

DMD #30635

Molecular cloning of the baboon UDP-glucuronosyltransferase 2B gene family and their activity in conjugating morphine

Kirsten Abildskov, Piper Weldy, and Marianne Garland

Division of Neonatology, Department of Pediatrics, Columbia University, New York, New York

(K.A., P.W., M.G.)

DMD #30635

Running Title: Baboon UGT2B gene family

Address Correspondence to:

Kirsten Abildskov
Columbia University
Irving Institute for Clinical and Translational Research
630 West 168th Street, PH 10-106
New York, NY 10032

FAX: 212-305-3213
Phone: 212-305-9306
EMail: kma1@columbia.edu

Number of:

Text Pages: 37
Tables: 4
Figures: 3
Words in Abstract: 246
Words in Introduction: 497
Words in Discussion: 1843

List of Nonstandard Abbreviations:

UGT, UDP-glucuronosyltransferase; M3G, morphine-3- β -glucuronide; M6G, morphine-6- β -glucuronide; DEPC, diethyl pyrocarbonate; PCR, polymerase chain reaction; HEK293, human embryonic kidney 293; DMEM, Dulbecco's modified Eagle's medium; FBS, fetal bovine serum; HPLC, high performance liquid chromatography; N-glycosylation, asparagine-linked glycosylation.

ABSTRACT

Glucuronidation by UDP-glucuronyltransferase 2B enzymes (UGT2Bs) is a major pathway for the elimination of endobiotics and xenobiotics, including therapeutic drugs. Morphine, a probe drug for UGT2B7, is metabolized to morphine-3- β -glucuronide (M3G) and morphine-6- β -glucuronide (M6G) in humans. Morphine has been used in a series of experiments in the baboon to characterize developmental changes in fetal glucuronidation. This study identifies the baboon UGT2B family of enzymes, compares them to that of the human and the monkey (*Macaca fascicularis*), and measures the activity of the individual baboon UGT2Bs toward morphine. UGT2B cDNAs were cloned from the liver of adult and newborn baboons and expressed in HEK293 cells. The UGT activity toward morphine was assessed by the rate of formation of M3G and M6G by HPLC. Eight baboon UGT2Bs were cloned and identified: UGT2B41 and UGT2B42, that are 90% homologous to human UGT2B4; UGT2B43 that is 93% homologous to human UGT2B15; and UGT2B39, 40, 44, 45, and 46 that are 89 to 91% homologous to human UGT2B7. Homology between baboon and monkey UGT2B ranged from 92.6 to 99.1%, with the primary protein structure of UGT2B43 being 99.1% identical to monkey UGT2B20 including a unique R96I substitution. Gene conversion interfered with the phylogenetic signal in the baboon UGT2B7-like and the monkey UGT2B4-like groups and led to concerted evolution of these enzymes. All baboon UGT2Bs metabolized morphine to both M3G and M6G. This study lays the foundation for investigating the regulation of UGT2B enzymes during fetal and neonatal development in the baboon.

INTRODUCTION

The major clearance pathway for morphine in humans is glucuronidation. UDP-glucuronosyltransferases (UGTs), the superfamily of genes encoding enzymes that catalyze glucuronidation, undergo considerable regulation during development in humans as well as other species (Dutton, 1978; Dutton and Leakey, 1981; de Wildt et al., 1999; Strassburg et al., 2002; Caspersen et al., 2007). Morphine has been used in a series of experiments in the baboon to investigate the impact of the developmental changes in glucuronidation on fetal drug and metabolite disposition (Garland et al., 2005; Garland et al., 2006). To support the use of this model in extrapolating results from the baboon to humans, it is important to delineate the differences in these drug metabolizing enzymes between the species.

UDP-glucuronosyltransferases have been classified into two major families, UGT1 and UGT2, based on the similarity of their primary structures. Currently, a total of 19 functional human UGT1 and UGT2 isoforms have been identified and mainly categorized into the two sub-families, UGT1A and UGT2B (Mackenzie et al., 2005). UGTs from both of these sub-families of enzymes have been identified as major pathways for the elimination of drugs from all therapeutic classes in humans (Green et al., 1998; Barbier et al., 2000a; Court et al., 2001; Court et al., 2003; Soars et al., 2003).

Morphine is considered a probe drug for UGT2B7. In humans, morphine is metabolized to morphine-3- β -glucuronide (M3G) and morphine-6- β -glucuronide (M6G) (Coffman et al., 1997; Court et al., 2003). Although human UGT2B4, UGT2B15 and UGT2B17 have also been shown

DMD #30635

to glucuronidate morphine to M3G, this is to a much lesser extent than human UGT2B7. Human UGT2B4 was also found to glucuronidate morphine to M6G, but again to a lesser extent than UGT2B7 (Court et al., 2003; Ohno et al., 2008).

The recent molecular cloning of the UGT1A family of genes in the baboon revealed considerable structural and functional conservation between the baboon and human UGT1As (Caspersen et al., 2007). In addition, earlier studies have demonstrated that both the baboon and the cynomolgus monkey (*Macaca fascicularis*) show similar patterns of glucuronidation of drugs and steroids (Garland et al., 1996; Garland et al., 1998a; Garland et al., 1998b; Barbier and Belanger, 2003). Together, these data provide evidence that the baboon is a clinically relevant model for studying perinatal drug metabolism. However, more recent studies in the baboon show surprisingly high metabolic clearance of morphine by the fetus and seemed at odds with data from human infants warranting a closer examination of species differences (Chay et al., 1992; Hartley et al., 1993; Scott et al., 1999; Garland et al., 2005; Garland et al., 2006).

The identification of the baboon UGT2B enzymes involved in morphine glucuronidation is an important step toward the understanding of the species differences in drug metabolism. The baboon UGT2B isoforms have therefore been cloned, the primary protein structures compared to that of human and monkey (*Macaca fascicularis*) UGT2Bs, and the phylogenetic relationship between these species examined. Finally, the activity of the individual baboon UGT2Bs toward morphine has been assessed.

MATERIALS AND METHODS

Liver tissue collection

Tissues from adult and newborn baboons (*Papio anubis*) were made available for this study from a breeding colony maintained by the investigators and the Institute of Comparative Medicine at Columbia University, New York, NY. All collections were done in accordance with animal care protocols approved by the Institutional Animal Care and Use Committee of the Columbia University and in accordance with guidelines of the US Department of Agriculture and the National Institutes of Health. Liver tissues were quickly cut into smaller fragments, immediately immersed in liquid nitrogen, and stored at -80°C until use.

RNA extraction

Total RNA was extracted from 50-100 mg liver tissue using the Trizol reagent (Invitrogen, Carlsbad, CA, USA) according to the manufacturer's instructions and the resulting pellet was re-suspended in water treated with diethyl pyrocarbonate (DEPC, Invitrogen). The RNA was quantified by absorbance at 260 nm, and sample integrity was assessed by the appearance of 28S and 18S rRNA on agarose gels.

cDNA synthesis and PCR

DMD #30635

cDNA was prepared from 3 μg total RNA, in 10 μl total reaction volume, using the Superscript III first strand synthesis kit (Invitrogen) and Oligo-dT primers. The cDNA (1 μl) was amplified by polymerase chain reaction (PCR) on a PTC-100 (MJ Research, Waltham, MA, USA) using PfuUltra polymerase (Stratagene, La Jolla, CA, USA) in a total reaction volume of 50 μl containing 0.5 μM primers (Invitrogen), 250 μM dNTPs (Amersham Biosciences, GE Healthcare, NJ, USA), and PfuUltra reaction buffer. A modified hotstart and stepdown PCR protocol was used to minimize nonspecific reactions. The thermal cycling protocol was as follows: pre-incubation at 95°C for 10 min, hotstart at 80°C, 35 cycles of denaturing at 94°C for 1 min, annealing stepwise down from 60 to 54°C for 1 min, elongation at 72°C for 1 min, followed by a final elongation step at 72°C for 10 min.

PCR primers were designed to amplify UGT2B genes in the baboon based on published mRNA sequences of human and *Macaca fascicularis* UGT2B genes. The following GenBank accession numbers were used: human UGT2B4 (Accession# BC026264), human UGT2B7 (Accession# BC030974), human UGT2B10 (Accession# NM001075), human UGT2B11 (Accession# NM001073), human UGT2B28 (Accession# NM053039), human UGT2B15 (Accession# AF180322), *Macaca fascicularis* UGT2B19 (Accession# AF112112), *Macaca fascicularis* UGT2B30 (Accession# AF401657), *Macaca fascicularis* UGT2B9 (Accession# U91582), *Macaca fascicularis* UGT2B18 (Accession# AF016310), and *Macaca fascicularis* UGT2B23 (Accession# AF112113).

