 described above, commercial supersomes expressing UGT1A and UGT2B isoenzymes (Corning, Woburn, MA, USA), and microsomes of UGT2B11-expressing HEK293 prepared in-house as described (Lepine et al., 2004). Glucuronidation assays were conducted at 37°C using 20 μ g membrane proteins and a final concentration of 200 μ M cycloheximide (Sigma-Aldrich). Cycloheximide-G quantification was conducted as specified above for the UGT2B10 substrates, using a linear gradient with 10% methanol/90% water/1 mM ammonium formate as the initial conditions followed by a linear gradient to 85% methanol/15% water/ 1 mM ammonium formate in 5 min.

Immunofluorescence – Subcellular distribution of UGT2B10 enzyme and alternative isoforms was carried out in HEK293 cells stably expressing either protein and detected with the anti-UGT2B10 antibody #1845 (1:1000) as described (Rouleau et al., 2016).

DMD #79221

Immunoprecipitation - Stable HEK293 *UGT2B10_v1* cells (2×10^6 cells) were seeded in 10 cm-plates and transiently transfected with alt. UGT2B10 V5-tagged constructs (3 μ g) using Effectene as per manufacturer's instructions (Qiagen). 36 h post-transfection, cells were processed for cross-linking and immunoprecipitation with the polyclonal anti-V5 (1:600; NB600-380; Novus Biologicals) as described (Rouleau et al., 2016).

Protein and mRNA stability in cell models - HEK293 cell models were seeded in 6-well plates (4×10^5 cells/well) and grown for 48 h. HepG2 cell models were seeded in 6-well plates (1.5×10^6 cells/well) and grown for 24 h. Cells were rinsed and incubated for 16h with media containing 1 μ M MG132 (Calbiochem, EMD Millipore, Etobicoke, ON, Canada) or vehicle (DMSO). Cells were harvested and protein extracts were prepared in Lysis Buffer. Cells were lysed for 30 min on a rotation unit, homogenized by pipetting up and down through fine needles (18G and 20G, 10 times each) on ice and cleared by centrifugation for 15 min at 13,000 *g* prior to analysis by immunoblotting using anti-UGT2B10 #1845. Duplicate cell samples were harvested for RNA extraction and RT-PCR analysis of UGT2B10 variants as described (Rouleau et al., 2016). Assays were done twice. UGT2B10 proteins half-lives were determined by treatment with cycloheximide (20 μ g/ml) for 0-16 h as described (Turgeon et al., 2003). Cells were washed in PBS, collected and lysed by scraping in Lysis Buffer, prior to analysis by immunoblotting.

Assessment of the glycosylation status of UGT2B10 proteins expressed in HEK293 cells - Microsomes (20 μ g proteins) prepared from HEK293 stably expressing each UGT2B10 isoform were treated with Endo H and O-glycosidase obtained from New England

DMD #79221

Biolabs Inc. (Ipswich, MA, USA) as described (Girard-Bock et al., 2016). Assays were performed in two independent experiments.

DMD #79221

Results

UGT2B10 gene expression predominates in human liver and is regulated by alternative splicing

For a quantitative assessment of UGT2B10 expression, we performed a realignment of public GTEX and TCGA RNA-seq data from human tissue samples to the fully annotated UGT variant sequence database. This database was created based on RNA-seq experiments previously conducted with several human tissues (Tourancheau et al., 2016). Data indicated that the UGT2B10 transcriptome is comprised of one canonical and 10 alt. transcripts arising from the single *UGT2B10* gene. The UGT2B10 alt. transcripts are created by partial intronization of exon 1-2, exon skipping of exon 4-6 and inclusion of novel terminal exons 6b or 6c. UGT2B10 transcripts and encoded isoforms are depicted in **Figure 1**. Data revealed *UGT2B10* as one of the highest expressed UGTs in normal liver samples whereas variants arising from alternative splicing represented nearly 50% of the *UGT2B10* hepatic transcriptome (**Figure 2A**). Some variant classes were remarkably abundant, at levels comparable to those of the canonical *UGT2B10_v1* transcript encoding the UGT2B10 enzyme, namely those with N-terminal truncations and with novel C-terminal sequences (**Figure 2A**). Total *UGT2B10* expression was 2.15-fold ($P=0.013$) higher in hepatocellular carcinoma (HCC; $n=44$) relative to normal livers ($n=50$), and was highly variable among individuals (coefficients of variation of 93 to 117%) (**Figure 2B**). In contrast to the high hepatic levels, expression of *UGT2B10* in other normal tissues surveyed was much lower, with values below 1 TPM in the bladder, breast, colon, kidney, lung, prostate, and skin (data not shown).

Alternative splicing creates UGT2B10 variants with novel in-frame C-terminal sequences detected in human liver at the protein level

Splicing events generating *UGT2B10* transcripts (*n9* and *n10*) harbouring novel C-terminal sequences were of particular interest for this study. These alt. variants are produced by intronization of parts of the canonical exon 6 and inclusion of a novel exon 6c (**Figure 1**). These AS events appeared specific to human and were not noted in other species based on BLAST searches in the non-redundant nucleotide collection and RNA Reference sequence databases. Transcripts were abundant in human livers and were 3.5 fold higher in hepatic tumors relative to normal tissues. They represented on average 15% of total *UGT2B10* expression in normal liver tissues and 25% in HCC tissues (**Figure 2A-B**). The considerable interindividual variability for these alt. transcripts was higher than for the total and canonical *UGT2B10* hepatic expression. We also noted a significant positive correlation between the expression of the canonical and alt. transcripts ($\rho=0.854-0.915$, $P\leq 0.001$) (**Figure 2C**).

The putative UGT2B10 proteins encoded by these alt. transcripts are referred to as UGT2B10 isoforms 4 and 5, named i4 and i5 respectively. They are predicted to retain both the substrate and the co-substrate (UDP-GlcA) binding domains coupled to a novel alt. C-sequence (**Figure 1**). Isoform 4 lacks the 43 C-terminal amino acids including the transmembrane domain and the positively charged C-terminal tail of the UGT2B10 enzyme that are replaced by 10 novel amino acids encoded by exon 6c (**Figure 3A**). As for isoform 5, because a smaller portion of exon 6 is intronized, the encoded protein is predicted to retain 18 of the 24 transmembrane domain residues and to be extended by

DMD #79221

65 novel amino acids, of which half are encoded by a frame-shift in exon 6, and half by exon 6c (**Figure 1, 3A**).

Validation of endogenous protein expression was possible for UGT2B10_i5 through the identification of a unique peptide sequence by targeted mass spectrometry-multiple reaction monitoring (MRM) of liver UGT2B10 immunoprecipitated from multiple donors (**Figure 3B,C**). This result is in line with the detection of the corresponding transcripts by PCR in human livers (**Supplemental Figure 2A**). The endogenous expression of UGT2B10_i4 could not be addressed by this approach given the short and hydrophobic nature of the C-terminal unique sequence.

Alternative isoforms with novel C-terminal sequences co-localized with the UGT2B10 enzyme and modified its activity *in vitro*

The alt. UGT2B10 isoforms were stably expressed in the embryonic kidney cell line HEK293, devoid of endogenous UGT expression, alone or with the canonical UGT2B10 enzyme. As well, their expression was examined in the liver cell model HepG2 that expresses endogenously the UGT2B10 enzyme and conjugates substrates of the enzyme such as amitriptyline and levomedetomidine. We initially confirmed protein expression using immunoblot and immunofluorescence experiments using an anti-UGT2B10 antibody (#1845) that targets amino acids encoded by exon 1 and therefore recognizes the three UGT2B10 isoforms (**Figure 4A**). Indeed, we detected UGT2B10_i4 and UGT2B10_i5, near their predicted molecular weights (MW) of 57 and 66 kDa in cell models (**Figure 4A**).

The subcellular distribution of each protein was examined in the HEK293 models where the canonical and alt. proteins were detected and largely restricted to an endoplasmic

DMD #79221

reticulum (ER) localization. This was confirmed by the co-localization with the ER marker protein disulfide isomerase (PDI) (**Figure 4B**). Each isoform also displayed minor perinuclear and Golgi localization (**Supplemental Figure 3**). The glycosylation status of each protein was studied by subjecting microsomes from HEK293 cell models to Endo H glycosidase, which cleaves N-linked sugars on asparagine acquired by ER-resident enzymes, and to O-glycosidase, which removes serine and threonine O-linked complex sugars acquired in the Golgi. Each UGT2B10 protein was sensitive to Endo H treatment, revealed by the shift to a higher mobility protein band upon treatment (**Figure 4C**). In contrast, their mobility was not affected by a treatment with O-glycosidase.

