J. Membrane Biol. 208, 213–227 (2005)
DOI: 10.1007/s00232-005-7004-x

The Journal of
Membrane
Biology

Organic Anion Transporting Polypeptides of the OATP/*SLCO* Superfamily: Identification of New Members in Nonmammalian Species, Comparative Modeling and a Potential Transport Mode

Fabienne Meier-Abt*, Younes Mokrab*, Kenji Mizuguchi*,†

*Department of Biochemistry, University of Cambridge, Old Addenbrookes Site, 80 Tennis Court Road, Cambridge CB1 1GA, United Kingdom

†Department of Applied Mathematics and Theoretical Physics, Centre for Mathematical Sciences, Wilberforce Road, Cambridge CB3 0WA, United Kingdom

Received: 15 September 2005/Revised: 5 December 2005

Abstract. Organic anion-transporting polypeptides (human, OATPs; other animals, Oatps; gene symbol, *SLCO*/*Slco*) form a transport protein superfamily that mediates the translocation of amphipathic substrates across the plasma membrane of animal cells. So far, OATPs/Oatps have been identified in human, rat and mouse tissues. In this study, we used bioinformatic tools to detect new members of the OATP/*SLCO* superfamily in nonmammalian species and to build models for the three-dimensional structure of OATPs/Oatps. New OATP/*SLCO* superfamily members, some of which form distinct novel families, were identified in chicken, zebrafish, frog, fruit fly and worm species. The lack of OATP/*SLCO* superfamily members in plants, yeast and bacteria suggests the emergence of an ancient Oatp protein in an early ancestor of the animal kingdom. Structural models were generated for the representative members OATP1B3 and OATP2B1 based on the known structures of the major facilitator superfamily of transport proteins. A model was also built for the large extracellular region between transmembrane helices 9 and 10, following the identification of a novel homology with the Kazal-type serine protease inhibitors. Along with the electrostatic potential and the conservation of key amino acid residues, we propose a common transport mechanism for all OATPs/Oatps, whereby substrates are translocated through a central, positively charged pore in a rocker-switch type of mechanism. Several amino acid residues were identified that may play crucial roles in the proposed transport mechanism.

Correspondence to: Kenji Mizuguchi; email: Kenji@cryst.bioc.cam.ac.uk

Key words: Organic anion-transporting polypeptide (OATP) — Phylogeny — Structural model — Transport mode

Introduction

Organic anion-transporting polypeptides (OATPs in humans, Oatps in all other species; gene symbol *SLCO*/*Slco*, for solute carrier OATP) are transmembrane solute carriers that mediate the sodium-independent transport of a wide range of amphipathic endogenous and exogenous organic compounds across the plasma membrane of animal cells. So far, 39 members of the OATP/*SLCO* superfamily have been identified in mammalian species such as human, rat and mouse (Hagenbuch & Meier 2003, 2004). Whereas most OATPs/Oatps are expressed in multiple tissues, including the blood-brain barrier, choroid plexus, lung, heart, liver, intestine, kidney, placenta and testes, a few show organ-specific expression, such as Oatp1b2, OATP1B1 and OATP1B3, which are selectively expressed at the basolateral plasma membrane of liver parenchymal cells (hepatocytes). Typical OATP/Oatp substrates include bile salts, organic dyes, steroid conjugates, thyroid hormones, anionic oligopeptides and numerous drugs (Hagenbuch & Meier, 2003).

Important functions of OATPs/Oatps are the transport of amphipathic substances across the blood-brain barrier (Hagenbuch, Gao & Meier, 2002) and the uptake of potentially toxic amphipathic substances from the blood into liver cells (St. Pierre et al., 2001; Meier & Stieger, 2002; Kullak-Ublick, Stieger & Meier, 2004). Within hepatocytes, the amphipathic endo- and xenobiotics are biotransformed into more

water-soluble products that can then be excreted into bile. Given their central role in the elimination of potentially toxic amphipathic substances, it is not surprising that genetic polymorphisms leading to the decreased expression and/or function of hepatocellular OATPs/Oatps (Tirona & Kim, 2002) and the co-administration of OATP/Oatp inhibitors (Fattinger et al., 2000) are both associated with increased systemic bioavailability of drugs. Conversely, decreased OATP/Oatp-mediated uptake can prevent toxic liver injury such as amatoxin- or microcystin-induced hepatotoxicity (Meier-Abt, Faulstich & Hagenbuch, 2004; Fischer et al., 2005). Because of these important physiological roles, the elucidation of the mechanism(s) of OATP/Oatp-mediated substrate transport is of considerable interest to both basic physiology and clinical medicine. So far, extensive transport studies in *Xenopus laevis* oocytes have identified many transport substrates of individual OATPs/Oatps (Hagenbuch & Meier, 2003, 2004). However, very little is known about the actual mechanism of OATP/Oatp-mediated transport.

Several conserved structural features are believed to be important for the transport function of OATPs/Oatps. These features include the number of transmembrane (TM) helices, the so-called superfamily signature and the large extracellular region between TM helices 9 and 10. Based on hydropathy analyses, all OATPs/Oatps are thought to have 12 TM helices, referred to as H1-H12 hereafter (Kullak-Ublick et al., 1995). The number of TM helices in OATPs/Oatps is consistent with the observation of a distant evolutionary relationship with the major facilitator superfamily (MFS) of transport proteins, the members of which also have 12 TM helices (Chang et al., 2004). However, whereas there is considerable information on the structure of the MFS proteins (Abramson et al., 2003; Huang et al., 2003), little is known about the arrangement of the TM helices relative to each other in OATPs/Oapts. The superfamily signature is a stretch of 13 amino acids, which are well conserved between human, rat and mouse OATPs/Oatps. The signature is located at the extracellular border of H6 and has been used to search protein databases for OATP/Oatp-related proteins (Hagenbuch & Meier, 2003). The large extracellular region between H9 and H10 has drawn attention to itself due to the presence of 11 conserved cysteine residues, all of which are disulfide-bonded (Haenggi et al., 2005). Mutation of just one cysteine residue is enough to destroy any transport function of the OATPs/Oatps (Haenggi et al., 2005), indicating a central role of the large extracellular region in transporting amphipathic substances across the plasma membrane.

A recent paper analyzed the amino acid sequence conservation within the 36 human, rat and mouse OATP/Oatp members that had been identified at the time (Hagenbuch & Meier, 2004). A phylogenetic tree was generated, and a new nomenclature was introduced (HUGO gene nomenclature committee website, www.gene.ucl.ac.uk/nomenclature/). The whole superfamily is classified into families and each family, into subfamilies. The individual families contain proteins with amino acid sequence identities of ≥40% and are indicated by numbers (e.g., OATP1, OATP2, OATP3, etc.). Subfamilies contain proteins with amino acid sequence identities of ≥60% and are denoted by letters (e.g., OATP1A, OATP1B, OATP1C, OATP2A, etc.). Individual paralogues within a subfamily are indicated by different final numbers (e.g., Oatp1a1, OATP1A2, Oatp1a3, Oatp1a4, etc.). The 36 human, rat and mouse OATPs/Oatps have been classified into six families. Whereas in some families, such as OATP3, there is a strict 1:1 relationship between proteins found in human, rat and mouse, this is not the case in other families. The subfamily OATP1A, for example, has only one human member but five rat members. OATP1B contains two human members but only one rat and one mouse member. Such inequalities are believed to be the result of gene-duplication events after the species divergence of humans and rodents. Not surprisingly, the conservation level is much higher in the OATP3 family than in the OATP1 family.

The first objective of the present study involved an extended identification and classification of potential Oatps in nonmammalian species, such as chicken, zebrafish, frog, fruit fly and worm. Together with human, rat and mouse, these species are representatives of animal evolution. We present OATP/*SLCO* homologues in all of these species, including the frog. This discovery may have implications on the transport studies carried out in the *X. laevis* expression system. We further show that while chicken, zebrafish and frog Oatps fall into families that also contain mammalian members, fruit fly and worm Oatps form new families. This correlates well with the proposed evolution of the named species and implies the presence of an ancient Oatp in the ancestor of animals.

The second aim of this study was to create structural models of representative OATP/*SLCO* superfamily members based on their distant relationship with the MFS. Combined with the analysis of amino acid conservation patterns across all known OATP/*SLCO* superfamily members, the structural models provide clues about critical residues for (1) the transport activity of OATPs/Oatps and (2) the substrate specificity of certain OATPs/Oatps. The models are validated by the calculation of electrostatic potentials, the examination of conserved hydrogen bonds and conserved breaks in helices and the investigations of known polymorphisms. On the basis of these analyses, we propose that OATPs/Oatps function in a rocker-switch type of mechanism,

translocating the substrate through a central, positively charged pore.

The third aim of the study concerned the large extracellular region. We report a distant homology of this region to the Kazal-type serine protease inhibitors. We show possible functions of the large extracellular region suggested from the model building.

Methods

Identification of New Members of the OATP/SLCO Superfamily and their Phylogenetic Classification

The accession codes for human, rat and mouse OATPs/Oatps were taken from the paper reporting the new nomenclature (Hagenbuch & Meier, 2004). In addition to the proteins reported by Hagenbuch and Meier (2004), the mammalian sequences used in our analyses included the rat proteins rOatp4c1 and rOatp5a1 and the mouse protein mOatp5a1. (However, mOatp5a1 is severely truncated and was discarded after examining the multiple sequence alignments below.) The amino acid sequences were retrieved from the National Center of Biotechnology Information (NCBI) website (www.ncbi.nlm.nih.gov). The human, rat and mouse OATP/Oatp sequences were aligned using the program MAFFT (Katoh et al., 2002). Different options of MAFFT were applied, including an iterative refinement option. The program TMHMM (Krogh et al., 2001) was then used to predict the TM regions of every sequence. A profile hidden Markov model (HMM) was created from the best mammalian OATP/Oatp alignment, using the HMMER program (Eddy, 1998). Profile HMMs represent sets of probability factors that provide information on the likelihood of a certain protein belonging to a specific family/superfamily. This HMM was used to search the chicken, zebrafish, frog, fruit fly and worm peptide and protein databases, which were downloaded from the EBI and the ENSEMBL websites (EBI, www.ebi.org; ENSEMBL, www.ensembl.org).

