The human UDP-glucuronosyltransferases: studies on substrate binding and catalytic mechanism

Anne-Sisko Patana

Research Program in Structural Biology and Biophysics
Institute of Biotechnology, University of Helsinki

Department of Biological and Environmental Sciences, Faculty of Biosciences,
Division of Biochemistry, University of Helsinki

and

National Graduate School in Informational and Structural Biology,
Åbo Akademi University

ACADEMIC DISSERTATION

To be presented for public examination with the permission of the Faculty of Biosciences of the University of Helsinki, in lecture room 2402 (Telkänpönttö) at Viikki Biocenter 3 (Viikinkaari 1), Helsinki, on February 6th , 2009, at 12 noon.

Helsinki 2009

Supervised by
Research Director Adrian Goldman
Institute of Biotechnology
Structural Biology and Biophysics
University of Helsinki, Finland
and
Docent Moshe Finel
Faculty of Pharmacy
Centre for Drug Research (CDR)
University of Helsinki, Finland

Reviewed by
Professor Risto Juvonen
Faculty of Pharmacy
Department of Pharmacology and Toxicology
University of Kuopio, Finland
and
Docent Tuomas Haltia
Faculty of Biosciences
Department of Biological and Environmental Sciences
Division of Biochemistry
University of Helsinki, Finland

Opponent
Professor Michael W. H. Coughtrie
Biochemical Pharmacology
Division of Molecular and Translational Medicine
Ninewells Hospital and Medical School
University of Dundee, Scotland, UK

Custos
Professor Carl G. Gahmberg
Faculty of Biosciences
Department of Biological and Environmental Sciences
Division of Biochemistry
University of Helsinki, Finland

ISBN 978-952-10-5184-5 (pbk.)
ISBN 978-952-10-5185-2 (PDF)
ISSN 1795-7079

Helsinki University Print
Helsinki 2008

To my beloved husband Harri

Contents

Abbreviations

4-ABP	4-aminobiphenyl
ALG13	Yeast *N*-acetylglucosamine transferase
AzMC	7-Azido- 4-methylcoumarin
AZT	Azidothymidine (Zidovudine)
BGT	β-glucosyltransferase
BMH	Bismaleimidohexane
CalG3	Calicheamicinone-4,6-dideoxy-4-hydroxylamino-α-D-glucosyltransferase from *Micromonospora echinospora*
CAZy	Carbohydrate-Active enZymes database
CN-I	Crigler-Najjar syndrome type I
CN-II	Crigler-Najjar syndrome type II
COS	Cells CV-1, simian in origin and carrying SV40 genetic material
CPT-11	Irinotecan
CYP	Cytochrome P450
DEPC	Diethyl pyrocarbonate
ER	Endoplasmic reticulum
GmIF7GT	UDP-isoflavone 7-*O*-glucosyltransferase from the roots of soybean (*Glycine max*) seedlings
GT	Glycosyltransferase
GtfB	TDP/UDP-glucose:aglycosyl-vancomycin glucosyltransferase from *Amycolatopsis orientalis* A82846
GtfA	dTDP-β-L-4-epi-epivancosamine: epivancosaminyltransferase from *Amycolatopsis orientalis* A82846
GtfD	UDP-β-L-4-epi-vancosamine: vancomycin-pseudoaglycone vancosaminyltransferase from *Amycolatopsis orientalis* ATCC19795
HA	Hemagglutinin
HEK293	Human embryonic kidney cells line 293
HPLC	High performance liquid chromatography
IMAC	Immobilised metal ion affinity chromatography
K_m	Michaelis constant
MBP	Maltose binding protein
4-MU	4-methylumbelliferone
MurG	UDP-GlcNAc: *N*-acetylmuramyl-(pentapeptide)-PP-C55 *N*-acetylglucosaminyltransferase from *Eschericia coli*
NDP	Nucleotide diphosphate
OleD	Oleandomycin glycosyltransferase D
OleI	Oleandomycin glycosyltransferase I
RMSD	Root mean square deviation
Sf9	*Spodoptera frugiperda*-derived cell line
SN-38	7-Ethyl-10-hydroxycamptothecin
SDS-PAGE	Sodium dodecyl sulfate polyacrylamide gel electrophoresis

TLC	Thin layer chromatography
TCC	Tetrachlorocatechol
TDP	Thymidine diphosphate
UDP	Uridine diphosphate
UDPGA	Uridine diphosphoglucuronic acid
UGT	UDP-glucuronosyltransferase
UGT71G1	Triterpene/flavonoid glycosyltransferase from the legume *Medicago truncatula*
UrdGT2	C-C bond-forming dTDP-D-olivose-transferase from *Streptomyces fradiae*
UGT85H2	Multifunctional UDP-glucose : (iso)flavonoid β-glucosyltransferase from *Medicago truncatula*
UGT72B1	2,5-dihydroxybenzoic acid-3,4-dihydroxybenzoic acid-glucosyltransferase from *Arabidopsis thaliana*
UPLC	Ultra performance liquid chromatography
*Vv*GT1	UPD-glucose: flavonoid 3-*O*-glucosyltransferase from red grape (*Vitis vinifera*)
V_{max}	Maximum velocity
1YD	Y486D mutant form of UGT1A1
6YD	Y485D mutant form of UGT1A6
Å	Ångström, 10^{-10} m

Amino acids

A	Ala	Alanine
R	Arg	Arginine
N	Asn	Asparagine
D	Asp	Aspartate
C	Cys	Cysteine
E	Glu	Glutamate
Q	Gln	Glutamine
G	Gly	Glycine
H	His	Histidine
I	Ile	Isoleucine
L	Leu	Leucine
K	Lys	Lysine
M	Met	Methionine
F	Phe	Phenylalanine
P	Pro	Proline
S	Ser	Serine
T	Thr	Threonine
W	Trp	Tryptophan
Y	Tyr	Tyrosine
V	Val	Valine

List of original publications

This thesis is based on the following publications:

I Kurkela M, **Patana A-S**, Mackenzie PI, Court MH, Tate CG, Hirvonen J, Goldman A and Finel M (2007) Interactions with other human UDP-glucuronosyltransferases attenuate the consequences of the Y485D mutation on the activity and substrate affinity of UGT1A6. *Pharmacogenet Genomics* **17:**115-126.

II **Patana A-S**, Kurkela M, Goldman A and Finel M (2007) The human UDP-glucuronosyltransferase: identification of key residues within the nucleotide-sugar binding site. *Mol Pharmacol* **72:**604-611.

III Xiong Y, **Patana A-S**, Miley MJ, Zielinska AK, Bratton SM, Miller GP, Goldman A, Finel M, Redinbo MR and Radominska-Pandya A (2008) The first aspartic acid of the DQxD motif for human UDP-glucuronosyltransferase 1A10 interacts with UDP-glucuronic acid during catalysis. *Drug Metab Dispos* **36:**517-522.

IV **Patana A-S**, Kurkela M, Finel M and Goldman A (2008) Mutation analysis in UGT1A9 suggests a relationship between substrate and catalytic residues in UDP-glucuronosyltransferases. *Protein Eng Des Sel* **21:**537-543.

The publications are referred to in the text by their roman numerals (I-IV).

Abstract

UDP-glucuronosyltransferases (UGTs, EC 2.4.1.17) are membrane bound glycosyltransferases located in the endoplasmic reticulum (ER). They catalyse the transfer of glucuronic acid from UDP-glucuronic acid (UDPGA) to endogenic or exogenic compounds, usually highly lipophilic molecules. As a consequence the aqueous solubility of the substrate (aglycone) is increased and it is more easily excreted from the body. UGTs consist of two domains located on the lumenal side of the ER membrane, a single-pass transmembrane helix and a short cytosolic tail. Currently, 19 different UGTs are known in man and, based on sequence and gene organisation, they are divided into three subfamilies: UGT1A, UGT2A, and UGT2B. The functional unit of UGTs probably consists of two or more monomers.

Bilirubin, a degradation product of haemoglobin, is glucuronidated by UGT1A1 and depressed activity of this enzyme can lead to hyperbilirubinemia. The UGT1A1 Y486D mutation causes hyperbilirubinemia and, due to exon sharing in the UGT1A subfamily, individuals that are homozygous for the 1A1-Y486D mutation carry the corresponding mutation in all their UGT1As. Recombinant UGT1A6 carrying this mutation, 6YD, had drastically decreased glucuronidation activity and increased K_m for both the aglycone and UDPGA. However, co-infection of the 6YD mutant with UGT1A4 partially restored the glucuronidation activity and decreased the K_ms of substrates to close to wild-type levels. Strong evidence for the formation of a tight hetero-oligomer was provided by the co-purification experiment.

The purification of UGTs in an active form has turned out to be difficult and tortuous, as is their expression in the quantities required for crystallisation purposes. In the absence of a three-dimensional structure of the full-length enzyme, I used other methods to obtain structural information. Binding of the co-substrate UDPGA was studied in UGT1A6, UGT1A9, UGT1A10 and UGT2B7 by site-directed mutagenesis, modelling, single-point activity and kinetic measurements with several substrates. The results indicated that residues H371, E379, D395 and Q396 (UGT1A6 numbering) are involved in UDPGA binding. The effects of the mutations on the glucuronidation rate were substrate- and isoform-dependent.

UGTs belong to the GT1 family of glycosyltransferases. Many enzymes in this family use a serine protease-type catalytic mechanism, where histidine and aspartate form a “catalytic dyad”. Such a pair in human UGT1A9 could be H37 and either D143 or D148. This histidine is not totally conserved among human UGTs, being replaced by proline in UGT1A4 and leucine in UGT2B10. Interestingly, while most human UGTs tend to glucuronidate phenolic compounds leading to the formation of *O*-glucuronides, both UGT1A4 and UGT2B10 are more specialised for *N*-glucuronidation. The roles of H37, D143 and D148 were investigated in UGT1A9 by site-directed mutagenesis, single-point activity and kinetic measurements. H37 was shown to be much more important in *O*- than in *N*-glucuronidation. D148 also played a different role in these two types of glucuronidation reactions. The results implied that H37-D143 is the catalytic dyad in UGT1A9, but could not exclude the possibility that D148 is the aspartate in the catalytic dyad.

1. Introduction

Mammalian UDP-glucuronosyltransferases (UGTs) play a major role in converting both exogenic and endogenic substances to more water soluble forms by conjugating them with the glucuronic acid from UDP-glucuronic acid (UDPGA). This enzymatic reaction is a biotransformation reaction, the purpose of which in general is to remove foreign lipophilic compounds from the body to bile or urine. These elimination processes are divided into phase I and phase II metabolism. Cytochrome P450 mono-oxygenases (CYPs) are the major phase I enzymes and UDP-glucuronosyltransferases are the major phase II enzymes. Quite often, the molecules undergo modifications in both phases before they can be excreted (Ioannides, 2002).

The type of enzymatic reaction catalysed by UGTs, conjugation of a carbohydrate group to another molecule, is one of the most common reactions in nature and the enzymes catalysing these reactions are called glycosyltransferases (GTs) (Davies, *et al.*, 2005). UGTs belong to this enzyme class and, based on sequence and predicted structure, they are further classified as members of the GT-B family and the GT1 subfamily. Unlike most other GT1 enzymes, which are soluble (Ross, *et al.*, 2001) and function as monomers (Finel and Kurkela, 2008), mammalian UGTs are bound to the endoplasmic reticulum (ER) and most probably function as dimers or higher oligomers (Finel and Kurkela, 2008; Coleman, *et al.*, 1997). One purpose of the studies in the present thesis was to shed light on the homo- or hetero-oligomerisation of UGTs and its possible physiological significance (Study I).

Due to the nature of the reaction, UGTs have a wide variety of different kinds of substrates that prior to glucuronidation, are referred to as aglycones. Each of the currently known 19 human UGT-isoforms glucuronidates several substrates; in addition one substrate is usually glucuronidated by several isoforms. This kind of promiscuity is advantageous for enzymes involved in detoxification, but understanding the mechanism behind it is still one of the major challenges in the UGT research field. Since 1953, when UDPGA was first characterised as a donor nucleotide sugar in glucuronidation reactions (Dutton and Storey, 1953), our knowledge of these enzymes and reactions has increased considerably. However, due to the inherent difficulties of expressing and purifying these membrane-bound enzymes as active proteins, no complete three-dimensional structures of any of the isoforms have yet been determined. Only after solving some eukaryotic GT1 structures during the last few years (Shao, *et al.*, 2005; Offen, *et al.*, 2006; Brazier-Hicks, *et al.*, 2007; Li, *et al.*, 2007b) and obtaining a partial crystal structure of the carboxy-terminal domain of human UGT2B7 (Miley, *et al.*, 2007), have we managed to gain some insight into the possible structural details of human UGTs. In the present thesis I used homology modelling, site-directed mutagenesis, single point activity and kinetic measurements (Studies II, III and IV) to identify which residues are involved in UDPGA and aglycone binding and to examine the catalytic mechanism.

The most common type of glucuronidation reaction is *O*-glucuronidation, where the glucuronic acid is conjugated to a hydroxyl group of the aglycone, but amino groups can also serve as conjugation targets leading to the formation of *N*-glucuronides. Certain UGT isoforms, namely UGT1A4 and UGT2B10, seem to be more specialised in less studied *N*-glucronidation (Green, *et al.*, 1995; Green, *et al.*, 1998; Green and Tephly, 1998; Kaivosaari,

et al., 2002; Sorich, *et al*., 2006; Chen, *et al*., 2007; Kaivosaari, *et al*., 2007). The mechanistic differences between *O*- and *N*-glucuronidation are presented and discussed in the present thesis (Study IV).

2. Review of the literature

2.1. Xenobiotic metabolism

It has been estimated that during one day humans consume 5,000 – 10,000 different chemicals, of which quite a few are potential carcinogens. Most of these chemicals come from our diet and are natural substances originating from plants or generated during cooking (Ames and Gold, 1998; Ames, *et al*., 1990). They do not naturally occur in our bodies and we cannot utilise them; thus they are recognised as foreign compounds that should be eliminated, and are frequently called xenobiotics (Greek: xenos = "foreign", bios = "life"). However, the chemicals that finally reach our bodies are usually lipophilic and cannot be excreted *via* bile or urine without modification. For this purpose, the body has efficient enzyme systems to convert these hydrophobic substances into more hydrophilic ones and thus facilitate their elimination. This xenobiotic biotransformation process protects us from hazardous effects in several ways: it shortens our exposure to the foreign chemicals and prevents their accumulation in our bodies. Usually, the biological activity of the compound is abolished during this biotransformation process, but in some cases the chemical metabolite may be more toxic than the parent compound (toxification or bioactivation). The cellular localisation of the phase I and phase II biotransformation enzymes is mainly the endoplasmic reticulum (ER) and cytosol. They are found in many different tissues, but the liver is the main site for xenobiotic metabolism (Ioannides, 2002).

2.1.1. Phase I metabolism

Three types of Phase I metabolism exist: oxidation, reduction and hydrolysis. The oxidation reactions are catalysed by different enzymes, such as cytochrome P450 (CYP), flavin mono-oxygenases (FMO) and amine, xanthine and aldehyde oxidases (Ioannides, 2002; Zhang, *et al*., 2006). Of these, the cytochrome P450 enzymes form the most important group. They catalyse reactions, in which one atom of oxygen is incorporated into the organic substrate and the other oxygen atom is reduced to water. The aromatic and aliphatic hydroxylation reactions are common examples of such oxidation reactions (Guengerich, 2002; Sheweita, 2000). It is estimated that CYP superfamily enzymes are involved in the metabolism of about 80% of current drugs (Fisher, *et al*., 2001). The other possible phase I reactions are epoxidation, dealkylation, oxidative deamination, nitroreduction, azoreduction, reductive dehalogenation and hydrolysis reactions. Thus, in phase I reactions some functional group (*e.g.* –OH, -COOH, $-NH_2$ or -SH) is typically introduced (Zhang, *et al*., 2006). Often these reactions are followed by phase II reactions (Fig. 2.1).

2.1.2. Phase II metabolism

Phase II metabolic reactions are conjugation reactions primarily catalysed by different kinds of transferases. In these reactions a molecule is added to a functional group, often that which was added during phase I. Phase II conjugation reactions include sulfonation catalysed by cytosolic sulfotransferases, gluthatione conjugation by glutathione *S*-transferases, methylation by methyltransferases, acetylation by acetyltransferases and amino acid conjugation by amino acid acyltransferases. However, the most important phase II reaction is glucuronidation catalysed by UDP-glucuronosyltransferases. They are responsible for about 35% of all drugs metabolised by phase II enzymes (Kiang, *et al*., 2005). I discuss UGTs and their relationship to other enzymes in what follows.

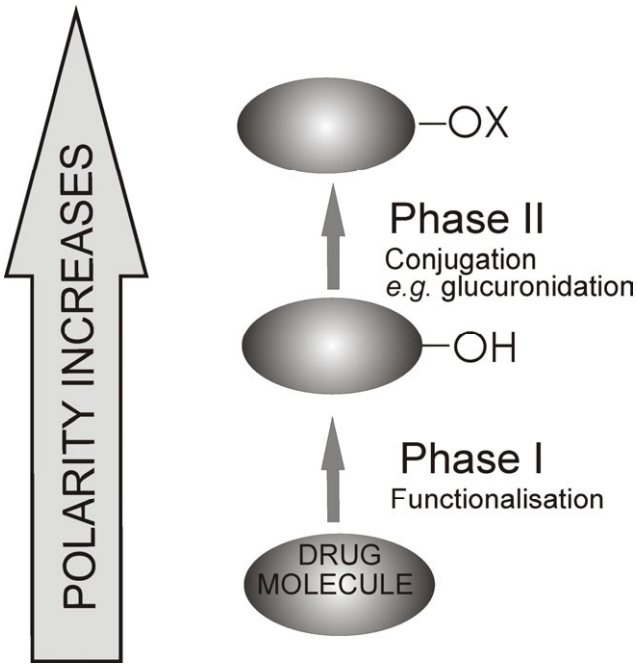


Figure 2.1 *Schematic representation of phase I and phase II metabolic reactions. Molecules can undergo both phase I and phase II reactions or only one of the two. In the figure, oxygen (O) can be also replaced by nitrogen (N) or sulfur (S).*

2.2. Glycosyltransferases

Based on amino acid sequence and predicted structure, human UDP-glucuronosyltransferases belong to the glycosyltransferase (GT) superfamily, which forms one of the largest enzyme groups in nature. They catalyse the transfer of a glycosyl moiety from a sugar donor to a saccharide, protein, lipid, DNA or small molecule acceptor (Unligil and Rini, 2000) and are thus involved in the synthesis of oligosaccharides, polysaccharides, and glycoconjugates (Breton, *et al*., 2006). Like human UGTs, the majority of GTs use an activated nucleotide sugar as donor. Glycosyltransferases have a significant role in many

important biological processes, like recognition and signalling (Breton and Imberty, 1999; Campbell, *et al*., 1997). Based on sequences and donor sugars, glycosyltransferases have been classified into different families (91 at the moment). The most recent information and classification of these enzymes is available from the continuously updated carbohydrate-active enzyme database (CAZy) at http://afmb.cnrs-mrs.fr/CAZY/ (Campbell, *et al*., 1997; Coutinho, *et al*., 2003).

Even though the diversity of the reactions catalysed and the amino acid sequences is high, almost all the glycosyltransferases so far have been found to adopt either the GT-A or GT-B folds (Unligil and Rini, 2000; Bourne and Henrissat, 2001; Hu and Walker, 2002; Lairson, *et al*., 2008). The GT-A fold consists of two dissimilar β/α/β domains, with the amino-terminal domain binding the nucleotide sugar and the highly variable carboxy-terminal domain binding the acceptor. Almost all the GT-A family members have a common DxD motif, which is involved in coordinating a divalent cation, usually a manganese or magnesium ion, in the catalytic centre. The metal ion is required for the binding of the nucleotide sugar (Breton and Imberty, 1999; Wiggins and Munro, 1998). The GT-A fold was originally detected in the nucleotide-diphospho-sugar transferase SpsA from *Bacillus subtilis* (Charnock and Davies, 1999) (Fig. 2.2A).

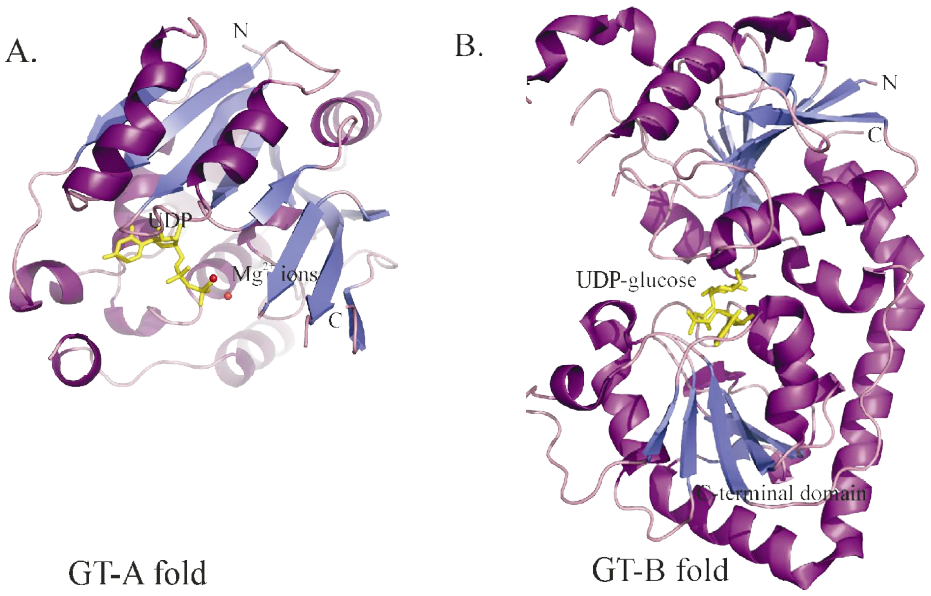


Figure 2.2 *Cartoon representation of glycosyltransferase GT-A and GT-B folds. Beta sheets are coloured blue, alpha helices purple, nucleotide sugar yellow and magnesium ions in the GT-A fold red.* ***A.*** *Nucleotide-diphospho-sugar transferase SpsA from* Bacillus subtilis. ***B.*** *T4 phage β-glucosyltransferase.*

The GT-B fold consists of two rather similar Rossman fold domains separated by a linker region, with the catalytic site placed between the two domains (Fig. 2.2B). The amino-terminal domain binds the acceptor and the carboxy-terminal domain the nucleotide-sugar

(Breton, *et al.*, 2006; Bourne and Henrissat, 2001). Even though the carboxy-terminal domains between different GT-B family members are highly conserved, no strictly conserved residues have been found (Hu and Walker, 2002). In contrast to the GT-A family, there is no evidence of a bound metal ion, although divalent cations may be needed for full activity. The first GT-B structure solved was phage T4 β-glucosyltransferase (Vrielink, *et al.*, 1994) (Fig. 2.2B).