Enzymatic assays in intact cells were subsequently conducted in the hepatic cell model HepG2 that expresses the endogenous UGT2B10 enzyme. Compared to the reference cell line (stably transfected with pcDNA6), expression of alt. UGT2B10_i4 isoform enhanced glucuronidation of the UGT2B10 enzyme with two drug substrates, amitriptyline and levomedetomidine, by 1.5 to 2 fold (**Figure 5A,B**). In turns, the presence of alt. UGT2B10_i5 significantly impaired the glucuronidation of amitriptyline and tended to reduce that of levomedetomidine as well, by 20-25%. Glucuronidation assays conducted with the UGT negative cells HEK293 stably expressing either alt. UGT2B10 isoforms revealed no transferase activity for amitriptyline and levomedetomidine. When co-expressed with the UGT2B10 enzyme, we observed a significant inhibition by 23 to 65% of glucuronidation activity by HEK293 cells in the presence of the alt. i4 or i5 proteins (**Figure 5A,B**). Since we observed a co-localization of alt. isoforms and the UGT2B10 enzyme in the ER, we addressed their potential interaction as a possible regulatory mechanism. Immunoprecipitations were conducted with an anti-V5 epitope antibody using cell models stably expressing the UGT2B10

DMD #79221

enzyme and transiently expressing either alt. isoform tagged with the V5 epitope. The UGT2B10 enzyme was immunoprecipitated with each alt. isoform, indicating their ability to form complexes (**Figure 5C**).

Alternative isoforms have shorter half-lives than the UGT2B10 enzyme and are targeted for degradation by the proteasome

Protein stability was evaluated in both cell models. The UGT2B10 enzyme displayed a half-life over 16 hours in both HEK293 (exogenous expression) and HepG2 (endogenous expression) cell models whereas the alt. isoforms displayed superior turnover rates. The alt. UGT2B10_i5 was the least stable, with short half-lives of 1.9 and 0.7 hours in HEK293 and HepG2, respectively (**Figure 6A**). The turnover rate of isoform i4 differed between the two cell models, and was 11.5 hours in HEK293 but much shorter in HepG2 (1.5 hours). We noted in HepG2 cells a significant recovery of alt. proteins expression, even rising above those of non-treated cells by 16h after initiation of the cycloheximide treatment. This was not observed for the UGT2B10 enzyme, nor for any of the UGT2B10 proteins in HEK293 cells. This observation implied a possible inactivation in HepG2 cells, which was confirmed by the detection of two glucuronides of cycloheximide (G1 and G2) with HepG2 microsomes (**Supplemental Figure 5**). This was further validated using microsomal fractions of pooled human livers and UGT supersomes. In these experiments, UGT1A9 and UGT2B7 most efficiently glucuronidated cycloheximide (G1 and G2), whereas UGT1A3 and UGT1A4 (G1) as well as UGT1A1 and UGT2B4 (G1 and G2) also conjugated some cycloheximide (**Supplemental Figure 5**), with some of them detected in HepG2 cells and not in the

DMD #79221

HEK293 cell model. As a consequence, this may lead to an inaccurate assessment of UGT proteins half-lives using this approach in the HepG2 model.

Accordingly, the difference in protein stability between the UGT2B10 enzyme and alt. proteins was further addressed by proteasomal inhibition. Whereas the enzyme levels were nearly unperturbed by inhibition of the ubiquitin-proteasome system, alt. UGT2B10 and more particularly isoform i5, were stabilized, indicating that they were degraded via the ubiquitin-proteasome system. Proteasomal inhibition for 16 h increased the ratio of alt. isoform / UGT2B10 enzyme in HepG2 liver cell models, whereas stabilization was more modest in HEK293 (**Figure 6B**). Increased levels of alt. isoforms were not derived from enhanced transcription, verified at the mRNA level (**Supplemental Figure 2B**).

Differential induction of UGT2B10 alternative transcripts in liver cells by phenobarbital and a CAR agonist

HepaRG cells constitute a good surrogate system to study hepatic functions and response to drug treatments. An analysis of public HepaRG RNA-seq data (Li et al., 2015) with the exhaustive UGT transcriptome revealed an expression of canonical and alt. *UGT2B10* in comparable proportions to the human liver (not shown). In HepaRG cells treated with the constitutive androstane receptor (CAR) agonist CITCO, a significant and preferential induction of the alt. UGT2B10 transcripts (1.4 fold, $P=0.003$) occurred, whereas the enzyme-coding *v1* transcript was not significantly altered (**Figure 7**). Phenobarbital did not significantly induce expression of *UGT2B10* in wild type HepaRG cells. In contrast, in HepaRG cells with a CAR knock-out, alt. transcripts were significantly upregulated by 1.8 fold ($P\leq 0.001$) by phenobarbital, and 1.4 fold ($P=0.011$) by the CAR agonist, whereas *v1* was not significantly perturbed (**Figure 7**).

DMD #79221

Discussion

The recent expansion of the pharmacogene transcriptome by AS has shed light on a novel mechanism regulating drug metabolism and clearance (Bellemare et al., 2010a; Bellemare et al., 2010b; Guillemette et al., 2010; Guillemette et al., 2014; Rouleau et al., 2014; Chhibber et al., 2016; Rouleau et al., 2016; Tourancheau et al., 2016; Tourancheau et al., 2017). Our study of the *UGT2B10* transcriptome, encoding a key detoxification enzyme specialized in N-glucuronidation of multiple harmful xenobiotics (Kaivosaaari et al., 2007; Kaivosaaari et al., 2011; Kato et al., 2013), demonstrated that AS accounts for a large proportion of *UGT2B10* gene expression and especially in the liver. This observation held true in liver tumors, where *UGT2B10* expression was enhanced by two-fold in hepatocellular carcinoma. Indeed, our analysis of next-generation sequencing data revealed that *UGT2B10* expression prevails in the liver whereas in all other tissues surveyed, including the lung, its expression was low to undetected. This is consistent with the expression determined at the RNA level in several human tissues, including those of the aerodigestive tract (Ohno and Nakajin, 2011; Court et al., 2012; Jones and Lazarus, 2014). It supports that a main detoxification site of UGT2B10-dependent N-glucuronidation is the liver, where UGT2B10 is one of the most abundant UGT enzyme based on proteomics data (Fallon et al., 2013; Sato et al., 2014; Margaillan et al., 2015). Besides, AS also provides an explanation for the multiple observations reporting a lack of correlation between mRNA and protein expression, such as in hepatocellular carcinoma where the RNA expression remained equivalent between tumor tissues and adjacent normal tissues whereas glucuronidation activity was drastically decreased (Lu et al., 2015).

DMD #79221

With a focus on one abundant class of hepatic alt. *UGT2B10* variants containing a novel 3' terminal exon that were confirmed at the protein level in human liver samples and in heterologous expression models, we expose their regulated expression and influence on *UGT2B10* enzyme activity *in vitro*. In fact, the transcriptional regulation of *UGT2B10* has been poorly studied. A response element for the bile acid sensing farnesoyl X receptor (FXR) was recently uncovered in the *UGT2B10* promoter region and participated in the induction of *UGT2B10* by the FXR agonists GW4064 and chenodeoxycholic acid (Lu et al., 2017a). In the cell model HepaRG, a surrogate to human primary hepatocytes in drug metabolism studies (Antherieu et al., 2012), RNA-seq data revealed drug-induced regulation of the *UGT2B10* transcriptome by both the constitutive androstane receptor (CAR) and pregnane X receptor (PXR). The superior induction of alt. *UGT2B10* by CAR and PXR agonists observed herein, especially in HepaRG cells devoid of CAR expression, further raises the possibility of a PXR-dependent remodelling of splicing events at the *UGT2B10* locus that may significantly influence *UGT2B10* detoxification activity. Whether other receptors such as FXR, previously reported to regulate *UGT2B10* transcription (Lu et al., 2017a), also influence splicing remain to be addressed. When expressed in human cells, alt. *UGT2B10* acted as regulators of the glucuronidation activity of the *UGT2B10* enzyme, possibly conveyed by heterologous complexes formed between the enzyme and alternative proteins; a regulatory mechanism documented for other human UGTs (Bushey and Lazarus, 2012; Menard et al., 2013; Rouleau et al., 2014; Rouleau et al., 2016). Our findings further suggest a cell-specific influence given that in HepG2 cells, an increased N-glucuronidation of the *UGT2B10* substrates amitriptyline and levomedetomidine by the endogenous enzyme was observed, in contrast with a repression of enzyme activity in HEK293 cells. The

DMD #79221

endogenous expression of additional UGTs other than UGT2B10 in HepG2 cells and different protein turnover rates could be among factors influencing their functions. Likewise, the impact of AS on UGT2B10 activity has been documented previously by findings of a common polymorphism (rs116294140) that disrupts a splice site in exon 3 and introduces a premature stop codon possibly triggering non-sense mRNA decay (Fowler et al., 2015). This polymorphism, particularly more frequent among African Americans, significantly reduced N-glucuronidation of drugs such as RO5263397 as well as nicotine and cotinine (Murphy et al., 2014; Fowler et al., 2015). In fact, occurrence of this splice site variant as well as the coding variant Asp67Tyr (rs61750900) were estimated to collectively explain over 24% of interindividual variability in cotinine glucuronidation (Patel et al., 2015). Our results support that alternative splicing at the *UGT2B10* locus may be a major factor contributing to this variability in the constitutive expression of the gene, with a potential impact on responses to substrates of the UGT2B10 pathway.