To retrieve only the true orthologues of OATPs/Oatps from the HMMER hits, a suitable e-value cut-off had to be chosen. The HMMER hits were examined by performing reciprocal BLAST (Altschul et al., 1997) searches against the human and rat genomes on the NCBI website. As noted previously, OATPs/Oatps are distantly related to the MFS proteins. A lenient e-value cut-off in the HMMER search process is thus expected to result in hits to protein superfamilies other than OATP/*SLCO*. In contrast, a too stringent cut-off will lead to true OATP/*SLCO* superfamily members being missed. Whether a reported hit is a true orthologue or a false-positive was determined by examining the highest-scoring BLAST hit in the human and rat genomes. The HMMER threshold was finally set to an e value of 0.01.

The hits retrieved by the searches were clustered so that no two sequences in the remaining representative group shared more than 85% sequence identity. The clustering cut-off was chosen to ensure that the remaining sequence pool included at least one translational product from every gene identified initially. Nevertheless, some alternative splicing products may have been lost in this process. In order to remove incorrectly predicted gene products, translation was checked manually and truncated sequences that did not span all 12 TM regions were discarded. Some true protein sequences may also have been removed in the latter process.

The mammalian and the newly identified sequences were aligned using MAFFT (Katoh et al., 2002), and TM regions were predicted using TMHMM (Krogh et al., 2001). The best alignment of all mammalian and nonmammalian sequences was chosen according to the criteria mentioned in Results.

The full alignment was used as input to the tree-construction program KITSCH (Felsenstein, 1989).

Structural Models

Homologues of OATP/Oatp proteins, the structure of which has been solved, were searched for using HMMER (Eddy, 1998), PSI-BLAST (Altschul et al., 1997), FUGUE (www-cryst.bioc.cam.ac.uk/fugue; Shi, et al., 2001) and a fold recognition meta server (bioinfo.pl/meta; Ginalski et al., 2003). Meta servers comprise many different programs, allowing errors associated with specific programs to be cancelled out. For PSI-BLAST and the meta server, the individual full-length sequences of OATP1B3 and OATP2B1 and the sequence of the large extracellular region of OATP1B3 (residues 430–530) were used as input. For FUGUE, a multiple sequence alignment of the mammalian OATP/Oatp sequences, clustered at the 60% sequence identity level, was used.

The structural homologues identified (the glycerol-3-phosphate transporter Protein Data Bank (PDB) 1pw4 [Huang et al., 2003] and the lactose permease PDB 1pv6 [Abramson et al., 2003]) were aligned using the COMPARER program (Sali & Blundell, 1990). Structural profiles were calculated for the individual structures separately and for the structure-based alignment. These profiles represent family-specific scoring systems and are required for the alignment of sequences against structures using FUGUE (Shi et al., 2001). Using the FUGUEALI program (Shi et al., 2001), the clustered mammalian OATP/Oatp sequences were aligned against the profiles derived from the individual structures or from the superposition. The alignments were analyzed using the same criteria as before. If necessary, long insertions were excised from the OATP/Oatp alignment in order to improve the alignment quality.

Once the decision was made to use both structures as templates for the comparative modeling of OATP1B3 and OATP2B1, the sequences other than that of OATP1B3 or OATP2B1 were removed from the full alignment and a new sequence-structure alignment was generated for modeling. The structures of OATP1B3 and OATP2B1 were modeled using the MODELLER program (Sali & Blundell, 1993; Fiser & Sali, 2003). MODELLER creates a set of distance and torsion angle restraints from the coordinates of the template(s) and the alignment. For each alignment, five models were built that satisfied as many restraints as possible.

The resulting models were evaluated on the basis of the energy values calculated by MODELLER (Sali & Blundell, 1993; Fiser & Sali, 2003) and the sequence-structure compatibility scores computed by PROSA2003 (Sippl, 1993) and VERIFY3D (Luthy, Bowie & Eisenberg, 1992). The alignments were manually altered using the ViTO editor (Catherinot & Labesse, 2004), and models were rebuilt to improve the quality scores. Finally, the single best model was chosen for OATP1B3 and OATP2B1.

The program JOY (Mizuguchi et al., 1998a) was employed to display the structural features of both the model and template structure(s) in a multiple sequence alignment. PyMOL (http://www.pymol.org) and RASMOL (Sayle, & Milner-White, 1995) were used to visualize the models three-dimensionally (3D). The electrostatic potential of the models was calculated by APBS (Baker et al., 2001) using the PyMOL plug-in with the default parameter settings.

The coordinates of the cardiac glycoside digoxin were taken from the PDB entry 1igj (Rose et al., 1983) and docked onto the

putative active site (Arg^{181}) of OATP1B3 using GOLD (Verdonk et al., 2003).

For the large extracellular region, the existing profile of the Kazal-type serine protease inhibitors in the HOMSTRAD database (www-cryst.bioc.cam.ac.uk/homstrad/; Mizuguchi et al., 1998b) was used to align a clustered multiple sequence alignment of OATPs/Oatps with these structures using FUGUE. After examining the alignments generated by the programs on the fold recognition meta server (Ginalski et al., 2003), the final alignment with porcine pancreatic trypsin inhibitor (PDB 1tgs, chain I; Bolognesi et al., 1982) was produced and manually adjusted. The rest of the procedure was the same as that for the TM domain modeling.

Results

Identification of New Members of the OATP/SLCO Superfamily and their Phylogenetic Classification

HMM-based database searches (Eddy, 1998) coupled with reciprocal BLAST searches (*see* Methods) identified novel members of the OATP/*SLCO* superfamily in chicken (*Gallus gallus*), zebrafish (*Danio rerio*), frog (*Xenopus tropicalis*), fruit fly (*Drosophila melanogaster*) and worm (*Caenorhabditis elegans*) species. These species were chosen because their genomes have been fully sequenced and because of their representative positions in animal evolution.

Figure 1 shows an extract of a multiple sequence alignment for all identified OATP/Oatp sequences including the newly discovered ones. The alignment was produced using MAFFT with an iterative refinement option (Katoh et al., 2002) because this program produced the best alignment in terms of the conservation of (1) the TM helices, (2) the superfamily signature (Hagenbuch & Meier, 2003), (3) the disulfide-bonded cysteine residues (Haenggi et al., 2005) and (4) the potential *N*-glycosylation sites in extracellular regions (Jacquemin et al., 1994). The full alignment is given in the electronic supplement.

Figure 2 displays a phylogenetic tree constructed from the alignment shown in Figure 1. After reexamining the existing classification system for mammalian OATP/*SLCO* superfamily members (Hagenbuch & Meier, 2004), the newly identified members were properly placed into the larger phylogenetic tree and assigned consistent new names. Table 1 shows the protein identifiers and names of the newly identified OATP/*SLCO* superfamily members. These names were generated based on the clustering indicated by the two red lines in Figure 2: the first red line (corresponding to the 40% sequence identity level) defines the family number, and the second red line (corresponding to the 60% sequence identity level) defines the subfamily letter; specific branching within a subfamily defines the subfamily member number, assigning the same number to orthologues and different numbers to paralogues.

A total of seven OATP families with two or more members can be distinguished, including the multi-species families OATP1, OATP2, OATP3, OATP4, OATP5 and OATP6 and the worm-specific family Oatp14. In addition, there are eight single-member families, ranging from the fly dOatp7a1 to the worm cOatp16a1.

Structure-Based Alignment of OATPS and MFS Confirms their Distant Relationship

Searching the PDB with HMMER (Eddy, 1998) and BLAST (Altschul et al., 1997) for OATP/*SLCO* homologues resulted in no hits. FUGUE (Shi et al., 2001), however, with its environment-specific substitution scores, found two significant hits to members of the MFS, the *Escherichia coli* glycerol-3-phosphate transporter (PDB 1pw4; Huang et al., 2003) with a Z score of 18.45 (> 99% confidence) and the *E. coli* lactose permease (PDB 1pv6; Abramson et al., 2003) with a Z score of 12.76 (> 99% confidence). The same two proteins were also identified by several programs on the fold recognition meta server (Ginalsky et al., 2003). This is consistent with the previously noted distant relationship between OATPs/Oatps and MFS proteins (Chang et al., 2004). (This relationship is also documented in the Pfam database [Bateman et al., 2004]). Two human proteins, OATP1B3 and OATP2B1, were chosen for model building because of their distinct expression and transport properties, which are representative of the OATP/*SLCO* superfamily and because they belong to two different families, both of which have been studied extensively in several laboratories.

Using FUGUE, mammalian OATP/Oatp sequences, including OATP1B3 and OATP2B1, were aligned against the structure of either the glycerol-3-phosphate transporter or the lactose permease or against a structure-based alignment of the two. Comparisons of the resulting alignments suggested that the glycerol-3-phosphate transporter and the lactose permease were roughly equally close to the OATP/*SLCO* superfamily; thus, we decided to use both MFS proteins as the templates for the modeling of OATPs/Oatps. To obtain alignments of good quality according to the criteria mentioned above, the N-terminal variable region, the large extracellular loop between H9 and H10 and the extracellular region between H3 and H4 had to be removed. Furthermore, most of the extracellular and intracellular protein regions in the final models were judged to be unreliable and excised.

A structure-based alignment of the final model of OATP1B3 and the two templates is shown in Figure 3. A corresponding alignment for OATP2B1 is given in the electronic supplement. Also shown in Figure 3 is the full-length sequence of OATP1B3, illustrating the unmodeled extracellular and intracellular portions.