DMD #30635

Primers were designed from the untranslated regions of a consensus sequence using Primer3 software (<http://frodo.wi.mit.edu>). All forward primers contained the CACC sequence at the 5' end to allow for directional cloning. The primers used were:

F-UGT2B4/2B7:	CACCTTGCATTGCAMCAGGATG
F-UGT2B15:	CACCYGMRTAAGACCAGGATG
R-UGT2B4:	YSCAGCTTCCARCCTCA
R-UGT2B7:	GTTTTCCAGCTTCAAATCTC
R-UGT2B15:	ATCCACTTCAGGCTTTTGA

where M= C+A, Y= C+T, S= C+G, and R= A+G.

cDNA cloning

The PCR products were extracted from agarose gels and purified using the QIAEX II gel extraction kit (Qiagen, Valencia, CA, USA). Purified PCR products were cloned into the pcDNA6.2/V5/GW/D-TOPO vector and transformed into TOP10-E.coli competent cells using the pcDNA Gateway Directional TOPO Expression Kit (Invitrogen). Plasmids were purified using the Plasmid Miniprep Spin Kit (Qiagen). Inserts were sequenced in the 5' and 3' directions at the Columbia University DNA Analysis and Sequencing Facility. cDNA sequences were first screened for quality with the Chromas Lite sequence viewer (<http://www.technelysium.com.au>) and then assembled with CodonCode Aligner software (CodonCode Corporation, Dedham, MA,

DMD #30635

USA). The UGT Nomenclature Committee approved the baboon UGT2B names and the sequences were submitted to GenBank.

Stable expression of UGT2B cDNA

HEK293 cells (American Type Culture Collection, ATTC, Manassas, VA, USA) were grown to 50% confluence in 24 well plates containing Dulbecco's Modified Eagle's Medium (DMEM) with 10% Fetal Bovine Serum (FBS, Invitrogen). The cells were transfected using Fugene (Roche, Indianapolis, IN, USA) and 1 µg of empty (control) or UGT2B cDNA containing plasmid. Transfected cells were allowed to grow for 2 days after which the cells were passed to 10 cm diameter plates and continued to be grown in DMEM with 10% FBS, 1% Penicillin/Streptomycin (Invitrogen), and 5 µg/ml Blastidicin (Invitrogen) for selective growth of transfected cells. Cells were passed until several plates of viable cells were at least 80% confluent, at which point the cells were harvested.

Heavy membrane fraction isolation

The cells were harvested into ice cold lyzing buffer consisting of 150 mM KCl, pH 7.2, 10 mM HEPES, 0.5 mM dithiothreitol (Sigma-Aldrich, MO, USA), and with 1 mini-tablet Complete Protease Inhibitor Cocktail (Roche, Indianapolis, IN, USA) per 25 ml lyzing buffer. The harvested cells were lyzed with an Ultra-Turrax T8 homogenizer (IKA Works, Inc., Wilmington, NC, USA) in four 1 sec bursts on ice. Lyzed cells were centrifuged for 7 min at 800 x g and 4°C. The supernatant fraction was saved, and the pellet re-suspended in 5 ml of fresh cold lyzing

DMD #30635

buffer. The lysis procedure and centrifugation were repeated as above. The combined supernatant fractions were then centrifuged for 30 min at 17,000 x g at 4°C and all supernatant was removed. The pellet was re-suspended in 300-500 µl of 150 mM KCl, pH 7.2 containing 10 mM HEPES. Protein concentrations were measured using the BioRad Protein Assay micro-titerplate protocol (BioRad, Hercules, CA, USA) and bovine serum albumin (Sigma-Aldrich) as a standard.

Morphine glucuronidation assay

A glucuronidation assay using morphine as substrate and measuring the metabolites M3G and M6G was adapted from that described by Fisher (Fisher et al., 2000). Glass incubation tubes 12 x 17 mm (Fisher Scientific, Pittsburgh, PA, USA) were prepared with 10 µg of the pore-forming polypeptide alamethicin (Sigma-Aldrich) just before use by drying down 4 µl of 5 mg alamethicin in 2 ml methanol under nitrogen. 100 µg of recombinant UGT2B protein in 100 µl of 0.1 M phosphate buffer, pH 7.4, was added to each tube and placed on ice for 15 min. 50 µl of 4 mM MgCl₂, 20 mM morphine (Spectrum, New Brunswick, NJ, USA), and 0.1 M phosphate buffer, pH 7.4 were then added and the tubes were pre-incubated in a water bath at 37°C for 3 min.

The reaction was initiated by the addition of 50 µl of 20 mM UDP-glucuronic acid (Sigma-Aldrich) in 0.1 M phosphate buffer, pH 7.4 and stopped after exactly 30 min for each sample by the addition of 150 µl ice-cold 100% acetonitrile, HPLC grade (Fisher Scientific). The tubes were placed on ice for 30 min. After addition of 10 µl of the internal standard nalorphine

DMD #30635

(Sigma-Aldrich) at 1 mg/ml, the samples were mixed and transferred to 1.5 ml Eppendorf centrifuge tubes and spun at 10,000 x g for 5 min at 4°C. The supernatants were removed and diluted 1:10 with 24% acetonitrile, HPLC grade prior to being assayed by HPLC. All incubations were performed in duplicate. Preliminary studies indicated that the formation of both M3G and M6G were linear up to 140 µg of protein and up to 50 min of incubation.

HPLC assay for M3G and M6G

The samples were analyzed for M3G and M6G by direct injection into a Waters Alliance 2695 HPLC System (Waters, Milford, MA, USA) with fluorometric and coulochemical detection as previously described (Garland et al., 2006). Briefly, M3G was measured by fluorescence detection on a FluorLC305 Fluorescence Detector (Linear Instruments, ESA, Bedford, MA, USA) with an excitation wavelength of 210 nm and an emission wavelength of 340 nm. M6G was measured by coulochemical detection on a Coulochem II Electrochemical Detector with a 5011 analytical cell with potentials of +225mV and +350mV for E1 and E2 respectively. The internal standard, nalorphine, was measured by both detectors. Separation was performed on a Spherisorb C18 (ODS2, 3µm, 4.6 x 100 mm ID) column (Waters) at ambient temperature. The injection volume was 50µl. The isocratic elution was performed with a mobile phase consisting of 10 mM sodium phosphate monobasic (Fisher Scientific), 1.5 mM sodium dodecyl sulfate (Invitrogen), pH 2.1 containing 24% acetonitrile HPLC grade at the rate of 1.5 ml/min.

Eight (8) point external standard curves ranging from 100 - 15,000 ng/100 µl were prepared from M3G and M6G obtained from the National Institute on Drug Abuse, Division of Neuroscience

DMD #30635

and Behavioral Research (Bethesda, MD). The Waters system used Empower Software for data collection and analysis. Standard curves were generated using the metabolite/internal standard area ratios. The formation of M3G and M6G was then quantified by comparison of the area ratio (metabolite/internal standard) of the incubated samples to that of the standard curves. Standard curve correlation coefficients (r) were > 0.999 for both M3G and M6G. The UGT activity was calculated by dividing the metabolite concentration by the protein concentration and the incubation time and expressed as nmoles/minute/mg of protein.

Sequence analysis.

The cloned baboon UGT2B cDNAs were compared to those of both human and monkey (*Macaca fascicularis*) UGT2B cDNAs, which were obtained from the GenBank, by multiple alignments of the coding regions using the software suite MEGA4.0 (Tamura et al., 2007). The protein sequences were translated and aligned. The distance matrixes were calculated using the pair-wise distance calculation. A distance based neighbor-joining phylogenetic tree was generated. The reliability of the phylogeny was tested by bootstrap analysis with 1000 replicates. The bootstrap test produced a final reconstructed neighbor-joining consensus tree. Multiple alignments of the baboon UGT2B protein sequences for Fig. 2 were performed using MultAlin software (<http://bioinfo.genopole-toulouse.prd.fr/multalin/multalin.html>).

The baboon, human, and monkey UGT2B cDNA sequences were analyzed for gene conversion events using GENECONV version 1.81 (<http://www.math.wustl.edu/~sawyer/geneconv/>) (Sawyer, 1989). GENECONV analyzes multiple alignments and uses statistical methods for

DMD #30635

inferring conversion events based on the distribution of substitutions. Fragments that potentially originated from gene conversion were identified and statistical significance ($p < 0.05$) of the fragments given based on 10,000 permutations. GENCONV determined both global and pairwise fragments. The global P-values were corrected for multiple comparisons and are a stronger predictor of gene conversion than the pairwise.

RESULTS

cDNA cloning

Eight novel UGT2Bs clones were isolated from the livers of adult and newborn baboons. The GenBank accession numbers and the official names are listed in Table 1. The cDNAs encoding the baboon UGT2Bs varied in length from 1626 to 1634 bp with open reading frames of 1584 to 1590 bp, which translates into proteins of 528 to 530 amino acids. A comparison of the baboon UGT2B cDNAs to human UGT2B homologs showed amino acid identities ranging from 89% to 93%.