GT1 enzymes

Within the superfamily of glycosyltransferases, human UGTs belong to the GT1 family according to CAZy database. Enzymes of that family adopt the GT-B fold and typically transfer sugars to lipophilic small-molecule acceptors (Bowles, *et al.*, 2006). About 50% of the GT1 family enzymes contain a signature motif (Bowles, *et al.*, 2006; Mackenzie, *et al.*, 1997; Paquette, *et al.*, 2003) (see Fig. 2.6). So far, the GT1 family contains seven full three-dimensional structures from bacteria and four from plants. Bacterial TDP-epi-vancosaminyltransferase GtfA, UDP-glucosyltransferase GtfB and vancosaminyltransferase GtfD are involved in vancomycin synthesis (Mulichak, *et al.*, 2003; Mulichak, *et al.*, 2001; Mulichak, *et al.*, 2004); oleandomycin glycosyltransferases OleD and OleI inactivate oleandomycin and diverse macrolide antibiotics (Bolam, *et al.*, 2007); CalG3 is calicheamicinone 4,6-dideoxy-4-hydroxylamino-α-D-glucosyltransferase (Zhang, *et al.*, 2008); and glycosyltransferase UrdGT2 catalyses the formation of a C-C bond between a polyketide aglycone and D-olivose (Mittler, *et al.*, 2007). Three of the plant GT1 structures, UGT71G1, *Vv*GT1 and UGT85H2 are flavonoid glucosyltransferases (Shao, *et al.*, 2005; Offen, *et al.*, 2006; Li, *et al.*, 2007b) and UGT72B1 is a bifunctional *N*- and *O*-glucosyltransferase (Brazier-Hicks, *et al.*, 2007). In addition there are two partial structures: the carboxy-terminal domains of human UGT2B7 (Miley, *et al.*, 2007) and yeast *N*-acetylglucosamine transferase (ALG13) (Wang, *et al.*, 2008). Although the sequence identity between human UGTs and those GT1 enzymes whose structures have been solved is typically below 20%, the fold is predicted to be the same.

2.3. UDP-glucuronosyltransferases

2.3.1. Overview

UDP-glucuronosyltransferases (UGTs, EC 2.4.1.17) are major phase II enzymes that play an important role in the detoxification of different endogenous and exogenous compounds, *e.g.* drugs and environmental toxins, in humans. In the glucuronidation reaction glucuronic acid from the donor UDP-glucuronic acid (UDPGA) is conjugated to a functional group, mostly a hydroxyl, carboxyl, amino or thiol group on the activated target molecule (aglycone), leading to the formation of *O*-, *N*-, or *S*-glucuronides, respectively (Fig. 2.3)

(Dutton, 1980). Some compounds, *e.g.* sulfinpyrazone or phenylbutazone, are metabolised *via* uncommon *C*-glucuronidation pathway (Fig. 2.3) (Richter, *et al.*, 1975; Kerdpin, *et al.*, 2006; Nishiyama, *et al.*, 2008).

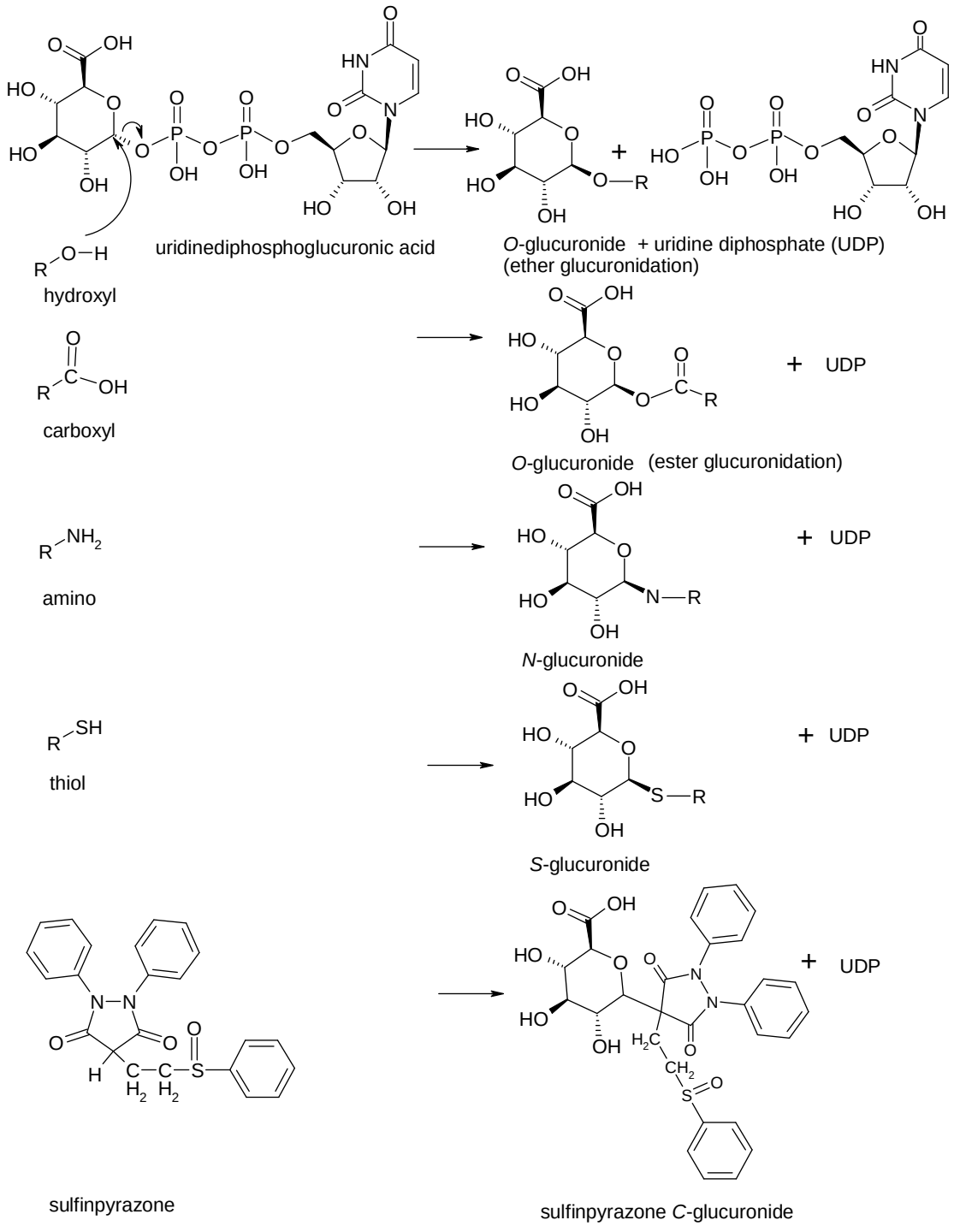


Figure 2.3 *The* O-, N-, S- *and an example of* C*-glucuronidation reactions catalysed by UDP-glucuronosyltransferases.*

By forming the β-D-glucuronide, the polarity of the aglycone is increased and it can be more easily excreted to bile or urine. This glucuronidation reaction is an important detoxification pathway not only in humans, but also in other vertebrates. Usually the β-D-glucuronides are not biologically active, but in some cases increased toxicity of the glucuronide has been reported (Radominska-Pandya, *et al.*, 1999; King, *et al.*, 2000; Tukey and Strassburg, 2000; Fisher, *et al.*, 2001; Ouzzine, *et al.*, 2003). In addition to various exogenous substrates, UDP-glucuronosyltransferases have an important role in the elimination of several endogenous substrates, such as bile acids, steroid hormones and especially bilirubin, the degradation product of haemoglobin. It is removed from the body by the liver isoform UGT1A1. Mutations leading to non-functional or less active UGT1A1 can cause hyperbilirubinemia, classified either as Crigler-Najjar or Gilbert syndrome. Untreated, the severe form of hyperbilirubinemia can lead to death (Hirschfield and Alexander, 2006). UGTs are mainly found in the liver, but they have also a very significant role in extra-hepatic metabolism in the gastrointestinal tract and kidney (Fisher, *et al.*, 2001; Tukey and Strassburg, 2001; Soars, *et al.*, 2002).

Human UGTs are around 530 amino acids in length, consisting of two approximately similar-sized domains, like other GT-B fold glycosyltransferases (Radominska-Pandya, *et al.*, 1999) (Fig. 2.2B). The amino-terminal domain binds the aglycone and the carboxy-terminal domain the UDPGA (Mackenzie, 1990). The first 25 residues or so form a signal sequence that directs the enzyme to the endoplasmic reticulum (ER) and is later cleaved off (Kurkela, *et al.*, 2003). The length of the mature protein is thus between 500 and 510 residues. The enzymes are 50 - 60 kDa in size and most of their mass is located in the ER lumen; they have a single-pass 17-residue long transmembrane helix near the carboxy-terminus and a 19 - 26 residue long cytosolic tail (Meech and Mackenzie, 1998; Radominska-Pandya, *et al.*, 1999). Mammalian UGTs may function as dimers or higher oligomers (Radominska-Pandya, *et al.*, 1999; Finel and Kurkela, 2008) (Fig. 2.4).

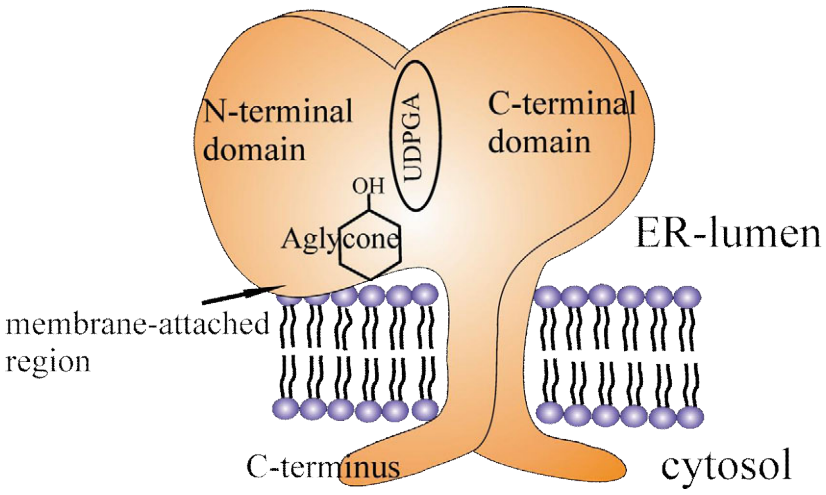


Figure 2.4 *Schematic representation of human UGT topology. UGTs consist of two domains and are predicted to function as dimers or higher oligomers. The amino-terminal domain binds the aglycone and the carboxy-terminal domain the UDPGA cosubstrate; the catalytic site is placed between the two domains. Most of the enzyme mass is located on the lumenal side of the endoplasmic reticulum and the carboxy-terminal tail is on the cytosolic side of the membrane.*

In addition to the carboxy-terminal transmembrane helix, the amino-terminal domain may also contain a membrane-attached region (Ciotti, *et al*., 1998). UGTs can be retained in the ER even if the transmembrane and cytosolic parts are removed (Meech and Mackenzie, 1998; Ouzzine, *et al*., 1999a; Ouzzine, *et al*., 1999b). This amino-terminal membrane region may be needed to help highly lipophilic substances to reach the active site (Radominska-Pandya, *et al*., 2005). The region has also been suggested to play a role in dimerisation. This conclusion was arrived at based on studies where residues 152 – 180 (the predicted membrane embedded region) were deleted, almost completely abolishing the dimerisation of UGT1A1 (Ghosh, *et al*., 2001).

Based on amino acid sequence and gene organisation, human UGTs can be divided into two major subfamilies, UGT1 and UGT2, or alternatively into three subfamilies, UGT1A, UGT2A and UGT2B (Mackenzie, *et al*., 1997; Mackenzie, *et al*., 2005) (Fig. 2.5). In the UGT1 subfamily, the carboxy-terminal half of the protein is encoded by the shared exons 2 to 5 and is thus identical between all human UGT1-enzymes, whereas the variable amino-terminal half, encoded by exon 1, is responsible for substrate specificity (Ritter, *et al*., 1992a; Gong, *et al*., 2001). UGT2 enzymes are encoded by separate genes, except UGT2A1 and UGT2A2, which probably use exon-sharing (Mackenzie, *et al*., 2005). 19 different human UGTs have been recognised and thus far cDNAs for 17 functional human UGT-isoforms have been isolated; several pseudogenes are also known (Mackenzie, *et al*., 2005; Sorich, *et al*., 2008) (Fig. 2.5).

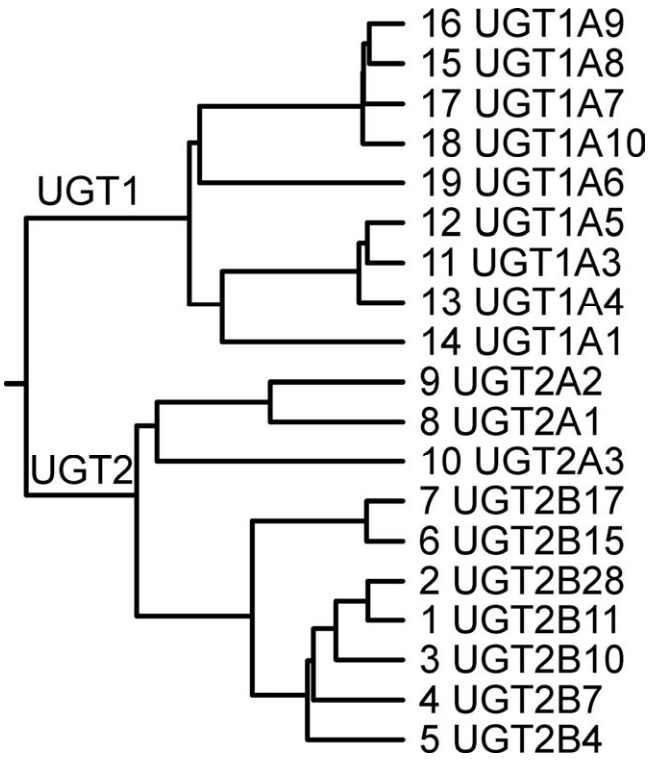


Figure 2.5 *Phylogenetic tree of human UGT-enzymes showing evolutionary distances between different isoforms. The alignment was made by ClustalW* (Larkin, *et al*., 2007) *and the tree was constructed with the evolutionary Trace Server, TraceSuite II* (Innis, *et al*., 2000).

Because they are xenobiotic-metabolising enzymes, human UGTs have a wide variety of substrates. Several hundred compounds have been identified as substrates for these enzymes (Tukey and Strassburg, 2000) and the list is growing continually. Usually one isoform can glucuronidate several substrates and one substrate can be glucuronidated by several isoforms (Mackenzie, *et al*., 2000), though several isoform-specific substrates are also known (Court, 2005). Because of the difficulties in expressing these enzymes and purifying them in active form, there is no complete three-dimensional structure of any mammalian UGT. The only mammalian UGT crystal structure so far is a segment of the carboxy-terminal domain of human UGT2B7 (Miley, *et al*., 2007). UGTs are suggested to function as dimers, but in the absence of pure active enzyme this cannot be definitively proven. The structure of the amino-terminal domain is not known. Structural information on how aglycone or UDPGA binds is also not available.

2.3.2. Hyperbilirubinemia

250 to 350 mg of bilirubin (see Fig. 2.9), a degradation product of haemoglobin, is formed in the human body per day. It is highly insoluble in water and therefore bound to albumin in the circulation (Harb and Thomas, 2007). This unconjugated bilirubin is then transported to the liver, where it is conjugated with glucuronic acid before excretion into bile. UGT1A1 is the only significant bilirubin glucuronidating isoform in human (Bosma, *et al*., 1994; Bosma, 2003). Elevated levels of unconjugated bilirubin in the serum result in jaundice and a condition known as unconjugated hyperbilirubinemia.

Three forms of unconjugated hyperbilirubinemia have been identified in man: Gilbert syndrome and two types of Crigler-Najjar syndrome, CN-I and CN-II (Crigler and Najjar, 1952; Kadakol, *et al*., 2000; Bosma, 2003). Gilbert syndrome is a mild and very common form of hyperbilirubinemia, leading to transiently elevated bilirubin levels (Strassburg, 2008). Five to ten percent of the Caucasian population are estimated to have Gilbert syndrome (Owens and Evans, 1975; Kraemer and Klinker, 2002; Bosma, 2003). It is commonly caused by additional TA nucleotides in the TATA box of the UGT1A1 gene promoter region, leading to reduced UGT1A1 expression levels (Bosma, *et al*., 1995). This allele, UGT1A1*28, reduces gene transcription to about 20% of the normal level (Bosma, 2003) and the hepatic UGT1A1 bilirubin activity is usually around 30% of normal (Yamamoto, *et al*., 1998; Miners, *et al*., 2002; Rouits, *et al*., 2004). Gilbert syndrome can also be caused by point mutations in the coding region of UGT1A1.

Crigler-Najjar syndrome (CN) leads to moderately or highly elevated unconjugated bilirubin levels. CN-I is a rare, but very severe, form of hyperbilirubinemia, in which UGT1A1 activity is undetectable (Francoual, *et al*., 2002). In CN-I patients serum bilirubin values are usually above 350 µM and the amount of bilirubin conjugates in bile is extremely low (Jansen, 1999). Without treatment, CN-I disease leads to death. The condition can be completely cured only by liver transplantation. CN-II is a milder form of hyperbilirubinemia. Serum bilirubin levels are usually below 350 µM and are lowered by phenobarbital treatment (Jansen, 1999). The amount of bilirubin monoconjugates in bile is measurable, and low

amounts of di-conjugates can also be detected (Jansen, 1999). UGT1A1 activity is detectable, but below 10% of the normal level (Koiwai, *et al*., 1996).

Both CN-type diseases are inherited and caused by mutations in UGT1A1 coding exons. Mutations are either in the amino-terminal domain, coded by exon 1, or in the carboxy-terminal domain, coded by exons 2 – 5. As exons 2 – 5 are common to all members of the UGT1 subfamily, mutations in these could lead to impaired activity of all the isoforms of the subfamily (Jansen, 1999). The first mutations leading to CN-I syndrome were identified in 1992 (Bosma, *et al*., 1992; Ritter, *et al*., 1992b) and to CN-II syndrome in 1993 (Aono, *et al*., 1993; Bosma, 2003). Several mutations leading to either Gilbert, CN-I or CN-II type disorders have been detected in the human UGT1A1 gene (113 UGT1A1 variant forms have been reported so far, of which 74 are CN variants). The UDP-Glucuronosyltransferase Alleles Nomenclature page (http://www.pharmacogenomics.pha.ulaval.ca/sgc/ugt_alleles/) and references therein contain the most up-to-date summary of the reported mutations.

Y486D mutation

One of the first mutations identified in patients is Y486D, which leads to a mild form of hyperbilirubinemia (Aono, *et al*., 1993; Maruo, *et al*., 1998). In the homozygous form it has been found in both Gilbert- and CN-II-type syndrome patients, the variation probably being due to other genetic differences (Takeuchi, *et al*., 2004). Normally these conditions are not accompanied by many symptoms, but the mutation may contribute to the toxic effects of certain medications. One example is the antitumor agent irinotecan (CPT-11), which is metabolised to 7-ethyl-10-hydroxycamptothecin (SN-38) (Jinno, *et al*., 2003). SN-38 is the active metabolite and is further glucuronidated by UGT1A1. The effect of the Y486D mutation in UGT1A1 on SN-38 glucuronidation was investigated with COS-1 cells expressing the mutant protein. The glucuronidation efficiency was drastically decreased (Jinno, *et al*., 2003). The equivalent mutation was also studied using the UGT1A1 and UGT1A6 (Y485D mutation in UGT1A6) isoforms expressed in COS-7 cells. The glucuronidation of 2-amino-5-nitro-4-trifluoromethylphenol (a major metabolite of flutamide, a nonsteroidal anti-androgenic agent) was about 12% of the wild type activity in the UGT1A1 mutant, but below 1% in the UGT1A6 mutant (Ito, *et al*., 2002). Overall, the effect of the Y→D mutation on the glucuronidation rate varies depending on the substrate and the isoform studied (Udomuksorn, *et al*., 2007). Usually the glucuronidation rate is substantially reduced, except for UGT1A9, where even increased glucuronidation rates have been reported (Kurkela, *et al*., 2004).