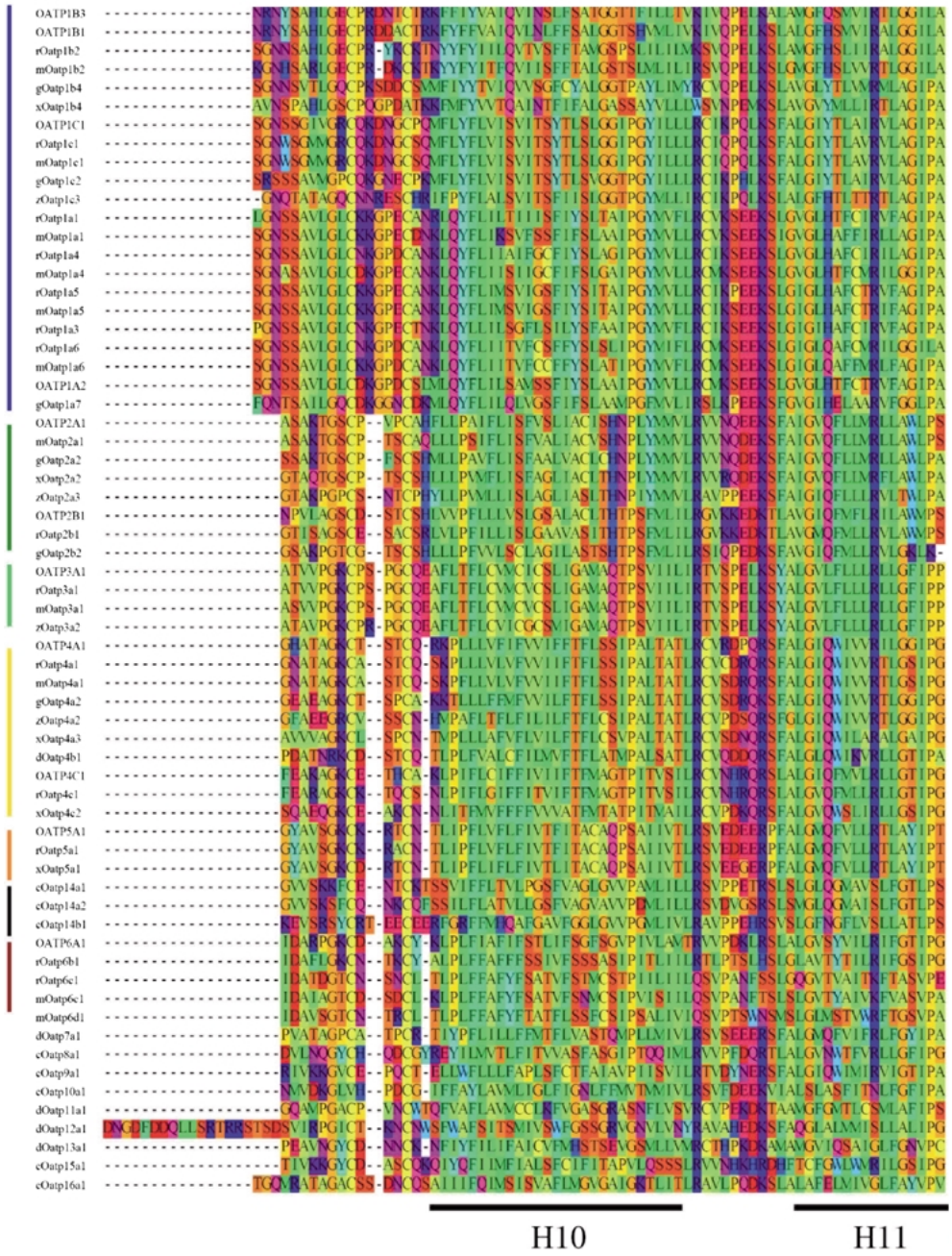

Fig. 1. Extract from the best alignment created for all OATP/Oatp members identified in the human, rat, mouse, chicken, zebrafish, frog, fruit fly and worm genomes. (This region includes the predicted TM helix 10 and part of helix 11, as indicated at the bottom.) Human proteins are denoted by capital letters, rat proteins by the prefix *r*, mouse proteins by the prefix *m*, chicken proteins by the prefix *g*, zebrafish proteins by the prefix *z*, frog proteins by the prefix *x* and worm proteins by the prefix *c*. The database accession codes for the human, rat and mouse proteins are found in Hagenbuch and Meier (2004), and the accession codes for the newly identified nonmammalian proteins are shown in Table 1. Multimember OATP families are denoted by the colored lines on the left (blue, OATP1; olive, OATP2; green, OATP3; yellow, OATP4; orange, OATP5; violet, OATP6; black, OATP14). The black line designates the newly identified OATP family OATP14 in worms. Amino acid residues are colored according to the Taylor scheme (Taylor, 1986). The alignment was generated with MAFFT (Katoh et al., 2002) and its quality assessed according to the criteria mentioned in the text. The full alignment can be found in the electronic supplement.

Structural Models of OATP1B3 and OATP2B1

The well-aligned TM regions (Fig. 3 and electronic supplement) suggested that the OATP models for these regions should be of reasonable quality. This is also expected from the previous observation that all the MFS proteins would adopt highly similar structures despite their diverse sequences (Vardy et al.,

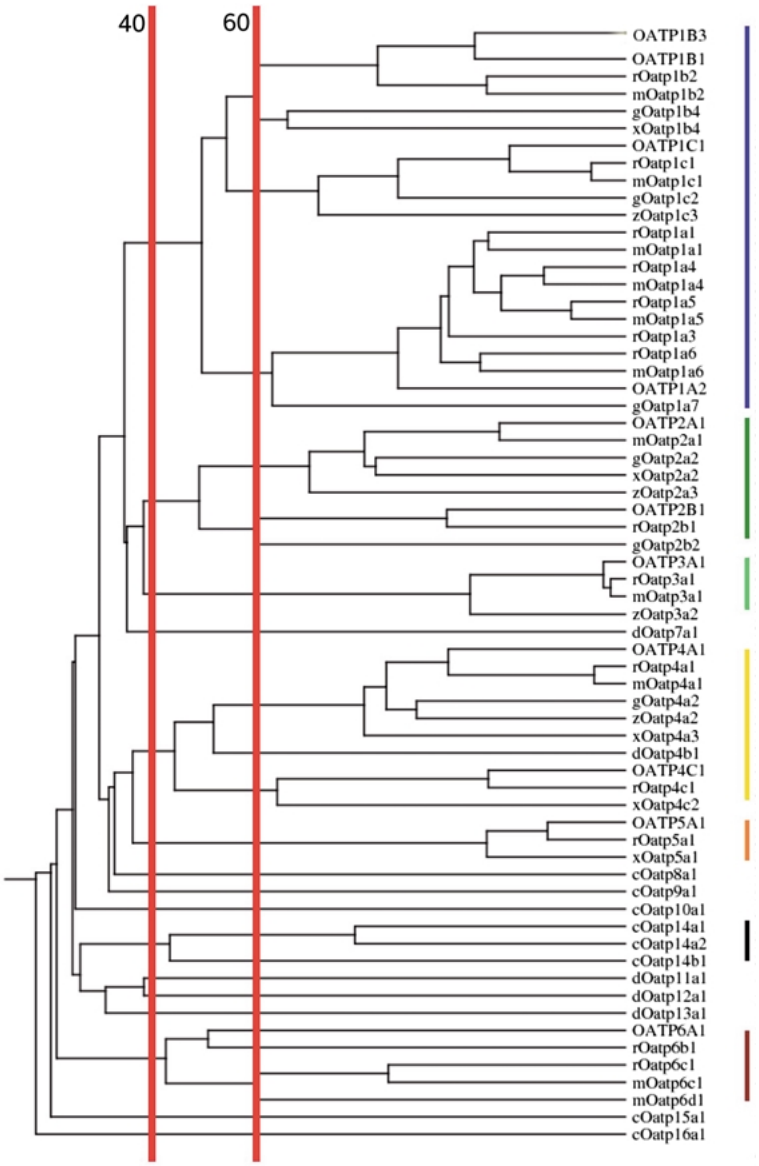

Fig. 2. A new phylogenetic tree and annotation of chicken, zebrafish, frog, fruit fly and worm members of the OATP/*SLCO* superfamily. Chicken Oatps are indicated by *g*, zebrafish Oatps by *z*, frog Oatps by *x*, fruit fly Oatps by *d* and worm Oatps by *c*. The first and second red lines correspond to 40% and 60% amino acid sequence identity within the mammalian OATP/*SLCO* superfamily members and were used to define the family and subfamily membership, respectively. Multimember OATP families are indicated by the rainbow-colored lines on the right (blue, OATP1; olive, OATP2; green, OATP3; yellow, OATP4; orange, OATP5; violet, OATP6). The black line designates the newly identified OATP family OATP14 in worms. The tree was generated using KITSCH (Felsenstein, 1989).

2004). Indeed, according to the sequence-structure compatibility programs PROSA2003 (Sippl, 1993) and VERIFY3D (Luthy et al., 1992), the final models of OATP1B3 and OATP2B1 were of relatively good quality (although these assessment programs were designed for globular proteins and thus the model evaluation may comprise some errors). The PROSA Z-score values were -3.46 for the OATP1B3 model and even better for the OATP2B1 model. The VERIFY3D scores were positive for most of the modeled regions except for the areas near the excised loops, which show negative scores also in the template structures.

Figure 4 shows a cartoon representation of the OATP1B3 model. As mentioned above, the overall structure is similar to those of the known MFS proteins, with common features including a central pseudo twofold symmetry axis, perpendicular to the membrane plane, and a central pore. The symmetry axis relates the six N-terminal helices to the six C-terminal helices around the central pore, which is closed at one side of the membrane and open at the other. Helices 1, 2, 4 and 5 of the N-terminal half and helices 7, 8, 10 and 11 of the C-terminal half face the pore, whereas helices 3, 6, 9 and 12 are largely embedded in the bilayer. The presence of a central pore is likely to be a conserved feature and functionally significant (*see below*). Because both the glycerol-3-phosphate transporter and the lactose permease were solved in a conformation with the pore opening toward the intracellular side (Abramson et al., 2003; Huang et al., 2003), the pore of the OATP1B3 model faces the intracellular fluid, i.e., the cytoplasm. Very similar features can be observed in the OATP2B1 model (*data not shown*).