Phylogenetic analysis

The amino acid sequences that were deduced from nucleotides of the corresponding baboon cDNAs were aligned with those of the cynomolgus monkey and human UGT2Bs and distance matrixes were calculated (Table 2). The baboon UGT2B enzymes share a greater than 74% sequence similarity, which is typical for the UGT2 family of enzymes (Mackenzie et al., 2005). The baboon UGT2Bs fall into three groups: UGT2B43 is the most highly conserved baboon UGT2B with a 92.1% and 90% identity to human UGT2B15 and UGT2B17 respectively; UGT2B41 and UGT2B42, are 88.3 and 88.5% identical to human UGT2B4; and UGT2B39, UGT2B240, UGT2B44, UGT2B45, and UGT2B46, are from 88.8 to 90.2% identical to human UGT2B7 (Table 2). These homologies ranging from 87% to 92% are similar to those found in a comparison between monkey and human UGT2B genes (Belanger et al., 1997; Beaulieu et al.,

DMD #30635

1998; Belanger et al., 1999; Barbier et al., 1999a; Barbier et al., 1999b; Girard et al., 2002). No baboon UGT2Bs similar to human UGT2B10, UGT2B11 or UGT2B28 are obvious.

An overall higher level of identity was found, as expected, between the baboon and the cynomolgus monkey UGT2Bs. The average sequence identity was 95.4% (ranging from 92.6 to 99.1%) with the highest level of identity (99.1%) found between baboon UGT2B43 and monkey UGT2B20. The same groupings also exist in the baboon and the monkey UGT2Bs. A comparison between the baboon and the two UGT2B isoforms isolated to date from another closely related *Macaca* species, the rhesus monkey (*Macaca mulatta*), show a similar high identity (Dean et al., 2004). These isoforms, UGT2B9*2 and UGT2B33, show an identity ranging from 93% to 94.3% and 95.3% to 99% to the baboon UGT2B7 like isoforms, respectively. Mouse and rat Ugt2b isoforms on the other hand are more distantly related with identities to the baboon UGT2Bs ranging from 61.5% to 72%. (Sequences for rhesus monkey, mouse, and rat proteins were all obtained from the GenBank).

A phylogenetic tree was reconstructed to further illustrate the relationship between the baboon, cynomolgus monkey, and human UGT2B enzymes (Fig. 1). Overall the baboon and the monkey UGT2Bs interleaf on the phylogenetic tree. Multiple duplications have occurred in both the baboon and the monkey within the UGT2B4-like and the UGT2B7-like groups. This is in contrast to the single branches formed by the human UGT2B4 and UGT2B7. On the other hand, the UGT2B15-like group has a single duplication in the human that is not replicated in the baboon or the monkey. The human UGT2B10, UGT2B11, and UGT2B28 remain a

DMD #30635

monophyletic clade of paralogs, even though the cloning primers were designed to probe for all possible baboon UGT2B enzymes.

The high level of sequence similarity of the UGT2Bs, which is generally true of the UGT2 gene family, complicates identifying true orthologs between species (Table 2). The UGT2 genes are, as a result of this, sequentially numbered based on the chronological order of discovery (Mackenzie et al., 2005). The within species duplication/deletion events make it not possible to suggest orthologous relationships between human and old world monkey UGT2Bs (Fig. 1).

While it is possible to suggest an orthologous relationship within the old world monkeys between baboon UGT2B43 and monkey UGT2B20 in the UGT2B15-like group, the phylogenetic relationship for the other groups of UGT2Bs is more complex.

The multiple duplications and the high level of similarity within the baboon UGT2Bs and that reported for the monkey UGT2Bs, suggest a recombination within the old world monkey UGT2B gene families (Barbier et al., 1999b; Girard et al., 2002). Gene conversion is a recombination process between highly homologous sequences involving the nonreciprocal transfer of genetic material from one gene to another. The sequences of the two genes have become more similar as a result and this can lead to concerted evolution (Drouin et al., 1999; Chen et al., 2007). The human, baboon, and monkey UGT2Bs were therefore analyzed for gene conversion events using the GENECONV software (Table 3). Global P-values for potential converted segments reached statistical significance for baboons in the UGT2B7-like group and for monkeys in the UGT2B4-like group. No statistical significance was found for the human UGT2Bs. This indicates that gene conversion is a likely mechanism operating within the baboon

DMD #30635

and the monkey, but not in human UGT2Bs. The presence of gene conversion within the old world monkey UGT2B4-like and UGT2B7-like genes means that orthologous relationships between these genes cannot formally be inferred because different parts of their sequences have different evolutionary histories (Posada et al., 2002).

Protein sequence analysis

The multiple alignments of the deduced baboon UGT2B protein sequences reveal primary protein structures characteristic of UGT2B enzymes (Fig. 2). The carboxyl termini of the baboon UGT2B proteins are, as for all UGT proteins, the most highly conserved region and therefore more homologous than the amino termini. The protein sequences contain hydrophobic signal peptides from amino acids 1 to 23, as well as hydrophobic trans-membrane regions in amino acids ranging from 493 to 511. The consensus sequence for UGT enzymes is also present in amino acids ranging from 371 to 400 (Mackenzie et al., 1997).

Binding studies and NMR spectroscopy as well as molecular modeling has demonstrated that the morphine binding site in human UGT2B7 is located within amino acids 96 and 101 (Coffman et al., 2003). Furthermore, the substitution of the negatively charged aspartic acid in position 99 with the neutral alanine led to a marked decrease in morphine binding. The baboon UGT2B7-like proteins all show a substitution in position 99, where aspartic acid has been substituted with the only other negatively charged amino acid, glutamic acid. A substitution in position 98 of the neutral serine with the neutral isoleucine in UGT2B39 and UGT2B44 is the only other difference between the baboon and the human UGT2B7s in this region.

DMD #30635

In addition, the alignments of the deduced UGT2B protein sequences show a single amino acid substitution at position 96 in UGT2B43, where a highly conserved arginine has been substituted with isoleucine. The alignment of baboon, monkey, and human amino acid sequences used in the phylogenetic analysis above (sequences not shown) indicate that the arginine in position 96 is conserved in all the other UGT2B proteins except for monkey UGT2B20, which also contains an isoleucine in position 96. Furthermore, the primary protein structures of baboon UGT2B43 and monkey UGT2B20 are identical, except for a difference in only 3 amino acids (Y16C, G50R, and Q217H), none of which are located in the proposed substrate-binding region between positions 60 to 120 (Tephly and Burchell, 1990).

Post-translational asparagine-linked glycosylation (N-glycosylation) is known to regulate protein function. It has also been found to be essential for the activity of several UGT2B enzymes (Barbier et al., 2000b). The mammalian UGT2B proteins that have been characterized to date have between one and four potential N-glycosylation sites with the sequences of NX(S/T). However, not all UGT2Bs are glycosylated, nor are all potential sites actually used. To date, only the human UGT2B7, UGT2B15, and UGT2B17 and the monkey UGT2B20 have been shown to be glycosylated. It has also been demonstrated that UGT2B15 and UGT2B20 are glycosylated at position 65 only (Barbier et al., 2000b). Potential N-glycosylation sites in the primary protein structures of human, baboon, and monkey UGT2Bs were compared (Table 4). Several potential N-glycosylation sites are found in the baboon UGT2B proteins. The location and number of these sites are identical to those found in monkey UGT2B proteins. This includes the unique fourth site at position 103 that is only found in monkey UGT2B20 and baboon UGT2B43.

DMD #30635

UGT2B activity toward morphine

The heavy membrane fractions of the expressed baboon UGT2B cDNAs were tested for activity toward morphine. All of the cloned baboon UGT2Bs have the ability to glucuronidate morphine to both M3G and M6G (Fig. 3). Heavy membrane fractions from cells transfected with empty plasmid did not show any activity. This indicates that the background activity of UGT2Bs in HEK293 cells is either absent or undetectable. While the assay is only semi-quantitative in view of the variable transfection efficiency among clones, the activity was by far highest in the UGT2B7-like group of the baboon UGT2Bs. Of special note is the ability of the UGT2B15-like UGT2B43 to metabolize morphine, not only to M3G, but also to M6G.

DISCUSSION

Eight novel functioning UGT2B isoforms were expressed in the liver of the baboon that all exhibit some capacity to metabolize morphine to both M3G and M6G. Although there are considerable duplications and deletions among the primates, a close relationship exists among the human, the baboon, and the monkey (*Macaca fascicularis*) UGT2B enzyme families.

The enzyme activity toward morphine of the baboon UGT2Bs shows that all the baboon cDNA clones encode functional enzymes. In addition, as in humans, all UGT2B7-like and UGT2B4-like enzymes in the baboon metabolize morphine to M3G and M6G. In contrast, human UGT2B15 and UGT2B17 have only been shown to metabolize morphine to M3G (Court et al., 2003; Ohno et al., 2008). However, the baboon UGT2B43 metabolizes morphine to M3G and to M6G, although to a lesser extent than the other baboon UGT2B enzymes. The predominant UGT2B isoforms responsible for morphine glucuronidation in the baboon were, as in the human, the UGT2B7-like enzymes. These results indicate that morphine can also be considered an isoform-selective probe drug in the baboon even though more UGT2B7-like isoforms are present in this species. Interestingly, the two most active baboon UGT2B7-like isoforms, UGT2B39 and UGT2B44, both have a S98I substitution compared to that of human UGT2B7 and other baboon UGT2B7-like enzymes. However, both serine and isoleucine are neutral amino acids, so this substitution would not be expected to alter the binding capacity of the enzymes.