2.3.3. Oligomerisation of mammalian UGTs

Mammalian UGTs, like many other single-pass membrane proteins, are thought to function as dimers or oligomers. Several different methods have been used in attempts to clarify the oligomeric state of these enzymes. Based on gel filtration and polyacrylamide gradient gel

electrophoresis, it was proposed at the beginning of the 1980s that rat liver UGTs exist as units larger than monomers (Matern, *et al*., 1982). Radiation inactivation experiments in 1984 suggested that UGTs may exist as monomers or oligomers consisting of up to four subunits (Peters, *et al*., 1984). In this method, membranes containing the protein are exposed to ionizing radiation; the degree of inactivation is proportional to the radiation dose and the size of the molecule (Radominska-Pandya, *et al*., 2005; Finel and Kurkela, 2008). Later studies based on this method suggested that monoglucuronide formation of phenols is catalysed by dimeric UGTs, and diglucuronide formation by tetrameric UGTs (Gschaidmeier and Bock, 1994).

In 1997 Meech and Mackenzie proposed that UGTs form catalytically active dimers *via* their amino-terminal domains. This conclusion was derived from mutation, co-expression and SDS-PAGE analyses (Meech and Mackenzie, 1997). They showed that two inactive units of UGT2B1 can become active through oligomerisation. Evidence for oligomerisation also came from work in which a human UGT1A1 nonsense mutation led to the formation of inactive enzyme and caused Crigler-Najjar syndrome type II (CN-II). The inactive mutant form was shown to inhibit the activity of the full-length UGT1A1, indicating an interaction between them (Koiwai, *et al*., 1996).

Immunopurification studies with rat hepatic microsomes suggested that there were protein-protein interactions between UGT1s and UGT2B1 (Ikushiro, *et al*., 1997). In the study, UGT1 isoforms were bound to a Sepharose column which was conjugated with UGT1 carboxy-terminal-specific anti-peptide antibodies. UGT2B1 was shown to co-elute with UGT1-isoforms. Chemical cross-linking studies with bismaleimidohexane (BMH), which is reactive towards sulfhydryl groups, suggested that UGTs form heterodimers (Ikushiro, *et al*., 1997). Studies using gel permeation chromatography of solubilised microsomes implied that UGT1A1 is a dimer in solution and the yeast two-hybrid system indicated that it has intermolecular interactions only with itself (Ghosh, *et al*., 2001). Co-immunoprecipitation studies using human liver microsomes suggested interactions between UGT1A1, UGT1A6 and UGT2B7 (Fremont, *et al*., 2005) and the same method, combined with *in vivo* fluorescence resonance energy transfer (FRET), was used to demonstrate protein-protein interactions in live COS cells (Operana and Tukey, 2007). In this study, Operana and Tukey (2007) showed that all the UGT1 subfamily members self-oligomerise (homodimerise) and, in contrast to a previous study (Ghosh, *et al*., 2001), UGT1A1 was also shown to form hetero-oligomers with UGT1A3, UGT1A4, UGT1A6, UGT1A7, UGT1A8, UGT1A9 and UGT1A10.

Oligomerisation of UGTs has also been studied with recombinant enzymes expressed in Sf9 insect cells (Kurkela, *et al*., 2003). Both His- and HA-tagged human UGT1A9-isoforms were produced and the solubilised membrane extract was loaded onto an IMAC-column, which binds His-tagged proteins. When His-tagged proteins were eluted with imidazole, HA-tagged enzymes were also detected in the same fractions, demonstrating an interaction between the monomers (Kurkela, *et al*., 2003).

Overall, it is now quite clear that the active units of UGTs are complexes larger than monomers. Nowadays they are often assumed to be dimers (Kurkela, *et al*., 2003; Kurkela, *et al*., 2004a; Kurkela, *et al*., 2004b; Operana and Tukey, 2007), even though higher oligomers cannot be excluded (Finel and Kurkela, 2008). Nonetheless, the question is more

complicated if we try to investigate the role and physiological significance of homo- or hetero-oligomerisation *in vivo*.

The formation of hetero-oligomers has been reported in several studies (Meech and Mackenzie, 1997; Ikushiro, *et al.*, 1997; Ishii, *et al.*, 2004; Operana and Tukey, 2007), even though in one study the UGT1A1 isoform was shown to form only homo-oligomers (Ghosh, *et al.*, 2001). The first study indicating that hetero-oligomerisation can affect substrate specificity was done on morphine with guinea pig enzymes UGT2B21 and UGT2B22 (Ishii, *et al.*, 2001). The same group confirmed the results later by showing that hetero-oligomerisation enhances chloramphenicol glucuronidation (Ishii, *et al.*, 2004). Hetero-oligomerisation between UGT1A4 and a truncated form of UGT1A9 (UGT1A9 from which the carboxy-terminal transmembrane helix and the following cytosolic tail has been removed - referred to as 1A9Sol) was shown to affect substrate specificity (Kurkela, *et al.*, 2004a). In this work, UGT1A4, which glucuronidates entacapone at nonsignificant rates, was co-expressed with 1A9Sol in Sf9 insect cells. The decrease in the K_m of entacapone was related to the probability of hetero-oligomer formation, thus revealing important monomer-monomer interactions within the UGTs (Kurkela, *et al.*, 2004a). Hetero-oligomerisation was also reported to affect the kinetic parameters (K_m and V_{max}) in studies where UGT1A1, UGT1A4 and UGT1A6 were co-expressed in HEK293 cells. K_m and V_{max} increased or decreased depending on the isoform and substrate used (Fujiwara, *et al.*, 2007a). Fujiwara *et al*. (2007a; 2007b) expressed double isoforms of UGT1A1/UGT1A9, UGT1A4/UGT1A9 and UGT1A6/UGT1A9 in HEK293 cells and analysed the interactions. UGT1A1, UGT1A4 and UGT1A6 acquired thermal stability and resistance to detergent when co-expressed with UGT1A9 and their kinetic parameters were also affected (Fujiwara, *et al.*, 2007b). Quite recently, alternative splicing variants of human UGT1A1 have been found in several tissues. These inactive variants decrease the activity of normal UGT1A1 (UGT1A1_i1) through direct protein-protein interactions (Lévesque, *et al.*, 2007; Girard, *et al.*, 2007). Thus hetero-oligomerisation may affect UGT function, enzymatic characteristics and substrate selectivity *in vivo* (Operana and Tukey, 2007; Fujiwara, *et al.*, 2007b), and could also play a negative regulatory role in the human glucuronidation pathway (Lévesque, *et al.*, 2007; Girard, *et al.*, 2007).

2.3.4. Co-substrate binding

The functional unit of human UGTs consists of two or more monomers, and the monomeric unit of mammalian UGTs consists of two domains. The more conserved carboxy-terminal domain of UGTs binds the co-substrate UDPGA and the more variable amino-terminal domain the aglycone. The 44-residue long consensus motif (PS00375), which is thought to be involved in binding the UDP moiety of the nucleotide sugar, is located in the carboxy-terminal half of all UDP-glycosyltransferases (Mackenzie, *et al.*, 1997) (Fig. 2.6). It can be found in mammalian, plant and bacterial species (Mackenzie, *et al.*, 1997; Ross, *et al.*, 2001). It also should be noted that, in the human UGT1A subfamily, the carboxy-terminal domain is encoded by the shared exons 2 to 5, and is thus totally identical among the different isoforms of this subfamily. The carboxy-terminal domains of UGT2 enzymes, except for UGT2A1

UGT2A2, are encoded by separate exons, but the amino acid sequence identity between UGT1 and UGT2 enzymes in this domain is still about 60% and the similarity over 70%.

[FW] - x(2) - [QL] - x(2) - [LIVMYA] - [LIMV] - x(4,6) - [LVGAC] - [LVFYAHM] - [LIVMF] - [STAGCM] - [HNQ] - [STAGC] - G - x(2) - [STAG] - x(3) - [STAGL] - [LIVMFA] - x(4,5) - [PQR] - [LIVMTA] - x(3) - [PA] - x(2,3) - [DES] - [QEHNR]

Figure 2.6 *Consensus sequence found in all UDP-glycosyltransferases (Prosite code PS00375). This motif is suggested to be involved in co-substrate binding and is part of the carboxy-terminal domain.*

The glucuronidation reaction takes place in the lumen of the endoplasmic reticulum. The hydrophilic UDPGA sugar synthesised in the cytosol has to reach the active site. Current data suggest that independent UDPGA transporters supply UDPGA to the catalytic site (Kobayashi, *et al.*, 2006).

The first suggestion that the UDPGA binding site is in the carboxy-terminal domain was based on studies with chimeric UGTs (Mackenzie, 1990). Later, results from photoaffinity labelling with UGT2B4 fusion peptides confirmed that the UDPGA binding site is between residues 299 and 446 (Pillot, *et al.*, 1993). This was further narrowed to residues 350 - 400 using photolabelled UGT1A6, and sequence and structure comparisons (Radominska-Pandya, *et al.*, 1999). However, photoaffinity labelling, inhibitor studies and sequence alignments suggest that UDPGA interacts not only with the carboxy-terminal but also with the amino-terminal domains of UGTs (Radominska-Pandya, *et al.*, 1999). The same conclusion was drawn from studies where a homology model of UGT1A1 was built using the plant protein UGT71G1 as a template and the UDPGA co-substrate was docked to the model (Li and Wu, 2007; Locuson and Tracy, 2007).

Comparison of the carboxy-terminal domain structure of UGT2B7 to the other GT1 family enzyme structures suggests that UDPGA binds to the same site as co-substrate in these enzymes (Miley, *et al.*, 2007) (Fig. 2.7). The predicted UDPGA binding site in UGTs seems to be remarkably similar to the UDP-glucose binding site in other GT1 enzymes (Mulichak, *et al.*, 2001; Shao, *et al.*, 2005; Offen, *et al.*, 2006; Brazier-Hicks, *et al.*, 2007; Li, *et al.*, 2007b; Mittler, *et al.*, 2007). This is reasonable, because UDP-glucose and UDP-glucuronic acid are very similar, differing only at the 6-position of the sugar (Fig. 2.8). The RMSD of the carboxy-terminal segment of human UGT2B7 compared to those GT1 structures that use UDP-glucose as a co-substrate is around 2 Å (Miley, *et al.*, 2007). The other co-substrates used in the GT1 family are UDP-rhamnose (Jones, *et al.*, 2003) and various TDP-derivatives (Mulichak, *et al.*, 2003; Mulichak, *et al.*, 2004; Mittler, *et al.*, 2007). CalG3, calicheamicinone-4,6-dideoxy-4-hydroxylamino-α-D-glucosyltransferase, is quite promiscuous for sugar donors and has been demonstrated to accept ten alternative NDP-sugar donors (Zhang, *et al.*, 2008).

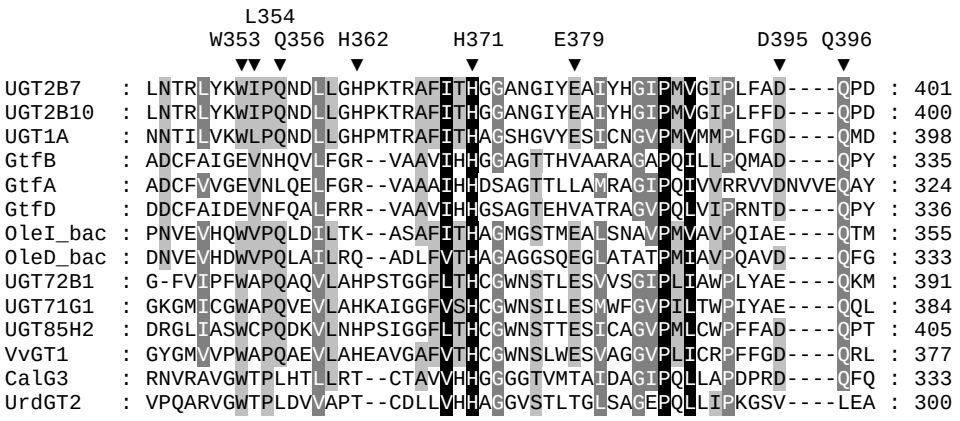


Figure 2.7 *Partial sequence alignment of 14 GT1 family enzymes. UGT2B7, UGT2B10 and UGT1A are human enzymes and the other eleven are crystallised glycosyltransferases whose complete structures have been solved. Shown are residues from 300 to 400, which is the most conserved region in all GT1 enzymes. The carboxy-terminal domain is totally identical between all UGT1A enzymes (UGT1A6 numbering is used for the UGT1A subfamily).*

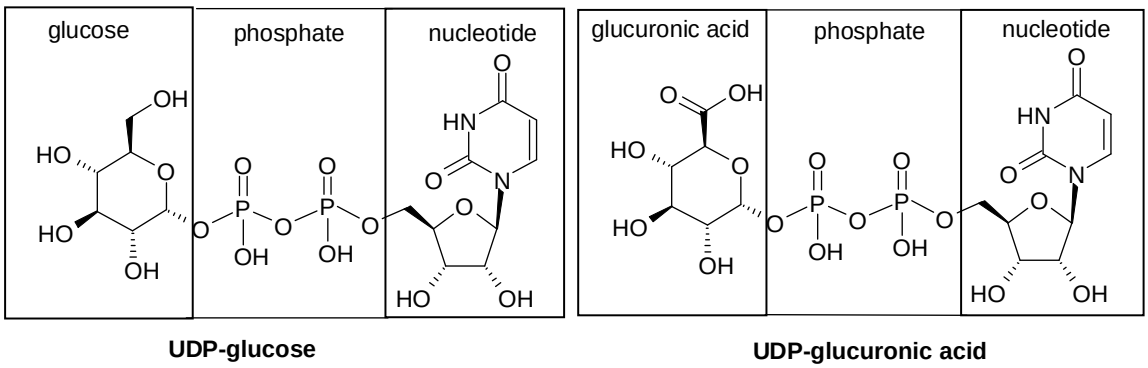


Figure 2.8 *Chemical structures of UDP-glucose and UDP-glucuronic acid.*

The key conclusions from the different modelling studies concerning the specific residues making contacts with UDPGA are quite consistent with each other (Miley, *et al.*, 2007; Li and Wu, 2007; Locuson and Tracy, 2007). In addition, the UDPGA binding site has been studied by chemical modification (Ouzzine, *et al.*, 2000) and by site-directed mutagenesis (Miley, *et al.*, 2007). Miley *et al.* (2007) mutated several residues in the predicted UDPGA binding region of UGT2B7 and studied the activities of mutants in single-point activity measurements. They divided the predicted interacting residues into three different groups: residues interacting with 1) nucleotide, 2) phosphate, or 3) glucuronic acid (Fig. 2.8). The residues interacting with phosphate or glucuronic acid seemed to be more critical for activity than those interacting with the nucleotide. Q356 is predicted to make contact with the uracil base; W353 is predicted to stack over the uracil ring and E379 is predicted to hydrogen bond to the ribose 2'-OH. The main chain carbonyl and -NH of L354 (Locuson and Tracy, 2007)

and side chain of R335 have been suggested to make contacts to the uridine group (Miley, *et al.*, 2007; Li and Wu, 2007; Locuson and Tracy, 2007). These interactions have been found in several other GT-B enzyme structures (Shao, *et al.*, 2005; Offen, *et al.*, 2006; Brazier-Hicks, *et al.*, 2007; Li, *et al.*, 2007b; Bolam, *et al.*, 2007). Mutations in residues W356, Q359 and R338 (W353, Q356 and R335 in UGT1A6 numbering), predicted to make contacts to uridine, had only small or moderate effects on the UGT2B7 glucuronidation activity (Miley, *et al.*, 2007).

The most conserved residue in the predicted UDPGA binding area is H371 (UGT1A6 numbering) (Fig. 2.7). It has been suggested to interact with the diphosphate moiety of UDPGA (Miley, *et al.*, 2007; Li and Wu, 2007; Locuson and Tracy, 2007) and may have a role in stabilising the leaving group during catalysis (Radominska-Pandya, *et al.*, 2005). Based on mutational studies, H362 has also been suggested to interact with UDPGA (Ouzzine, *et al.*, 2000), but this has not been confirmed by other studies. Residues D395 and Q396, which are highly conserved among GT1 enzymes, are predicted to interact with the 3'O and 4'O of the glucuronic acid. Mutating the equivalent residues in UGT2B7 (D398, Q399) had a drastic effect on enzyme activity (Miley, *et al.*, 2007). The high conservation of the corresponding residues among all vertebrate UGTs led to the suggestion that they play an important role in the recognition of the donor molecule (Li and Wu, 2007). In a very recent paper, based on biochemical experiments and homology-based modelling using *E. coli* UDP-galactose 4-epimerase as a template, residues K314 and K404 in UGT1A10 were suggested to be involved in UDPGA interactions (Banerjee, *et al.*, 2008). Nonetheless, the specific role of these residues has not been discussed in any other papers.

Based on photoaffinity labelling studies, the residues in the amino-terminal domain have also been proposed to make contacts to UDPGA (Radominska-Pandya, *et al.*, 1999). This view was supported by the structures of UDP-glucose:GT1-enzyme complexes. Suggestions that the amino-terminal domain residues G36, S37 and R195 in UGT1A1 are involved in UDPGA binding have only recently been presented (Li and Wu, 2007; Locuson and Tracy, 2007), and their roles have yet to be confirmed by mutational or biochemical studies. Generally the effect of mutations in the UDPGA binding area on activity seems to depend on substrate, implying that positioning of the UDPGA is more critical for some aglycones than for others.

2.3.5. Aglycone binding

As described above, the carboxy-terminal domain binds UDPGA. Conversely, the amino-terminal domain, formed from the first 260 residues of the mature protein, binds the aglycone (Mackenzie, 1990) and so is responsible for substrate specificity. The substrates typically contain aromatic ring structures and are highly hydrophobic; some are even classified as lipids (Fig. 2.9). An amino-terminal membrane-attached region, suggested to be located between residues 140 – 240 (Radominska-Pandya, *et al.*, 2005), may be needed to help highly lipophilic substances reach the active site. The more specific location of this region is not known, but based on analysis by the computer program RAOARGOS (Argos, *et al.*, 1982) the membrane-attached region has been suggested to be located between residues

159 – 177 (UGT1A1 numbering) (Ciotti, *et al.*, 1998). However, this needs to be confirmed by further studies.

An unusual feature of enzymes that metabolise xenobiotics is their promiscuity, and hence the aglycone binding site cannot be highly selective. Instead, each isoform has to be able to accept a wide range of molecules ranging from single phenols to long chain fatty acids and multi-ring structures (Fig. 2.9). In addition, all UGTs have different and usually partly overlapping substrate specificities, implying that the most variable residues may have an important role in substrate specificity and selection. The most diverse region for all UGTs is the area between residues 90 and 190 (Radominska-Pandya, *et al.*, 1999). By constructing UGT2B7-UGT2B15 chimeras, Lewis *et al.* have in fact shown that residues 61 – 194 of UGT2B15 are responsible for substrate binding and selectivity (Lewis, *et al.*, 2007).

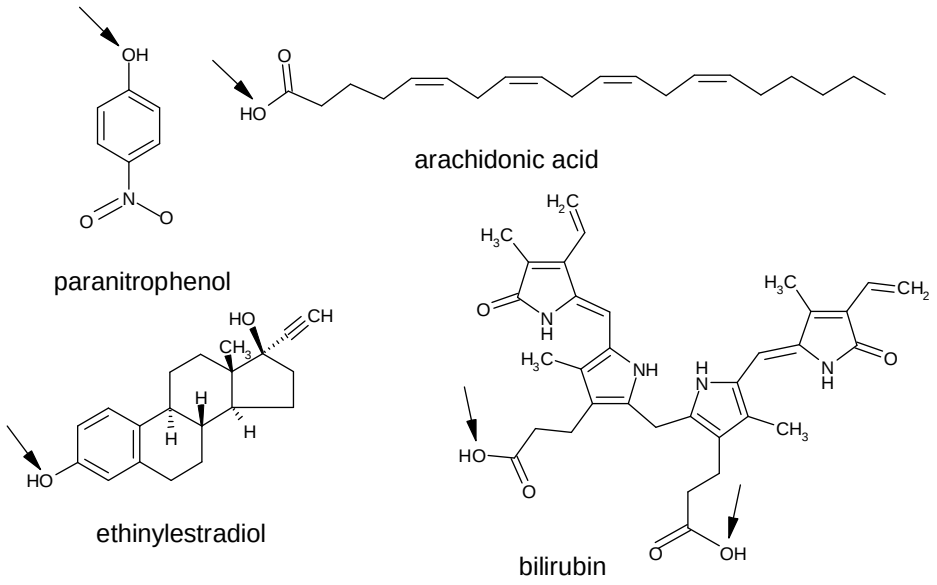


Figure 2.9 *Examples of the chemical structures of UGT1A1 substrates showing the structural diversity of aglycones glucuronidated by one UGT isoform. Arrows indicate the sites of the glucuronidation.*

Photoaffinity labelling, NMR, mutational studies, studies with chimeric constructs and modelling have been used to define the aglycone binding site and the specific residues involved in binding. The fluorescent photoactive compound 7-azido-4-methylcoumarin (AzMC), structurally related to the UGT1A6 substrate 4-methylumbelliferone (4-MU), was used to examine the phenol binding site in human UGT1A6. The results suggested that, a conserved KxxPxP motif, residues 76 – 81, was involved in aglycone binding (Battaglia, *et al.*, 1998; Senay, *et al.*, 1999). Dubois *et al.* (1999) studied two highly homologous human isoforms, UGT2B15 and UGT2B17, which differ by only 29 residues. Both enzymes can glucuronidate steroids at the 17β-OH position, but only UGT2B17 can glucuronidate C_{19} steroids at the 3α-OH position. Several individual mutations in UGT2B17 abolished activity completely, whereas, in UGT2B15, only mutations at residues 121 and 181 impaired

catalytic function. Dubois *et al.* (1999) suggested that these residues are in the active site and interact with the substrate molecule. The S121W mutation in UGT2B17 altered the specificity towards the 3α-OH in C_{19} steroids, whereas the W121S mutation in UGT2B15 did not confer 3α-OH C_{19} steroid specificity indicating that, despite the high similarity between these isoforms, different residues are involved in determining the steroid specificity (Dubois, *et al.*, 1999).