The alternate UGT2B10 proteins with novel in-frame C-terminal sequences are predicted to include the entire putative catalytic domain (Radomska-Pandya et al., 2010), although their enzyme activity could not be confirmed in standard *in vitro* assay conditions when expressed in HEK293 human cells. This could be due to an inadequate topology of the alt. proteins essential for enzyme function. *In silico* analysis of the novel exonic sequence with NCBI BLAST tools did not reveal a match with other nucleotide or amino acid sequences of any organism. The splicing event at the *UGT2B10* locus appears specific to humans. This is consistent with the low to undetectable activity towards preferred UGT2B10 substrates such as N-heterocyclic amines and aliphatic

DMD #79221

tertiary amines in most primates and other mammals, suggesting that *UGT2B10* expression occurs preferentially in humans (Kaivosaari et al., 2007; Kaivosaari et al., 2008; Zhou et al., 2010; Lu et al., 2017b). The novel appended amino acid sequence encoded in the alternative frame by exon 6 is however related to that of putative alternative variants of several other human proteins including CLCN3 (chloride exchange transporter 3), ALG9 (alpha-1, 2-mannosyltransferase), and CHFR (E3-ubiquitin ligase) supporting that the new sequence may encode a conserved domain (**Supplemental Figure 6**).

As for subcellular localization, our immunofluorescence data indicated that the alt. isoforms, partially or completely lacking the transmembrane domain and devoid of the positively charged lysine tail (**Figure 3**) are nonetheless ER-resident proteins when expressed in HEK293 human cells. Although studied in a limited set of UGTs, the high sequence similarity of ER retention elements among UGTs supports a common ER retention mechanism. ER residency of UGT enzymes is mediated by at least four protein regions: the N-terminal signal peptide, a short hydrophobic patch in the N-terminal substrate binding domain, the C-terminal transmembrane region and the cytoplasmic dilysine motif containing the ER retention signal "KXKXX" (K, lysines, X, any amino acids) (Jackson et al., 1993; Ciotti et al., 1998; Ouzzine et al., 1999; Barre et al., 2005). In line with our observations, these structural features may be partially redundant, as none appears strictly essential to ER retention. The Endo H sensitivity and lack of sensitivity towards O-glycosidase for each of the UGT2B10 enzyme and alt. UGT2B10 also support ER residency. Consistent with their co-localization, the potential of alt.

DMD #79221

UGT2B10 to interact with the UGT2B10 enzyme suggests that this may be a mechanism underlying their regulatory function.

The preferential stabilization of alt. isoforms by proteasomal inhibition and their shorter half-lives relative to the UGT2B10 enzyme suggest that distinct pathways may govern the turnover of alt. isoforms and the enzyme. This also suggests that a small variation in RNA levels has the potential to affect AS protein expression level. The significant abundance of alt. transcripts encoding these proteins may indicate that hepatic cells are poised to adapt levels of regulatory isoforms in response to various endogenous or exogenous stimuli. This is supported by the preferential mRNA expression of alternates in HepaRG cells treated with nuclear receptors agonists. While the structural determinants of UGT protein stability are scarcely known, mechanisms modulating UGT turnover may significantly contribute to the regulation of their detoxifying functions, a well-documented aspect for some drug metabolizing CYPs (Zhukov and Ingelman-Sundberg, 1999; Kim et al., 2016 and references therein). When expressed in the HEK293 model, half-lives of UGTs were more than 12-16 hours for UGT2B4, UGT2B7, UGT2B10 and UGT2B15 whereas UGT1A1 and UGT2B17 were more labile with half-lives less than 3 hours (Turgeon et al., 2001; Rouleau et al., 2016); this study). However, given that cycloheximide is glucuronidated by several UGT enzymes as demonstrated herein, alt. UGT protein half-lives established by translational inhibition with this compound in cells expressing UGTs likely constitute an inaccurate estimate that will be influenced by the enzymes expressed.

DMD #79221

A limitation of our study is the lack of detection of alt. UGT2B10 by immunoblotting in pooled liver microsomes from 50 individuals using the polyclonal anti-UGT2B10 antibody #1845. Their expression varied widely among individuals (CV of 155%), with some having no or barely detectable hepatic alt. *UGT2B10* variants. Levels of alt. proteins in our human liver pool may be too low for detection by immunoblotting. In contrast, the detection of alt. UGT2B10 by MS-MRM was performed following an immunoprecipitation step that enriched UGT2B10, thus improving the sensitivity of detection. The high turnover rate of the alt. proteins, as observed in the two cell models, may also influence our ability to detect them in human livers by immunoblotting. Additional experiments are required to ascertain the expression of alt. UGT2B10 proteins in individual human liver samples and their expression ratio relative to the UGT2B10 enzyme.

In conclusion, our study reveals that AS creates a diversified *UGT2B10* transcriptome and represents half of the UGT2B10 expression in the human liver, with a wide interindividual variability. Alternate UGT2B10 proteins may significantly influence the UGT2B10-dependent detoxification of amine-containing drugs such as anti-psychotic and tobacco metabolites and are expected to modulate endogenous substrates of the UGT2B10 enzyme, which are currently unknown. Our study further highlights a long-term stability of the UGT2B10 enzyme that contrasts with the lability of alt. proteins, the latter being regulated by proteasomal degradation. Most interestingly, we further expose a preferential induction by PXR and CAR inducers of alt. *UGT2B10* with novel in-frame C-terminal sequences in hepatic cells, implying a fine regulation of the AS process by xenosensing transcription factors. Our study highlights an important regulatory role of

DMD #79221

AS in UGT2B10 expression and detoxification functions that may explain part of the significant variability in N-glucuronidation, largely mediated by the UGT2B10 pathway in the liver. We thus believe that interindividual differences in the clinical response to UGT2B10 substrates are likely to be understood through the AS process affecting both the constitutive and inducible expression of UGT2B10.

DMD #79221

Acknowledgements

Authors wish to thank Dr H. Wang and B. Mackowiak (U. of Maryland) for their collaboration with the HepaRG data and Patrick Caron, Véronique Turcotte, Anne-Marie Duperré, Camille Girard-Bock and Andréa Fournier for their technical assistance. We also thank the Proteomics Platform of the CHU de Québec Research Center for their services.

Authorship contribution

Participated in research design: Labriet, Allain, Rouleau, Audet-Delage, Guillemette

Conducted experiments: Labriet, Allain, Audet-Delage, Villeneuve

Performed data analysis: Labriet, Allain, Rouleau, Audet-Delage, Villeneuve, Guillemette

Wrote or contributed to the writing of the manuscript: Labriet, Rouleau, Guillemette

Financial disclosure

All authors declare they have no competing interest

DMD #79221

References

- Antherieu S, Chesne C, Li R, Guguen-Guillouzo C, and Guillouzo A (2012) Optimization of the HepaRG cell model for drug metabolism and toxicity studies. *Toxicol In Vitro* **26**:1278-1285.
- Barre L, Magdalou J, Netter P, Fournel-Gigleux S, and Ouzzine M (2005) The stop transfer sequence of the human UDP-glucuronosyltransferase 1A determines localization to the endoplasmic reticulum by both static retention and retrieval mechanisms. *FEBS J* **272**:1063-1071.
- Beaulieu M, Levesque E, Hum DW, and Belanger A (1998) Isolation and characterization of a human orphan UDP-glucuronosyltransferase, UGT2B11. *Biochem Biophys Res Commun* **248**:44-50.
- Bellemare J, Rouleau M, Girard H, Harvey M, and Guillemette C (2010a) Alternatively spliced products of the UGT1A gene interact with the enzymatically active proteins to inhibit glucuronosyltransferase activity in vitro. *Drug Metab Dispos* **38**:1785-1789.
- Bellemare J, Rouleau M, Harvey M, and Guillemette C (2010b) Modulation of the human glucuronosyltransferase UGT1A pathway by splice isoform polypeptides is mediated through protein-protein interactions. *J Biol Chem* **285**:3600-3607.
- Berg JZ, Mason J, Boettcher AJ, Hatsukami DK, and Murphy SE (2010) Nicotine metabolism in African Americans and European Americans: variation in glucuronidation by ethnicity and UGT2B10 haplotype. *J Pharmacol Exp Ther* **332**:202-209.