Table 1. Annotation of newly identified Oatp members in chicken (prefix *g*), zebrafish (prefix *z*), frog (prefix *x*), fruit fly (prefix *d*) and worm (prefix *c*)

Peptide/protein ID	Name
ENSGALP00000021470	gOatp1a7
ENSGALP00000021446	gOatp1b4
ENSGALP00000021444	gOatp1c2
ENSGALP00000010438	gOatp2a2
ENSGALP00000027904	gOatp2b2
ENSGALP00000008750	gOatp4a2
ENSDARP00000020391	zOatp1c3
ENSDARP00000050498	zOatp2a3
ENSDARP00000014829	zOatp3a2
ENSDARP00000014471	zOatp4a2
ENSXETP00000026397	xOatp1b4
ENSXETP00000025839	xOatp2a2
ENSXETP00000024581	xOatp4a3
ENSXETP00000051447	xOatp4c2
ENSXETP00000033334	xOatp5a1
Q8SY02	dOatp4b1
Q9VLB3	dOatp7a1
Q9W269	dOatp11a1
Q9VK84	dOatp12a1
Q9VVH9	dOatp13a1
Q9XWC5	cOatp8a1
Q20702	cOatp9a1
Q93550	cOatp10a1
Q20538	cOatp14a1
Q95ZT3	cOatp14a2
O62421	cOatp14b1
Q21157	cOatp15a1
Q9VK82	cOatp16a1

The protein IDs refer to ENSEMBL in the case of chicken, zebrafish and frog and to UniProt in the case of fruit fly and worm (see references). The criteria used to derive the protein names are denoted in the text.

Figure 5 shows the electrostatic potential mapped onto the surface of the OATP1B3 model. Viewing the model from the intracellular side, the electrostatic potential is very positive within the putative pore. This is consistent with a pore that is exposed to an aqueous phase. Moreover, the positive electrostatic pore potential should facilitate binding and transport of negatively charged compounds, which conforms with the anionic nature of the majority of OATP/Oatp substrates. Viewing the model from the extracellular side, a small portion of the positively charged pore can be seen, and it seems plausible that the pore could be exposed to the extracellular fluid upon a conformational change. Viewed from the lateral side, the surface of the OATP1B3 model is neutral as it is required for its interposition within the lipid bilayer. Very similar features were observed in the electrostatic potential of the OATP2B1 model (*data not shown*).

Conserved Charged/Polar Residues Cluster around the Putative Pore

We identified several amino acid residues that are conserved within the OATP1 family and may characterize their function. In Figure 6, conserved Arg, Lys, Asn, Gln and His residues that were predicted to be solvent-exposed and face the pore are shown as stick models. These residues are responsible for the positive potential of the pore. An analogous pattern of conserved charged/polar residues was observed for OATP2B1 (*data not shown*). Some of these residues are also conserved in several other multimember families; however, none of them was found to be fully conserved throughout the OATP/*SLCO* superfamily, and no "family-specific" residue was identified, as defined by evolutionary trace analysis (Lichtarge, Bourne & Cohen, 1996) (i.e., residues invariant within each family but that change between families). However, one position is fully conserved in OATP1 and occupied by arginine (Arg^{181}); but at the equivalent position, no arginine is found in any other families. Arg^{181} is the only charged/polar, pore-facing residue unique to OATP1 in this sense. Similarly in OATP2, His^{579} is the only charged/polar residue unique to OATP2 and is found spatially close to Arg^{181} of OATP1. We hypothesize that Arg^{181} contributes to the substrate-binding site of the OATP1 family, whereas His^{579} is part of the substrate-binding site of the OATP2 family (*see below*).

Other Conserved Residues are Likely to Play Key Structural Roles

Further analysis of conservation patterns identified three hydrogen bonds that are likely to be conserved throughout the OATP1 and the OATP2 families. (1) The side chain of Asn^{77} is hydrogen-bonded to the main chain carbonyl of Phe^{73} in OATP1B3. (The equivalent residues in OATP2B1 are Asn^{98} and Asn^{94}, respectively.) This interaction may be important in stabilizing the pore-facing H2. (2) The side chain of Arg^{93} is hydrogen-bonded to the main chain carbonyl of Val^{189} in OATP1B3 (Arg^{114} and Gln^{207} in OATP2B1). This interaction may be important in fixing the orientation of the pore-facing H4. (3) The side chain of Ser^{228} is hydrogen-bonded to the main chain carbonyl of Ala^{225} in OATP1B3 (Ser^{246} and Gly^{243} in OATP2B1). Similar to the first case, this interaction may serve to stabilize the pore-facing H5. None of these conserved hydrogen bonds occurs in the template structures, suggesting that they may be important for the specific activity of OATPs/Oatps. Site-specific mutagenesis of the residues in question should provide further information on their exact roles.

Several conserved proline and glycine residues have been identified that result in breaks and bends in

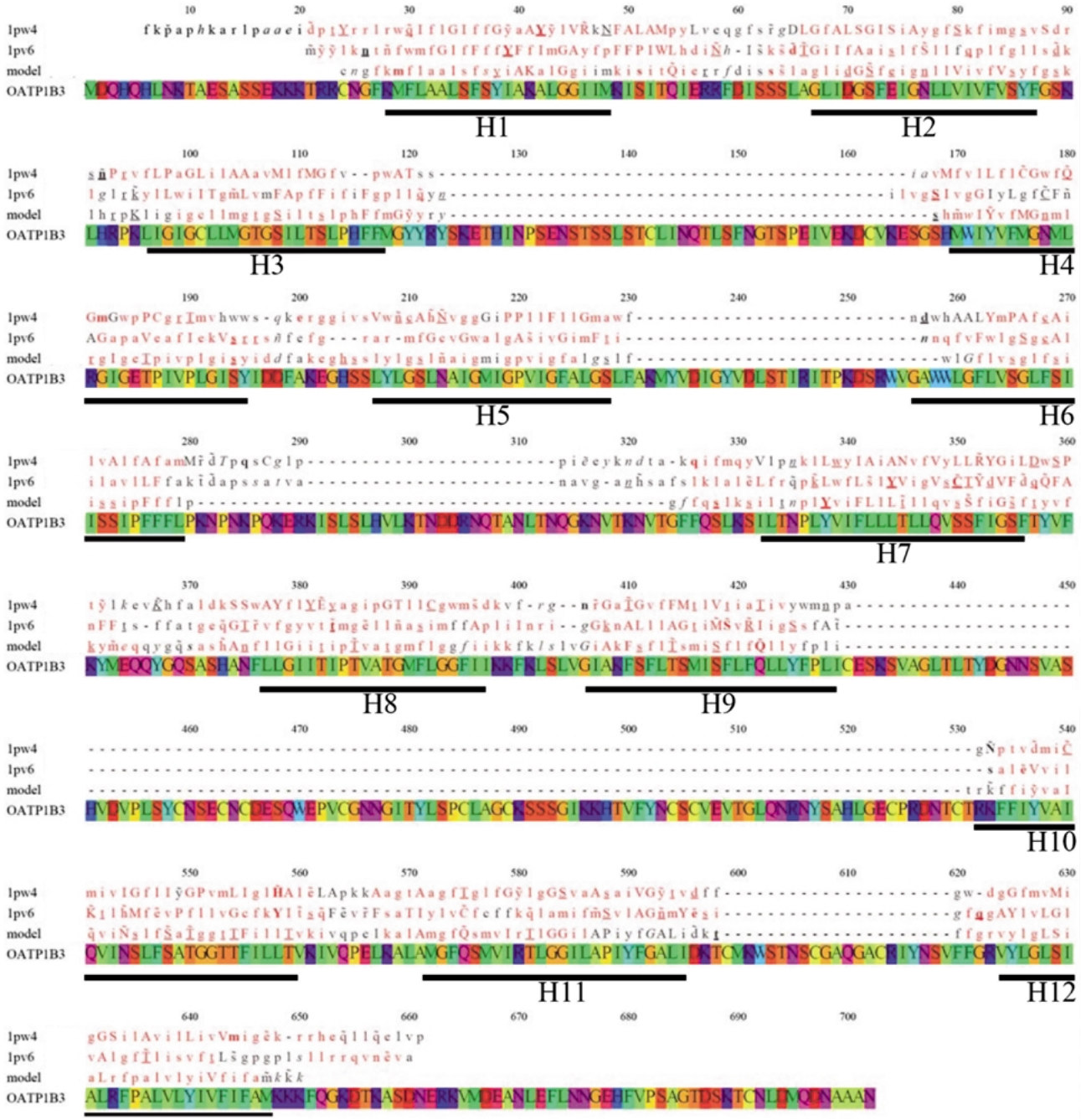

Fig. 3. Structure-based alignment of the *E. coli* glycerol-3-phosphate transporter (PDB 1pw4; Huang et al., 2003), the *E. coli* lactose permease (PDB 1pv6; Abramson et al., 2003) and the model of OATP1B3. The two template structures (1pw4 and 1pv6) were aligned by COMPARER (Sali & Blundell, 1990); then, the sequence of OAPT1B3 was added by FUGUE (Shi et al., 2001), followed by manual adjustments. The model was built by MODELLER (Sali & Blundell, 1993; Fiser & Sali, 2003). In order to illustrate the positions of the unmodeled extracellular and intracellular portions (*see text*), the full-length amino acid sequence of OATP1B3 is also shown. The alignment was formatted and analyzed with JOY (Mizuguchi et al., 1998a) to display key structural features in one dimension. The formatting convention of JOY is as follows: red, α-helix; blue, β-strand; maroon, 3_{10} helix; lower case, solvent-accessible; upper case, solvent-inaccessible; bold, hydrogen bond to main-chain amide; underlined, hydrogen bond to main-chain carbonyl; tilde, hydrogen bond to another side chain; cedilla, disulfide bond; italic, positive phi torsion angle. The positions of the TM helices, as predicted by TMHMM (Krogh, 2001) for the OATP1B3 sequence, are denoted by the black lines and the H1-H12 captions. Above the alignment, the amino acid numbers of the full-length OATP1B3 sequence are given as in UniProt (Bairoch et al., 2005).