The total morphine clearance in adult humans, which is almost entirely metabolic, is similar to that found in adult baboons. On the other hand, the metabolic clearance in the baboon fetus is

DMD #30635

greater than what might have been expected based on enzyme activity comparisons (Garland et al., 2005). The multiple copies of UGT2B7-like isoforms that metabolize morphine could be one possible explanation for the enhanced metabolism of morphine observed in the fetal baboon. In the adult baboon, they would have less of an impact as morphine metabolism is flow limited.

The phylogenetic analysis indicates that the encoded UGT2B proteins are highly conserved among humans, baboons, and monkeys although several duplications have taken place in both the human and the old world monkeys (Table 2, Fig. 1.) The major differences are the multiple duplications of the UGT2B7-like genes and a duplication of the UGT2B4-like gene in the old world monkeys and the single duplication in the human UGT2B15-like genes that is absent in the baboon and the monkey. The UGT2B10, UGT2B11, and UGT2B28 human paralogs are all absent in the old world monkeys indicating that these genes most likely have been deleted. Of note, both UGT2B17 and UGT2B28 are frequent deletion polymorphisms among human populations (Menard et al., 2009).

The multiple duplications make it impossible to infer one-to-one orthologous relationships between the human and any of the baboon or the monkey UGT2Bs. Furthermore, the detection of gene conversion events operating on the primate UGT2B genes interferes with the phylogenetic signal between several of these genes. However based on bootstrap analysis and the absence of gene conversion, an orthologous relationship is suggested between baboon UGT2B43 and monkey UGT2B20 in the UGT2B15-like group. The UGT2B4-like group shows intermediate bootstrap support indicating a partial loss of the phylogenetic signal due to gene

DMD #30635

conversions. The human, baboon, and monkey UGT2B within the UGT2B7-like group share a high level of sequence identity, but the phylogenetic signal is obscured by gene conversion events detected in all the baboon genes of this group.

The families of drug metabolizing enzymes have developed under selective pressure from exogenous substrates throughout the evolution of different species (Nebert and Dieter, 2000; Thomas, 2007). This is well illustrated in the cytochrome P450 family where enzymes involved in endogenous pathways have retained strong orthologous relationships across species but enzymes involved in xenobiotic metabolism show considerable within species diversity (Thomas, 2007). In addition, this study concluded that the xenobiotic cytochrome P450 enzymes have evolved primarily through frequent gene duplications and losses, so called birth-and-death evolution, in a species specific manner. The duplication, diversification, and deletion of the UGT2B enzymes may therefore have occurred in a species-specific manner. The comparison of human, baboon, and monkey UGT2Bs reveals that these enzymes are subject to both species-specific duplications and deletions as well as gene conversion. The gene conversion within the old world monkeys was extensive enough to break down the orthologous relationship in the UGT2B4-like and the UGT2B7-like groups. This indicates that duplication and gene conversion in the old world monkey UGT2B4-like and UGT2B7-like groups most likely happened after speciation. Furthermore, the detection of gene conversion operating within the baboon and monkey UGT2Bs supports that a mixture of concerted and birth-and-death evolution likely contributed to the evolution of these genes in the baboon and the monkey, but not in the human (Nei et al., 2005). This is consistent with an earlier study of UGT2B genes in several species indicating that most human, mouse, and rat UGT2Bs were duplicated after speciation and where

DMD #30635

gene conversion were detected in rats, and zebrafish, but not in humans (Li and Wu, 2007).

However, this study did not include any old world monkeys in their analysis.

The protein sequence analysis indicated a high level of structural conservation between the human, the baboon, and the monkey UGT2Bs. The sequence identity was especially high between the baboon and the monkey UGT2B proteins. The 99.1% identical protein structure between the baboon UGT2B43 and monkey UGT2B20 is the highest reported with any other UGT2B protein (human or primate) characterized to date (Belanger et al., 1997; Beaulieu et al., 1998; Belanger et al., 1999; Barbier et al., 1999a; Barbier et al., 1999b; Girard et al., 2002). As with UGT2B20, baboon UGT2B43 has an isoleucine substitution of a highly conserved arginine at position 96, not found in any other UGT2B (Barbier et al., 2000b). UGT2B20 has been found to be a more labile enzyme than human UGT2B15 due to this isoleucine substitution which reduces the length of the predicted α -helix between amino acids 84 and 100. The reduced length of the α -helix correlated with decreased enzyme stability (Barbier et al., 2000b).

The identical number and location of N-glycosylation sites known to be essential for the activity of several UGT2B enzymes, was found in the baboon and the monkey UGT2Bs. UGT2B20 is the only monkey UGT2B characterized to date that has been shown to be glycosylated (Barbier et al., 2000b). While the extent of glycosylation or other enzyme functions cannot be predicted by the similarities of protein sequence, the especially high level of identity (99.1%) between baboon UGT2B43 and UGT2B20, makes it reasonable to expect that baboon UGT2B43 would also be glycosylated. Furthermore, the R96I substitution in UGT2B43 and in UGT2B20 makes it very likely that both these enzymes are more labile than human UGT2B15. This suggests that

DMD #30635

these two enzymes are indeed true orthologs. This high level of conservation may reflect the specific role of this enzyme in androgen metabolism (Barbier et al., 1999a; Thomas, 2007).

In order to fully understand the origin of the multiple duplications in the baboon and the monkey UGT2B7-like genes, a comparison of the human, the baboon and the monkey genomes would be necessary. The genomes for the baboon and the *Macaca fascicularis* are not yet available for comparison. However, the rhesus monkey (*Macaca mulatta*) genome is available in GenBank, although only two UGT2B7 isoforms have been isolated and characterized to date. The remaining UGT2B isoforms on the rhesus genome are predicted from modeling.

An alignment of the monkey (*Macaca fascicularis*) UGT2Bs, which have been isolated to date, with the *Macaca mulatta* genome shows a more than 99% identity. However, it did leave one predicted UGT2B7-like gene on the genome without any known expressed isoform (data not shown). The *Macaca fascicularis* UGT2Bs have all been isolated from cDNA libraries, which leaves the possibility that not all *Macaca fascicularis* UGT2B isoforms have been isolated as of yet. So far, it appears there is at least one more UGT2B7-like isoform in the baboon than in macaca species consistent with species specific diversification.

The UGT2B genes are located on the long arm of chromosome 4 below the centromere in humans. In the *Macaca mulatta*, the UGT2B genes are located on chromosome 5 above the centromere. The recent constructions of a genetic linkage map of the *Macaca mulatta* to the human genome, shows a pericentric inversion of this region between the macaca and the human chromosomes (Karere et al., 2008). A pericentric inversion has also been shown in a genetic

DMD #30635

linkage map of the baboon (*Papio hyrnadryas*) chromosome 5 to the human chromosome 4, as well as on chromosome 4 between the human and the chimpanzee (Yunis and Prakash, 1982; Rogers et al., 2000; Mikkelsen et al., 2005; Cox et al., 2006).

The chromosomal rearrangement of human chromosome 4 compared to that of the chimpanzee, such as the pericentric inversion, has been associated with a higher divergence of genes located within the inversion (Marques-Bonet et al., 2007). In addition, this study found that although chromosomal speciation is not common between the human and the chimpanzee lineage, chromosome 4 may be an exception and has potentially been associated with a speciation episode. The pericentric inversion in chromosome 4 of hominoids contains genes involved with biotic stimulus response, such as UGT2Bs, that are likely to be related to adaptation, and hence, may also be more likely associated with chromosomal speciation. This is consistent with the findings in a comparative study of the chromosomes from the human, the chimpanzee, the gorilla, and the orangutan indicating that similar, but distinct pericentric inversions occurred on chromosome 4 in each of the speciation events of the orangutan, the gorilla, and the chimpanzee compared to that of the human (Yunis and Prakash, 1982). Thus, it is possible to suggest that the pericentric inversions in the baboon and the monkey chromosome 5 which contain the UGT2B genes occurred as a result of the speciation event from the common ancestral line. Moreover, the subsequent divergence of the baboon, and the monkey from that of the human UGT2B genes would have happened after speciation and under selective pressure from exogenous substances (Nebert and Dieter, 2000).

DMD #30635

In summary, this study reports the isolation of eight novel UGT2B isoforms expressed in the baboon liver. These enzymes all demonstrated activity toward morphine and the ability to glucuronidate morphine to both M3G and M6G. In addition, the similarities between the primary structures of the human, the baboon, and the monkey UGT2B proteins provided further evidence that there is a close relationship among the UGT2B families of enzymes in these species. Although a one-to-one orthologous relationship cannot be established for human UGT2B7, the UGT2B7-like isoforms in the baboon all have substantial activity toward morphine. This, together with the close relationship demonstrated previously between the human and the baboon UGT1A enzyme families, further substantiates that the baboon is an excellent model for studying clinically relevant aspects of drug metabolism. The data obtained in this study, lays the foundation for investigating the regulation of UGT2B enzymes during fetal and neonatal development in the baboon, and facilitates the extrapolation of results from the baboon to humans.