DMD #79221

- Bolger AM, Lohse M, and Usadel B (2014) Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* **30**:2114-2120.
- Bray NL, Pimentel H, Melsted P, and Pachter L (2016) Near-optimal probabilistic RNA-seq quantification. *Nat Biotechnol* **34**:525-527.
- Bushey RT and Lazarus P (2012) Identification and functional characterization of a novel UDP-glucuronosyltransferase 2A1 splice variant: potential importance in tobacco-related cancer susceptibility. *J Pharmacol Exp Ther* **343**:712-724.
- Chen G, Blevins-Primeau AS, Dellinger RW, Muscat JE, and Lazarus P (2007) Glucuronidation of nicotine and cotinine by UGT2B10: loss of function by the UGT2B10 Codon 67 (Asp>Tyr) polymorphism. *Cancer Res* **67**:9024-9029.
- Chen G, Dellinger RW, Gallagher CJ, Sun D, and Lazarus P (2008a) Identification of a prevalent functional missense polymorphism in the UGT2B10 gene and its association with UGT2B10 inactivation against tobacco-specific nitrosamines. *Pharmacogenet Genomics* **18**:181-191.
- Chen G, Dellinger RW, Sun D, Spratt TE, and Lazarus P (2008b) Glucuronidation of tobacco-specific nitrosamines by UGT2B10. *Drug Metab Dispos* **36**:824-830.
- Chen G, Giambrone NE, and Lazarus P (2012) Glucuronidation of trans-3'-hydroxycotinine by UGT2B17 and UGT2B10. *Pharmacogenet Genomics* **22**:183-190.
- Chhibber A, French CE, Yee SW, Gamazon ER, Theusch E, Qin X, Webb A, Papp AC, Wang A, Simmons CQ, Konkashbaev A, Chaudhry AS, Mitchel K, Stryke D, Ferrin TE, Weiss ST, Kroetz DL, Sadee W, Nickerson DA, Krauss RM, George AL, Schuetz EG, Medina MW, Cox NJ, Scherer SE, Giacomini KM, and Brenner

DMD #79221

- SE (2016) Transcriptomic variation of pharmacogenes in multiple human tissues and lymphoblastoid cell lines. *Pharmacogenomics J. In press*.
- Chouinard S, Pelletier G, Belanger A, and Barbier O (2006) Isoform-specific regulation of uridine diphosphate-glucuronosyltransferase 2B enzymes in the human prostate: differential consequences for androgen and bioactive lipid inactivation. *Endocrinology* **147**:5431-5442.
- Ciotti M, Cho JW, George J, and Owens IS (1998) Required buried alpha-helical structure in the bilirubin UDP-glucuronosyltransferase, UGT1A1, contains a nonreplaceable phenylalanine. *Biochemistry* **37**:11018-11025.
- Court MH, Zhang X, Ding X, Yee KK, Hesse LM, and Finel M (2012) Quantitative distribution of mRNAs encoding the 19 human UDP-glucuronosyltransferase enzymes in 26 adult and 3 fetal tissues. *Xenobiotica* **42**:266-277.
- de Jonge HJ, Fehrmann RS, de Bont ES, Hofstra RM, Gerbens F, Kamps WA, de Vries EG, van der Zee AG, te Meerman GJ, and ter Elst A (2007) Evidence based selection of housekeeping genes. *PLoS One* **2**:e898.
- Fallon JK, Neubert H, Hyland R, Goosen TC, and Smith PC (2013) Targeted quantitative proteomics for the analysis of 14 UGT1As and -2Bs in human liver using NanoUPLC-MS/MS with selected reaction monitoring. *J Proteome Res* **12**:4402-4413.
- Fowler S, Kletzl H, Finel M, Manevski N, Schmid P, Tuerck D, Norcross RD, Hoener MC, Spleiss O, and Iglesias VA (2015) A UGT2B10 splicing polymorphism common in african populations may greatly increase drug exposure. *J Pharmacol Exp Ther* **352**:358-367.

DMD #79221

- Girard-Bock C, Benoit-Biancamano MO, Villeneuve L, Desjardins S, and Guillemette C (2016) A Rare UGT2B7 Variant Creates a Novel N-Glycosylation Site at Codon 121 with Impaired Enzyme Activity. *Drug Metab Dispos* **44**:1867-1871.
- Guillemette C, Levesque E, Harvey M, Bellemare J, and Menard V (2010) UGT genomic diversity: beyond gene duplication. *Drug Metab Rev* **42**:24-44.
- Guillemette C, Levesque E, and Rouleau M (2014) Pharmacogenomics of human uridine diphospho-glucuronosyltransferases and clinical implications. *Clin Pharmacol Ther* **96**:324-339.
- Haakensen VD, Biong M, Lingjaerde OC, Holmen MM, Frantzen JO, Chen Y, Navjord D, Romundstad L, Luders T, Bukholm IK, Solvang HK, Kristensen VN, Ursin G, Borresen-Dale AL, and Helland A (2010) Expression levels of uridine 5'-diphospho-glucuronosyltransferase genes in breast tissue from healthy women are associated with mammographic density. *Breast Cancer Res* **12**:R65.
- Jackson MR, Nilsson T, and Peterson PA (1993) Retrieval of transmembrane proteins to the endoplasmic reticulum. *J Cell Biol* **121**:317-333.
- Jones NR and Lazarus P (2014) UGT2B gene expression analysis in multiple tobacco carcinogen-targeted tissues. *Drug Metab Dispos* **42**:529-536.
- Kaivosaari S, Finel M, and Koskinen M (2011) N-glucuronidation of drugs and other xenobiotics by human and animal UDP-glucuronosyltransferases. *Xenobiotica* **41**:652-669.
- Kaivosaari S, Toivonen P, Aitio O, Sipila J, Koskinen M, Salonen JS, and Finel M (2008) Regio- and stereospecific N-glucuronidation of medetomidine: the differences between UDP glucuronosyltransferase (UGT) 1A4 and UGT2B10 account for the complex kinetics of human liver microsomes. *Drug Metab Dispos* **36**:1529-1537.

DMD #79221

- Kaivosaari S, Toivonen P, Hesse LM, Koskinen M, Court MH, and Finel M (2007) Nicotine glucuronidation and the human UDP-glucuronosyltransferase UGT2B10. *Mol Pharmacol* **72**:761-768.
- Kato Y, Izukawa T, Oda S, Fukami T, Finel M, Yokoi T, and Nakajima M (2013) Human UDP-glucuronosyltransferase (UGT) 2B10 in drug N-glucuronidation: substrate screening and comparison with UGT1A3 and UGT1A4. *Drug Metab Dispos* **41**:1389-1397.
- Kazmi F, Barbara JE, Yerino P, and Parkinson A (2015) A long-standing mystery solved: the formation of 3-hydroxydesloratadine is catalyzed by CYP2C8 but prior glucuronidation of desloratadine by UDP-glucuronosyltransferase 2B10 is an obligatory requirement. *Drug Metab Dispos* **43**:523-533.
- Kerdpin O, Mackenzie PI, Bowalgaha K, Finel M, and Miners JO (2009) Influence of N-terminal domain histidine and proline residues on the substrate selectivities of human UDP-glucuronosyltransferase 1A1, 1A6, 1A9, 2B7, and 2B10. *Drug Metab Dispos* **37**:1948-1955.
- Kim SM, Wang Y, Nabavi N, Liu Y, and Correia MA (2016) Hepatic cytochromes P450: structural deignons and barcodes, posttranslational modifications and cellular adapters in the ERAD-endgame. *Drug Metab Rev* **48**:405-433.
- Lepine J, Bernard O, Plante M, Tetu B, Pelletier G, Labrie F, Belanger A, and Guillemette C (2004) Specificity and regioselectivity of the conjugation of estradiol, estrone, and their catecholestrogen and methoxyestrogen metabolites by human uridine diphospho-glucuronosyltransferases expressed in endometrium. *J Clin Endocrinol Metab* **89**:5222-5232.

DMD #79221

- Levesque E, Beaulieu M, Green MD, Tephly TR, Belanger A, and Hum DW (1997) Isolation and characterization of UGT2B15(Y85): a UDP-glucuronosyltransferase encoded by a polymorphic gene. *Pharmacogenetics* **7**:317-325.
- Li D, Mackowiak B, Brayman TG, Mitchell M, Zhang L, Huang SM, and Wang H (2015) Genome-wide analysis of human constitutive androstane receptor (CAR) transcriptome in wild-type and CAR-knockout HepaRG cells. *Biochem Pharmacol* **98**:190-202.
- Lu D, Wang S, Xie Q, Guo L, and Wu B (2017a) Transcriptional Regulation of Human UDP-Glucuronosyltransferase 2B10 by Farnesoid X Receptor in Human Hepatoma HepG2 Cells. *Mol Pharm* **14**:2899-2907.
- Lu D, Xie Q, and Wu B (2017b) N-glucuronidation catalyzed by UGT1A4 and UGT2B10 in human liver microsomes: Assay optimization and substrate identification. *J Pharm Biomed Anal* **145**:692-703.
- Lu L, Zhou J, Shi J, Peng XJ, Qi XX, Wang Y, Li FY, Zhou FY, Liu L, and Liu ZQ (2015) Drug-Metabolizing Activity, Protein and Gene Expression of UDP-Glucuronosyltransferases Are Significantly Altered in Hepatocellular Carcinoma Patients. *PLoS One* **10**:e0127524.
- Margaillan G, Rouleau M, Klein K, Fallon JK, Caron P, Villeneuve L, Smith PC, Zanger UM, and Guillemette C (2015) Multiplexed Targeted Quantitative Proteomics Predicts Hepatic Glucuronidation Potential. *Drug Metab Dispos* **43**:1331-1335.
- Menard V, Collin P, Margaillan G, and Guillemette C (2013) Modulation of the UGT2B7 enzyme activity by C-terminally truncated proteins derived from alternative splicing. *Drug Metab Dispos* **41**:2197-2205.