the TM helices of the two OATP models. Pro^{114} in OATP1B3 (Pro^{135} in OATP2B1) is conserved throughout all known OATPs/Oatps, including the newly identified ones, with one nonconforming member. The proline residue in question is responsible for a major break in H3. The break in this

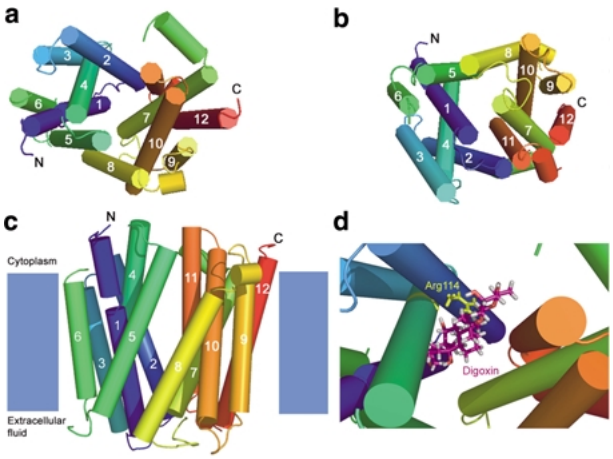

Fig. 4. Cartoon representation of the structural model of OATP1B3, viewed from (*a*) the intracellular side, (*b*) the extracellular side and (*c*) the lateral side. Helices are shown as cylinders and rainbow-colored (also in Figs. 6 and 7). The zooming is identical for (*a-c*). The approximate positions of the membrane bilayer and of the N and C termini are indicated in (*c*). (*d*) OATP1B3 substrate digoxin (shown in cyan as stick model) fitted into the pore of the OATP1B3 structural model, viewed from the intracellular side (in approximately the same orientation as *a* but zoomed in). The docking was performed by GOLD (Verdonk et al., 2003). The figure was generated with PyMOL (www.pymol.org).

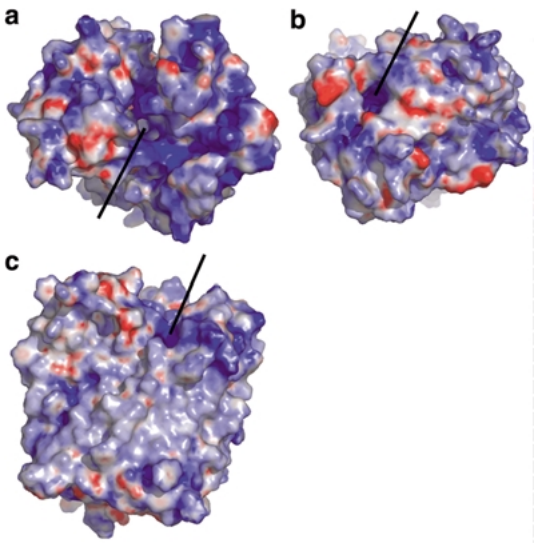

Fig. 5. Electrostatic potential of the structural model of OATP1B3, mapped onto its molecular surface and viewed from (*a*) the intracellular side, (*b*) the extracellular side and (*c*) the lateral side. (*a*) and (*b*) are in the same orientation as (*a*) and (*b*) from Figure 4, respectively. Regions of negative, positive and neutral potential are shown in red, blue and white, respectively. The position of the pore is indicated by a black line. The electrostatic potential was calculated by APBS (Baker et al., 2001) and visualized with PyMOL (www.pymol.org).

non-pore-facing helix occurs in both template structures and, thus, seems to be important for the general nature of MFS-related carrier proteins rather than

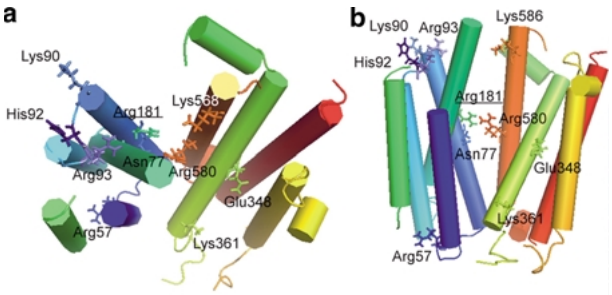

Fig. 6. Conserved amino acid residues that may be important for the transport function of OATP1B3. Solvent exposed and pore-facing Arg, Lys, Asn, Gln and His residues are shown as a stick model if they are conserved within the OATP1 family (with a maximum of one nonconforming member). Of these residues, Arg at 181 is fully conserved within the OATP1 family but does not occur in any other OATP/*SLCO* superfamily members at the equivalent position in the alignment (*underlined*). The TM helices, shown as cylinders, are rainbow-colored and viewed from (*a*) the intracellular side (as in Fig. 4*a*) and (*b*) the extracellular side (as in Fig. 4*b*). For clarity, helices 5, 8 and 10 are not displayed.

for OATP/Oatp-specific activity. The conserved residues Gly^{219} and Pro^{220} in OATP1B3 (Gly^{237} and Pro^{238} in OATP2B1), of which the former is conserved throughout all known OATPs/Oatps with one nonconforming member, lead to a break in H5, thereby allowing the major part of H5 to face the pore. The curved nature of H8 is caused by three conserved residues: Gly^{379}, Pro^{384} and Gly^{394} in OATP1B3 (Gly^{414}, Pro^{419} and Gly^{429} in OATP2B1), of which the latter glycine residue is conserved throughout all known OATPs/Oatps with one nonconforming member. H8 is curved also in both template structures, suggesting that the bend is important for the general nature of MFS-related transport proteins. The conserved residue Pro^{588} in OATP1B3 (Pro^{615} in OATP2B1) has a helix-breaking effect, which allows the major part of H11 to face the pore. There is no break in H11 in the two template structures. The break in H11 observed in the models could be an artefact from the superposition of the two templates. However, the conservation of the proline residue in question suggests that the break may be functionally important in maintaining the specific orientation of H11 in the OATP structures. Other conserved residues within the OATP/*SLCO* superfamily, responsible for slight kinks in pore-facing helices, include Pro^{187} and Gly^{573} in OATP1B3 (Pro^{205} and Gly^{600} in OATP2B1). All these residues are potential targets for site-directed mutagenesis.

The Large Extracellular Region would Adopt a Structure Similar to that of the Kazal-Type Serine Protease Inhibitors

Submission of the large extracellular region of OATP1B3 (residues 430–530) to the fold recognition

meta server (Ginalsky et al., 2003) resulted in consistent alignments with the Kazal-type serine protease inhibitors from most programs, some of which reported highly significant confidence values, such as a FASS03 score of −19.5 (>97% confidence) (Jaroszewski et al., 2005). The modest 3D jury (Ginalski et al., 2003) score (23.50) can be explained by the fact that these structures are small domains of about 56 amino acid residues and may still represent a significant value. After examining the alignments generated by the programs on the fold recognition meta server, porcine pancreatic trypsin inhibitor (PDB 1tgs, chain I; Bolognesi et al., 1982) was chosen as the single best template for modeling. This structure was the most frequently reported hit on the meta server.

The final sequence-structure alignment is displayed in Figure 7a. The Kazal-type serine protease inhibitors contain three disulfide bonds, all of which were aligned with conserved cysteine residues in the large extracellular region. The model assessment programs PROSA2003 (Sippl, 1993) and VERIFY3D (Luthy et al., 1992) reported high quality for the model, lending additional support to this structure prediction. (The PROSA2003 Z score was −4.59, as expected for high-quality models of this size, and the VERIFY3D scores were positive throughout all regions.)

The electrostatic potential of the large extracellular region is shown in Figure 7c. Surprisingly, it is not positive, indicating that the large extracellular region is probably not involved in the attraction of the substrate to the transporter.

Discussion

Identification of New Members of the OATP/SLCO Superfamily and their Phylogenetic Classification

Several Oatps have been identified in nonmammalian species, such as chicken, zebrafish, frog, fruit fly and worm. They have been classified into families and subfamilies and have been given new, consistent names (*see* Table 1). Figure 2 reveals that all chicken, zebrafish and frog Oatps fall into families that also contain mammalian members and that have thus been previously identified. In contrast, fruit fly and worm OATP/*SLCO* members form new families, such as OATP14. Whereas one fruit fly Oatp, dOatp4b1, is part of a family that also contains mammalian members, none of the worm Oatps belongs to previously defined families. These results conform well with the evolutionary distances between chicken, zebrafish, frog, fruit fly, worm and the mammalian species, the chicken being closest, followed by the zebrafish, frog, fruit fly and worm. No members of the OATP/*SLCO* superfamily have been found in plants, yeast or bacteria. Together with the above results, this suggests that an ancient Oatp evolved in the ancestor of animals.

The identification of Oatps in *Xenopus tropicalis* has potentially important implications on the transport studies that are routinely performed in the *X. laevis* expression system by numerous research groups (Hagenbuch & Meier, 2004). Based on the absence of sodium-independent bile salt uptake activity in *X. laevis* oocytes, it was assumed that no endogenous Oatps are present in *X. laevis* oocytes. However, it is highly likely that the *X. laevis* genome contains Oatp sequences, given the discovery of OATP/*SLCO* members in *X. tropicalis*. Many OATPs/Oatps are known to transport other substrates better than bile salts, and if such multispecific OATPs/Oatps are indeed expressed in the oocyte membrane, this could explain the occasionally observed very high uptake rates of potential OATP/Oatp substrates into water-injected oocytes (Tamai et al., 2001).