REFERENCES

- Barbier O and Belanger A (2003) The cynomolgus monkey (*Macaca fascicularis*) is the best animal model for the study of steroid glucuronidation. *J Steroid Biochem Mol Biol* **85**:235-245.
- Barbier O, Belanger A and Hum DW (1999a) Cloning and characterization of a simian UDP-glucuronosyltransferase enzyme UGT2B20, a novel C19 steroid-conjugating protein. *Biochem J* **337** (Pt 3):567-574.
- Barbier O, Girard C, Breton R, Belanger A and Hum DW (2000b) N-glycosylation and residue 96 are involved in the functional properties of UDP-glucuronosyltransferase enzymes. *Biochemistry* **39**:11540-11552.
- Barbier O, Levesque E, Belanger A and Hum DW (1999b) UGT2B23, a novel uridine diphosphate-glucuronosyltransferase enzyme expressed in steroid target tissues that conjugates androgen and estrogen metabolites. *Endocrinology* **140**:5538-5548.
- Barbier O, Turgeon D, Girard C, Green MD, Tephly TR, Hum DW and Belanger A (2000a) 3'-azido-3'-deoxythymidine (AZT) is glucuronidated by human UDP-glucuronosyltransferase 2B7 (UGT2B7). *Drug Metab Dispos* **28**:497-502.
- Beaulieu M, Levesque E, Barbier O, Turgeon D, Belanger G, Hum DW and Belanger A (1998) Isolation and characterization of a simian UDP-glucuronosyltransferase UGT2B18 active on 3-hydroxyandrogens. *J Mol Biol* **275**:785-794.
- Belanger G, Barbier O, Hum DW and Belanger A (1999) Molecular cloning, expression and characterization of a monkey steroid UDP-glucuronosyltransferase, UGT2B19, that conjugates testosterone. *Eur J Biochem* **260**:701-708.
- Belanger G, Beaulieu M, Levesque E, Hum DW and Belanger A (1997) Expression and characterization of a novel UDP-glucuronosyltransferase, UGT2B9, from cynomolgus monkey. *DNA Cell Biol* **16**:1195-1205.
- Caspersen CS, Reznik B, Weldy PL, Abildskov KM, Stark RI and Garland M (2007) Molecular cloning of the baboon UDP-glucuronosyltransferase 1A gene family: evolution of the

- primate UGT1 locus and relevance for models of human drug metabolism.
Pharmacogenet Genomics **17**:11-24.
- Chay PC, Duffy BJ and Walker JS (1992) Pharmacokinetic-pharmacodynamic relationships of morphine in neonates. *Clin Pharmacol Ther* **51**:334-342.
- Chen JM, Cooper DN, Chuzhanova N, Ferec C and Patrinos GP (2007) Gene conversion: mechanisms, evolution and human disease. *Nat Rev Genet* **8**:762-775.
- Coffman BL, Kearney WR, Goldsmith S, Knosp BM, Tephly TR, Dean B, Arison B, Chang S, Thomas PE and King C (2003) Opioids bind to the amino acids 84 to 118 of UDP-glucuronosyltransferase UGT2B7
- Coffman BL, Rios GR, King CD and Tephly TR (1997) Human UGT2B7 catalyzes morphine glucuronidation. *Drug Metab Dispos* **25**:1-4.
- Court MH, Duan SX, von Moltke LL, Greenblatt DJ, Patten CJ, Miners JO and Mackenzie PI (2001) Interindividual variability in acetaminophen glucuronidation by human liver microsomes: identification of relevant acetaminophen UDP-glucuronosyltransferase isoforms. *J Pharmacol Exp Ther* **299**:998-1006.
- Court MH, Krishnaswamy S, Hao Q, Duan SX, Patten CJ, Von Moltke LL and Greenblatt DJ (2003) Evaluation of 3'-azido-3'-deoxythymidine, morphine, and codeine as probe substrates for UDP-glucuronosyltransferase 2B7 (UGT2B7) in human liver microsomes: specificity and influence of the UGT2B7*2 polymorphism. *Drug Metab Dispos* **31**:1125-1133.
- Cox LA, Mahaney MC, Vandenberg JL and Rogers J (2006) A second-generation genetic linkage map of the baboon (*Papio hamadryas*) genome. *Genomics* **88**:274-281.
- de Wildt SN, Kearns GL, Leeder JS and van den Anker JN (1999) Glucuronidation in humans. Pharmacogenetic and developmental aspects. *Clin Pharmacokinet* **36**:439-452.
- Dean B, Arison B, Chang S, Thomas PE and King C (2004) Identification of UGT2B9*2 and UGT2B33 isolated from female rhesus monkey liver. *Arch Biochem Biophys* **426**:55-62.
- Drouin G, Prat F, Ell M and Clarke GD (1999) Detecting and characterizing gene conversions between multigene family members. *Mol Biol Evol* **16**:1369-1390.
- Dutton GJ (1978) Developmental aspects of drug conjugation, with special reference to glucuronidation. *Annu Rev Pharmacol Toxicol* **18**:17-35.
- Dutton GJ and Leakey JE (1981) The perinatal development of drug-metabolizing enzymes: what factors trigger their onset? *Prog Drug Res* **25**:189-273.

DMD #30635

- Fisher MB, Campanale K, Ackermann BL, VandenBranden M and Wrighton SA (2000) In vitro glucuronidation using human liver microsomes and the pore-forming peptide alamethicin. *Drug Metab Dispos* **28**:560-566.
- Garland M, Abildskov KM, Kiu TW, Daniel SS and Stark RI (2005) The contribution of fetal metabolism to the disposition of morphine. *Drug Metab Dispos* **33**:68-76.
- Garland M, Abildskov KM, Taylor S, Benzeroual K, Caspersen CS, Arroyo SE, Kiu TW, Reznik B, Weldy P, Daniel SS and Stark RI (2006) Fetal morphine metabolism and clearance are constant during late gestation. *Drug Metab Dispos* **34**:636-646.
- Garland M, Szeto HH, Daniel SS, Tropper PJ, Myers MM and Stark RI (1996) Zidovudine kinetics in the pregnant baboon. *J Acquir Immune Defic Syndr Hum Retrovirol* **11**:117-127.
- Garland M, Szeto HH, Daniel SS, Tropper PJ, Myers MM and Stark RI (1998a) Placental transfer and fetal metabolism of zidovudine in the baboon. *Pediatr Res* **44**:47-53.
- Garland M, Szeto HH, Daniel SS, Tropper PJ, Myers MM and Stark RI (1998b) Implications of the kinetics of zidovudine in the pregnant baboon following oral administration. *J Acquir Immune Defic Syndr Hum Retrovirol* **19**:433-440.
- Girard C, Barbier O, Turgeon D and Belanger A (2002) Isolation and characterization of the monkey UGT2B30 gene that encodes a uridine diphosphate-glucuronosyltransferase enzyme active on mineralocorticoid, glucocorticoid, androgen and oestrogen hormones. *Biochem J* **365**:213-222.
- Green MD, King CD, Mojarrabi B, Mackenzie PI and Tephly TR (1998) Glucuronidation of amines and other xenobiotics catalyzed by expressed human UDP-glucuronosyltransferase 1A3. *Drug Metab Dispos* **26**:507-512.
- Hartley R, Green M, Quinn M and Levene MI (1993) Pharmacokinetics of morphine infusion in premature neonates. *Arch Dis Child* **69**:55-58.
- Karere GM, Froenicke L, Millon L, Womack JE and Lyons LA (2008) A high-resolution radiation hybrid map of rhesus macaque chromosome 5 identifies rearrangements in the genome assembly. *Genomics* **92**:210-218.
- Li C and Wu Q (2007) Adaptive evolution of multiple-variable exons and structural diversity of drug-metabolizing enzymes. *BMC Evol Biol* **7**:69.
- Mackenzie PI, Bock KW, Burchell B, Guillemette C, Ikushiro S, Iyanagi T, Miners JO, Owens IS and Nebert DW (2005) Nomenclature update for the mammalian UDP glycosyltransferase (UGT) gene superfamily. *Pharmacogenet Genomics* **15**:677-685.