Coffman *et al.* (2001; 2003) made three different maltose binding protein (MBP) fusion constructs of UGT2B7 and analysed them by NMR spectroscopy. The binding properties were analysed by examining the effect of the protein constructs on the NMR spectra and relaxation rates of the morphine. The constructs containing the UGT2B7 residues 24 – 118 or 84 – 118 bound the morphine, whereas the construct containing residues 24 – 96 did not. Based on this, the authors modelled a binding pocket within residues 96 – 101, which was supported by mutational analysis.

Barre *et al.* (2007) studied residue 33 in UGT2B4 and UGT2B7. The residue in this position is not conserved; it is glycine in UGT1s, phenylalanine in UGT2B4 and tyrosine in UGT2B7. Mutational analysis indicated that the presence of an aromatic residue at position 33 is important for the activity of both UGT2B4 and UGT2B7. The substrate selectivities of UGT1A3 and UGT1A4, which share 93.4% sequence identity, were investigated by constructing chimeras and by mutational studies (Kubota, *et al.*, 2007; Li, *et al.*, 2007a). UGT1A3 glucuronidates phenolic compounds, whereas UGT1A4 tends to glucuronidate amine compounds (Green, *et al.*, 1995; Green, *et al.*, 1998; Green and Tephly, 1998). Within the first 44 residues, UGT1A3 and UGT1A4 differ only in positions 36 and 40, being isoleucine and histidine in UGT1A3 and threonine and proline in UGT1A4, respectively. Introducing the I36T and H40P mutations to UGT1A3 reduced the glucuronidation of phenolic compounds remarkably, whereas introducing the T36I and P40H mutations to UGT1A4 reduced the activity towards tertiary amines. In contrast, the H40P mutation in UGT1A3 increased the glucuronidation of tertiary amines and the T36I mutation in UGT1A4 increased the glucuronidation of planar phenols. These residues thus seem to have a critical role in the differing phenol and tertiary amine aglycone substrate selectivities (Kubota, *et al.*, 2007). Similar conclusions regarding the role of residue 40 were also obtained in another study (Li, *et al.*, 2007a).

R52 and H54, located in a conserved amino-terminal region in UGT1A6, were studied by site-directed mutagenesis. They seem to be important for function and integrity, but are not directly involved in aglycone binding (Senay, *et al.*, 1997). In UGT1A10, the $F^{90}MVF^{93}$ motif has been shown to be important in binding phenolic compounds (Xiong, *et al.*, 2006) and estrogen (Xiong, *et al.*, 2006; Starlard-Davenport, *et al.*, 2007). F90 was especially important and may form hydrophobic aromatic ring stacking interactions with the phenolic ring of the aglycone (Xiong, *et al.*, 2006; Miller, *et al.*, 2008). The motif was identified using photoaffinity labelling followed by LC-MS/MS analysis, and by site-directed mutagenesis (Xiong, *et al.*, 2006). The role of F90 is consistent with the model of Li and Wu (2007), which predicts that the equivalent residue in UGT1A1, F92, interacts with the aglycone.

A very recent paper suggested that UGT2B7 had two binding sites for each substrate studied (AZT, 4-MU, 1-NP) maybe reflecting the dimer formation. The results are based on an analysis using multisite and empirical kinetic models. In addition, the data suggest the

existence of multiple effector sites (Uchaipichat, *et al.*, 2008) thus making the kinetic and inhibitor studies of UGTs still more complicated.

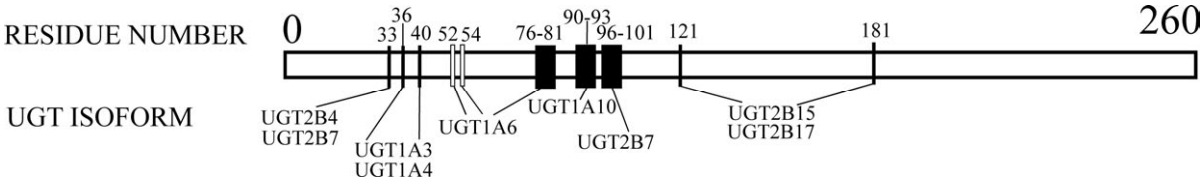


Figure 2.10 *Aglycone binding residues studied in human UGTs. All the residues or motifs presented here are suggested to interact with the aglycone except for residues 52 and 54 in UGT1A6.*

2.3.6. Binding order in UGTs

For an enzyme that takes two substrates (UDPGA and aglycone in case of UGTs) and forms two products (glucuronide and UDP), the enzyme mechanism can be classified either as sequential or ping-pong. Two-substrate and two-product reactions are usually termed bi-bi reactions. In a sequential mechanism, both substrates have to combine with the enzyme, forming a ternary complex, before the reaction can proceed. Sequential mechanisms can be either compulsory ordered, where a particular substrate always binds first, or random ordered, in which the binding order of the two substrates varies. In a ping-pong mechanism, the second substrate can bind only after the first substrate is released.

As early as 1972 bisubstrate kinetics and inhibition studies with guinea-pig bilirubin UGT, equivalent to the human UGT1A1 isoform, led to suggestions that the mechanism was sequential compulsory ordered bi-bi (Potrepka and Spratt, 1972). Potrepka and Patt suggested that UDPGA bound first and bilirubin afterwards, leading to formation of the ternary complex (Potrepka and Spratt, 1972). During the same year, another group came to a different conclusion, suggesting that the reaction is sequential random ordered bi-bi (Vessey and Zakim, 1972). Their conclusion was based on bisubstrate kinetics, inhibition and isotope exchange studies with bovine and guinea-pig liver microsomes using paranitrophenol as the aglycone (Vessey and Zakim, 1972). A few years later Sanchez and Tephly (1975) concluded that the plausible mechanism for morphine glucuronidation in rat hepatic microsomes is sequential ordered, with UDPGA binding first and morphine binding second. Koster and Noordhoek suggested the same mechanism (sequential ordered) for rat intestinal microsomes, but an opposite binding order: 1-naphthol binding first and UDPGA second (Koster and Noordhoek, 1983).

Bisubstrate kinetics, product inhibition and dead-end competitive inhibition studies were performed on two purified rat liver steroid UGTs. Testosterone and androsterone were the substrates. The results were consistent with a random ordered sequential kinetic mechanism (Falany, *et al.*, 1987). Interestingly, the same group had earlier suggested a compulsory ordered sequential mechanism for morphine glucuronidation in hepatic microsomes (Sanchez and Tephly, 1975). Two-substrate kinetic analysis of purified human liver hyodeoxylcholic-acid:UDP-glucuronosyltransferase was consistent with a sequential

mechanism (Matern, *et al.*, 1991), but these studies did not include experiments which could differentiate between compulsory or random ordered binding. Bisubstrate reaction kinetics and inhibition studies with partially purified UGT indicated that the reaction occurs *via* a random ordered sequential mechanism (Yin, *et al.*, 1994).

Most of the studies mentioned above were performed using rat hepatic microsomes. Fourteen different rat UGT-isoforms are known today (http://som.flinders.edu.au/FUSA/ClinPharm/UGT/) and thus several isoforms may contribute to the glucuronidation process, interfering with the interpretation of kinetic data. In man, bilirubin and morphine are quite specific substrates glucuronidated only by one UGT-isoform, whereas 1-naphthol and paranitrophenol are glucuronidated by several UGT-isoforms. In addition, the kinetics may be distorted in purified protein preparations in which the enzyme has been removed from its normal lipidic environment. These factors may explain the variable results.

The matter was finally laid to rest decisively by Luukkanen and co-workers (2005). They studied eight recombinant human UGTs of the 1A subfamily by bisubstrate kinetics and inhibition studies. This was the first study using recombinant enzymes, allowing measurements of individual isoforms in membranes. The results clearly showed that the mechanism is compulsory ordered bi-bi and that UDPGA binds first.

2.3.7. Catalytic mechanism and residues in GT1 enzymes

When a new glycosidic bond forms, two stereochemical outcomes are possible: the configuration of the anomeric carbon can either be retained or inverted (Fig. 2.11) (Lairson, *et al.*, 2008).

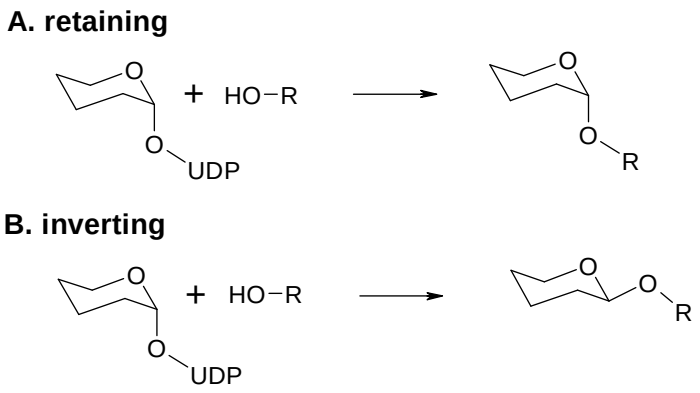


Figure 2.11 *Two main catalytic mechanisms of glycosyltransferases: retention or inversion of the anomeric configuration. Human UGTs are inverting glycosyltransferases, like all other members of the GT1 family.*

Human UGTs, as well as all the other GT1 family members, are inverting glycosyltransferases (Johnson and Fenselau, 1978) that use a direct displacement S_N2-reaction mechanism (Yin, *et al.*, 1994). An active site residue functions as a catalytic base

that deprotonates the substrate for a nucleophilic attack on the anomeric carbon of the sugar donor (Lairson, *et al.*, 2008).

However, the catalytic residues in GT1 family enzymes seem not to be conserved. The first three solved GT1 family structures, GtfA (Mulichak, *et al.*, 2003), GtfB (Mulichak, *et al.*, 2001) and GtfD (Mulichak, *et al.*, 2004) are bacterial enzymes involved in antibiotic vancomycin synthesis. Structural, kinetic and mutational analyses suggest that these enzymes use aspartate (D13) as the catalytic base (Fig. 2.12) (Mulichak, *et al.*, 2004). However, a number of plant and other glycosyltransferases use a serine protease-like mechanism instead, as can be seen from the sequence alignment (Fig. 2.12) (Shao, *et al.*, 2005; Offen, *et al.*, 2006; Li, *et al.*, 2007b; Brazier-Hicks, *et al.*, 2007; Bolam, *et al.*, 2007; Zhang, *et al.*, 2008). A histidine residue near the amino-terminus (H20 in *Vv*GT1) and an adjacent aspartate or glutamate (D119 in *Vv*GT1) function as a "catalytic dyad" (Fig. 2.12). The corresponding histidine residue (H18) is also found in *E. coli* MurG, which is a member of the distantly related GT28 family. The H18A mutation impairs the catalytic efficiency of MurG drastically (Hu, *et al.*, 2003; Crouvoisier, *et al.*, 2007). The authors, however, suggest that E125 plays a catalytic role in this enzyme (Crouvoisier, *et al.*, 2007).

Interestingly, UDP-isoflavone 7-*O*-glucosyltransferase, GmIF7GT, has a conserved His-Asp dyad, but these residues were shown to be unimportant for the enzyme activity. Instead, mutational analysis suggested that an acidic residue at position 392 is essential for catalysis in this enzyme (Noguchi, *et al.*, 2007). The C-C bond-forming glycosyltransferase UrdGT2 has been suggested to use aspartate D137 as the activator for C-C bond formation (Mittler, *et al.*, 2007).

```
                        D13
                        ▼
GtfB     : -MRVLLATCGSRGDTEPLVALA...-EGCAAVVTTGLLAAAIGVRSVAEKLGIPYF : 121
GtfA     : -MRVLITGCGSRGDTEPLVALA...-EGCDAVVTTGLLPAAVAVRSMAEKLGIPYR : 120
GtfD     : -MRVLLSVCGTRGDVEIGVALA...-EGCAAVVAVGDLAAATGVRSVAEKLGLPFF : 122
                        H20                     119
                        ▼                       ▼
VvGT1    : NPHVAVLAFPFSTHAAPLLAVV...--G-RPVSCLVADAFIWFAADMAAEMGVAWL : 137
UGT72B1  : TPHVAIIPSPGMGHLIPLVEFA...----RLPTALVVDLFGTDAFDVAVEFHVPPY : 135
UGT71G1  : NSELIFIPAPGIGHLASALEFA...------VVGLVLDFFCVSMIDVGNEFGIPSY : 139
OleI_bac : TPHISFFNIPGHGHVNPSLGIV...AYADDRPDLIVYDIASWPAPVLGRKWDIPFV : 134
OleD_bac : PAHIAMFSIAAHGHVNPSLEVI...AYADDIPDLVLHDITSYPARVLARRWGVPAV : 129
UGT85H2  : KPHVVMIPYPVQGHINPLFKLA...--NVPPVTCLVSDCCMSFTIQAAEEFELPNV : 143
CalG3    : HMRVLFVSSPGIGHLFPLIQLA...LVDDYRPDLVVYEQGATVGLLAADRAGVPAV : 155
```

Figure 2.12 *Sequence alignment of ten GT1 enzymes and their catalytic residues.*

2.3.8. Catalytic mechanism and residues in UGTs

It was suggested as early as 1958 that the glucuronidation mechanism is an S_N2 type substitution reaction with nucleophilic attack of the aglycone on the electrophilic C1 atom of glucuronic acid (Axelrod *et al.*, 1958). This was later confirmed by bisubstrate kinetics and inhibition studies (Johnson and Fenselau, 1978). More detailed suggestions of the residues

involved in catalysis came much later. In 1994, Battaglia *et al*. (1994a; 1994b) showed that histidine and a carboxyl group are involved in catalysis in human UGTs. These studies were based on experiments on UGT1A6 using chemical modification reagents. Later, based on inhibition of the enzyme activity by the histidyl-selective reagent diethyl pyrocarbonate (DEPC), H370 in UGT1A6 (correct numbering H371) was suggested to function in catalysis (Ouzzine, *et al.*, 2000).

The identity of the catalytic residue was revisited when two new plant GT1 structures became available (Shao, *et al.*, 2005; Offen, *et al.*, 2006). Miley *et al.* (2007) used the plant *Vv*GT1 structure to generate a homology model of UGT2B7. The authors suggested that UGTs, like several other GT1 enzymes (see previous section), use a serine hydrolase-like catalytic mechanism where histidine and aspartate function as a catalytic dyad. Hence H35 of UGT2B7 deprotonates a phenolic group of the acceptor ligand, assisting a nucleophilic attack at the C1 atom of glucuronic acid. The protonated histidine is stabilised by the adjacent D151. This view was supported by mutational analysis (Miley, *et al.*, 2007). Two different groups published a homology model of UGT1A1 at almost the same time and came to similar conclusions, namely that the active site residues are H39 and D151 in UGT1A1 (Li and Wu, 2007; Locuson and Tracy, 2007). However, Li *et al.* (2007a) suggested, based on extensive site-directed mutagenesis and kinetic measurements, that D150 (UGT1A6 numbering) plays a major catalytic role in the UGT1A family, and H38 is involved in defining the substrate specificity rather than in catalysis.

3. Aims of the study

The main aim of my studies was to obtain structural and functional information on human UDP-glucuronosyltransferases. My specific goals were:

1) To understand homo- and hetero-oligomerisation of human UGTs and its possible physiological significance.
2) To identify residues involved in UDPGA binding.
3) To identify residues involved in aglycone binding and recognition.
4) To clarify the mechanism of glucurondiation:
    a. To identify residues involved in catalysis and the catalytic mechanism.
    b. To understand the differences between *N*- and *O*-glucuronidation.

These questions were approached by co-expression, mutation, single-point activity, kinetics and modelling studies. I also tried to express and purify protein for crystallisation.

4. Methods

4.1. Methods used

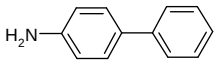
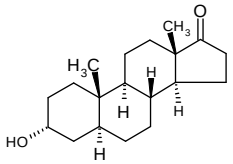
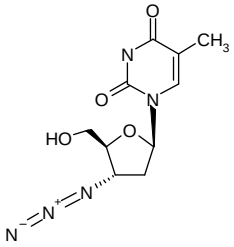
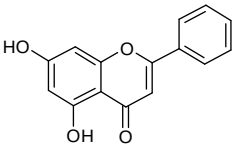
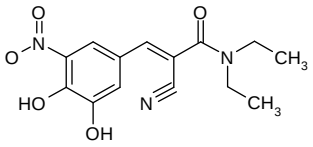
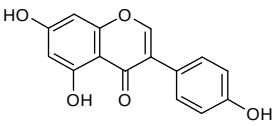
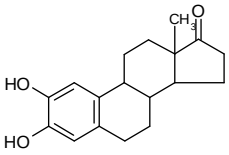
Experimental methods used in this thesis work are listed in Table I. Detailed descriptions of the methods in this study are found in the original publications I-IV or references therein.

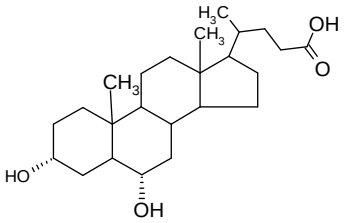
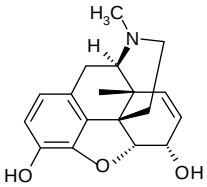
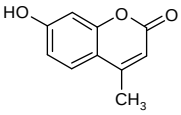
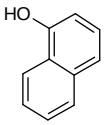
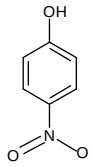
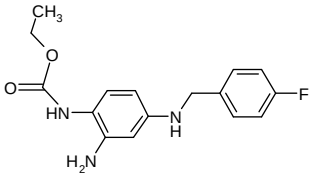
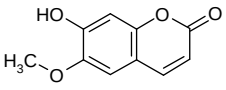
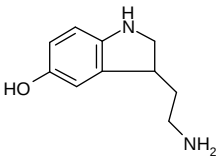
Table 1. *List of the methods used in this thesis. Numbers I-IV refer to the original publications.*

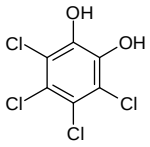
Method		Study
Molecular biology	DNA sequencing	I, II, III, IV
	PCR	II, III, IV
	Cloning	I, II, III, IV
	Expression	I, II, II, IV
	Co-expression	I
	Site-directed mutagenesis	I, II, III, IV
Enzymatic analysis	Single-point activity	I,II, III, IV
	Kinetics	I, II, III, IV
	Bisubstrate kinetics	II
	HPLC	I, II, IV
	UPLC	IV
	TLC	III
Protein	Dot blot	I, II, IV
	Protein purification	I
	Western blot	III
	Membrane preparation	I, II, III, IV
	Homology modelling	II, III
	Sequence comparisons	II, III, IV
Software	Clustal W	II, III, IV
	GeneDoc	II, III, IV
	GraphPad Prism	I, II, IV
	InsightII	II, III
	PyMol	II, III, IV
	Procheck	II, III

4.2. Substrates used

Table 2. *List of the substrates used in the thesis and the publications in which they were used.*

Substrate	Structure	Study	Method
4-Aminobiphenyl		IV	HPLC
Androsterone *		III	TLC
Azidothymidine * (AZT)(Zidovudine)		I	HPLC
Chrysin *		III	TLC
Entacapone		IV	HPLC
Genistein *		III	TLC
2-hydroxyestrone (2-OH-E1) *		III	TLC

Hyodeoxycholic acid *		III	TLC
Morphine *		I	HPLC
4-methylumbelliferone		II, IV	HPLC
1-naphthol		I,II, IV	HPLC
Paranitrophenol		III	HPLC
Retigabine		IV	UPLC
Scopoletin		I, II, IV	HPLC
Serotonin		I	HPLC

Tetrachlorocatechol *		III	TLC

The activity measurements with the substrates marked with a * were performed in a collaborator's laboratory

4.3. Equations used in fitting kinetic measurements

1) Michaelis-Menten equation:

$$v = \frac{V_{max}[S]}{K_m + [S]}$$

2) Bisubstrate kinetics

$$v = \frac{V_{max}[UDPGA][AGLY]}{(K_{d(UDPGA)}K_{m(AGLY)} + K_{m(UDPGA)}[AGLY] + K_{m(AGLY)}[UDPGA] + [UDPGA][AGLY])}$$

3) Substrate inhibition

$$v = \frac{V_{max}}{1 + K_m/[S] + [S]/K_i}$$

v = velocity
V_{max} = maximum velocity
[S] = substrate concentration
K_m = Michaelis constant
[UDPGA] = UDPGA concentration
[AGLY] = aglycone concentration
K_d(UDPGA) = dissocation constant for UDPGA
K_m(UDPGA) = Michaelis constant for UDPGA
K_m(AGLY) = Michaelis constant for aglycone
K_i = dissociation constant for inhibitor

5. Results and Discussion

5.1. Monomer-monomer interactions and their significance (Study I)

The conserved tyrosine residue (Y486 in UGT1A1) is located in the carboxy-terminal domain of human UGTs, just before the predicted trans-membrane helix. The Y486D mutation in human UGT1A1 is a pathological mutation leading to Crigler-Najjar syndrome type II (Aono, *et al.*, 1993). When this mutation is present all the UGT1A enzymes carry it because they all share exons 2 to 5 and have identical carboxy-terminal domains.