DMD #79221

Murphy SE (2017) Nicotine Metabolism and Smoking: Ethnic Differences in the Role of P450 2A6. *Chem Res Toxicol* **30**:410-419.

Murphy SE, Park SS, Thompson EF, Wilkens LR, Patel Y, Stram DO, and Le Marchand L (2014) Nicotine N-glucuronidation relative to N-oxidation and C-oxidation and UGT2B10 genotype in five ethnic/racial groups. *Carcinogenesis* **35**:2526-2533.

NCBI dbSNP build 151 accessed Dec. 8, 2017
https://www.ncbi.nlm.nih.gov/projects/SNP/snp_ref.cgi

Ohno S and Nakajin S (2011) Quantitative analysis of UGT2B28 mRNA expression by real-time RT-PCR and application to human tissue distribution study. *Drug Metab Lett* **5**:202-208.

Ouzzine M, Magdalou J, Burchell B, and Fournel-Gigleux S (1999) An internal signal sequence mediates the targeting and retention of the human UDP-glucuronosyltransferase 1A6 to the endoplasmic reticulum. *J Biol Chem* **274**:31401-31409.

Patel YM, Stram DO, Wilkens LR, Park SS, Henderson BE, Le Marchand L, Haiman CA, and Murphy SE (2015) The contribution of common genetic variation to nicotine and cotinine glucuronidation in multiple ethnic/racial populations. *Cancer Epidemiol Biomarkers Prev* **24**:119-127.

Pattanawongsa A, Nair PC, Rowland A, and Miners JO (2016) Human UDP-Glucuronosyltransferase (UGT) 2B10: Validation of Cotinine as a Selective Probe Substrate, Inhibition by UGT Enzyme-Selective Inhibitors and Antidepressant and Antipsychotic Drugs, and Structural Determinants of Enzyme Inhibition. *Drug Metab Dispos* **44**:378-388.

DMD #79221

- Radomska-Pandya A, Bratton SM, Redinbo MR, and Miley MJ (2010) The crystal structure of human UDP-glucuronosyltransferase 2B7 C-terminal end is the first mammalian UGT target to be revealed: the significance for human UGTs from both the 1A and 2B families. *Drug Metab Rev* **42**:133-144.
- Rath A, Glibowicka M, Nadeau VG, Chen G, and Deber CM (2009) Detergent binding explains anomalous SDS-PAGE migration of membrane proteins. *Proc Natl Acad Sci U S A* **106**:1760-1765.
- Rouleau M, Roberge J, Bellemare J, and Guillemette C (2014) Dual roles for splice variants of the glucuronidation pathway as regulators of cellular metabolism. *Mol Pharmacol* **85**:29-36.
- Rouleau M, Tourancheau A, Girard-Bock C, Villeneuve L, Vaucher J, Duperre AM, Audet-Delage Y, Gilbert I, Popa I, Droit A, and Guillemette C (2016) Divergent Expression and Metabolic Functions of Human Glucuronosyltransferases through Alternative Splicing. *Cell Rep* **17**:114-124.
- Sato Y, Nagata M, Tetsuka K, Tamura K, Miyashita A, Kawamura A, and Usui T (2014) Optimized methods for targeted peptide-based quantification of human uridine 5'-diphosphate-glucuronosyltransferases in biological specimens using liquid chromatography-tandem mass spectrometry. *Drug Metab Dispos* **42**:885-889.
- Tourancheau A, Margaillan G, Rouleau M, Gilbert I, Villeneuve L, Levesque E, Droit A, and Guillemette C (2016) Unravelling the transcriptomic landscape of the major phase II UDP-glucuronosyltransferase drug metabolizing pathway using targeted RNA sequencing. *Pharmacogenomics J* **16**:60-70.

DMD #79221

- Tourancheau A, Rouleau M, Guauque-Olarte S, Villeneuve L, Gilbert I, Droit A, and Guillemette C (2017) Quantitative profiling of the UGT transcriptome in human drug-metabolizing tissues. *Pharmacogenomics J. In press*.
- Turgeon D, Carrier JS, Levesque E, Hum DW, and Belanger A (2001) Relative enzymatic activity, protein stability, and tissue distribution of human steroid-metabolizing UGT2B subfamily members. *Endocrinology* **142**:778-787.
- Turgeon D, Chouinard S, Belanger P, Picard S, Labbe JF, Borgeat P, and Belanger A (2003) Glucuronidation of arachidonic and linoleic acid metabolites by human UDP-glucuronosyltransferases. *J Lipid Res* **44**:1182-1191.
- Ware JJ, Chen X, Vink J, Loukola A, Minica C, Pool R, Milaneschi Y, Mangino M, Menni C, Chen J, Peterson RE, Auro K, Lyytikainen LP, Wedenoja J, Stiby AI, Hemani G, Willemsen G, Hottenga JJ, Korhonen T, Heliovaara M, Perola M, Rose RJ, Paternoster L, Timpson N, Wassenaar CA, Zhu AZ, Davey Smith G, Raitakari OT, Lehtimäki T, Kahonen M, Koskinen S, Spector T, Penninx BW, Salomaa V, Boomsma DI, Tyndale RF, Kaprio J, and Munafo MR (2016) Genome-Wide Meta-Analysis of Cotinine Levels in Cigarette Smokers Identifies Locus at 4q13.2. *Sci Rep* **6**:20092.
- Zhou D, Guo J, Linnenbach AJ, Booth-Genthe CL, and Grimm SW (2010) Role of human UGT2B10 in N-glucuronidation of tricyclic antidepressants, amitriptyline, imipramine, clomipramine, and trimipramine. *Drug Metab Dispos* **38**:863-870.
- Zhukov A and Ingelman-Sundberg M (1999) Relationship between cytochrome P450 catalytic cycling and stability: fast degradation of ethanol-inducible cytochrome P450 2E1 (CYP2E1) in hepatoma cells is abolished by inactivation of its electron donor NADPH-cytochrome P450 reductase. *Biochem J* **340** (Pt 2):453-458.

DMD #79221

Footnotes

This work was supported by the Canadian Institutes of Health Research [FRN-42392]; and the Canada Research Chair in Pharmacogenomics (Tier I); AL and EPA were supported by graduate scholarships from the "Fonds d'enseignement et de recherche" (FER) of the Faculty of pharmacy, Laval University (graduate scholarship); YAD was supported by a graduate scholarship from the "Fonds de Recherche en Santé du Québec". The Genotype-Tissue Expression (GTEx) Project was supported by the Common Fund of the Office of the Director of the National Institutes of Health, and by National Cancer Institute (NCI), National Human Genome Research Institute (NHGRI), National Heart, Lung, and Blood Institute, National Institute on Drug Abuse, National Institute of Mental Health, and National Institute of Neurological Disorders and Stroke and The Cancer Genome Atlas (TCGA) is managed by the NCI and NHGRI (<http://cancergenome.nih.gov>).

DMD #79221

Figure legends

Figure 1. Schematic overview of *UGT2B10* mRNA transcripts and encoded proteins. *Top.* The *UGT2B10* gene is comprised of 6 canonical exons and 2 alternative exons 6b and 6c. *Bottom.* The canonical *UGT2B10_v1* transcript encodes the *UGT2B10_i1* enzyme. Alternative splicing produces 3 classes of variants, with N-terminal or C-terminal truncations, and novel C-terminal sequence. The N-terminally truncated *UGT2B10_n3* and *_n4* variants are reported in the NCBI RefSeq *UGT2B10* gene entry (<https://www.ncbi.nlm.nih.gov/gene/7365>) to encode, respectively, isoforms i3 and i2 in RefSeq. Consequently, the isoform encoded by the transcripts *n9* and *n10* were named i4 and i5 respectively. Exons are represented by colored boxes and skipped exons by dashed boxes. Thinner parts of exons 1 and 6 represent untranslated regions. *UGT2B10* is the *UGT2B* with the largest 3'untranslated region and only the first 279 of 1424 nucleotides of exon 6 are coding in the canonical mRNA. Note that the predicted molecular weight of mature proteins is smaller by 2.5 kDa due to the cleavage of the signal peptide.