DMD #30635

- Mackenzie PI, Owens IS, Burchell B, Bock KW, Bairoch A, Belanger A, Fournel-Gigleux S, Green M, Hum DW, Iyanagi T, Lancet D, Louisot P, Magdalou J, Chowdhury JR, Ritter JK, Schachter H, Tephly TR, Tipton KF and Nebert DW (1997) The UDP glycosyltransferase gene superfamily: recommended nomenclature update based on evolutionary divergence. *Pharmacogenetics* **7**:255-269.
- Marques-Bonet T, Sanchez-Ruiz J, Armengol L, Khaja R, Bertranpetit J, Lopez-Bigas N, Rocchi M, Gazave E and Navarro A (2007) On the association between chromosomal rearrangements and genic evolution in humans and chimpanzees. *Genome Biol* **8**:R230.
- Menard V, Eap O, Harvey M, Guillemette C and Levesque E (2009) Copy-number variations (CNVs) of the human sex steroid metabolizing genes UGT2B17 and UGT2B28 and their associations with a UGT2B15 functional polymorphism. *Hum Mutat* **30**:1310-1319.
- Mikkelsen T, Hillier L, Eichler E, Zody M, Jaffe D, Yang S, Ernard W, Hellman I, Lindblad-Toh K and Altheide T, et al. (2005) Initial sequence of the chimpanzee genome and comparison with the human genome. *Nature* **437**:69-87.
- Nebert DW and Dieter MZ (2000) The evolution of drug metabolism. *Pharmacology* **61**:124-135.
- Nei M, Rooney AP, Nikolaidis N, Piontkivska H, Su C, Rogozin IB, Gu X, Sitnikova T, Ota T, Lu BZ and Gojobori T (2005) Concerted and birth-and-death evolution of multigene families. *Annu Rev Genet* **39**:121-152.
- Ohno S, Kawana K and Nakajin S (2008) Contribution of UDP-glucuronosyltransferase 1A1 and 1A8 to morphine-6-glucuronidation and its kinetic properties. *Drug Metab Dispos* **36**:688-694.
- Posada D, Crandall KA and Holmes EC (2002) Recombination in evolutionary genomics. *Annu Rev Genet* **36**:75-97.
- Rogers J, Mahaney MC, Witte SM, Nair S, Newman D, Wedel S, Rodriguez LA, Rice KS, Slifer SH, Perelygin A, Slifer M, Palladino-Negro P, Newman T, Chambers K, Joslyn G, Parry P and Morin PA (2000) A genetic linkage map of the baboon (*Papio hamadryas*) genome based on human microsatellite polymorphisms. *Genomics* **67**:237-247.
- Sawyer S (1989) Statistical test for detecting gene conversion. *Molecular Biology and Evolution* **6**:526-638 (<http://www.math.wustl.edu/~sawyer>).
- Scott CS, Riggs KW, Ling EW, Fitzgerald CE, Hill ML, Grunau RV, Solimano A and Craig KD (1999) Morphine pharmacokinetics and pain assessment in premature newborns. *J Pediatr* **135**:423-429.
- Soars MG, Ring BJ and Wrighton SA (2003) The effect of incubation conditions on the enzyme kinetics of udp-glucuronosyltransferases. *Drug Metab Dispos* **31**:762-767.

DMD #30635

- Strassburg CP, Strassburg A, Kneip S, Barut A, Tukey RH, Rodeck B and Manns MP (2002) Developmental aspects of human hepatic drug glucuronidation in young children and adults. *Gut* **50**:259-265.
- Tamura K, Dudley J, Nei M and Kumar S (2007) MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. *Mol Biol Evol* **24**:1596-1599.
- Tephly TR and Burchell B (1990) UDP-glucuronosyltransferases: a family of detoxifying enzymes. *Trends Pharmacol Sci* **11**:276-279.
- Thomas JH (2007) Rapid birth-death evolution specific to xenobiotic cytochrome P450 genes in vertebrates. *PLoS Genet* **3**:e67.
- Yunis JJ and Prakash O (1982) The origin of man: a chromosomal pictorial legacy. *Science* **215**:1525-1530.

DMD #30635

FOOTNOTES

This research was funded through a National Institutes of Drug Abuse and Child Health and Development grant award R01 DA14215.

DMD #30635

LEGENDS FOR FIGURES

Fig. 1. Neighbor-joining (NJ) reconciled phylogenetic tree with baboon, human, and monkey (*Macaca fascicularis*) UGT2B enzymes. The sequences for human, and monkey UGT2B genes were obtained from the GenBank. The protein coding regions were aligned using MEGA4.0 and a NJ distance-based phylogenetic tree using p-distances was generated. Robustness of bifurcations is estimated by bootstrap analysis with 1000 replicates and the consensus tree is shown. Numbers at bifurcations indicate bootstrap values in percent. Annotation: bab (baboon, *Papio anubis*), h (human), and mon (cynomolgus monkey, *Macaca fascicularis*).

Fig. 2. Multiple alignments of deduced babUGT2B proteins amino acid sequences were aligned using the Multalin web page. Identical amino acids are shown in black background and >80% conservation is illustrated by boxes. Residue 96 is indicated by a star.

Fig. 3. Morphine glucuronidation in heavy membranes from HEK293 cells expressing individual baboon UGT2B isoforms or HEK293 vector control. Morphine-3-UGT and morphine-6-UGT activity were measured at a morphine concentration of 5mM. Each bar represents the mean of duplicate measurements.

TABLE 1

Baboon (Papio anubis) UGT2B cDNAs

Baboon UGT	GenBank Accession Number	Number of amino acids in coding region	Human UGT Homologs
UGT2B41	EU856383	528	UGT2B4 (480/528) 90%
UGT2B42	EU856382	528	UGT2B4 (480/528) 90%
UGT2B39	EU856384	529	UGT2B7 (479/529) 91%
UGT2B40	EU856385	529	UGT2B7 (478/529) 90%
UGT2B44	EU856388	528	UGT2B7 (470/528) 89%
UGT2B45	EU856387	529	UGT2B7 (475/529) 90%
UGT2B46	EU856386	528	UGT2B7 (474/528) 90%
UGT2B43	EU856381	530	UGT2B15 (492/530) 93%

The baboon UGT2B gene names are sequentially numbered based upon their chronological order of discovery and were approved by the UGT Nomenclature Committee. Human homologs list the number of amino acid identities and percent of amino acid identity.

DMD #30635

TABLE 2

Similarities between deduced amino acid sequences of the protein coding regions of UGT2Bs

Baboon versus Baboon								
	<i>bab2B41</i>	<i>bab2B42</i>	<i>bab2B39</i>	<i>bab2B40</i>	<i>bab2B44</i>	<i>bab2B45</i>	<i>bab2B46</i>	<i>bab2B43</i>
<i>bab2B41</i>	100							
<i>bab2B42</i>	93.8	100						
<i>bab2B39</i>	81.8	83.1	100					
<i>bab2B40</i>	82.0	83.0	97.5	100				
<i>bab2B44</i>	81.3	83.0	95.5	96.4	100			
<i>bab2B45</i>	80.9	82.4	94.7	95.5	95.6	100		
<i>bab2B46</i>	81.6	82.6	95.5	96.2	96.2	93.9	100	
<i>bab2B43</i>	74.1	75.4	76.4	76.2	75.6	76.0	76.5	100

Human versus Baboon								
	<i>bab2B41</i>	<i>bab2B42</i>	<i>bab2B39</i>	<i>bab2B40</i>	<i>bab2B44</i>	<i>bab2B45</i>	<i>bab2B46</i>	<i>bab2B43</i>
<i>h2B4</i>	88.3	88.5	86.2	86.2	86.0	85.6	85.6	76.9
<i>h2B7</i>	81.4	82.0	89.8	90.2	88.8	89.2	89.4	75.4
<i>h2B10</i>	80.3	81.0	86.7	87.1	86.7	87.5	86.0	75.6
<i>h2B11</i>	81.1	81.3	85.6	86.6	86.6	86.2	86.0	74.7
<i>h2B28</i>	78.8	79.4	83.9	84.9	85.2	84.9	84.7	73.2
<i>h2B15</i>	76.5	76.5	79.0	78.8	78.0	77.7	77.7	92.1
<i>h2B17</i>	76.1	76.1	77.7	78.1	77.5	77.1	77.3	90.0

Monkey versus Baboon								
	<i>bab2B41</i>	<i>bab2B42</i>	<i>bab2B39</i>	<i>bab2B40</i>	<i>bab2B44</i>	<i>bab2B45</i>	<i>bab2B46</i>	<i>bab2B43</i>
<i>mon2B19</i>	93.8	97.9	82.6	82.4	82.4	81.8	82.0	75.0
<i>mon2B30</i>	92.8	93.8	82.2	82.0	82.0	81.4	81.4	75.2
<i>mon2B9</i>	81.6	83.0	95.7	96.2	98.3	95.8	95.6	76.0
<i>mon2B18</i>	82.2	83.0	96.2	96.4	95.8	95.1	96.8	75.8
<i>mon2B23</i>	79.9	81.1	95.8	94.0	93.9	92.6	93.2	74.7
<i>mon2B20</i>	74.1	75.0	76.0	75.8	75.2	75.6	75.2	99.1

Annotations:

bab (baboon, *Papio anubis*), h (human), and mon (cynomolgus monkey, *Macaca fascicularis*).

DMD #30635

TABLE 3.