Structural Model of OATP1B3 and OATP2B1 and a Putative Transport Mode

OATP1B3 and OATP2B1 are representative human members of the OATP/*SLCO* superfamily. Whereas OATP1B3 is specifically expressed in human liver, OATP2B1 is found in liver cells as well as in placental cells (Tamai et al., 2000; St. Pierre et al., 2002). All of OATP2B1's substrates are transported also by other mammalian OATPs/Oatps. Similarly, OATP1B3 shares most of its transport substrates with other mammalian OATPs/Oatps. However, OATP1B3 also transports the cardiac glycoside digoxin and the intestinal octapeptide cholecystokinin 8 (Ismair et al., 2001), none of which is a substrate for any other mammalian OATP/Oatp. Because of these common as well as distinct expression and transport properties of OATP2B1 and OATP1B3 and because of the existence of studies on polymorphisms and artificial mutations in OATP1 and OATP2 family members, OATP1B3 and OATP2B1 were chosen for structural modeling.

The evolutionary relationship between OATPs/Oatps and the MFS has been suggested by several independent analyses; Chang et al. (2004) used PSI-BLAST (Altschul et al., 1997), and the same relationship is reported in the Pfam database (Bateman et al., 2004), the classification of which is based on a similar but distinct technique using HMMs (Eddy, 1998). Here, we have confirmed these results using the sequence-structure homology recognition software FUGUE (Shi et al., 2001) and the fold recognition meta server (Ginalsky et al., 2003). All these software programs report statistical scores and have been benchmarked extensively (e.g., at the CASP experiments [Moult et al., 2003]). As a control,

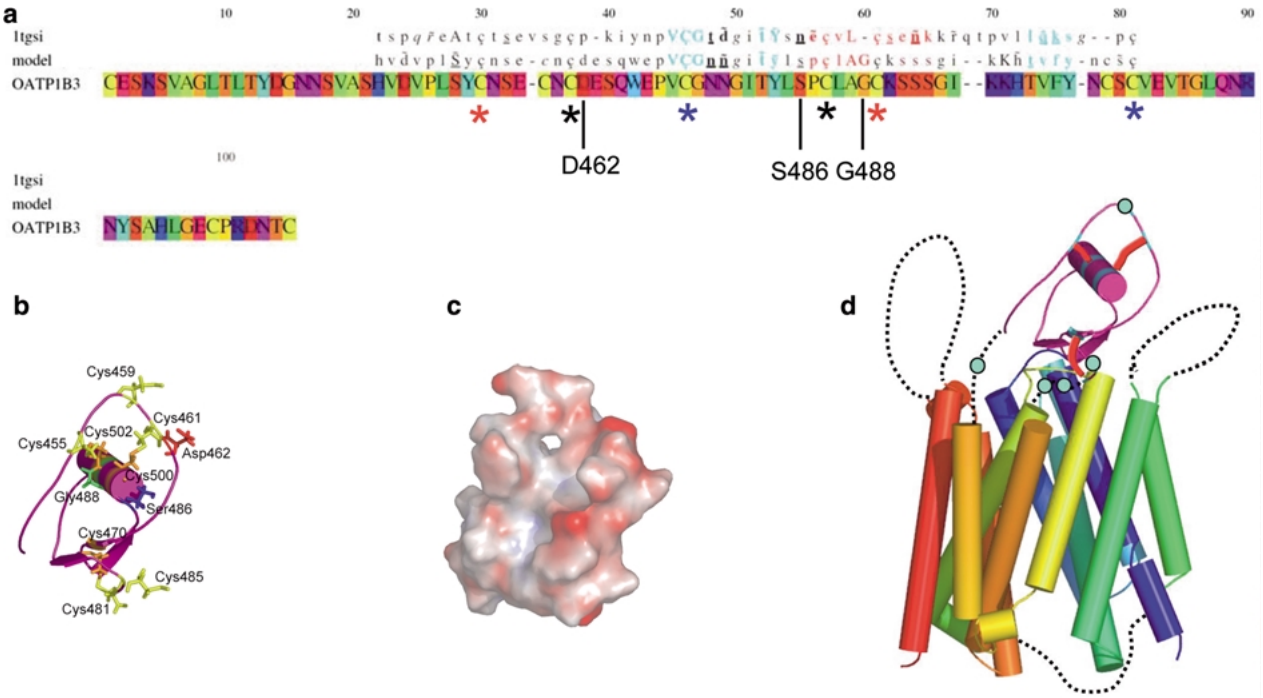

Fig. 7. (*a*) Structure-based alignment of the model of the large extracellular region of OATP1B3 and porcine pancreatic secretory trypsin inhibitor (PDB 1tgs, chain I; Bolognesi et al., 1982). Shown, too, is the amino acid sequence of the entire large extracellular region (residues 430-530) between H9 and H10. Actual residue numbers can be obtained by adding 429 to the alignment positions shown at the top. The three pairs of cysteine residues that would form putative disulfide bridges are marked with asterisks in three different colors. The alignment was generated with FUGUE (Shi et al., 2001) and formatted with JOY (Mizuguchi et al., 1998a). The formatting convention of JOY is given in the legend of Figure 3. (*b*) The three known polymorphisms that can be mapped onto the model of the large extracellular region: Ser^{486} to Phe (blue), Asp^{462} to Gly (red) and Gly^{488} to Ala (green). These are also highlighted in the alignment in (*a*). The three pairs of cysteine residues forming putative disulfide bridges in this region are shown in stick and colored in yellow-orange. The model was built with MODELLER (Sali & Blundell, 1993; Fiser & Sali, 2003) based on the alignment in (*a*). (*c*) Electrostatic potential calculated for the structural model of the large extracellular region. Regions of negative, positive and neutral potential are shown in red, blue and white, respectively. The molecular surface is viewed is in the same orientation as (*b*) and (*d*). The electrostatic potential was calculated by APBS (Baker et al., 2001) and visualized with PyMOL (www.pymol.org). (*d*) Hypothetical illustration of the whole OATP1B3 protein, including the model for the TM regions (in rainbow colors), the modeled part of the large extracellular region (in magenta) and the unmodeled loops (black discontinuous lines). Shown as red bars are the three disulfide bonds, which were modeled on the basis of the alignment of the large extracellular region with the Kazal-type serine protease inhibitors. The other cysteine residues in the large extracellular region are shown as cyan circles (two of which are within the modeled region, one between H9 and the model and two between the model and H10). No additional disulfide bonds are shown since no prediction can be made for the connectivity of these and the other cysteine residues on the extracellular loops.

we examined two recently solved crystal structures of transporters (Hunte et al., 2005; Yamashita et al., 2005), each containing 12 TM helices. In contrast to the original high scores ($>99\%$ confidence) for the MFS, our FUGUE analysis reported nonsignificant scores ($<50\%$ confidence), concluding that there was no evidence for the OATP/Oatp sequences adopting either of these new structures. Taken together, all the evidence strongly suggests that the relationship between OATPs/Oatps and the MFS is not simply due to sharing 12 TM helices. Our structural models have been validated further using several independent methods (*see below*).

The comparative modeling of OATP1B3 and OATP2B1 revealed an overall fold highly similar to that of MFS, with a pseudo twofold axis around a central pore. The presence of a central pore is likely to be a conserved feature among OATP/Oatps and functionally significant for the following reasons. (1) On the basis of the criteria such as the conservation of the TM helices, the sequence-structure alignments between OATP1B3/OATP2B1 and the two templates are reliable near the pore region. (2) The model qualities are good around the putative pore region according to PROSA2003 (Sippl, 1993) and VERIFY3D (Luthy et al., 1992). (3) The electrostatic potential of both OATP1B3 and OATP2B1 models is positive within the pore region (Fig. 5), consistent with the anionic nature of the majority of OATP/Oatp substrates. (4) The size of the pore in the models is large enough to accommodate typical substrates of OATPs/Oatps. This is illustrated in Figure 4d, which shows the OATP1B3 substrate digoxin docked into the pore, thereby confirming that digoxin easily fits

into the pore of the OATP1B3 model. (Currently, no experimental data exist to verify the details of the docked model.) These observations suggest that the positive pore is a common feature of all OATPs/Oatps.

Both the glycerol-3-phosphate transporter and the lactose permease are proposed to contain a single substrate-binding site and to operate by an alternating-access mechanism through a rocker-switch type of movement (Abramson et al., 2003; Huang et al., 2003). In this mechanism, the transport protein is believed to have two major alternating conformations: inward-facing and outward-facing. At any moment, a single binding site in a pore is accessible from only one side of the membrane. Interconversion between the two conformations is only possible via a substrate-bound form of the transport protein. Based on the structural models of OATP1B3 and OATP2B1, we propose that OATPs/Oatps might operate by a similar rocker-switch type of transport mechanism as reported for the glycerol-3-phosphate transporter and the lactose permease. The functional significance of this proposal will be further discussed below.

Potential Substrate-Binding Site

Mapping the conserved residues onto the structural models identified a single residue unique to the OATP1 family (Arg^{181}) and another to the OATP2 family (His^{579}), both of which face the pore and are likely to contribute to the positive potential. Because many OATP families, including OATP1 and OATP2, have partially overlapping substrate specificities, some active site residues are expected to be conserved across more than one family. However, OATP1 and OATP2 family members also show family-specific substrate transport (Hagenbuch & Meier, 2003). Since Arg^{181} and His^{579} are similarly distant from both sides of the membrane, which is consistent with the idea of a single binding site that is alternatively accessible from each side of the membrane, it can be hypothesized that Arg^{181} contributes to the substrate-binding site of OATP1 family members and His^{579} contributes to the substrate-binding site of OATP2 family members. Mutational analyses of Arg^{181} in OATP1 or His^{579} in OATP2 would provide further information and help test this hypothesis. Because most of the extracellular and intracellular regions have been excised from the structural models, no predictions can be made from these models on possible extracellular or intracellular docking sites for the substrates.