Figure 2. Alternative splicing diversifies the hepatic *UGT2B10* transcriptome. A. Relative levels of canonical (*v1*) and alt. transcript classes in the normal human liver and in hepatocellular carcinoma. The Δ C-term+new sequence class is predominantly comprised of *n9* and *n10* transcripts whereas transcript *n8* represented less than 2% in normal and tumor tissues. B. Interindividual variability in *UGT2B10* expression for all transcripts (Total), canonical (*v1*) and alt. transcripts (*n9/n10*). Boxes represent 25-75 percentiles, whiskers 10-90 percentiles. Median is indicated by the horizontal line and

DMD #79221

mean by a "+". CV, coefficient of variation; FC, fold change; N: normal tissues; T: tumor tissues; TPM: transcripts per million. C. Correlation between *UGT2B10_v1* and *n9/n10* expression in normal livers and hepatocellular carcinoma. All expression data were derived by a realignment of RNA-Seq data from GTEx (n = 50) and TCGA (n = 44) to the fully annotated UGT variant sequence (Tourancheau et al., 2016).

Figure 3. Alternate UGT2B10 protein is expressed in the human liver. A. C-terminal amino acid sequences of the UGT2B10 enzyme (i1) and alt. isoforms i4 and i5. Sequences unique to each alt. proteins are italicized. TMD: transmembrane domain; K^+ : ER retention signal. **B.** Experimental approach for the detection of UGT2B10_i5 by immunoprecipitation and multiple reaction monitoring (MRM). **C.** The common UGT2B10 peptide NSWNFK (left) and the alt. specific peptide LLGSSNNPPILASQR (right) were detected in tryptic digests of UGT2B10 immunoprecipitated from human liver samples (upper chromatograms). The chromatograms of control peptides (lower chromatograms) labelled with stable isotopes mixed with the immunopurified UGT2B10 confirmed the identity of i1 and i5 peptides. Representative chromatograms are shown (n = 2).

Figure 4. Interaction of UGT2B10 isoforms in the ER. A. Stable expression of UGT2B10 alternative isoforms in HEK293 and HepG2 cell models. UGT2B10 was detected in microsomal fractions of each cell model (20 µg proteins) and of human liver (L) using the anti-UGT2B10 antibody. The expression ratio of the UGT2B10 enzyme relative to alt. isoforms is estimated to more than 4 fold in each model. Note that the apparent molecular weight of the enzyme was lower than predicted, likely explained by the influence of the hydrophobic amino acids in the transmembrane domains (TMD) that

DMD #79221

perturb the interactions with SDS and the shape of the denatured protein (Rath et al., 2009). **B.** UGT2B10 enzyme and alt. isoforms (i4 and i5) co-localized with the ER marker protein disulfide isomerase (PDI). Merged images of UGT2B10 proteins labelled with anti-UGT2B10 (green), PDI (red), and nuclei (blue) are shown. Separate images and co-localization with other subcellular markers are shown in **Supplemental Figure 3**. Bar = 10 μ m. **C.** Glycosylation status of the UGT2B10 enzyme and alt. proteins. UGT2B10 proteins in microsomes from HEK293 cell models subjected to endoglycosidase H (Endo H) or O-glycosidase were detected by immunoblotting.

Figure 5. The alternative UGT2B10 isoforms modulate the glucuronidation activity of the UGT2B10 enzyme. A-B. *In situ* cell glucuronidation assays of amitriptyline and levomedetomidine in HEK293 and HepG2 cell models that co-express the alt. isoforms i4 and i5. Activity is presented relative to that in cells expressing only the UGT2B10 enzyme (control). Assays were conducted twice in triplicates. *t*, $P \leq 0.1$; *, $P \leq 0.05$; **, $P \leq 0.01$; ***, $P \leq 0.001$. **C.** Co-immunoprecipitation of the UGT2B10 enzyme with alt. isoforms i4 and i5 expressed in HEK293. Alternative isoforms were C-terminally tagged with the epitope V5 and immunoprecipitated (IP) with the anti-V5 antibody.

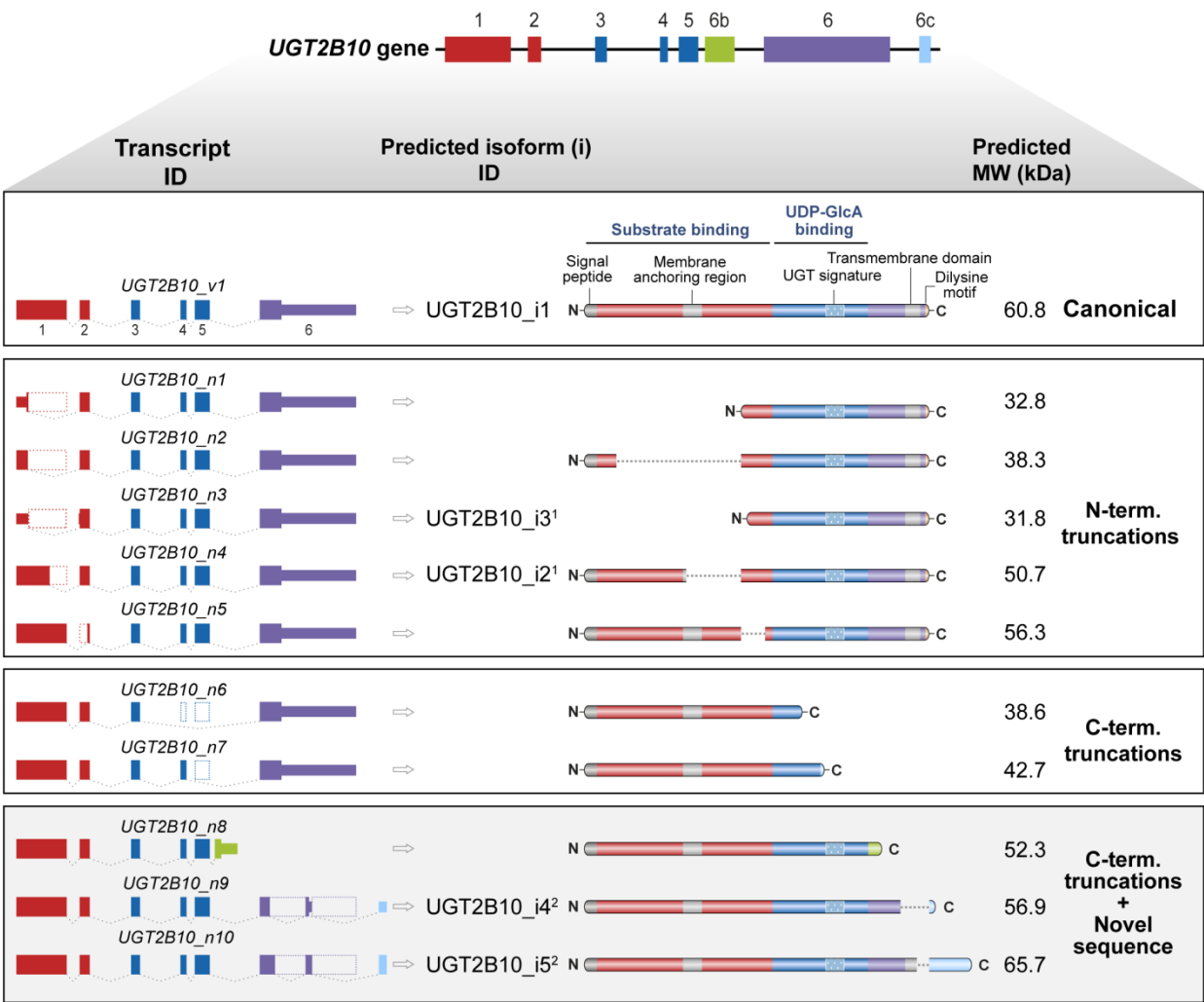
Figure 6. Stability of UGT2B10 proteins. A. The half-life of UGT2B10 proteins was determined by translational inhibition with cycloheximide (CHX) in HEK293 and HepG2 cell models. Protein levels were measured by densitometry scanning of immunoblots from two independent assays and are expressed as % of untreated cells. Half-lives were averaged from the two independent assays. A representative experiment is shown. An independent replicate assay is presented in **Supplemental Figure 4**. **B.** UGT2B10

DMD #79221

alternative isoforms are stabilized by proteasomal inhibition. HEK293 and HepG2 cell models were exposed to vehicle only (DMSO; lane "-") or MG132 (1 μ M; lanes "+") for 16 h. UGT2B10 was detected by immunoblotting total cell lysates (40 μ g) with anti-UGT2B10.

Figure 7. Differential induction of alt. UGT2B10 by phenobarbital and a CAR agonist in HepaRG cells. Expression data of *UGT2B10* canonical and alt. transcripts *n9* and *n10* (encoding i4 and i5) in HepaRG cells wild type (WT) or with a knock-out of the constitutive androstane receptor (CAR KO) were obtained from the public RNA-seq data GSE 71446. Cells were treated either with vehicle (DMSO), phenobarbital (PB, 1 mM) or the CAR agonist CITCO (1 μ M) for 24 h as described (Li et al., 2015). **, $P \leq 0.01$; ***, $P \leq 0.001$.