Gene conversion events indicated by GENECONV

Sequence Name	Global P-Value	Sequence	Pairwise P-Values	Sequence
babUGT2B39; UGT2B46	<0.0001 (0)	676 – 1193	<0.0001 (0)	676 – 1193
babUGT2B39; UGT2B40	0.0117 (0)	958 – 1587	0.0001 (0)	958 – 1587
babUGT2B39; UGT2B40			0.0121 (0)	291 – 650
babUGT2B39; UGT2B45	0.0343 (0)	1131 – 1305	0.0033 (0)	1131 – 1305
babUGT2B39; UGT2B45			0.0207 (0)	699 – 963
babUGT2B40; UGT2B44	0.0067 (0)	550 – 956	0.0006 (0)	550 – 956
babUGT2B40; UGT2B45			0.0321 (0)	1131 – 1305
babUGT2B44; UGT2B46	0.0117 (0)	1147 – 1587	0.0009 (0)	1147 – 1587
babUGT2B41; UGT2B42			0.0355 (1)	596 – 986
monUGT2B19; UGT2B30	0.0154 (0)	334 – 452	0.0095 (0)	334 – 452

Alignments of UGT2B4-like and UGT2B7-like sequences were analyzed with GENECONV. Human, baboon (*Papio anubis*), and cynomolgus monkey (*Macaca fascicularis*) sequences were divided into separate groups to limit analysis to within species gene conversion and reduce the multiple comparison penalty. P-values are based on 10,000 permutations and only significant (P<0.05) conversions are included. The different g-scales used to allow for mismatches are given in parentheses after the P-value. Only converted sequences >95nt are listed, and the position is relative to the start codon.

TABLE 4**Potential N-glycosylation sites in baboon, human, and monkey UGT2B enzymes**

<i>Mammalian UGT</i>	Number of potential N-glycosylation sites	Amino acid position and sequence of potential N-glycosylation sites				
h2B4	1				315-NTS	
bab2B41	2	68-NPS			315-NMS	
bab2B42	2	68-NPS			315-NMS	
mon2B19	2	68-NPS			315-NMS	
mon2B30	2	68-NPS			315-NMS	
h2B7	3	67-NNS	68-NSS		315-NMT	
bab2B39	2	67-NNS	68-NSS			
bab2B40	2	67-NNS	68-NSS			
bab2B44	2	67-NNS	68-NSS			
bab2B45	2	67-NNS	68-NSS			
bab2B46	2	67-NNS	68-NSS			
mon2B39	2	67-NNS	68-NSS			
mon2B18	2	67-NNS	68-NSS			
mon2B23	2	67-NNS	68-NSS			
h2B15	3	65-NAS			316-NMS	483-NLT
h2B17	3	65-NAS			316-NMS	483-NLT
bab2B43	4	65-NDS		103-NST	316-NMS	483-NLT
mon2B20	4	65-NDS		103-NST	316-NMS	483-NLT

Comparison of potential N-glycosylation sites in the primary structures of bab (baboon, *Papio anubis*), h (human) and mon (cynomolgus monkey, *Macaca fascicularis*) UGT2B enzymes. The sequences for human, and monkey UGT2B genes were obtained from the GenBank. Multiple alignments of deduced UGT2B proteins amino acid sequences were aligned using the MEGA4.0. Amino acid sequences were compared and the potential N-glycosylation sites identified and tabulated.

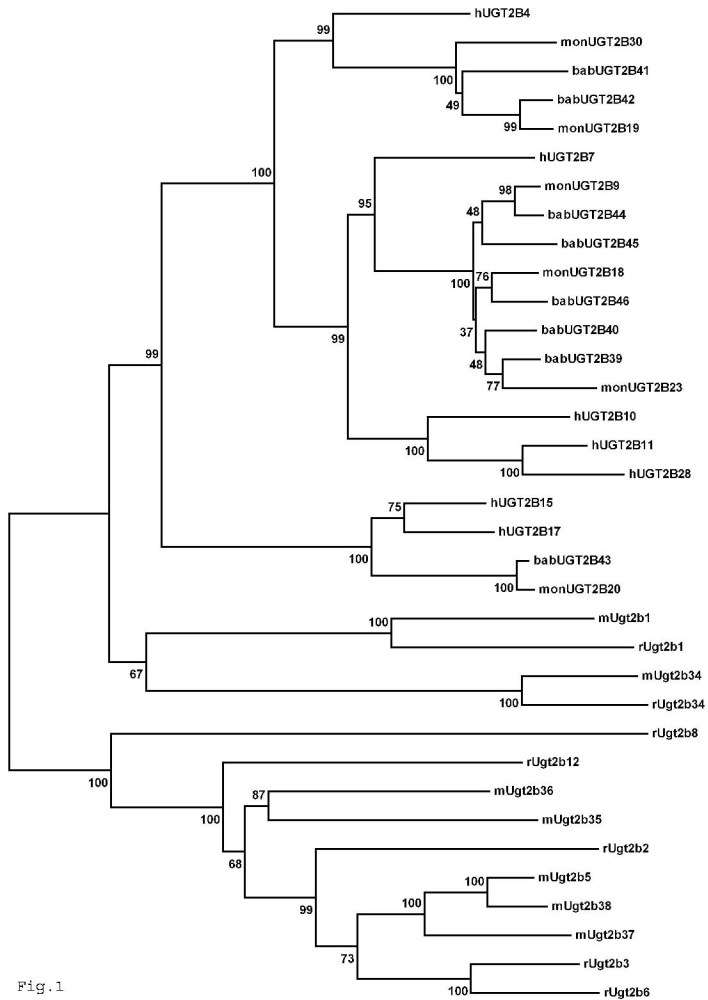


Fig. 1

1 10 20 30 40 50 60 70 80
babUG2B43 MSLKWTAVTLLQLSVPFSGGCGXVLVWVTPFVSHWNNKTLLELRLVQRGHEVTLASGASILFFDNNSSGALTEVFP7SLT
babUG2B41 MSLKWTAVTLLQLSVPFSGGCGXVLVWVTPFVSHWNNKTLLELRLVQRGHEVTLASGASILFFDNNSSGALTEVFP7SLT
babUG2B42 MSLKWTAVTLLQLSVPFSGGCGXVLVWVTPFVSHWNNKTLLELRLVQRGHEVTLASGASILFFDNNSSGALTEVFP7SLT
babUG2B39 MSLKWTAVTLLQLSVPFSGGCGXVLVWVTPFVSHWNNKTLLELRLVQRGHEVTLASGASILFFDNNSSGALTEVFP7SLT
babUG2B40 MSLKWTAVTLLQLSVPFSGGCGXVLVWVTPFVSHWNNKTLLELRLVQRGHEVTLASGASILFFDNNSSGALTEVFP7SLT
babUG2B45 MSLKWTAVTLLQLSVPFSGGCGXVLVWVTPFVSHWNNKTLLELRLVQRGHEVTLASGASILFFDNNSSGALTEVFP7SLT
babUG2B44 MSLKWTAVTLLQLSVPFSGGCGXVLVWVTPFVSHWNNKTLLELRLVQRGHEVTLASGASILFFDNNSSGALTEVFP7SLT
babUG2B46 MSLKWTAVTLLQLSVPFSGGCGXVLVWVTPFVSHWNNKTLLELRLVQRGHEVTLASGASILFFDNNSSGALTEVFP7SLT

90 100 110 120 130 140 150 160
babUG2B43 KMDNEDSLMLRLLDITVTSISNSTLSVYSXLRLLCGLQWYVYSRKLCDVSNKKMLKRLKRTHTPVLADALNCPGSELLABL
babUG2B41 ETEPEENIGTQLLKRNSDIKFKDAPVFLQIQRMVYTYDMIRRCRCDVSNKKMLKRLQDRSPVVLADVPSPGSELLABL
babUG2B42 ETEPEENIGTQLLKRNSDIKFKDAPVFLQIQRMVYTYDMIRRCRCDVSNKKMLKRLQDRSPVVLADVPSPGSELLABL
babUG2B39 KTEPENITIQGLKRNSELKFKDTLWFLYSQIRIMWRFDIIRNCCDVSNNKMLKRLQDRSPVVLADVPSPGSELLABL
babUG2B40 KTEPENITIQGLKRNSELKFKDTLWFLYSQIRIMWRFDIIRNCCDVSNNKMLKRLQDRSPVVLADVPSPGSELLABL
babUG2B45 KTEPENITIQGLKRNSELKFKDTLWFLYSQIRIMWRFDIIRNCCDVSNNKMLKRLQDRSPVVLADVPSPGSELLABL
babUG2B44 KTEPENITIQGLKRNSELKFKDTLWFLYSQIRIMWRFDIIRNCCDVSNNKMLKRLQDRSPVVLADVPSPGSELLABL
babUG2B46 KTEPENITIQGLKRNSELKFKDTLWFLYSQIRIMWRFDIIRNCCDVSNNKMLKRLQDRSPVVLADVPSPGSELLABL