We investigated the role of this mutation in recombinant human UGT1A1 and UGT1A6 (mutant forms 1YD and 6YD, respectively). Control and mutant UGTs were expressed as His-tagged proteins in baculovirus-infected insect cells. The Y486D mutation lowered the normalised glucuronidation rate in UGT1A1 to 13 - 17% and in UGT1A6 to 4 - 6% of wild-type, depending on the substrate used.

UGT1A4 is a human UGT that specialises in the glucuronidation of aglycones carrying an amine group (Green and Tephly, 1998; Tukey and Strassburg, 2000). Its rate of glucuronidation of phenolic compounds is very low and therefore HA-tagged UGT1A4 (HA-1A4) was used as a "silent partner" in co-expression studies. The interactions between UGT1A4 and the 6YD mutant increased the mutant activities remarkably: the normalised scopoletin glucuronidation activity in cells co-infected with HA-1A4 and His-6YD was over 20 times higher than in cells infected with 6YD alone. The scopoletin activity of HA-1A4 with respect to 6YD was negligible, so the increase in activity must be a consequence of the co-infection, not of intrinsic UGT1A4 activity. Co-expression of UGT1A4 with 1YD had a minor effect on activity towards scopoletin compared to cells infected with 1YD alone, possibly reflecting the tendency of UGT1A1 to form homo-oligomers (Ghosh, *et al.*, 2001).

The single-point activity measurements were followed by kinetic analyses using 1-naphthol as the aglycone. The Y485D mutation increased the K_m for both the UDPGA and the aglycone in recombinant UGT1A6. Surprisingly, the interaction of UGT1A4 with 6YD changed both K_m values in this mutant to values very close to the respective K_m values for wild-type UGT1A6. In addition, I demonstrated by affinity chromatography using an IMAC that HA-1A4 and His-6YD isoforms interact with each other. When His-tagged 6YD was eluted from the column, HA-tagged UGT1A4, which normally would not bind to the column, was present in the same fractions.

The activity increase by other UGTs was studied further: 6YD was co-expressed individually with 14 different UGT isoforms and the serotonin glucuronidation activity was measured, since serotonin is a rather selective UGT1A6 substrate (Court, 2005). All the UGTs studied, except UGT2B28, stimulated serotonin glucuronidation, although the impact of UGT1A5, UGT1A9 and UGT2B10 was quite small. Interestingly, even the 1YD and 4YD mutants, which have the same mutation as 6YD in the carboxy-terminal domain, stimulated serotonin glucuronidation when co-expessed with 6YD.

This study suggests that protein-protein interactions between individual UGT isoforms influence glucuronidation activity. Those interactions may attenuate the consequences of the Y486D mutation in all members of the UGT1 subfamily, though the study mostly focused on UGT1A6. To ensure that the observed interactions were not caused by the His- or HA-tags, we did co-infections with untagged UGTs. In these studies the co-infection either increased or decreased the activity, suggesting that interactions between different isoforms might have a regulatory role, controlling the enzyme activity *in vivo* as well. Other studies published after the submission of our manuscript have come to similar conclusions (Fujiwara, *et al.*, 2007a; Fujiwara, *et al.*, 2007b; Lévesque, *et al.*, 2007).

5.2. UDPGA binding (Study II+III)

While seeking a suitable low activity mutant for the hetero-oligomerisation studies, we also considered the H371A mutation, because it had been shown to decrease 4-MU activity considerably and was believed to have a catalytic role in UGT1A6 (Ouzzine, *et al.*, 2000). We therefore made the equivalent mutation in UGT1A9. However, when the 9H369A scopoletin glucuronidation activity was measured, the normalised activity appeared to be about 40% of the wild type, which is very high if the residue has a role in catalysis. This observation prompted us instead to study some other residues in this region of the enzyme. E379 was the next residue examined. It seemed to be involved in co-substrate binding rather than chemical catalysis. Thus, both Study II and Study III investigate a co-substrate binding by site-directed mutagenesis, modelling and single-point activity and kinetic measurements.

5.2.1. Kinetic studies

The following residues, H371 and E379 in UGT1A6, H369 and E377 in UGT1A9, D393, Q394, D396 and K399 in UGT1A10, and D398 and Q399 in UGT2B7 were mutated to alanine, the corresponding mutants were produced, and their roles investigated by single-point activity and kinetic measurements (Fig. 5.1).

```
UGT1A6  371: HAGSHGVYESICNGVPMVMMPLFGDQMDNAK : 401
UGT1A9  369: HAGSHGVYESICNGVPMVMMPLFGDQMDNAK : 399
UGT1A10 369: HAGSHGVYESICNGVPMVMMPLFGDQMDNAK : 399
UGT2B7  374: HGGANGIYEAIYHGIPMVGIPLFADQPDNIA : 404
```

Figure 5.1 *Sequence alignment of the carboxy-terminal portion of UGT1A6, 1A9, 1A10 and 2B7 showing the residue numbering and the mutated residues. This part of the sequence is totally identical in the UGT1A subfamily. The residues mutated to alanine in Study II and III are marked with a grey background.*

5.2.1.1. 6H371A, 6E379A, 9H369A, 9E377A

In Study II I first examined residues H371 and E379 (UGT1A6 numbering) as potential catalytic residues, but discovered that they are involved in UDPGA binding rather than catalysis. UGT1A6 and UGT1A9 exhibited different responses to the mutations, even though both mutations are in the identical carboxy-terminal domain (Fig. 5.1). The mutant activities in UGT1A6 were low with all the measured substrates, but variable in UGT1A9. Nonetheless, the kinetic studies revealed that the catalytic efficiency is severely impaired for both mutants and in both UGTs studied, and the total impairment of the catalytic activity was larger in UGT1A6, which is usually a highly active isoform (Table 5.1).

Table 5.1 *Catalytic efficiencies of UGT1A6, UGT1A9 and their mutants. V_{max} values are normalised relative to the corresponding unmutated control.*

	V_{max}/K_m $\left(min^{-1}mg^{-1}protein\right)$			
UGT	**UDPGA (scopoletin)**	**Scopoletin**	**1-naphthol**	**4-MU**
UGT1A9	0.042	0.076		
9H369A	0.0022	0.007		
9E377A	n.d	n.d		
UGT1A6	0.256	0.257	7.17	0.42
6H371A	0.0023	0.0076	0.036	0.0014
6E379A	0.0020	0.0083	0.11	0.0028

In the scopoletin kinetic assays at a constant UDPGA concentration, the K_m values for aglycone were very similar for 6H371A, 6E379A and 9H369A (1.5 – 1.7 mM) and, even though an accurate K_m could not be obtained for 9E377A, it was estimated to be about the same (around 2 mM). The aglycone kinetic curves were quite similar for all the mutants (Fig. 5.2 A, D).

I measured the K_m for UDPGA at a constant scopoletin concentration for all mutants except for 9E377A, where the aglycone was not soluble at high enough concentrations. For this mutant, I also could not detect the K_m for UDPGA, because the velocity was far below the V_{max} even at 20 mM UDPGA (Fig. 5.2F). The co-substrate K_m values varied from 3.9 to 7.6 mM when determined at a constant scopoletin concentration, but were lower when determined by bisubstrate kinetics, where both substrate concentrations are varied. UGT1A9 seemed to have higher UDPGA K_m values for both H and E mutants than UGT1A6 (Fig. 5.2 B, F). The dissociation constant value, K_d, determined by bisubstrate kinetics for the UGT1A6 mutants, was nine times higher for 6H371A and 33 times higher for 6E379A than for wild-type. This is the only study thus far where bisubstrate kinetics has been performed

for UGT-mutant forms, and it clearly indicates that both H371 and E379 are involved in UDPGA binding as both K_m and K_d values are highly affected (see also section 5.4).

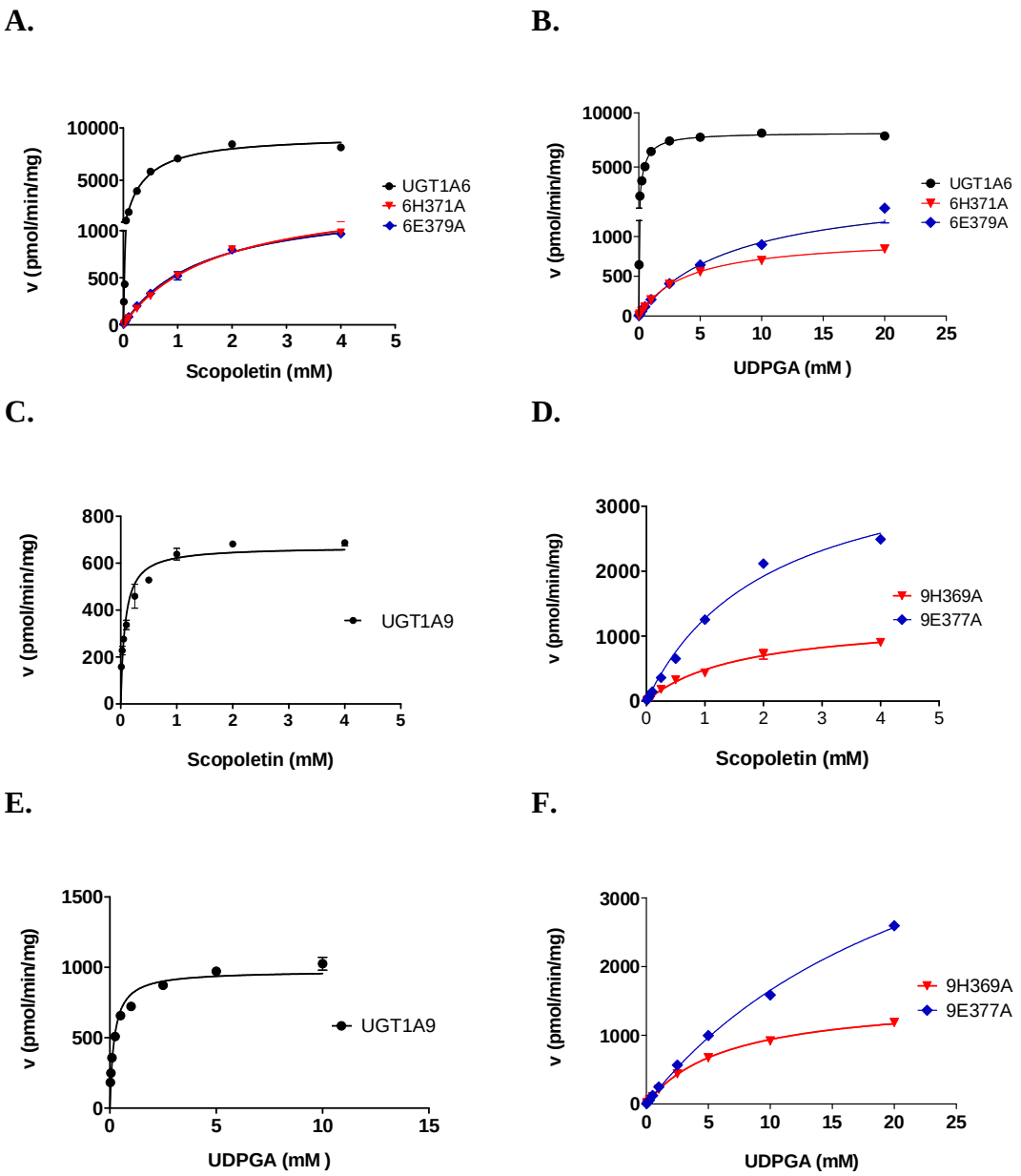


Figure 5.2 *Kinetic curves of H→A and E→A mutants of both UGT1A6 and UGT1A9 isoforms. The K_m values for scopoletin (aglycone) are quite similar between the UGT1A6 and UGT1A9 mutants, but the kinetic curves measuring UDPGA K_m values differ noticeably from each other, showing that the UDPGA K_m is more affected in the UGT1A9 than in the UGT1A6 mutant.*

5.2.1.2. 10D393, 10Q394, 10D396 and 10K399

In addition to H371 and E379 in UGT1A6, the D393 and Q394 residues in UGT1A10 and the equivalent residues, D398 and Q399, in UGT2B7 have been suggested to be

involved in UDPGA binding. They are both highly conserved among GT1 enzymes. In the solved GT1 structures, D or E with the subsequent Q is present in all except UrdGT2, which has V and L instead (Mittler, *et al.*, 2007) (Fig. 2.7). In the GT1 structures these residues make contacts to the sugar moiety of the nucleotide donor. UrdGT2 is not highly selective for the sugar donor: the main co-substrate is dNDP-D-olivose, but it can also accept dNDP-D-mycarose (Trefzer, *et al.*, 2002) or dNDP-rhodinose (Hoffmeister, *et al.*, 2003). The structure of the sugar moiety in some of these compounds does not allow the same interactions that D/E and Q form in most other GT1 enzymes. The D/E equivalent E392 in UDP-glucose:isoflavone 7-*O*-glucosyltransferase, GmIF7GT, has been suggested to play a catalytic role in this enzyme (Noguchi, *et al.*, 2007).

Surprisingly, mutating D393 and Q394 to alanine in UGT1A10 had different effects than mutating the corresponding residues in UGT2B7. The 10D393A mutant had no measurable activity with any of the five substrates studied, but the 10Q394A mutant had low to moderate activity with all of them. Conversely, the 2B7D398A mutant had some activity with three of the substrates studied, but 2B7Q399A was almost inactive. The activity measurements showed the importance of these residues, but kinetic analysis using paranitrophenol as the aglycone could be performed only on the D396 and K399 mutants; the other mutants were too inactive.

Mutating D396 and K399 to alanine had only minor effects on activity. The catalytic efficiency was not much affected, and all the other K_m values were close to the wild type value, except the paranitrophenol K_m for K399A, which increased three-fold. These two residues thus are not in direct contact with the substrates.

5.2.2. Modelling

The role of the mutated residues was further investigated by modelling studies. In both Study II and III, I constructed a homology model of a part of the carboxy-terminal domain. Due to the shared exons, the UGT1A6 (Study II) and UGT1A10 (Study III) carboxy-terminal sequences are totally identical; the residue numbering differs because the amino-terminal domains are not the same size (Fig. 5.1).

In Study II, the human UGT2B7 structure coordinates were not available and the most similar available GT1 structures, the triterpene/flavonoid glucosyltransferase UGT71G1 from the legume *Medicago truncatula* (PDB entries 2ACV and 2ACW) (Shao, *et al.*, 2005) and the UDP-glucose:flavonoid 3-*O*-glucosyltransferase from red grape (*Vv*GT1) (PDB entries 2C1X, 2C1Z and 2C9Z) (Offen, *et al.*, 2006) were used as templates. In Study III, the model was based on the partial structure of human UGT2B7 (PDB entry 2O6L) (Miley, *et al.*, 2007). Due to the different templates, I obtained two dissimilar models of the carboxy-terminal part of the human UGT1A subfamily. The RMSD difference between the models is 2.01 Å, but in the UDPGA–binding area, which was the main focus of these studies, it is only 0.7 Å. The main dissimilarities between the models are in the loop regions, which are flexible and so difficult to model. The positions of the α-helices and β-sheets and the main contacts to the UDPGA co-substrate, are very similar in both models.

My modelling in the carboxy-terminal region mainly focused on residues H371, E379, D395 and Q396. The models constructed support the activity and kinetic data, suggesting that these four residues make contacts to co-substrate. E379 makes hydrogen bonds to the ribose ring in uridine, H371 to the phosphate group and D395 and Q396 to the glucuronic acid moiety of UDPGA. All these residues and interactions are highly conserved among GT1 enzymes (Fig. 2.7). Residues 10D396 and 10K399 are not predicted to make any direct contacts to UDPGA. Based on both models, they are 8 – 10 Å away from UDPGA and they might rather interact with residues in the amino-terminal domain.

The two models predict some other contacts to the co-substrate. The main chain of L354 and the side chains of Q356 and R335 may hydrogen bond to the uridine group (Miley, *et al.*, 2007; Li and Wu, 2007; Locuson and Tracy, 2007). In both models these three residues surround the UDPGA binding pocket (Fig. 5.5). In UGT1A1, mutation of R336 (R335 in UGT1A6) to leucine, glutamine or tryptophan has been reported to cause the more severe form of hyperbilirubinemia, CN-I (Servedio, *et al.*, 2005).

Due to the different templates, the models also differ somewhat. One difference is W353, which has been suggested to form ring stacking contacts to the uracil ring (Miley, *et al.*, 2007; Li and Wu, 2007; Locuson and Tracy, 2007). In the UGT1A6 model (Study II) and in the plant proteins used as templates, this interaction exists (Shao, *et al.*, 2005; Offen, *et al.*, 2006). However, in the UGT1A10 model based on human UGT2B7, W353 would have to adopt a different rotamer conformation to be able to form π-stacking interactions to the uracil ring (Figs. 5.5 and 5.6). The structure of UGT2B7 does not have a bound co-substrate, and residue W353 probably adopts a different conformation upon UDPGA binding. This is true in UGT85H2, which also does not have a co-substrate bound (Shao, *et al.*, 2005; Offen, *et al.*, 2006; Li, *et al.*, 2007b). An equivalent tryptophan that makes ring stacking contacts to the uracil is found in most of the GT1 structures and is missing only in bacterial GtfA, GtfB and GtfD, which are close homologs involved in vancomycin synthesis (Figs. 2.7 and 5.3).

A second difference between the models is the position of S308. It is within hydrogen bonding distance of UDPGA in the UGT1A6 model (Study II), but a little further away in the UGT1A10 model (Study III). This serine residue and the preceding glycine are highly conserved in GT1s (Fig. 5.4). They are usually located in the loop region, except in CalG3 from *Micromonospora echinospora*, where they are in the α-helix (Zhang, *et al.*, 2008). In MurG, there were large changes in the corresponding "GGS" loop between the free and co-substrate (UDP-*N*-acetylglucosamine)-bound enzyme structures (Hu, *et al.*, 2003). In the plant UGT85H2 structure without bound UDP-glucose, the loop is not in the same conformation as in *Vv*GT1 and UGT71G1 with UDP-glucose bound (Shao, *et al.*, 2005; Offen, *et al.*, 2006; Li, *et al.*, 2007b). This might also be the case with human UGTs; the loop may adopt different conformations with and without UDPGA. So far this serine has not been mutated in any of the human UGTs, but mutating the equivalent residue to alanine in the plant UGT71G1 led to a total loss of activity (He, *et al.*, 2006) and mutation in the preceding glycine (G308E in UGT1A1) has been reported to lead to CN-I in humans (Labrune, *et al.*, 1994). Miley *et al*. (2007) mutated S308 in UGT2B7, which is equivalent to S305 in UGT1A6 and is not as conserved as the GS-motif (Fig. 5.4). The S308A mutation had only a minor effect on enzyme activity (Miley, *et al.*, 2007).

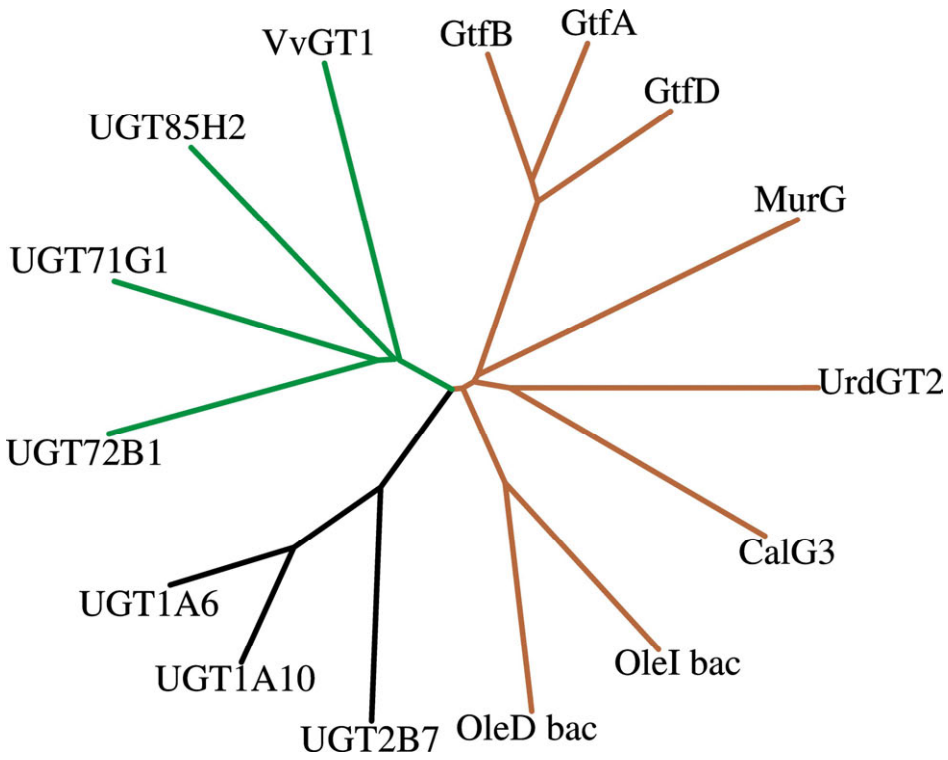


Figure 5.3 *Cladogram of 14 GT1 enzymes and one GT28 enzyme (MurG). It shows the relative evolutionary distances between the proteins. The complete structures of all the enzymes are available except the human UGTs. GtfA, GtfB, GtfD, MurG, UrdGT2, CalG3, OleI and OleD come from bacteria (brown), and VvGT1, UGT85H2, UGT71G1 and UGT72B1 (green) come from plants. UGT1A6, UGT1A10 and UGT2B7 (black) are human enzymes. The sequence alignment was made by ClustalW (Larkin, et al., 2007) and the cladogram by Phylogenetic tree plot (http://www.bioinformatics.nl/tools/plottree.html).*

```
                                  S305 S308
                                     ▼   ▼
UGT1A     : EAYINASGEHGIVVFSLGSMV : 310
UGT2B7    : EDFVQSSGENGVVVFSLGSMV : 313
UGT2B10   : EEFVQSSGENGVVVFSLGSMV : 312
GtfB      : ELAAFLDAGPPPVYLGFGSLG : 249
GtfA      : ELEAFLAAGSTPVYVGFGSSS : 232
GtfD      : ELEAFLAAGSPPVHIGFGSSS : 248
OleI_bac  : GTWEGPGDGRPVLLIALGSAF : 266
OleD_bac  : GGWQRPAGAEKVVLVSLGSAF : 243
UGT72B1   : LKWLDNQPLGSVLYVSFGSGG : 279
UGT71G1   : LKWLDEQPDKSVVFLCFGSMG : 287
UGT85H2   : LDWLESKEPGSVVYVNFGSTT : 306
VvGT1     : LQWLKERKPTSVVYISFGTVT : 282
CalG3     : GDRLPPVPARPEVAITMGTIE : 243
UrdGT2    : EPWMYTRDTRQRVLVTSGSRV : 208
```

Figure 5.4 *Alignment of GT1 enzymes around GS/GT residues. The V/I/LxxxxGS/GT –motif is highly conserved among glycosyltransferases. The serines are numbered according to UGT1A6.*

The third difference is that the side chain of S374 points in opposite directions in the two models (Figs. 5.5, 5.6). It is hydrogen bonded to the glucuronic acid group in the UGT1A10 model, but not in the UGT1A6 model. This residue is not conserved among GT1 enzymes and not even among UGTs (Fig. 2.7). The existence of this interaction and its possible effect on the glucuronidation differences between UGT2B7 and UGT1A10 could be studied further by mutational and kinetic analysis.