Figure 1



¹Nomenclature from RefSeq *Homo sapiens* annotation release 108 (<https://www.ncbi.nlm.nih.gov/gene/7365>)

²Nomenclature from Tourancheau et al. (2017)


 Focus of this study

Figure 2

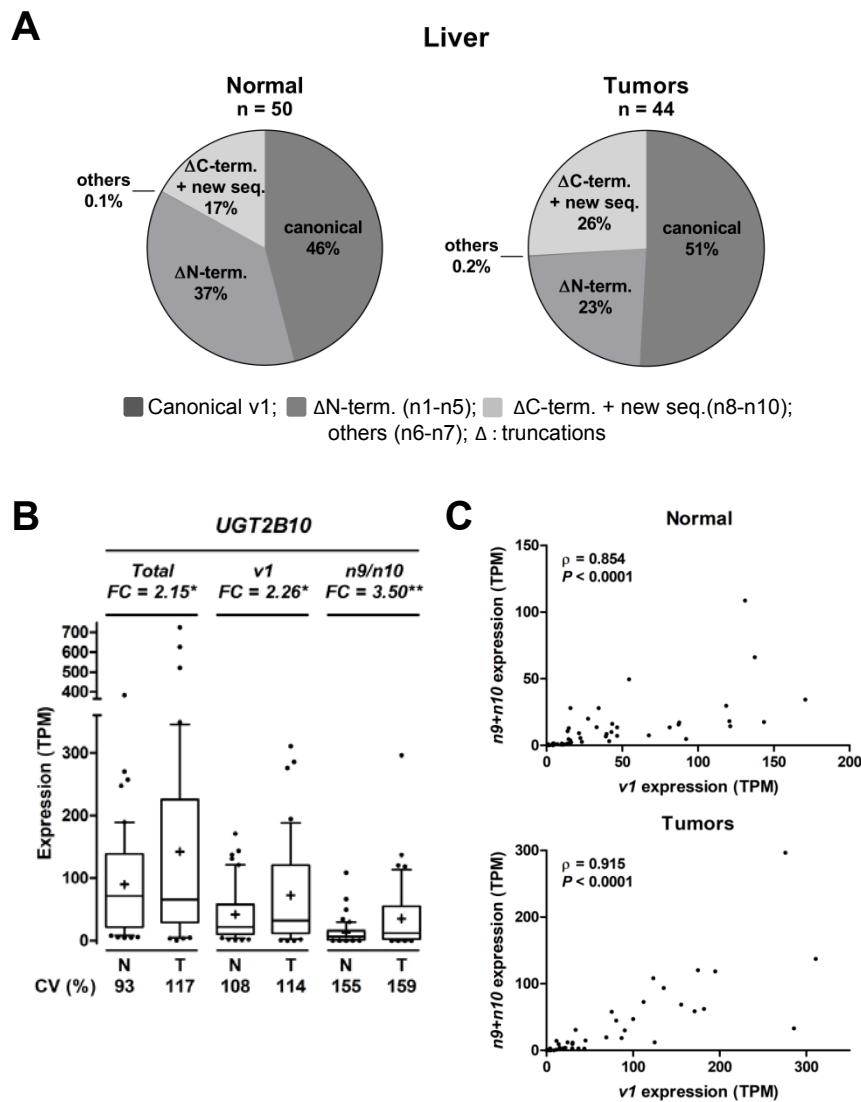