Impact of Single-Nucleotide Polymorphisms on the Proposed Transport Mode and Model Evaluation

Further evidence for the proposed transport mode of substrate translocation through a central pore comes from studies on polymorphisms and site-directed mutagenesis. Three polymorphisms and one artificial mutation have been identified and investigated in OATP1B3 (Letschert, Keppler & Konig, 2004). Substrates like bromosulfophthalein and estrone sulfate are transported by all mutated OATP1B3 proteins with comparable activities and, as shown for bromosulfophthalein, with comparable K_m values. With cholyltaurine as substrate, however, the polymorphism G522C and the artificial mutation G583E show no significant transport activity compared to the reference OATP1B3 protein. These mutations thus abolish transport function for some substrates (Letschert et al., 2004). Whereas the polymorphism G522C cannot be mapped onto the model of OATP1B3, the artificial mutation G583E can. According to the OATP1B3 model, G583 is buried and does not face the pore, so it is unlikely to be directly involved in substrate binding. It may be involved in a helix-stabilizing interaction that is important for the translocation of certain substrates.

Tirona et al. (2001) identified several polymorphisms in the OATP1 family member OATP1B1, three of which can be mapped onto the model of OATP1B3. They include F73L, I353T and V174A. For F73L and for I353T, Tirona et al. (2001) measured an 11-fold increase and a fourfold increase, respectively, in the K_m of estrone sulfate transport. V_{max} for the transport of estrone sulfate remained comparable between both mutants and the reference OATP1B1 protein. The altered K_m values in the presence of consistent V_{max} values suggest that these residues may be directly involved in the process of substrate recognition and substrate binding. Mapping these polymorphisms onto the model of OATP1B3 shows that the amino acid residues in positions 73 and 353 of OATP1B1 face the pore. According to Tirona et al. (2001), OATP1B1-V174A transports estrone sulfate with a K_m similar to the one observed for the reference OATP1B1 protein but with a decreased V_{max}. The latter may be due to reduced cell surface expression of the mutant proteins. No decrease in V_{max} was observed when Nozawa et al. (2002) repeated the experiment in HEK293 cells rather than in HeLa cells. Hence, the amino acid residue in position 174 of OATP1B1 appears to be less relevant to substrate transport, and it seems unlikely that this position is involved in ligand interactions and/or helix orientations. If mapped onto the OATP1B3 model, position 174 in OATP1B1 faces the membrane. The experimental observations are thus consistent with the proposed model and transport mechanism of OATP1B3.

Polymorphisms have been also reported for OATP2B1 (Nozawa et al., 2002). Of the two polymorphisms identified, one leads to a decrease in V_{max} for estrone sulfate transport to 70% of the value measured for the reference OATP2B1 protein, while

the K_m value is maintained. Thereby, membrane sorting is comparable to standard OATP2B1; thus, the residue in question (T392I) appears to be important for the process of substrate translocation. Such a role is consistent with the predicted pore-facing position of residue 392 within our model.

Studies aimed at explaining the different transport properties of rOatp1a4 and rOatp1a1 involve the site-directed mutagenesis of rOatp1a4 (Haenni et al., 2005). Six nonconserved sites have been mutated individually, and the effects on transport efficiency have been investigated. Three mutations, F328V, S334K and M361S, result in marked decreases in the amount of substrate taken up over a given time. Thereby, increases in K_m values of up to 45-fold have been measured. A fourth mutation, V354I, also leads to a reduction in transport efficiency, though to a less pronounced extent. The remaining two mutations, A330G and V353A, have no effect on the transport efficiency of rOatp1a4. The four sites whose mutation affects transport efficiency (F328V, S334K, V354I, M361S) face the pore in the models and may thus be expected to affect transport activity directly. Moreover, the two sites which have no impact on transport efficiency if mutated (A330G, V353A) are buried in the models and are thus expected not to affect transport activity directly. Hence, the models of OATPs/Oatps and the proposed mode of substrate translocation through a central pore correspond well with the results of the transport analysis for the mutated proteins.

Structural Model of the Large Extracellular Region

Several lines of evidence suggest a crucial role of the large extracellular region in the transport activity of OATPs/Oatps (Tirona et al., 2001; Pizzagalli et al., 2003). Given the absence of very close homologues in the structural database, the large extracellular region was selected for a separate structural modeling attempt.

A novel homology between this region of OATPs/Oatps and the Kazal-type serine protease inhibitors has been detected. This prediction was further reinforced by the high quality of the comparative model built on the basis of the structure of porcine pancreatic trypsin inhibitor (PDB 1tgs; Bolognesi et al., 1982).

The Kazal-type serine protease inhibitors are one of several serine protease inhibitor families. They are characterized by common structural features, including the characteristic cysteine distribution patterns and highly conserved 3D structures. They play a critical role in the precise regulation of several serine proteases and affect many physiological events, such as blood coagulation, development, food digestion, inflammation and immune response. More than 100 Kazal-type inhibitors have been identified in vertebrates, arthropods, nematodes and bacteria. In vertebrates, the Kazal-type inhibitors are mainly found in blood plasma, saliva, pancreatic secretions and egg whites (Zhu et al., 2005). To our knowledge, however, our observation is the first example of this structure found as a domain in an integral membrane protein. The precise function of this domain in OATPs/Oatas is not known.

Figure 7d shows a hypothetical representation of the whole OATP1B3 protein relative to the lipid bilayer. Thereby, putative disulfide bonds are also displayed. It was shown using *N*-biotinoylaminoethyl methanethiosulfonate (MTSEA-biotin) that all cysteine residues in the extracellular loops of OATP2B1 are disulfide-bonded (Haenggi et al., 2005). Three of the eight putative disulfide bonds in the large extraceullar region are predicted from the model based on the Kazal-type serine protease inhibitors. They are displayed in red. The remaining disulfide bonds, are not shown since no prediction can be made.

As illustrated in Figure 7a, only part of the large extracellular region can be modeled on the basis of the Kazal-type serine protease inhibitors, leaving the conformation of both ends of the modeled region undetermined. Hence, the exact orientation of the modeled part of the large extracellular region with respect to the TM helix model cannot be resolved, and the arrangement shown in Figure 7d is speculative.

Two of the 14 polymorphisms identified and analyzed by Tirona et al. (2001) can be mapped onto the partial model of the large extracellular region. Tirona et al. (2001) report reductions in V_{max} values and/or increased K_m values for these single-nucleotide polymorphisms. A similar reduction in V_{max} (to less than 50% for estrone sulfate transport) was observed for the second polymorphism in OATP2B1 identified by Nozawa et al. (2002). Such findings are in distinct contrast to single-nucleotide polymorphisms located in other extracellular loops, which are associated with minimal or very modest changes in the observed transport activity. Tirona et al. (2001) conclude that the large extracellular region is likely to play an important role in substrate transport. Given such an involvement of the large extracellular region in substrate transport, together with its mostly negative electrostatic potential (Fig. 7c) and its approximate location with respect to the positively charged pore (Fig. 7d), it seems plausible that the large extracellular region may cover the pore in the absence of substrates, moving away upon substrate binding. However, until direct experimental evidence is obtained, this proposition remains highly speculative.

Conclusion

This study presented for the first time 3D structural models of OATPs/Oatps that led to a hypothesis for

the mode of transport for OATP/*SLCO* superfamily members. Direct suggestions for site-directed mutagenesis and other laboratory-based experimentation were made to test the proposed rocker-switch type of mechanism of substrate transport. The identification of novel members of the OATP/*SLCO* superfamily established its ancient origin in animal evolution. Given the central role of OATPs/Oatps in detoxification processes, their conservation throughout the animal kingdom may mirror the fact that animals can only survive if they have a mechanism to eliminate endogenous and exogenous toxins from their body.

We thank Dr. Bruno Hagenbuch (Department of Pharmacology, Toxicology and Therapeutics, University of Kansas Medical Center) and Dr. Bruno Stieger (Department of Clinical Pharmacology and Toxicology, University Hospital Zurich) for their advice and provision of unpublished data. Y. M. is supported by a scholarship from the Algerian Ministry of Higher Education and Scientific Research.

References

Abramson, J., Smirnova, I., Kasho, V., Verner, G., Kaback, H.R., Iwata, S. 2003. Structure and mechanism of the lactose permease of *Escherichia coli*. *Science*. **301**:610–615

Altschul, S.F., Madden, T.L., Schaffer, A.A., Zhang, J., Zhang, Z., Miller, W., Lipman, D.J. 1997. Gapped BLAST and PSI-BLAST: A new generation of protein database search programs. *Nucleic Acids Res*. **25**:3389–3402

Bairoch, A., Apweiler, R., Wu, C.H., Barker, W.C., Boeckmann, B., Ferro, S., Gasteiger, E., Huang, H., Lopez, R., Magrane, M., Martin, M.J., Natale, D.A., O'Donovan, C., Redaschi, N., Yeh, L.S. 2005. The universal protein resource (UniProt). *Nucleic Acids Res*. **33**:D154–D159

Baker, N.A., Sept, D., Joseph, S., Holst, M.J., McCammon, J.A. 2001. Electrostatics of nanosystems: Application to microtubules and the ribosome. *Proc. Natl. Acad. Sci. U.S.A.* **98**:10037–10041

Bateman, A., Coin, L., Durbin, R., Finn, R.D., Hollich, V., Griffiths-Jones, S., Khanna, A., Marshall, M., Moxon, S., Sonnhammer, E.L.L., Studholme, D.J., Yeats, C., Eddy, S.R. 2004. The Pfam protein families database. *Nucleic Acids Res*. **32**:D138–D141

Bolognesi, M., Gatti, G., Menegatti, E., Guarneri, M., Marquart, M., Papamokos, E., Huber, R. 1982. Three-dimensional structure of the complex between pancreatic secretory inhibitor (Kazal type) and trypsinogen at 1.8 angstroms resolution. Structure solution, crystallographic refinement and preliminary structural interpretation. *J. Mol. Biol*. **162**:839–868

Catherinot, V., Labesse G. 2004. ViTO:tool for refinement of protein sequence-structure alignments. *Bioinformatics* **20:**3694–3696