The predicted UDPGA binding site in the models of UGT1A6 and UGT1A10 seems to be very similar to that seen in the UGT2B7 structure; only a few residues within hydrogen bonding distance to the UDPGA are not identical. In addition to S374, which is alanine in UGT2B7, the subsequent residue is also different; it is histidine in UGT1s and asparagine in UGT2B7. This residue is somewhat conserved among the GT1 subfamily, being asparagine in plant proteins UGT71G1, UGT85H2, UGT72B1 and *Vv*GT1 (Fig. 2.7). The structures of UGT71G1, UGT72B1 and *Vv*GT1 show that it hydrogen bonds to the oxygen of the α-phosphate of UDP. The structure of UGT85H2 does not have a co-substrate bound. Histidine can make the same contacts, but mutagenesis could be used to study whether changing these residues significantly affects the kinetic parameters of human UGTs.

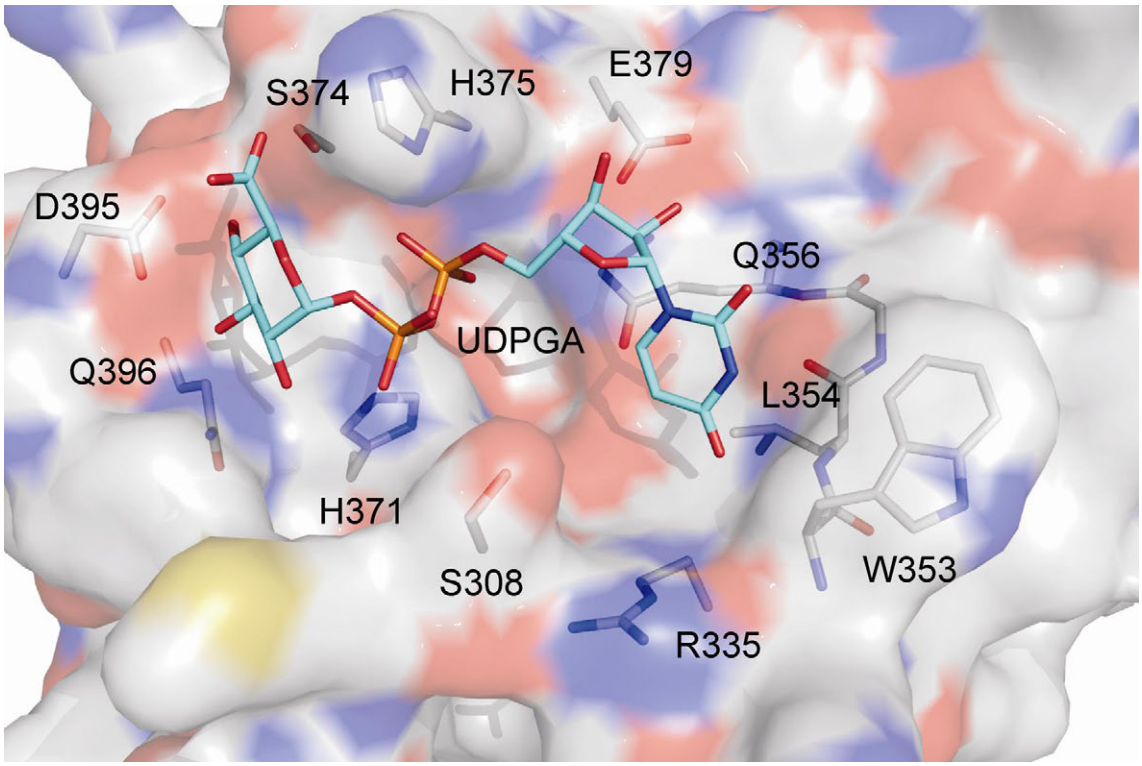


Figure 5.5 *The UDPGA binding pocket of the human UGT1A6 model based on the UGT2B7 structure.*

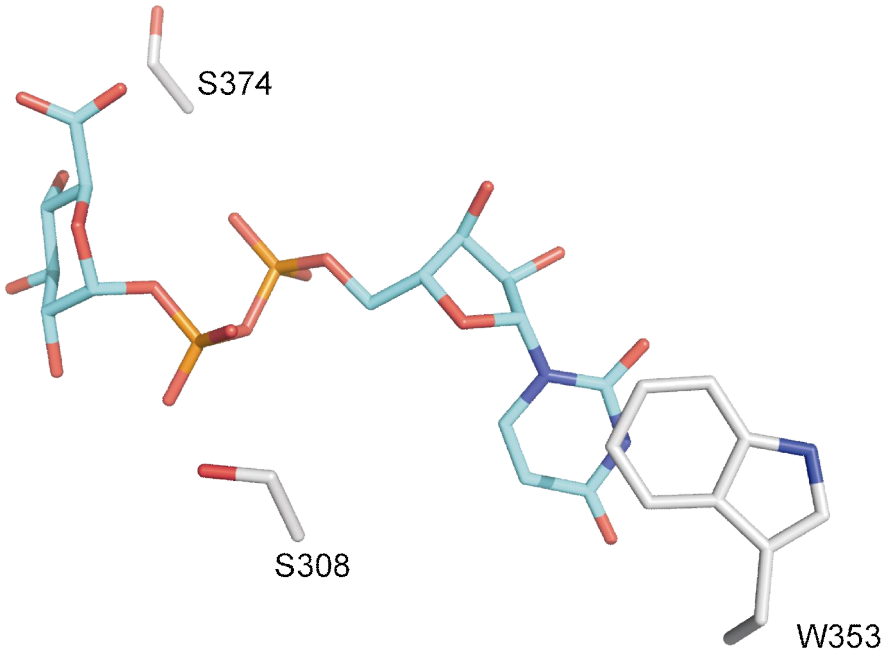


Figure 5.6 *A ball and stick representation of the model based on the plant enzymes, VvGT1 and UGT71G. UGT1A6 residue numbering is used. The main differences between the two models in the UDPGA binding pocket are at S308, W353 and S374.*

5.3. Aglycone binding and catalytic mechanism (Study IV)

When it turned out from my previous study that H371 in UGT1A6 is involved in UDPGA binding rather than playing a role as a catalytic base, we started to look for other candidate residues. Many GT1 enzymes seem to utilise a serine-protease-like catalytic mechanism, where histidine abstracts a proton from the aglycone hydroxyl for nucleophilic attack at the C1 carbon of glucuronic acid in UDPGA (Shao, *et al.*, 2005; Offen, *et al.*, 2006; Li, *et al.*, 2007b; Bolam, *et al.*, 2007; Zhang, *et al.*, 2008). By sequence comparison to the other GT1 structures, H37 in UGT1A9 seemed to be a good candidate for a catalytic base, even though the histidine in this position is not totally conserved among human UGTs and has earlier been proposed to have a structural rather than a catalytic role (Ouzzine, *et al.*, 2000; 2003). The assisting aspartate in UGT1A9 could be either D143 or D148 since they both are highly conserved. Thus, four different mutants, H37A, H37D, D143A and D148A, were generated in UGT1A9 and their behaviour was investigated by performing single-point activity and kinetic measurements.

As UGT1A4 and UGT2B10 are the only two human UGTs that lack the "catalytic" histidine and are more specialised for *N*-glucuronidation (Green and Tephly, 1998; Kaivosaari, *et al.*, 2007), we decided to study the role of H37 in UGT1A9 in both *O*- and *N*-glucuronidation (Fig. 5.7).

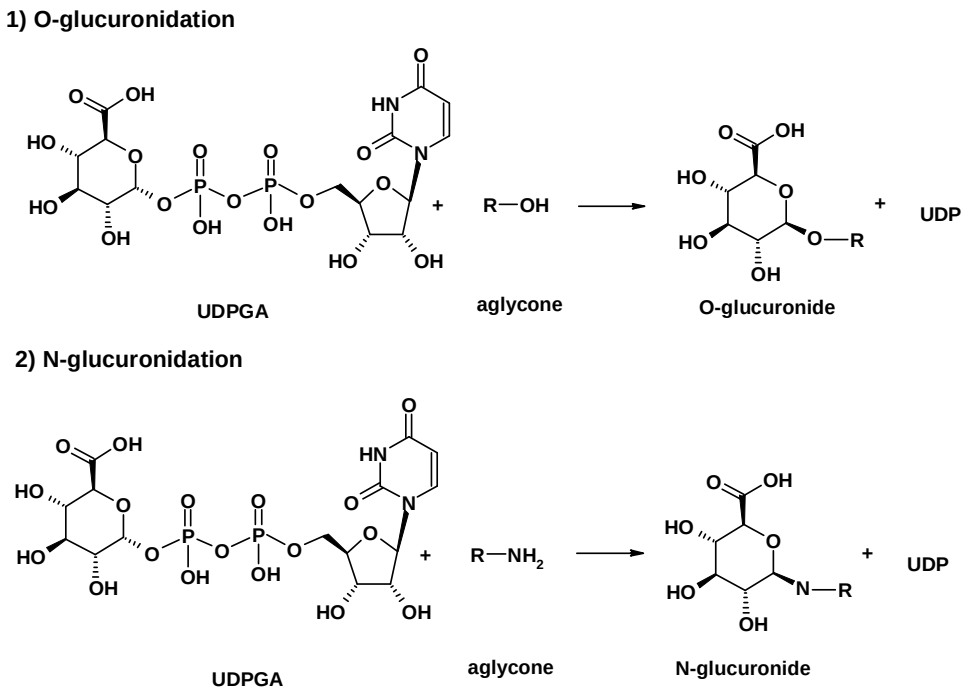


Figure 5.7 *Two main types of glucuronidation. Glucuronidation of phenolic compounds leads to the formation of* O*-glucuronides, and glucuronidation of amine compounds to* N*-glucuronides.*

As we hypothesised, mutating this histidine to alanine decreased the V_{max} with all the studied phenolic compounds that lead to the formation of *O*-glucuronides, but had only a moderate effect on the glucuronidation of the amine compounds that give rise to *N*-glucuronides (Fig. 5.7). The 9H37D mutation decreased the activity with both types of aglycone more than the 9H37A mutation. The K_m value for the 9H37A mutant with *N*-aglycones was similar to that of wild type, but 11 times higher than wild type with the *O*-aglycone, scopoletin. (Scopoletin was the only *O*-aglycone where the mutant had enough residual activity for kinetic studies). A similar effect was observed with 9D148A, although the activities for both the aspartate mutants were very low. The aglycone K_m for the 9D148A mutant changed for the *O*-aglycones, scopoletin and 4-methylumbelliferone, but not for the *N*-aglycone, 4-aminobiphenyl. The activity of the 9D143A mutant was low or unmeasurable with all the substrates studied and the aglycone K_m values were similar to that of the wild type. We could not detect any changes in UDPGA K_m values for any of the mutants studied.

My data clearly show that a different mechanism is used in *N*- than in *O*-glucuronidation. H37 and D148 have different roles with phenolic and amine substrates. H37 is not very important with amine compounds, but is essential with phenolic compounds (Fig. 5.9). However, in contrast to other GT1 enzymes that lose their activity totally upon mutating the catalytic histidine or aspartate (Shao, *et al.*, 2005; Offen, *et al.*, 2006; Li, *et al.*, 2007b), human UGTs retain some residual activity. Mutation of the equivalent histidine in UGT1A1 (H39D) has been reported to lead to the milder form of hyperbilirubinemia, CN-II (Maruo, *et al.*, 2006). My results are in agreement with studies on the bifunctional *O*- and *N*- plant glucosyltransferase UGT72B1, where it was shown that the corresponding histidine is needed for *O*- but not for *N*-glycosylation (Brazier-Hicks, *et al.*, 2007). Kubota *et al.* (2007)

have shown that this histidine has an important role in differentiating between phenolic and amine compounds. However, in those studies kinetic data were not presented. In addition, Li *et al.* (2007a) reported that the P40H mutation conferred phenolic acitivity on UGT1A4. I studied the same mutation in UGT1A4 before their paper was published, and in my preliminary measurements could not detect any marked increase in glucuronidation of phenolic compounds (unpublished results).

My work did not clearly show which of the mutated aspartates, D143 or D148 of UGT1A9, functions in the putative His-Asp dyad. The V_{max} of both mutants was low and clear differences between *O*- and *N*-glucuronidation could not be detected. However, D143A in UGT1A9, unlike D148A, did not change the aglycone K_m, which would be expected from a residue that does not make any direct contacts to the aglycone. On the other hand, all the other studies so far have suggested that D148 is the aspartate in the catalytic dyad (Miley, *et al.*, 2007; Li and Wu, 2007; Locuson and Tracy, 2007; Li, *et al.*, 2007a). The alignment also supports that view (Fig. 5.8). Li *et al.* (2007a) have also measured higher activity for the 9D143A mutant than we detected. In principle, the aspartate could influence the orientation of the histidine and thus change the K_m values as observed in the 9D148A mutant. This kind of behaviour has been reported in Ribonuclease A (Schultz, *et al.*, 1998).

Our results thus show that H37 is the catalytic residue in UGT1A9 and has a different role in *N*- and *O*-glucuronidation. Based on kinetic data, D143 functions in the His-Asp catalytic dyad; however other studies as well as the sequence alignment suggest that D148 fullfills this role. Clarification of the role of these aspartates will require further studies.

```
                 D143 D148
                   ▼    ▼
UGT1A4   : HLNATSFDVVLTDPVNLCGAVLAKYLSIPAV : 170
UGT1A9   : YLKESSFDAVFLDPFDNCGLIVAKYFSLPSV : 166
UGT2B7   : KVQESRFDVIFADAIFPCSELLAELFNIPFV : 169
UGT2B10  : KLQESRFDIVFADAYLPCGELLAELFNIPFV : 168
UGT72B1  : ----RLPTALVVDLFGTDAFDVAVEFHVPPY : 135
UGT71G1  : ------VVGLVLDFFCVSMIDVGNEFGIPSY : 139
OleI_bac : AYADDRPDLIVYDIASWPAPVLGRKWDIPFV : 134
OleD_bac : AYADDIPDLVLHDITSYPARVLARRWGVPAV : 129
GtfB     : -EGCAAVVTTGLLAAAIGVRSVAEKLGIPYF : 121
GtfA     : -EGCDAVVTTGLLPAAVAVRSMAEKLGIPYR : 120
GtfD     : -EGCAAVVAVGDLAAATGVRSVAEKLGLPFF : 122
UGT85H2  : --NVPPVTCLVSDCCMSFTIQAAEEFELPNV : 143
VvGT1    : --G-RPVSCLVADAFIWFAADMAAEMGVAWL : 137
UrdGT2   : FSRAWRPDLIVGGTMSYVAPLLALHLGVP-- : 117
CalG3    : LVDDYRPDLVVYEQGATVGLLAADRAGVPAV : 155
```

Figure 5.8 *Sequence alignment of four human UGTs and eleven GT1 enzymes showing the two aspartates mutated in UGT1A9.*

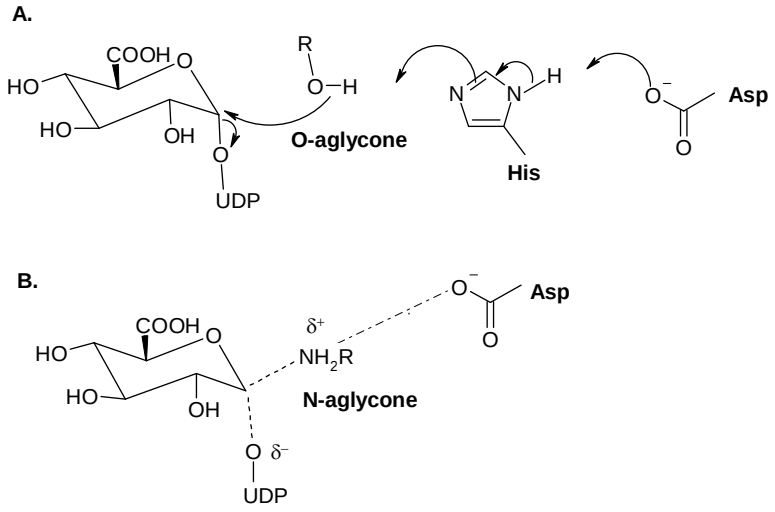


Figure 5.9 *The different catalytic mechanisms for* O- *and* N-*glucuronidation in UGT1A9.* ***A.*** *Proposed catalytic mechanism for* O-*glucuronidation, and* ***B.*** *for* N-*glucuronidation.*

5.4. Substrate binding order (Study II, III, IV)

Studies of the binding order of the substrates over the past 30-plus years have led to different conclusions as discussed in section 2.3.6. My data imply that the UDPGA binds first, as concluded previously by Luukkanen *et al.* (2005). Mutations in the carboxy-terminal domain that affect the K_m for UDPGA also affect the K_m for the aglycone, but the reverse is not true for the amino-terminal mutations, some of which affect only the aglycone K_m (see sections 5.2 and 5.3). So far we have not found any mutations that change only the K_m for UDPGA but not for the aglycone; on the other hand we have found several mutations that affect the aglycone K_m, but not that for UDPGA (this work and unpublished results). In addition, three types of GT1 structures are available: 1) those without any substrate bound, 2) those with nucleotide or nucleotide sugar bound, and 3) those with the ternary complex; however, no structures exist that contain only the aglycone bound (Mulichak, *et al*., 2001; Mulichak, *et al*., 2003; Mulichak, *et al*., 2004; ; Shao, *et al*., 2005; Offen, *et al*., 2006; Bolam, *et al*., 2007; Li, *et al*., 2007b; Brazier-Hicks, *et al*., 2007; Mittler, *et al*., 2007). This is only consistent with a compulsory ordered bi-bi mechanism where UDPGA binds first.

In the GT-B superfamily, MurG has also been shown to follow a compulsory ordered bi-bi mechanism in which the sugar donor binds first (Chen, *et al.*, 2002) and in BGT, UDP-glucose binding is reported to increase the affinity for DNA (Morera, *et al.*, 2001; Lariviere and Morera, 2004). The UDPGA binding in UGTs probably induces a conformational change that facilitates aglycone binding, as observed in some other glycosyltransferases (Qasba, *et al.*, 2005). The loop where the GS-motif (Fig. 5.4) is located may be part of this closure mechanism.

6. Conclusions

In the absence of a crystallographic structure for any mammalian UGTs, other methods have been used to obtain structural information on these enzymes. In my Ph.D. work I have studied both the quaternary and the tertiary structures of UGTs, concentrating on hetero-oligomerisation, substrate binding sites and on catalytic mechanism.

Co-infection studies revealed interactions between wild type UGTs and mutant forms, thus confirming that UGTs function as dimers or higher oligomers. Interactions between monomers were shown to have an important compensating role in some pathological mutations by increasing the severely depressed activity. The interactions between wild type UGT isoforms were shown to increase or decrease the activity, thus suggesting an important regulatory role for hetero-oligomerisation.

I also studied the substrate binding site. My data show that H371, E379, D395 and Q396 are involved in UDPGA binding and that other residues may have a role in co-substrate binding. The effects of these carboxy-terminal mutations depend on both substrate and UGT-isoform and are variable even within the UGT1 subfamily, where the carboxy-terminal half of the protein is totally identical. Our work also supports the compulsory ordered bi-bi mechanism, with UDPGA binding first.