Figure 3

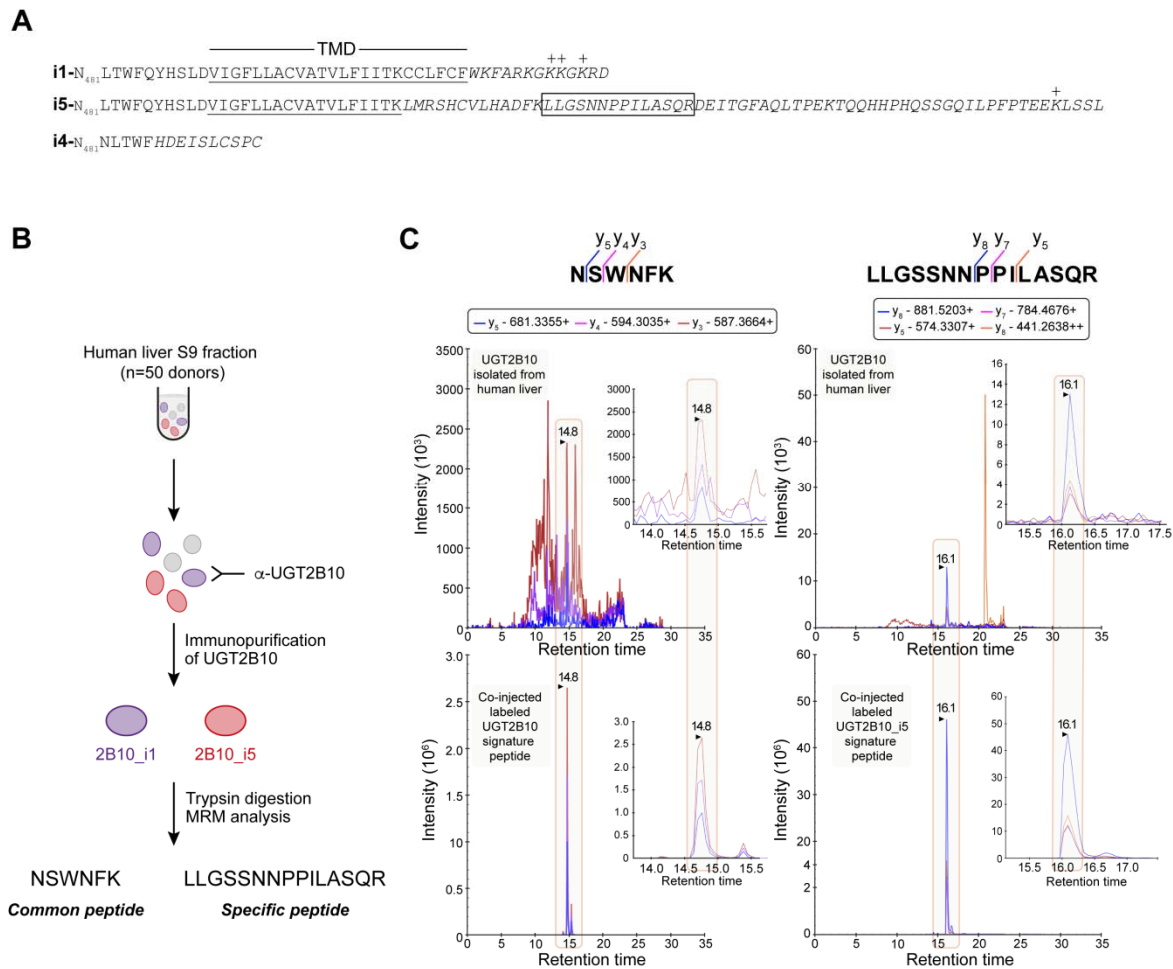


Figure 4

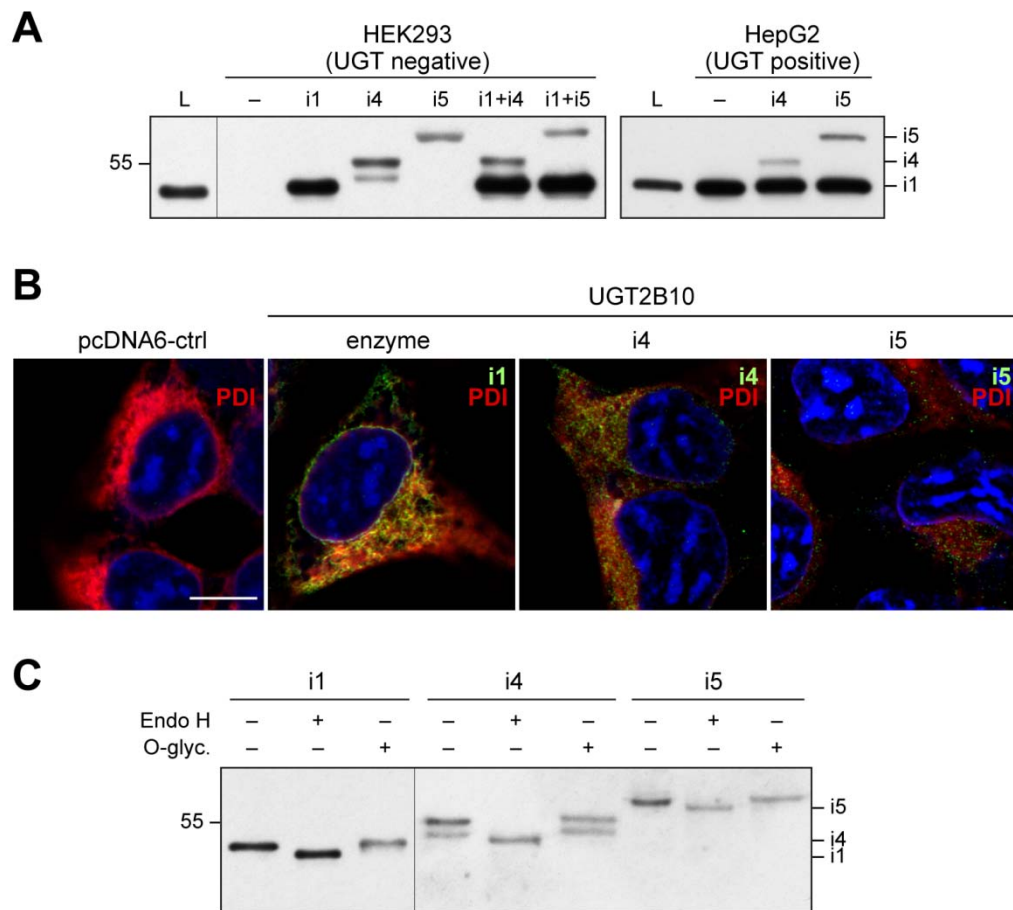


Figure 5

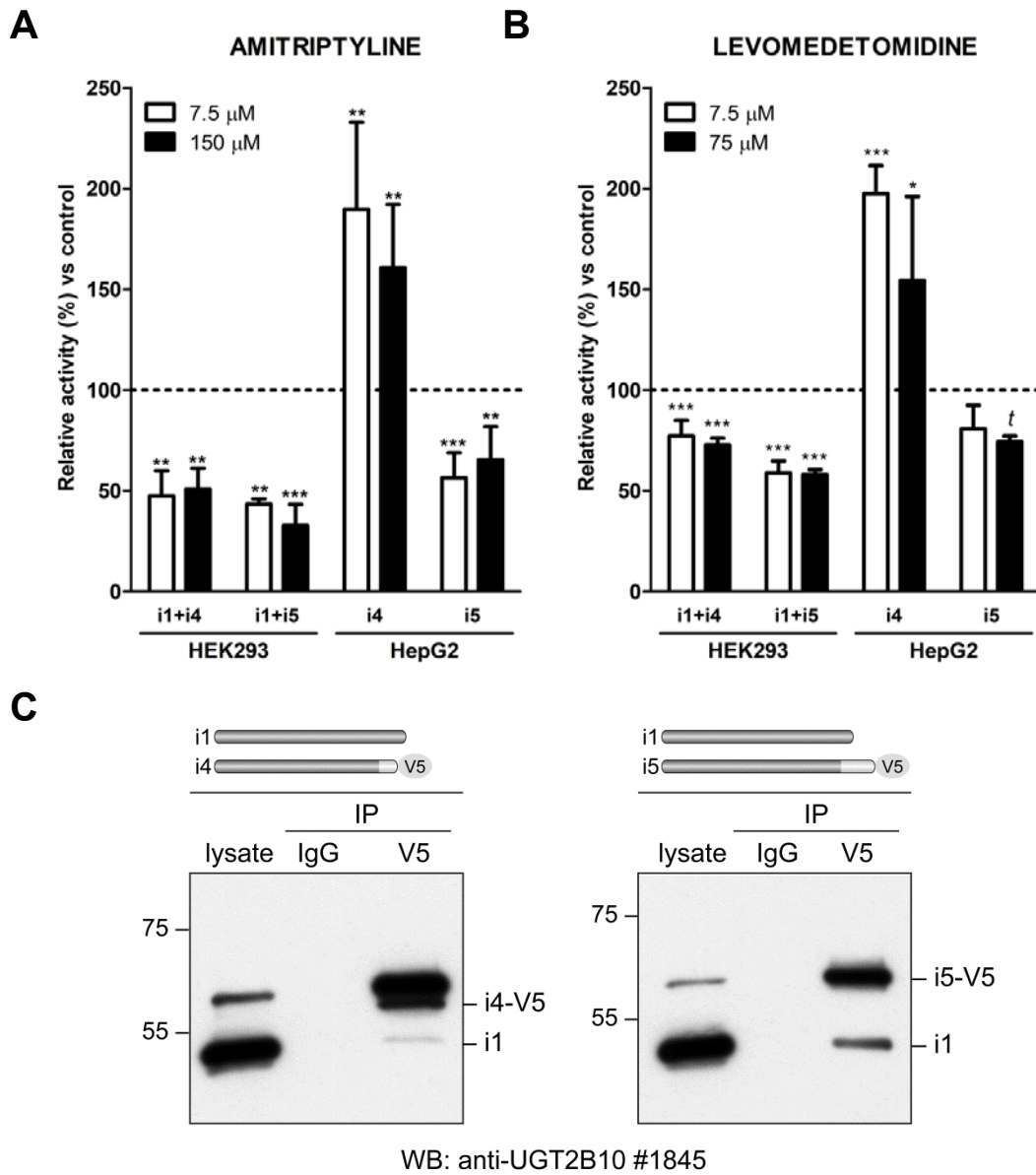


Figure 6

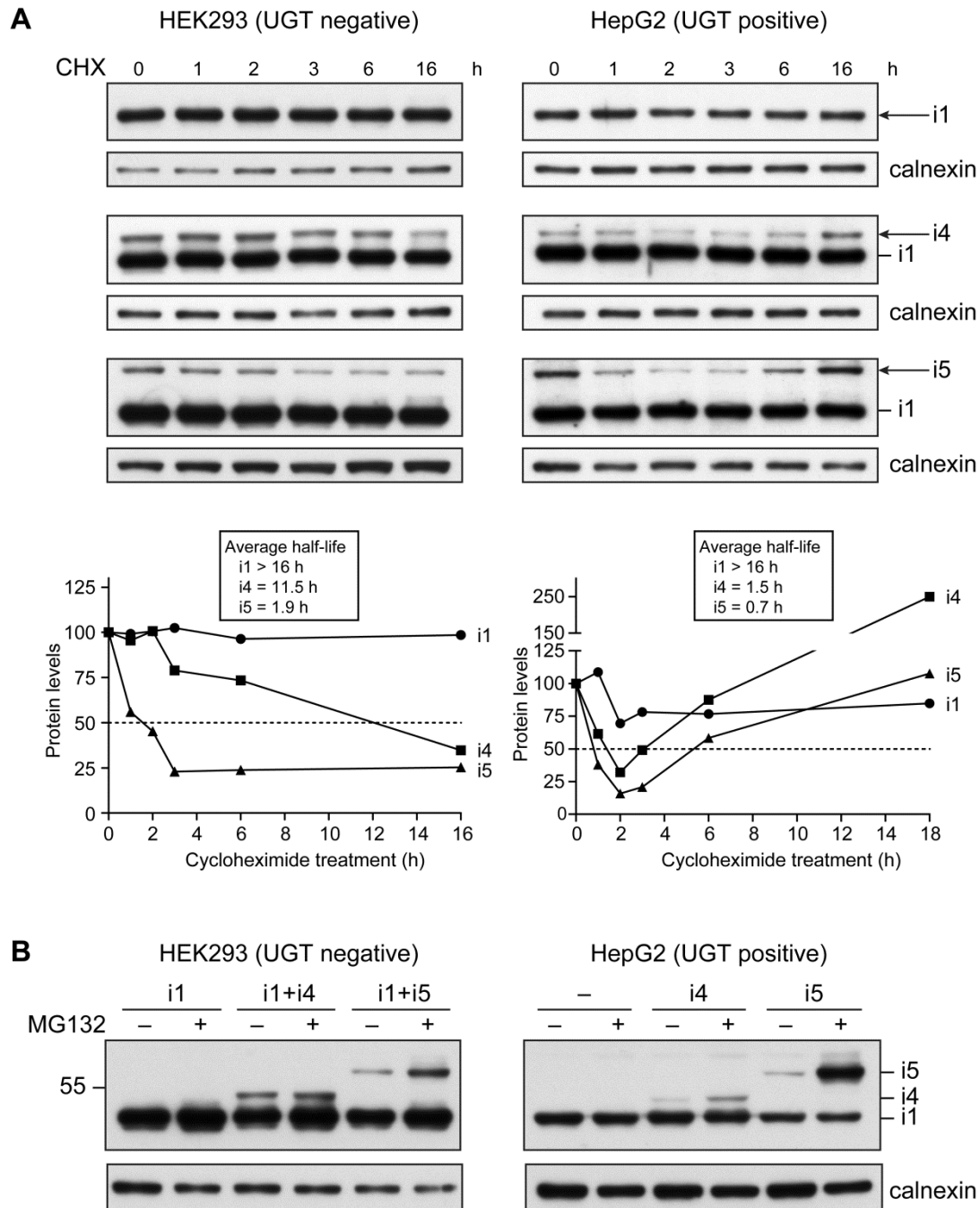


Figure 7