Chang, A.B., Lin, R., Studley, W.K., Tran, C.V., Milton, H.S. 2004. Phylogeny as a guide to structure and function of membrane transport proteins. *Mol. Membr. Biol*. **21**:171–181

DeLano, W.L. 2002. The PyMOL molecular graphics system, http://www.pymol.org

Eddy, S.R. 1998. HMMER: Profile hidden Markov models. *Bioinformatics* **14**:755–763

Fattinger, K., Cattori, V., Hagenbuch, B., Meier, P.J., Stieger, B. 2000. Rifamycin SV and rifamycin exhibit differential inhibition of the hepatic rat organic anion transporting polypeptides, Oatp1 and Oatp2. *Hepatology* **32**:82–86

Felsenstein, J. 1989. PHYLIP–phylogeny inference package (version 3.2). *Cladistics* **5**:164–166

Fischer, W.J., Altheimer, S., Cattori, V., Meier, P.J., Dietrich, D.R., Hagenbuch, B. 2005. Organic anion transporting polypeptides expressed in liver and brain mediate uptake of microcystin. *Toxicol. Appl. Pharmacol*. **203**:257–263

Fiser, A., Sali, A. 2003. MODELLER: Generation and refinement of homology-based protein structure models. *Methods Enzymol*. **374**:461–491

Ginalski, K., Elofsson, A., Fischer, D., Rychlewski, L. 2003. 3D-jury: A simple approach to improve protein structure predictions. *Bioinformatics* **19**:1015–1018

Hagenbuch, B., Gao, B., Meier, P.J. 2002. Transport of xenobiotics across the blood-brain barrier. *News Physiol. Sci*. **17**:231–234

Hagenbuch, B., Meier, P.J. 2003. The superfamily of organic anion transporting polypeptides. *Biochim. Biophys. Acta* **1609**:1–18

Hagenbuch, B., Meier, P.J. 2004. Organic anion transporting polypeptides of the OATP/SLC21 family: Phylogenetic classification as OATP/SLCO superfamily, new nomenclature and molecular/functional properties. *Pfluegers Arch*. **447**:653–665

Huang, Y., Lemieux, M.J., Song, J., Auer, M., Wang, D. 2003. Structure and mechanism of the glycerol-3-phosphate transporter from *Escherichia coli*. *Science* **301**:616–620

Hunte, C., Screpanti, E., Venturi, M., Rimon, A., Padan, E., Michel, H. 2005. Structure of a Na^+/H^+ antiporter and insights into mechanism of action and regulation by pH. *Nature* **435**:1197–1202

Ismair, M.G., Stieger, B., Cattori, V., Hagenbuch, B., Fried, M., Meier, P.J., Kullak-Ublick, G.A. 2001. Hepatic uptake of cholecystokinin octapeptide (CCK-8) by organic anion transporting polypeptides Oatp4 (Slc21a6) and OATP8 (SLC21A8) from rat and human liver. *Gastroenterology* **121**:1185–1190

Jacquemin E., Hagenbuch B., Stieger B., Wolkoff A.W., Meier P.J. 1994. Expression Cloning of a Rat Liver Na^+-Independent Organic Anion Transporter. *Proc. Natl. Acad. Sci. U.S.A.* 91:133–137

Jaroszewski, L., Rychlewski, L., Li, Z., Li, W., Godzik, A. 2005. FFAS03: A server for profile-profile sequence alignments. *Nucleic Acids Res*. **33**:W284–W288

Katoh, K., Misawa, K., Kuma, K., Miyata, T. 2002. MAFFT: A novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Res*. **30**:3059–3066

Krogh, A., Larsson, B., von Heijne, G. von , Sonnhammer, E.L.L. 2001. Predicting transmembrane protein topology with a hidden Markov model: Application to complete genomes. *J. Mol. Biol*. **305**:567–580

Kullak-Ublick, G.A., Hagenbuch, B., Stieger, B., Schteingart, C.D., Hoffmann, A.F., Wolkoff, A.W., Meier, P.J. 1995. Molecular and functional characterization of an organic anion transporting polypeptide cloned from human liver. *Gastroenterology* **109**:1274–1282

Kullak-Ublick, G.A., Stieger, B., Meier, P.J. 2004. Enterohepatic bile salt transporters in normal physiology and in liver disease. *Gastroenterology* **126**:322–342

Letschert, K., Keppler, D., Konig, J. 2004. Mutations in the SLCO1B3 gene affecting the substrate specificity of the hepatocellular uptake transporter OATP1B3 (OATP8). *Pharmacogenetics* **14**:441–452

Lichtarge, O., Bourne, H.R., Cohen, F.E. 1996. An Evolutionary Trace Method Defines Binding Surfaces Common to Protein Families. *J. Mol. Biol*. 257:342–358

Luthy, R., Bowie, J.U., Eisenberg, D. 1992. Assessment of protein models with three-dimensional profiles. *Nature* **356**:83–85

Maiden, M.C., Davis, E.O., Baldwin, S.A., Moore, D.C., Henderson, P.J. 1987. Mammalian and bacterial sugar transport proteins are homologous. Nature 325:641

Meier, P.J., Stieger, B. 2002. Bile salt transporters. *Annu. Rev. Physiol.* **64**:635–661

Meier-Abt, F., Faulstich, H., Hagenbuch, B. 2004. Identification of phalloidin uptake systems of rat and human liver. *Biochim. Biophys. Acta* **1664**:64–69

Mizuguchi, K., Deane, C.M., Blundell, T.L., Johnson, M.S., Overington, J.P. 1998a. JOY: Protein sequence-structure representation and analysis. *Bioinformatics* **14**:617–623

Mizuguchi, K., Deane, C.M., Blundell, T.L., Overington, J.P. 1998b. HOMSTRAD: A database of protein structure alignments for homologous families. *Protein Sci.* **7**:2469–2471

Moult, J., Fidelis, K., Zemla, A., Hubbard, T. 2003. Critical assessment of methods of protein structure prediction (CASP)-round V. *Proteins* **53**(Suppl. 6):334–339

Nozawa, T., Nakajima, M., Tamai, I., Noda, K., Nezu, J.-I., Sai, Y., Tsuji, A., Yokoi, T. 2002. Genetic polymorphisms of human organic anion transporters OATP-C (SLC21A6) and OATP-B (SLC21A9): Allele frequencies in the Japanese population and functional analysis. *J. Pharmacol. Exp. Ther.* **302**:804–813

Pizzagalli, F., Varga, Z., Huber, R.D., Folkers, G., Meier, P.J., St. Pierre, M.V. 2003. Identification of steroid sulfate transport processes in the human mammary gland. *J. Clin. Endocrinol. Metab.* **88**:3902–3912

Rose, D.R., Seaton, B.A., Petsko, G.A., Novotný, J., Margolies, M.N., Locke, E., Haber, E. 1983. Crystallization of the Fab fragment of a monoclonal anti-digoxin antibody and its complex with digoxin. *J. Mol. Biol.* **165**:203–206

Sali, A., Blundell, T.L. 1990. Definition of general topological equivalence in protein structures. A procedure involving comparison of properties and relationships through simulated annealing and dynamic programming. *J. Mol. Biol.* **212**:403–428

Sali, A., Blundell, T.L. 1993. Comparative protein modeling by satisfaction of spatial restraints. *J. Mol. Biol.* **234**:779–815

Sayle, R.E., Milner-White J. 2005. RasMol: Biomolecular graphics for all. *Trend in Biochemical Sciences (TIBS)* **20**:374.

Shi, J., Blundell, T.L., Mizuguchi, K. 2001. FUGUE: Sequence-structure homology recognition using environment-specific substitution tables and structure-dependent gap penalties. *J. Mol. Biol.* **310**:243–257

Sippl, M.J. 1993. Recognition of errors in three-dimensional structures of proteins. *Proteins* **17**:355–362

St. Pierre, M.V., Hagenbuch, B., Ugele, B., Meier, P.J., Stallmach, T. 2002. Characterization of an organic anion transporting polypeptide (OATP-B) in human placenta. *J. Clin. Endocrinol. Metab.* **87**:1856–1863

St. Pierre, M.V., Kullak-Ublick, G.A., Hagenbuch, B., Meier, P.J. 2001. Bile acid transport in hepatic and non-hepatic tissues. *J. Exp. Biol.* **204**:1673–1686

Tamai, I., Nezu, J., Uchino, H., Sai, Y., Oku, A., Shimane, M., Tsuji, A. 2000. Molecular identification and characterization of novel members of the human organic anion transporter (OATP) family. *Biochem. Biophys. Res. Commun.* **273**:251–260

Taylor, W.R. 1986. The classification of amino acid conservation. *J. Theor. Biol.* **119**:205–218

Tirona, R.G., Kim, R.B. 2002. Pharmacogenomics of organic anion transporting polypeptides (OATPs). *Arch. Drug Deliv. Rev.* **54**:1343–1352

Tirona, R.G., Leake, B.F., Merino, G., Kim, R.B. 2001. Polymorphisms in OATP-C: Identification of multiple allelic variants associated with altered transport activity among European- and African-Americans. *J. Biol. Chem.* **276**:35669–35675

Vardy, E., Arkin, I.T., Gottschalk, K.E., Kaback, H.R., Schuldiner, S. 2004. Structural conservation in the major facilitator superfamily as revealed by comparative modeling. *Protein Sci.* **13**:1832–1840

Verdonk, M.L., Cole, J.C., Hartshorn, M.J., Murray, C.W., Taylor, R.D. 2003. Improved protein-ligand docking using GOLD. *Proteins* **52**:609–623

Yamashita, A., Singh, S.K., Kawate, T., Jin, Y., Gouaux, E. 2005. Crystal structure of a bacterial homologue of Na^+/Cl^--dependent neurotransmitter transporters. *Nature* **437**:215–223

Zhu, L., Song, L., Chang, Y., Xu, W., Wu, L. 2005. Molecular cloning, characterization and expression of a novel serine proteinase inhibitor gene in bay scallops. *Fish Shellfish Immunol.* in press