By comparing the UGT sequence to those of the solved GT1 structures, I identified possible catalytic residues in human UGTs. Many other GT1 family members utilise a catalytic mechanism, in which a His-Asp pair forms a catalytic dyad. This mechanism was also suggested for human UGTs. The proposed catalytic base is H37 in UGT1A9 but, interestingly, UGT1A4 and UGT2B10, which tend to catalyse *N*-glucuronidation more, have proline and leucine instead of the histidine. I was able to obtain kinetic data with one *O*- and two *N*-aglcyones and showed that H37 was crucial in *O*-glucuronidation, but not in *N*-glucuronidation. I also showed that D148 in UGT1A9 had different roles in *N*- and in *O*-glucuronidation. My results suggest different catalytic mechanisms for these two types of reactions. The data imply that the catalytic dyad in UGT1A9 is D143-H37, but cannot clearly differentiate between the two aspartates studied.

Human UGTs are very important detoxifying enzymes and the impairment of their function as a consequence of polymorphism or mutation can lead to severe drug side effects. It is thus essential to learn more about the possible interactions between UGT isoforms and the physiological significance of these interactions. Studying the mutant or polymorphic forms with several substrates gives valuable information and can help to choose the right medication and prevent serious side effects. This will be more important in the future as personalised medicine develops. It is also important to be aware of possible protein-protein interactions between isoforms *in vivo* before drawing too many far-reaching conclusions based on *in vitro* studies with recombinant isoforms.

To be able to predict if a drug candidate is a UGT-inhibitor which might cause unwanted side effects, we need more structural information on different UGT-isoforms, and their differential substrate binding. Screening costs would be reduced if unintended inhibitor molecules could be eliminated in the early phases of drug discovery. In addition, potential glucuronidation or bioactivation of the new drug candidates has to be carefully examined, and the pharmaceutical industry could benefit from this structural information. This kind of

knowledge is also needed when designing specific inhibitors. In particular, the observed mechanistic difference between *N*- and *O*- glucuronidation opens up interesting possibilities to devise specific mechanism-based inhibitors for different UGT-isoforms.

Acknowledgements

This thesis work has been carried out between 2003 and 2008 in the Macromolecular Crystallography group at the Institute of Biotechnology, University of Helsinki. The work was funded by the Academy of Finland, Sigrid Juselius Foundation and National Graduate School in Informational and Structural Biology.

This project has been done in close collaboration with two groups at the University of Helsinki: the Macromolecular Crystallography group headed by Research Director Adrian Goldman and the Molecular and Biochemical Pharmacology group in the Centre for Drug Research (CDR) headed by Docent Moshe Finel.

My warmest thanks go to my both supervisors: Research Director Adrian Goldman and Docent Moshe Finel. I am deeply grateful to Adrian, who was open-minded enough to hire me to his group, even though I had not worked in a laboratory for several years due to my work in marketing. I cannot express how grateful I am for this opportunity to be back at the university and in the laboratory. When I started this project I could not imagine how much fun it would offer for me; I have enjoyed every moment. I am also highly indebted to Moshe, who has always been interested in my experiments and results and has always had time for discussions. You both have guided, supported and encouraged me, and tried to keep me busy during all these years. It has been a great pleasure to work with you.

I would like to thank professor Mart Saarma, Head of the Institute of Biotechnology, and Professor Carl G. Gahmberg, Head of the Division of Biochemistry, for providing excellent facilities and a positive working environment.

I thank both the reviewers, Risto Juvonen and Tuomas Haltia, for their constructive criticism and valuable comments during the pre-examination of this thesis. I am grateful to Professor Liisa Holm, member of my thesis committee, for her comments and advice and especially for originally guiding me back to the academic world. I also thank Mary Metzler for revising the language of this thesis.

I am thankful to all the current and previous lab members, especially Jack and Marika who have updated my knowledge in molecular biology, and Esko who has always been ready to answer my questions concerning chemistry, crystallography, NMR *etc*. Thanks also to Andrzej, Arnab, Chiara, Hanna, Heidi, Heli, Igor, Jaana, Juho, Katja, Lari, Maria, Mika, Pirkko, Sanjay, Saurabh, Seija, Sirpa, Tommi, Veli-Pekka, Veli-Matti and Vimal for all your help and for providing a nice and friendly working environment in the lab. My special thanks go to the lab members in the Molecular and Biochemical Pharmacology group, especially to Johanna Mosorin and Sanna Sistonen, who have guided and helped me with different laboratory techniques.

I wish to acknowledge all my co-authors for their contribution to the articles of this thesis, particularly Mika Kurkela, who has helped me with several methods and taught me a lot. It has been a nice experience to work with you. I want to thank Timo Hildén, vice president and general manager of Thermo Fisher Scientific's liquid handling consumables business, for providing me with first-class pipettes when I started in this group.

I sincerely thank all my friends, especially Kirsi, Titta, Lea and Klaus and our cell group: Ulla, Kirsti, Marjut and Pirjo for discussions, prayers and listening during these years.

My deepest gratitude goes to my parents, to my late father Veikko and to my mother Lea.

You have taught me to work hard already in early childhood. I thank my mother also for helping with childcare during these years. I thank my sister Hannele and my brother Janne, I have learned a lot just by growing up with you. Without you, I would be quite a different person. I am also thankful to my father-in-law Antero and his wife Sinikka for supporting me and being always interested in my work.

Finally I want to thank my family, who have taken care that I have also had some life outside the lab. I owe my deepest gratitude to my dearest sons Jesse and Eetu and apologise for some late working hours. I love you both so much and I always will. You have brought more joy into my life than I could have imagined before you were born. Above all, from the bottom of my heart I thank my beloved husband Harri, who has encouraged and supported me during all these years. He has taken good care of our sons, Jesse and Eetu, when I have been studying or been at work. There are no words to express my gratitude. Without you, I would not have even started this project. You have never questioned my work, even though our life would have been easier in many ways if I had just continued in my previous position. I dedicate this thesis to you.

References

Ames BN and Gold LS (1998) The prevention of cancer. *Drug Metab Rev* **30:**201-223.

Ames BN, Profet M and Gold LS (1990) Dietary pesticides (99.99% all natural). *Proc Natl Acad Sci U S A* **87:**7777-7781.

Aono S, Yamada Y, Keino H, Hanada N, Nakagawa T, Sasaoka Y, Yazawa T, Sato H and Koiwai O (1993) Identification of defect in the genes for bilirubin UDP-glucuronosyl-transferase in a patient with Crigler-Najjar syndrome type II. *Biochem Biophys Res Commun* **197:**1239-1244.

Argos P, Rao JK and Hargrave PA (1982) Structural prediction of membrane-bound proteins. *Eur J Biochem* **128:**565-575.

Axelrod J, Inscoe JK and Tomkins GM (1958) Enzymatic synthesis of *N*-glucosyluronic acid conjugates. *J Biol Chem* **232:**835-841.

Banerjee R, Pennington MW, Garza A and Owens IS (2008) Mapping the UDP-glucuronic acid binding site in UDP-glucuronosyltransferase-1A10 by homology-based modeling: confirmation with biochemical evidence. *Biochemistry* **47:**7385-7392.

Barré L, Fournel-Gigleux S, Finel M, Netter P, Magdalou J and Ouzzine M (2007) Substrate specificity of the human UDP-glucuronosyltransferase UGT2B4 and UGT2B7. Identification of a critical aromatic amino acid residue at position 33. *FEBS J* **274:**1256-1264.

Battaglia E, Pritchard M, Ouzzine M, Fournel-Gigleux S, Radominska A, Siest G and Magdalou J (1994a) Chemical modification of human UDP-glucuronosyltransferase UGT1*6 by diethyl pyrocarbonate: possible involvement of a histidine residue in the catalytic process. *Arch Biochem Biophys* **309:**266-272.

Battaglia E, Senay C, Fournel-Gigleux S, Herber R, Siest G and Magdalou J (1994b) The chemical modification of human liver UDP-glucuronosyltransferase UGT1*6 reveals the involvement of a carboxyl group in catalysis. *FEBS Lett* **346:**146-150.

Battaglia E, Terrier N, Mizeracka M, Senay C, Magdalou J, Fournel-Gigleux S and Radominska-Pandya A (1998) Interaction of periodate-oxidized UDP-glucuronic acid with recombinant human liver UDP-glucuronosyltransferase 1A6. *Drug Metab Dispos* **26:**812-817.

Bolam DN, Roberts S, Proctor MR, Turkenburg JP, Dodson EJ, Martinez-Fleites C, Yang M, Davis BG, Davies GJ and Gilbert HJ (2007) The crystal structure of two macrolide glycosyltransferases provides a blueprint for host cell antibiotic immunity. *Proc Natl Acad Sci U S A* **104:**5336-5341.

Bosma PJ (2003) Inherited disorders of bilirubin metabolism. *J Hepatol* **38:**107-117.

Bosma PJ, Chowdhury JR, Bakker C, Gantla S, de Boer A, Oostra BA, Lindhout D, Tytgat GN, Jansen PL and Oude Elferink RP (1995) The genetic basis of the reduced expression of bilirubin UDP-glucuronosyltransferase 1 in Gilbert's syndrome. *N Engl J Med* **333:**1171-1175.

Bosma PJ, Chowdhury NR, Goldhoorn BG, Hofker MH, Oude Elferink RP, Jansen PL and Chowdhury JR (1992) Sequence of exons and the flanking regions of human bilirubin-UDP-glucuronosyltransferase gene complex and identification of a genetic mutation in a patient with Crigler-Najjar syndrome, type I. *Hepatology* **15:**941-947.

Bosma PJ, Seppen J, Goldhoorn B, Bakker C, Oude Elferink RP, Chowdhury JR, Chowdhury NR and Jansen PL (1994) Bilirubin UDP-glucuronosyltransferase 1 is the only relevant bilirubin glucuronidating isoform in man. *J Biol Chem* **269:**17960-17964.

Bourne Y and Henrissat B (2001) Glycoside hydrolases and glycosyltransferases: families and functional modules. *Curr Opin Struct Biol* **11:**593-600.

Bowles D, Lim EK, Poppenberger B and Vaistij FE (2006) Glycosyltransferases of lipophilic small molecules. *Annu Rev Plant Biol* **57:**567-597.

Brazier-Hicks M, Offen WA, Gershater MC, Revett TJ, Lim EK, Bowles DJ, Davies GJ and Edwards R (2007) Characterization and engineering of the bifunctional *N*- and *O*-glucosyltransferase involved in xenobiotic metabolism in plants. *Proc Natl Acad Sci U S A*

Breton C and Imberty A (1999) Structure/function studies of glycosyltransferases. *Curr Opin Struct Biol* **9:**563-571.

Breton C, Snajdrova L, Jeanneau C, Koca J and Imberty A (2006) Structures and mechanisms of glycosyltransferases. *Glycobiology* **16:**29R-37R.

Campbell JA, Davies GJ, Bulone V and Henrissat B (1997) A classification of nucleotide-diphospho-sugar glycosyltransferases based on amino acid sequence similarities. *Biochem J* **326 (Pt 3):**929-939.

Charnock SJ and Davies GJ (1999) Structure of the nucleotide-diphospho-sugar transferase, SpsA from Bacillus subtilis, in native and nucleotide-complexed forms. *Biochemistry* **38:**6380-6385.

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 L, Men H, Ha S, Ye XY, Brunner L, Hu Y and Walker S (2002) Intrinsic lipid preferences and kinetic mechanism of *Escherichia coli* MurG. *Biochemistry* **41:**6824-6833.

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.

Coffman BL, Kearney WR, Goldsmith S, Knosp BM and Tephly TR (2003) Opioids bind to the amino acids 84 to 118 of UDP-glucuronosyltransferase UGT2B7. *Mol Pharmacol* **63:**283-288.

Coffman BL, Kearney WR, Green MD, Lowery RG and Tephly TR (2001) Analysis of opioid binding to UDP-glucuronosyltransferase 2B7 fusion proteins using nuclear magnetic resonance spectroscopy. *Mol Pharmacol* **59:**1464-1469.

Coleman J, Blake-Kalff M and Davies E (1997) Detoxification of xenobiotics by plants: chemical modification and vacuolar compartmentation. *Trends in Plant Science,* **2:**144-151.

Court MH (2005) Isoform-selective probe substrates for *in vitro* studies of human UDP-glucuronosyltransferases. *Methods Enzymol* **400:**104-116.

Coutinho PM, Deleury E, Davies GJ and Henrissat B (2003) An evolving hierarchical family classification for glycosyltransferases. *J Mol Biol* **328:**307-317.

Crigler JF, jr. and Najjar VA (1952) Congenital familial nonhemolytic jaundice with kernicterus. *Pediatrics* **10:**169-180.

Crouvoisier M, Auger G, Blanot D and Mengin-Lecreulx D (2007) Role of the amino acid invariants in the active site of MurG as evaluated by site-directed mutagenesis. *Biochimie,* **89:**1498-1508.

Davies GJ, Gloster TM and Henrissat B (2005) Recent structural insights into the expanding world of carbohydrate-active enzymes. *Curr Opin Struct Biol* **15:**637-645.

Dubois SG, Beaulieu M, Lévesque E, Hum DW and Belanger A (1999) Alteration of human UDP-glucuronosyltransferase UGT2B17 regio-specificity by a single amino acid substitution. *J Mol Biol* **289:**29-39.

Dutton GJ (1980) *Glucuronidation of Drugs and Other Compounds*. CRC Press Inc., Boca Raton, FL.

Dutton GJ and Storey ID (1953) The isolation of a compound of uridine diphosphate and glucuronic acid from liver. *Biochem J* **53:**37-38.

Falany CN, Green MD and Tephly TR (1987) The enzymatic mechanism of glucuronidation catalyzed by two purified rat liver steroid UDP-glucuronosyltransferases. *J Biol Chem* **262:**1218-1222.

Finel M and Kurkela M (2008) The UDP-glucuronosyltransferases as oligomeric enzymes. *Curr Drug Metab* **9:**70-76.

Fisher MB, Paine MF, Strelevitz TJ and Wrighton SA (2001) The role of hepatic and extrahepatic UDP-glucuronosyltransferases in human drug metabolism. *Drug Metab Rev* **33:**273-297.

Francoual J, Trioche P, Mokrani C, Seboui H, Khrouf N, Chalas J, Clement M, Capel L, Tachdjian G and Labrune P (2002) Prenatal diagnosis of Crigler-Najjar syndrome type I by single-strand conformation polymorphism (SSCP). *Prenat Diagn* **22:**914-916.

Fremont JJ, Wang RW and King CD (2005) Coimmunoprecipitation of UDP-glucuronosyltransferase isoforms and cytochrome P450 3A4. *Mol Pharmacol* **67:**260-262.

Fujiwara R, Nakajima M, Yamanaka H, Katoh M and Yokoi T (2007a) Interactions between human UGT1A1, UGT1A4, and UGT1A6 affect their enzymatic activities. *Drug Metab Dispos* **35:**1781-1787.

Fujiwara R, Nakajima M, Yamanaka H, Nakamura A, Katoh M, Ikushiro S, Sakaki T and Yokoi T (2007b) Effects of coexpression of UGT1A9 on enzymatic activities of human UGT1A isoforms. *Drug Metab Dispos* **35:**747-757.

Ghosh SS, Sappal BS, Kalpana GV, Lee SW, Chowdhury JR and Chowdhury NR (2001) Homodimerization of human bilirubin-uridine-diphosphoglucuronate glucuronosyltransferase-1 (UGT1A1) and its functional implications. *J Biol Chem* **276:**42108-42115.

Girard H, Lévesque E, Bellemare J, Journault K, Caillier B and Guillemette C (2007) Genetic diversity at the UGT1 locus is amplified by a novel 3' alternative splicing mechanism leading to nine additional UGT1A proteins that act as regulators of glucuronidation activity. *Pharmacogenet Genomics* **17:**1077-1089.

Gong QH, Cho JW, Huang T, Potter C, Gholami N, Basu NK, Kubota S, Carvalho S, Pennington MW, Owens IS and Popescu NC (2001) Thirteen UDPglucuronosyltransferase genes are encoded at the human UGT1 gene complex locus. *Pharmacogenetics* **11:**357-368.

Green MD, Bishop WP and Tephly TR (1995) Expressed human UGT1.4 protein catalyzes the formation of quaternary ammonium-linked glucuronides. *Drug Metab Dispos* **23:**299-302.

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.

Green MD and Tephly TR (1998) Glucuronidation of amine substrates by purified and expressed UDP-glucuronosyltransferase proteins. *Drug Metab Dispos* **26:**860-867.

Gschaidmeier H and Bock KW (1994) Radiation inactivation analysis of microsomal UDP-glucuronosyltransferases catalysing mono- and diglucuronide formation of 3,6-dihydroxybenzo(a)pyrene and 3,6-dihydroxychrysene. *Biochem Pharmacol* **48:**1545-1549.

Guengerich FP (2002) Cytochrome P450, in *Enzyme systems that metabolise drugs and other xenobiotics* (C. Ioannides, ed.) pp 33-65, John Wiley & Sons, Ltd., West Sussex, UK.

Harb R and Thomas DW (2007) Conjugated hyperbilirubinemia: screening and treatment in older infants and children. *Pediatr Rev* **28:**83-91.

He XZ, Wang X and Dixon RA (2006) Mutational analysis of the Medicago glycosyltransferase UGT71G1 reveals residues that control regioselectivity for (iso)flavonoid glycosylation. *J Biol Chem* **281:**34441-34447.

Hirschfield GM and Alexander GJ (2006) Gilbert's syndrome: an overview for clinical biochemists. *Ann Clin Biochem* **43:**340-343.

Hoffmeister D, Drager G, Ichinose K, Rohr J and Bechthold A (2003) The C-Glycosyltransferase UrdGT2 is unselective toward D- and L-configured nucleotide-bound rhodinoses. *J Am Chem Soc* **125:**4678-4679.

Hu Y, Chen L, Ha S, Gross B, Falcone B, Walker D, Mokhtarzadeh M and Walker S (2003) Crystal structure of the MurG:UDP-GlcNAc complex reveals common structural principles of a superfamily of glycosyltransferases. *Proc Natl Acad Sci U S A* **100:**845-849.

Hu Y and Walker S (2002) Remarkable structural similarities between diverse glycosyltransferases. *Chem Biol* **9:**1287-1296.

Ikushiro S, Emi Y and Iyanagi T (1997) Protein-protein interactions between UDP-glucuronosyltransferase isozymes in rat hepatic microsomes. *Biochemistry* **36:**7154-7161.

Innis CA, Shi J and Blundell TL (2000) Evolutionary trace analysis of TGF-β and related growth factors: implications for site-directed mutagenesis. *Protein Eng* **13:**839-847.

Ioannides C (2002) Xenobiotic Metabolism: An Overview, in *Enzyme systems that metabolise drugs and other xenobiotics* (C. Ioannides, ed.) pp 1-32, John Wiley & Sons, Ltd., West Sussex, UK.

Ishii Y, Miyoshi A, Maji D, Yamada H and Oguri K (2004) Simultaneous expression of guinea pig UDP-glucuronosyltransferase 2B21 (UGT2B21) and 2B22 in COS-7 cells enhances UGT2B21-catalyzed chloramphenicol glucuronidation. *Drug Metab Dispos* **32:**1057-1060.

Ishii Y, Miyoshi A, Watanabe R, Tsuruda K, Tsuda M, Yamaguchi-Nagamatsu Y, Yoshisue K, Tanaka M, Maji D, Ohgiya S and Oguri K (2001) Simultaneous expression of guinea pig UDP-glucuronosyltransferase 2B21 and 2B22 in COS-7 cells enhances UDP-glucuronosyltransferase 2B21-catalyzed morphine-6-glucuronide formation. *Mol Pharmacol* **60:**1040-1048.

Ito M, Yamamoto K, Maruo Y, Sato H, Fujiyama Y and Bamba T (2002) Effect of a conserved mutation in uridine diphosphate glucuronosyltransferase 1A1 and 1A6 on glucuronidation of a metabolite of flutamide. *Eur J Clin Pharmacol* **58:**11-14.

Jansen PL (1999) Diagnosis and management of Crigler-Najjar syndrome. *Eur J Pediatr* **158 Suppl 2:**89-94.

Jinno H, Tanaka-Kagawa T, Hanioka N, Saeki M, Ishida S, Nishimura T, Ando M, Saito Y, Ozawa S and Sawada J (2003) Glucuronidation of 7-ethyl-10-hydroxycamptothecin (SN-38), an active metabolite of irinotecan (CPT-11), by human UGT1A1 variants, G71R, P229Q, and Y486D. *Drug Metab Dispos* **31:**108-113.

Johnson LP and Fenselau C (1978) Enzymatic conjugation and hydrolysis of [18O]isoborneol glucuronide. *Drug Metab Dispos* **6:**677-679.

Jones P, Messner B, Nakajima J, Schaffner AR and Saito K (2003) UGT73C6 and UGT78D1, glycosyltransferases involved in flavonol glycoside biosynthesis in *Arabidopsis thaliana*. *J Biol Chem* **278:**43910-43918.

Kadakol A, Ghosh SS, Sappal BS, Sharma G, Chowdhury JR and Chowdhury NR (2000) Genetic lesions of bilirubin uridine-diphosphoglucuronate glucuronosyltransferase (UGT1A1) causing Crigler-Najjar and Gilbert syndromes: correlation of genotype to phenotype. *Hum Mutat* **16:**297-306.

Kaivosaari S, Salonen JS and Taskinen J (2002) *N*-Glucuronidation of some 4-arylalkyl-1H-imidazoles by rat, dog, and human liver microsomes. *Drug Metab Dispos* **30:**295-300.

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.

Kerdpin O, Elliot DJ, Mackenzie PI and Miners JO (2006) Sulfinpyrazone C-glucuronidation is catalyzed selectively by human UDP-glucuronosyltransferase 1A9. *Drug Metab Dispos* **34:**1950-1953.