*

170 180 190 200 210 220 230 240
babUG2B43 FMIIEFVYSLELRFVQVLEEKGGFLPPPSVVPVMSLSLDMQMTFERVQKNIWMLYVDFPFQVYDMKRWQDFYSGVLRGRIT
babUG2B41 LKLEFVYSLELRFVQVLEEKGGFLPPPSVVPVMSLSLDMQMTFERVQKNIWMLYVDFPFQVYDMKRWQDFYSGVLRGRIT
babUG2B42 LKLEFVYSLELRFVQVLEEKGGFLPPPSVVPVMSLSLDMQMTFERVQKNIWMLYVDFPFQVYDMKRWQDFYSGVLRGRIT
babUG2B39 FMIIEFVYSLELRFVQVLEEKGGFLPPPSVVPVMSLSLDMQMTFERVQKNIWMLYVDFPFQVYDMKRWQDFYSGVLRGRIT
babUG2B40 FMIIEFVYSLELRFVQVLEEKGGFLPPPSVVPVMSLSLDMQMTFERVQKNIWMLYVDFPFQVYDMKRWQDFYSGVLRGRIT
babUG2B45 LMIIEFVYSLELRFVQVLEEKGGFLPPPSVVPVMSLSLDMQMTFERVQKNIWMLYVDFPFQVYDMKRWQDFYSGVLRGRIT
babUG2B44 LMIIEFVYSLELRFVQVLEEKGGFLPPPSVVPVMSLSLDMQMTFERVQKNIWMLYVDFPFQVYDMKRWQDFYSGVLRGRIT
babUG2B46 LMIIEFVYSLELRFVQVLEEKGGFLPPPSVVPVMSLSLDMQMTFERVQKNIWMLYVDFPFQVYDMKRWQDFYSGVLRGRIT

250 260 270 280 290 300 310 320
babUG2B43 ILEEMKARIWILIRNSNDFPHPLLNVEFVQGLHCKPAKPLPKEMRESPVQSSGNGVVVPSLGSMTNMSSEERANNIASA
babUG2B41 ILEEMKARIWILIRNSNDFPHPLLNVEFVQGLHCKPAKPLPKEMRESPVQSSGNGVVVPSLGSMTNMSSEERANNIASA
babUG2B42 ILEEMKARIWILIRNSNDFPHPLLNVEFVQGLHCKPAKPLPKEMRESPVQSSGNGVVVPSLGSMTNMSSEERANNIASA
babUG2B39 ILEEMKADIWILIRNSNDFPHPLLNVEFVQGLHCKPAKPLPKEMRESPVQSSGNGVVVPSLGSMTNMSSEERANNIASA
babUG2B40 ILEEMKADIWILIRNSNDFPHPLLNVEFVQGLHCKPAKPLPKEMRESPVQSSGNGVVVPSLGSMTNMSSEERANNIASA
babUG2B45 ILEEMKADIWILIRNSNDFPHPLLNVEFVQGLHCKPAKPLPKEMRESPVQSSGNGVVVPSLGSMTNMSSEERANNIASA
babUG2B44 ILEEMKADIWILIRNSNDFPHPLLNVEFVQGLHCKPAKPLPKEMRESPVQSSGNGVVVPSLGSMTNMSSEERANNIASA
babUG2B46 ILEEMKADIWILIRNSNDFPHPLLNVEFVQGLHCKPAKPLPKEMRESPVQSSGNGVVVPSLGSMTNMSSEERANNIASA

330 340 350 360 370 380 390 400 410
babUG2B43 LACFPQKVLWRFDCKKEDTLGNTRLYKWPQNDLLGHPKTRAFITGGANGIYEAIYHGVPMVGPPLFADQDNNIAHKKX
babUG2B41 LACFPQKVLWRFDCKKEDTLGNTRLYKWPQNDLLGHPKTRAFITGGANGIYEAIYHGVPMVGPPLFADQDNNIAHKKX
babUG2B42 LACFPQKVLWRFDCKKEDTLGNTRLYKWPQNDLLGHPKTRAFITGGANGIYEAIYHGVPMVGPPLFADQDNNIAHKKX
babUG2B39 LACFPQKVLWRFDCKKEDTLGNTRLYKWPQNDLLGHPKTRAFITGGANGIYEAIYHGVPMVGPPLFADQDNNIAHKKX
babUG2B40 LACFPQKVLWRFDCKKEDTLGNTRLYKWPQNDLLGHPKTRAFITGGANGIYEAIYHGVPMVGPPLFADQDNNIAHKKX
babUG2B45 LACFPQKVLWRFDCKKEDTLGNTRLYKWPQNDLLGHPKTRAFITGGANGIYEAIYHGVPMVGPPLFADQDNNIAHKKX
babUG2B44 LACFPQKVLWRFDCKKEDTLGNTRLYKWPQNDLLGHPKTRAFITGGANGIYEAIYHGVPMVGPPLFADQDNNIAHKKX
babUG2B46 LACFPQKVLWRFDCKKEDTLGNTRLYKWPQNDLLGHPKTRAFITGGANGIYEAIYHGVPMVGPPLFADQDNNIAHKKX

420 430 440 450 460 470 480 490
babUG2B43 GAAVRLDFDIFMSSGDLNALKVINDPVEYENMKLSRHHQDPVPLDRAVPWFTEFVMMHKGAKHLLVAADHLTWQYHSL
babUG2B41 GAAVRLDFDIFMSSGDLNALKVINDPVEYENMKLSRHHQDPVPLDRAVPWFTEFVMMHKGAKHLLVAADHLTWQYHSL
babUG2B42 GAAVRLDFDIFMSSGDLNALKVINDPVEYENMKLSRHHQDPVPLDRAVPWFTEFVMMHKGAKHLLVAADHLTWQYHSL
babUG2B39 GAAVRLDFDIFMSSGDLNALKVINDPVEYENMKLSRHHQDPVPLDRAVPWFTEFVMMHKGAKHLLVAADHLTWQYHSL
babUG2B40 GAAVRLDFDIFMSSGDLNALKVINDPVEYENMKLSRHHQDPVPLDRAVPWFTEFVMMHKGAKHLLVAADHLTWQYHSL
babUG2B45 GAAVRLDFDIFMSSGDLNALKVINDPVEYENMKLSRHHQDPVPLDRAVPWFTEFVMMHKGAKHLLVAADHLTWQYHSL
babUG2B44 GAAVRLDFDIFMSSGDLNALKVINDPVEYENMKLSRHHQDPVPLDRAVPWFTEFVMMHKGAKHLLVAADHLTWQYHSL
babUG2B46 GAAVRLDFDIFMSSGDLNALKVINDPVEYENMKLSRHHQDPVPLDRAVPWFTEFVMMHKGAKHLLVAADHLTWQYHSL

500 510 520 530
babUG2B43 DVI@PLLCVAIVTPIK@K@LFC@W@F@V@R@G@K@E@D
babUG2B41 DVI@PLLCVAIVTPIK@K@LFC@W@F@V@R@G@K@E@D
babUG2B42 DVI@PLLCVAIVTPIK@K@LFC@W@F@V@R@G@K@E@D
babUG2B39 DVI@PLLCVAIVTPIK@K@LFC@W@F@V@R@G@K@E@D
babUG2B40 DVI@PLLCVAIVTPIK@K@LFC@W@F@V@R@G@K@E@D
babUG2B45 DVI@PLLCVAIVTPIK@K@LFC@W@F@V@R@G@K@E@D
babUG2B44 DVI@PLLCVAIVTPIK@K@LFC@W@F@V@R@G@K@E@D
babUG2B46 DVI@PLLCVAIVTPIK@K@LFC@W@F@V@R@G@K@E@D

Fig. 2

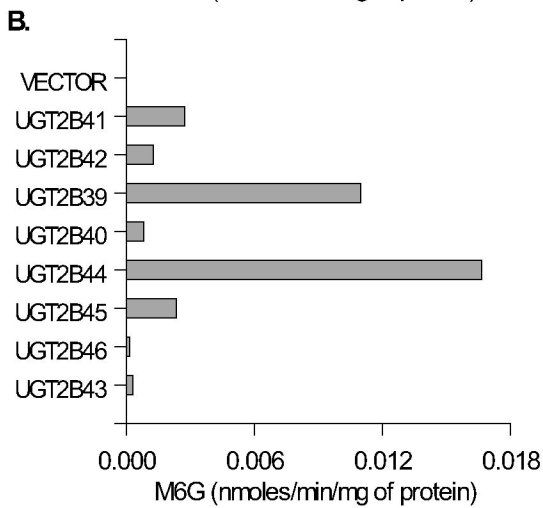
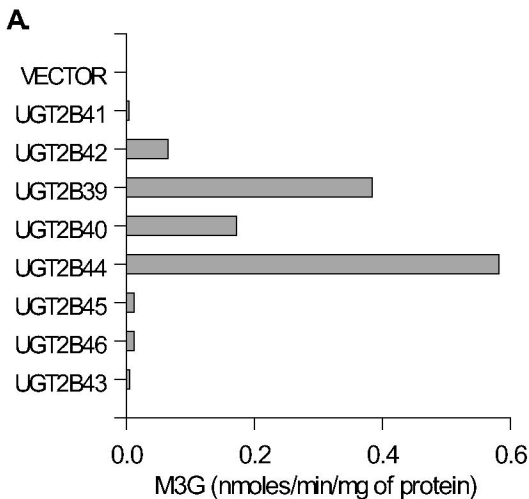


Fig. 3