Kiang TK, Ensom MH and Chang TK (2005) UDP-glucuronosyltransferases and clinical drug-drug interactions. *Pharmacol Ther* **106:**97-132.

King CD, Rios GR, Green MD and Tephly TR (2000) UDP-glucuronosyltransferases. *Curr Drug Metab* **1:**143-161.

Kobayashi T, Sleeman JE, Coughtrie MW and Burchell B (2006) Molecular and functional characterization of microsomal UDP-glucuronic acid uptake by members of the nucleotide sugar transporter (NST) family. *Biochem J* **400:**281-289.

Koiwai O, Aono S, Adachi Y, Kamisako T, Yasui Y, Nishizawa M and Sato H (1996) Crigler-Najjar syndrome type II is inherited both as a dominant and as a recessive trait. *Hum Mol Genet* **5:**645-647.

Koster AS and Noordhoek J (1983) Kinetic properties of the rat intestinal microsomal 1-naphthol:UDP-glucuronosyl transferase. Inhibition by UDP and UDP-N-acetylglucosamine. *Biochim Biophys Acta* **761:**76-85.

Kraemer D and Klinker H (2002) Crigler-Najjar syndrome type II in a caucasian patient resulting from two mutations in the bilirubin uridine 5'-diphosphate-glucuronosyltransferase (UGT1A1) gene. *J Hepatol* **36:**706-707.

Kubota T, Lewis BC, Elliot DJ, Mackenzie PI and Miners JO (2007) Critical roles of residues 36 and 40 in the phenol and tertiary amine aglycone substrate selectivities of UDP-glucuronosyltransferase 1A3 and 1A4. *Mol Pharmacol* **72**:1054-62.

Kurkela M, García-Horsman JA, Luukkanen L, Mörsky S, Taskinen J, Baumann M, Kostiainen R, Hirvonen J and Finel M (2003) Expression and Characterization of Recombinant Human UDP-glucuronosyltransferases (UGTs). *J Biol Chem* **278:**3536-3544.

Kurkela M, Hirvonen J, Kostiainen R and Finel M (2004a) The interactions between the *N*-terminal and C-terminal domains of the human UDP-glucuronosyltransferases are partly isoform-specific, and may involve both monomers. *Biochem Pharmacol* **68:**2443-2450.

Kurkela M, Morsky S, Hirvonen J, Kostiainen R and Finel M (2004b) An active and water-soluble truncation mutant of the human UDP-glucuronosyltransferase 1A9. *Mol Pharmacol* **65:**826-831.

Labrune P, Myara A, Hadchouel M, Ronchi F, Bernard O, Trivin F, Chowdhury NR, Chowdhury JR, Munnich A and Odievre M (1994) Genetic heterogeneity of Crigler-Najjar syndrome type I: a study of 14 cases. *Hum Genet* **94:**693-697.

Lairson LL, Henrissat B, Davies GJ and Withers SG (2008) Glycosyltransferases: Structures, Functions, and Mechanisms. *Annu Rev Biochem* **77:**521-555.

Lariviere L and Morera S (2004) Structural evidence of a passive base-flipping mechanism for beta-glucosyltransferase. *J Biol Chem* **279:**34715-34720.

Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H, Valentin F, Wallace IM, Wilm A, Lopez R, Thompson JD, Gibson TJ and Higgins DG (2007) Clustal W and Clustal X version 2.0. *Bioinformatics* **23:**2947-2948.

Lévesque E, Girard H, Journault K, Lepine J and Guillemette C (2007) Regulation of the UGT1A1 bilirubin-conjugating pathway: role of a new splicing event at the UGT1A locus. *Hepatology* **45:**128-138.

Lewis BC, Mackenzie PI, Elliot DJ, Burchell B, Bhasker CR and Miners JO (2007) Amino terminal domains of human UDP-glucuronosyltransferases (UGT) 2B7 and 2B15 associated with substrate selectivity and autoactivation. *Biochem Pharmacol* **73:**1463-1473.

Li C and Wu Q (2007) Adaptive evolution of multiple-variable exons and structural diversity of drug-metabolizing enzymes. *BMC Evol Biol* **7:**69.

Li D, Fournel-Gigleux S, Barre L, Mulliert G, Netter P, Magdalou J and Ouzzine M (2007a) Identification of Aspartic Acid and Histidine Residues Mediating the Reaction Mechanism and the Substrate Specificity of the Human UDP-glucuronosyltransferases 1A. *J Biol Chem* **282:**36514-36524.

Li L, Modolo LV, Escamilla-Trevino LL, Achnine L, Dixon RA and Wang X (2007b) Crystal structure of Medicago truncatula UGT85H2--insights into the structural basis of a multifunctional (iso)flavonoid glycosyltransferase. *J Mol Biol* **370:**951-963.

Locuson CW and Tracy TS (2007) Comparative modelling of the human UDP-glucuronosyltransferases: insights into structure and mechanism. *Xenobiotica* **37:**155-168.

Luukkanen L, Taskinen J, Kurkela M, Kostiainen R, Hirvonen J and Finel M (2005) Kinetic characterization of the 1A subfamily of recombinant human UDP-glucuronosyltransferases. *Drug Metab Dispos* **33:**1017-1026.

Mackenzie PI (1990) Expression of chimeric cDNAs in cell culture defines a region of UDP glucuronosyltransferase involved in substrate selection. *J Biol Chem* **265:**3432-3435.

Mackenzie PI, Miners JO and McKinnon RA (2000) Polymorphisms in UDP glucuronosyltransferase genes: functional consequences and clinical relevance. *Clin Chem Lab Med* **38:**889-892.

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.

Mackenzie PI, Walter Bock K, 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.

Maruo Y, Sato H, Yamano T, Doida Y and Shimada M (1998) Gilbert syndrome caused by a homozygous missense mutation (Tyr486Asp) of bilirubin UDP-glucuronosyltransferase gene. *J Pediatr* **132:**1045-1047.

Maruo Y, Topaloglu AK, Takahashi H, Mori A, Iwai M, Duzovali O, Yamamoto K, Matui K, Sato H and Takeuchi Y (2006) Crigler-Najjar Syndrome Type II Caused by a Homozygous Triple Mutation [T-3279G, A(TA)7TAA, and H39D] of UGT1A1. *J Pediatr Gastroenterol Nutr* **42:**236-239.

Matern H, Lappas N and Matern S (1991) Isolation and characterization of hyodeoxycholic-acid: UDP-glucuronosyltransferase from human liver. *Eur J Biochem* **200:**393-400.

Matern H, Matern S and Gerok W (1982) Isolation and characterization of rat liver microsomal UDP-glucuronosyltransferase activity toward chenodeoxycholic acid and testosterone as a single form of enzyme. *J Biol Chem* **257:**7422-7429.

Meech R and Mackenzie PI (1998) Determinants of UDP glucuronosyltransferase membrane association and residency in the endoplasmic reticulum. *Arch Biochem Biophys* **356:**77-85.

Meech R and Mackenzie PI (1997) UDP-glucuronosyltransferase, the role of the amino terminus in dimerization. *J Biol Chem* **272:**26913-26917.

Miley MJ, Zielinska AK, Keenan JE, Bratton SM, Radominska-Pandya A and Redinbo MR (2007) Crystal Structure of the Cofactor-Binding Domain of the Human Phase II Drug-Metabolism Enzyme UDP-Glucuronosyltransferase 2B7. *J Mol Biol* **369:**498-511.

Miller GP, Lichti CF, Zielinska AK, Mazur A, Bratton SM, Gallus-Zawada AE, Finel M, Moran JH and Radominska-Pandya A (2008) Identification of hydroxywarfarin binding site in human UGT1A10: Phenylalanine90 is crucial for the glucuronidation of 6- and 7-hydroxy- but not 8- hydroxywarfarin. *Drug Metab Dispos*

Miners JO, McKinnon RA and Mackenzie PI (2002) Genetic polymorphisms of UDP-glucuronosyltransferases and their functional significance. *Toxicology* **181-182:**453-456.

Mittler M, Bechthold A and Schulz GE (2007) Structure and action of the C-C bond-forming glycosyltransferase UrdGT2 involved in the biosynthesis of the antibiotic urdamycin. *J Mol Biol* **372:**67-76.

Morera S, Lariviere L, Kurzeck J, Aschke-Sonnenborn U, Freemont PS, Janin J and Ruger W (2001) High resolution crystal structures of T4 phage beta-glucosyltransferase: induced fit and effect of substrate and metal binding. *J Mol Biol* **311:**569-577.

Mulichak AM, Losey HC, Lu W, Wawrzak Z, Walsh CT and Garavito RM (2003) Structure of the TDP-epi-vancosaminyltransferase GtfA from the chloroeremomycin biosynthetic pathway. *Proc Natl Acad Sci U S A* **100:**9238-9243.

Mulichak AM, Losey HC, Walsh CT and Garavito RM (2001) Structure of the UDP-glucosyltransferase GtfB that modifies the heptapeptide aglycone in the biosynthesis of vancomycin group antibiotics. *Structure* **9:**547-557.

Mulichak AM, Lu W, Losey HC, Walsh CT and Garavito RM (2004) Crystal structure of vancosaminyltransferase GtfD from the vancomycin biosynthetic pathway: interactions with acceptor and nucleotide ligands. *Biochemistry* **43:**5170-5180.

Nishiyama T, Fujishima M, Masuda Y, Izawa T, Ohnuma T, Ogura K and Hiratsuka A (2008) Amino acid positions 69-132 of UGT1A9 are involved in the C-glucuronidation of phenylbutazone. *Arch Biochem Biophys* **478:**75-80.

Noguchi A, Saito A, Homma Y, Nakao M, Sasaki N, Nishino T, Takahashi S and Nakayama T (2007) A UDP-glucose:isoflavone 7-O-glucosyltransferase from the roots of soybean (*Glycine max*) seedlings: Purification, gene cloning, phylogenetics, and an implication for an alternative strategy of enzyme catalysis. *J Biol Chem* **282:**23581-90.

Offen W, Martinez-Fleites C, Yang M, Kiat-Lim E, Davis BG, Tarling CA, Ford CM, Bowles DJ and Davies GJ (2006) Structure of a flavonoid glucosyltransferase reveals the basis for plant natural product modification. *EMBO J* **25:**1396-1405.

Operana TN and Tukey RH (2007) Oligomerization of the UDP-glucuronosyltransferase 1A proteins: homo- and heterodimerization analysis by fluorescence resonance energy transfer and co-immunoprecipitation. *J Biol Chem* **282:**4821-4829.

Ouzzine M, Antonio L, Burchell B, Netter P, Fournel-Gigleux S and Magdalou J (2000) Importance of histidine residues for the function of the human liver UDP-glucuronosyltransferase UGT1A6: evidence for the catalytic role of histidine 370. *Mol Pharmacol* **58:**1609-1615.

Ouzzine M, Barre L, Netter P, Magdalou J and Fournel-Gigleux S (2003) The human UDP-glucuronosyltransferases: structural aspects and drug glucuronidation. *Drug Metab Rev* **35:**287-303.

Ouzzine M, Magdalou J, Burchell B and Fournel-Gigleux S (1999a) Expression of a functionally active human hepatic UDP-glucuronosyltransferase (UGT1A6) lacking the *N*-terminal signal sequence in the endoplasmic reticulum. *FEBS Lett* **454:**187-191.

Ouzzine M, Magdalou J, Burchell B and Fournel-Gigleux S (1999b) 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.

Owens D and Evans J (1975) Population studies on Gilbert's syndrome. *J Med Genet* **12:**152-156.

Paquette S, Moller BL and Bak S (2003) On the origin of family 1 plant glycosyl-transferases. *Phytochemistry* **62:**399-413.

Peters WH, Jansen PL and Nauta H (1984) The molecular weights of UDP-glucuronyltransferase determined with radiation-inactivation analysis. A molecular model of bilirubin UDP-glucuronyltransferase. *J Biol Chem* **259:**11701-11705.

Pillot T, Ouzzine M, Fournel-Gigleux S, Lafaurie C, Tebbi D, Treat S, Radominska A, Lester R, Siest G and Magdalou J (1993) Determination of the human liver UDP-glucuronosyltransferase 2B4 domains involved in the binding of UDP-glucuronic acid using photoaffinity labeling of fusion proteins. *Biochem Biophys Res Commun* **197:**785-791.

Potrepka RF and Spratt JL (1972) A study on the enzymatic mechanism of guinea-pig hepatic-microsomal bilirubin glucuronyl transferase. *Eur J Biochem* **29:**433-439.

Qasba PK, Ramakrishnan B and Boeggeman E (2005) Substrate-induced conformational changes in glycosyltransferases. *Trends Biochem Sci* **30:**53-62.

Radominska-Pandya A, Czernik PJ, Little JM, Battaglia E and Mackenzie PI (1999) Structural and functional studies of UDP-glucuronosyltransferases. *Drug Metab Rev* **31:**817-899.

Radominska-Pandya A, Ouzzine M, Fournel-Gigleux S and Magdalou J (2005) Structure of UDP-glucuronosyltransferases in membranes. *Methods Enzymol* **400:**116-147.

Richter WJ, Alt KO, Dieterle W, Faigle JW, Kriemler HP, Mory H and Winkler T (1975) C-glucuronides, a novel type of drug metabolites. *Helv Chim Acta* **58:**2512-2517.

Ritter JK, Chen F, Sheen YY, Tran HM, Kimura S, Yeatman MT and Owens IS (1992a) A novel complex locus UGT1 encodes human bilirubin, phenol, and other UDP-glucuronosyltransferase isozymes with identical carboxyl termini. *J Biol Chem* **267:**3257-3261.

Ritter JK, Yeatman MT, Ferreira P and Owens IS (1992b) Identification of a genetic alteration in the code for bilirubin UDP-glucuronosyltransferase in the UGT1 gene complex of a Crigler-Najjar type I patient. *J Clin Invest* **90:**150-155.

Ross J, Li Y, Lim E and Bowles DJ (2001) Higher plant glycosyltransferases. *Genome Biol* **2:**reviews 3004.1-3004.6.

Rouits E, Boisdron-Celle M, Dumont A, Guerin O, Morel A and Gamelin E (2004) Relevance of different UGT1A1 polymorphisms in irinotecan-induced toxicity: a molecular and clinical study of 75 patients. *Clin Cancer Res* **10:**5151-5159.

Sanchez E and Tephly TR (1975) Morphine Metabolism: IV. Studies on the Mechanism of Morphine:Uridine Diphosphoglucuronyltransferase and Its Activation by Bilirubin. *Mol Pharmacol* **11:**613-620.

Schultz LW, Quirk DJ and Raines RT (1998) His...Asp catalytic dyad of ribonuclease A: structure and function of the wild-type, D121N, and D121A enzymes. *Biochemistry* **37:**8886-8898.

Senay C, Battaglia E, Chen G, Breton R, Fournel-Gigleux S, Magdalou J and Radominska-Pandya A (1999) Photoaffinity labeling of the aglycon binding site of the recombinant human liver UDP-glucuronosyltransferase UGT1A6 with 7-azido-4-methylcoumarin. *Arch Biochem Biophys* **368:**75-84.

Senay C, Ouzzine M, Battaglia E, Pless D, Cano V, Burchell B, Radominska A, Magdalou J and Fournel-Gigleux S (1997) Arginine 52 and histidine 54 located in a conserved amino-terminal hydrophobic region (LX2-R52-G-H54-X3-V-L) are important amino acids for the functional and structural integrity of the human liver UDP-glucuronosyltransferase UGT1*6. *Mol Pharmacol* **51:**406-413.

Servedio V, d'Apolito M, Maiorano N, Minuti B, Torricelli F, Ronchi F, Zancan L, Perrotta S, Vajro P, Boschetto L and Iolascon A (2005) Spectrum of UGT1A1 mutations in Crigler-Najjar (CN) syndrome patients: identification of twelve novel alleles and genotype-phenotype correlation. *Hum Mutat* **25:**325.

Shao H, He X, Achnine L, Blount JW, Dixon RA and Wang X (2005) Crystal structures of a multifunctional triterpene/flavonoid glycosyltransferase from *Medicago truncatula*. *Plant Cell* **17:**3141-3154.

Sheweita SA (2000) Drug-metabolizing enzymes: mechanisms and functions. *Curr Drug Metab* **1:**107-132.

Soars MG, Burchell B and Riley RJ (2002) *In vitro* analysis of human drug glucuronidation and prediction of *in vivo* metabolic clearance. *J Pharmacol Exp Ther* **301:**382-390.

Sorich MJ, McKinnon RA, Miners JO and Smith PA (2006) The importance of local chemical structure for chemical metabolism by human uridine 5'-diphosphate-glucuronosyltransferase. *J Chem Inf Model* **46:**2692-2697.

Sorich MJ, Smith PA, Miners JO, Mackenzie PI and McKinnon RA (2008) Recent advances in the *in silico* modelling of UDP glucuronosyltransferase substrates. *Curr Drug Metab* **9:**60-69.

Starlard-Davenport A, Xiong Y, Bratton S, Gallus-Zawada A, Finel M and Radominska-Pandya A (2007) Phenylalanine(90) and phenylalanine(93) are crucial amino acids within the estrogen binding site of the human UDP-glucuronosyltransferase 1A10. *Steroids* **72:**85-94.

Strassburg CP (2008) Pharmacogenetics of Gilbert's syndrome. *Pharmacogenomics* **9:**703-715.

Takeuchi K, Kobayashi Y, Tamaki S, Ishihara T, Maruo Y, Araki J, Mifuji R, Itani T, Kuroda M, Sato H, Kaito M and Adachi Y (2004) Genetic polymorphisms of bilirubin uridine diphosphate-glucuronosyltransferase gene in Japanese patients with Crigler-Najjar syndrome or Gilbert's syndrome as well as in healthy Japanese subjects. *J Gastroenterol Hepatol* **19:**1023-1028.

Trefzer A, Blanco G, Remsing L, Kunzel E, Rix U, Lipata F, Brana AF, Mendez C, Rohr J, Bechthold A and Salas JA (2002) Rationally designed glycosylated premithramycins: hybrid aromatic polyketides using genes from three different biosynthetic pathways. *J Am Chem Soc* **124:**6056-6062.

Tukey RH and Strassburg CP (2000) Human UDP-glucuronosyltransferases: Metabolism, expression and disease. *Annu Rev Pharmacol Toxicol* **40:**581-616.

Tukey RH and Strassburg CP (2001) Genetic multiplicity of the human UDP-glucuronosyltransferases and regulation in the gastrointestinal tract. *Mol Pharmacol* **59:**405-414.

Uchaipichat V, Galetin A, Houston JB, Mackenzie PI, Williams JA and Miners JO (2008) Kinetic modeling of the interactions between 4-methylumbelliferone, 1-naphthol and zidovudine glucuronidation by UDP-glucuronosyltransferase 2B7 (UGT2B7) provides evidence for multiple substrate binding and effector sites. *Mol Pharmacol* **74:**1152-62.

Udomuksorn W, Elliot DJ, Lewis BC, Mackenzie PI, Yoovathaworn K and Miners JO (2007) Influence of mutations associated with Gilbert and Crigler-Najjar type II syndromes on the glucuronidation kinetics of bilirubin and other UDP-glucuronosyltransferase 1A substrates. *Pharmacogenet Genomics* **17:**1017-1029.

Unligil UM and Rini JM (2000) Glycosyltransferase structure and mechanism. *Curr Opin Struct Biol* **10:**510-517.

Vessey DA and Zakim D (1972) Regulation of microsomal enzymes by phospholipids. V. Kinetic studies of hepatic uridine diphosphate-glucuronyltransferase. *J Biol Chem* **247:**3023-3028.

Vrielink A, Ruger W, Driessen HP and Freemont PS (1994) Crystal structure of the DNA modifying enzyme beta-glucosyltransferase in the presence and absence of the substrate uridine diphosphoglucose. *EMBO J* **13:**3413-3422.

Wang X, Weldeghiorghis T, Zhang G, Imperiali B and Prestegard JH (2008) Solution structure of alg13: the sugar donor subunit of a yeast *N*-acetylglucosamine transferase. *Structure* **16:**965-975.

Wiggins CA and Munro S (1998) Activity of the yeast MNN1 alpha-1,3-mannosyltransferase requires a motif conserved in many other families of glycosyltransferases. *Proc Natl Acad Sci U S A* **95:**7945-7950.

Xiong Y, Bernardi D, Bratton S, Ward MD, Battaglia E, Finel M, Drake RR and Radominska-Pandya A (2006) Phenylalanine 90 and 93 are localized within the phenol binding site of human UDP-glucuronosyltransferase 1A10 as determined by photoaffinity labeling, mass spectrometry, and site-directed mutagenesis. *Biochemistry* **45:**2322-2332.

Yamamoto K, Sato H, Fujiyama Y, Doida Y and Bamba T (1998) Contribution of two missense mutations (G71R and Y486D) of the bilirubin UDP glycosyltransferase (UGT1A1) gene to phenotypes of Gilbert's syndrome and Crigler-Najjar syndrome type II. *Biochim Biophys Acta* **1406:**267-273.

Yin H, Bennett G and Jones JP (1994) Mechanistic studies of uridine diphosphate glucuronosyltransferase. *Chem Biol Interact* **90:**47-58.

Zhang C, Bitto E, Goff RD, Singh S, Bingman CA, Griffith BR, Albermann C, Phillips GN,Jr and Thorson JS (2008) Biochemical and structural insights of the early glycosylation steps in calicheamicin biosynthesis. *Chem Biol* **15:**842-853.

Zhang JY, Wang Y and Prakash C (2006) Xenobiotic-metabolizing enzymes in human lung. *Curr Drug Metab* **7:**939-948.