Biochemical and Biophysical Research Communications **274,** 79–86 (2000)
doi:10.1006/bbrc.2000.3105, available online at http://www.idealibrary.com on IDEAL®

Cloning of the Full-Length Coding Sequence of Rat Liver-Specific Organic Anion Transporter-1 (rlst-1) and a Splice Variant and Partial Characterization of the Rat lst-1 Gene

Supratim Choudhuri, Kenichiro Ogura,[1] and Curtis D. Klaassen
Department of Pharmacology, Toxicology and Therapeutics, University of Kansas Medical Center, Kansas City, Kansas 66160-7417

Received June 15, 2000

The full-length coding sequence of rat liver-specific organic anion transporter-1 (lst-1) and its splice variant have been cloned. The full-length rat lst-1 (designated rlst-1a) encodes a protein containing 687 amino acids and has 12-putative transmembrane domains, multiple potential N-glycosylation and phosphorylation sites. Therefore, rat lst-1a has 35 additional amino acid residues compared to the previously reported rat lst-1. A splice variant (designated rlst-1c) reported in this communication encodes a protein containing 654 amino acids and has 10-putative transmembrane domains. PCR analysis suggests that rlst-1a is the most abundant form in liver. Phylogenetic analysis reveals that rat lst-1a is an ortholog of human LST-1 (hLST-1) and mouse lst-1 (mlst-1). The rlst-1 gene is composed of 15 exons and 14 introns. Analysis of exon–intron boundary reveals that the splice variant rlst-1c lacks the entire exon 7, while the previously reported rat lst-1 (designated herein as rlst-1b) lacks approximately half of exon 10, and the splicing has occurred through alternative usage of a splice donor site within exon 10.

Hepatic uptake of organic anions and bile acids can be mediated by both Na^+-dependent and Na^+-independent transport systems. For example, uptake of taurocholate is mediated preferentially by a Na^+-dependent transport system, known as the Na^+/taurocholate cotransporting polypeptide (ntcp) (1), but taurocholate is also transported by Na^+-independent transport systems. However, other bile acids (such as cholate and chenodeoxycholate) are mainly transported by Na^+-independent transport systems (2, 3).

At least three families of hepatic Na^+-independent organic anion transporters have been reported: organic anion transporting polypeptides (oatps) (4–6), prostaglandin transporter (PGT) (7, 8) and the most recently described liver-specific organic anion transporter (lst) (9, 10). The human liver-specific organic anion transporter-1 (hLST-1) (9) has also been termed human OATP2 (11, 12). Rat liver-specific organic anion transporter (rlst-1) has recently been cloned and shown to be expressed exclusively in liver (10). Its expression can be down-regulated by bile-duct ligation and also in sepsis. This suggests a negative feedback protective mechanism through lst-1 down-regulation.

Surprisingly, the coding sequence of rat lst-1 (rlst-1) cDNA reported by Kakyo *et al.* (10), was found to lack 105 nucleotides compared to human LST-1 cDNA (9) and mouse lst-1 cDNA recently cloned in our laboratory (13). In their report, Kakyo *et al.* (10) compared the rat lst-1 cDNA with human LST-1 cDNA and proposed that the lack of 105 nucleotides that encode 35 amino acids, results in a protein with 11 rather than 12 transmembrane domains. Based on the comparison of rat lst-1 with human LST-1 and mouse lst-1, we reasoned that the reported rat lst-1 cDNA (10) was possibly derived from an alternatively spliced form of rat lst-1 mRNA.

Thus, the current study was undertaken in an attempt to determine whether the rat has a full-length lst-1 comparable to the human and mouse lsts. In this communication, we report the cloning and tissue-specific expression of the full-length coding sequence of rat lst-1 and its splice variant. The relative abundance of the full-length form, the splice variant reported here, and the other splice variant reported earlier (10) has also been studied. In addition, we have partially characterized the rat lst-1 gene in order to understand the origin of the splice variants.

Full-length sequence information have been submitted to the GenBank nucleotide sequence database under the Accession Nos. AF208545, AF217450, and from AF272557 to AF272571.

[1] Present address: Department of Drug Metabolism and Molecular Toxicology, Tokyo University of Pharmacy and Life Science, 1432-1 Horinouchi, Hachioji, Tokyo 192-0392, Japan.

0006-291X/00 $35.00

```
rlst-1a   MDHTQQSRKAAEAQPSRSKQTRFCDGFKLFLAALSFSYICKALGGVVMKSSITQIERRFDIPSSISGLIDGGFEIGNLLVIVFVSYFGSKLHRPKLIGIGCFIMGIGSILTALPHFFMGY
rlst-1c   MDHTQQSRKAAEAQPSRSKQTRFCDGFKLFLAALSFSYICKALGGVVMKSSITQIERRFDIPSSISGLIDGGFEIGNLLVIVFVSYFGSKLHRPKLIGIGCFIMGIGSILTALPHFFMGY
rlst-1b   MDHTQQSRKAAEAQPSRSKQTRFCDGFKLFLAALSFSYICKALGGVVMKSSITQIERRFDIPSSISGLIDGGFEIGNLLVIVFVSYFGSKLHRPKLIGIGCFIMGIGSILTALPHFFMGY
hLST-1    MDQNQHLNKTAEAQPSENKKTRYCNGLKMFLAALSLSFIAKTLGAIIMKSSIIHIERRFEISSSLVGFIDGSFEIGNLLVIVFVSYFGSKLHRPKLIGIGCFIMGIGGVLTALPHFFMGY
          **:.*: .*:******..*:**:*:*:*:******:*:*.*:**.::***** :*****:*.**: *:***.***********************************.:**********

rlst-1a   YKYAKENDIGSLGNST---LTCFINQMTSPTGPSPEIVEKGCEKGLKSHMWIYVLMGNMLRGIGETPIVPLGISYLDDFAKEGHTSMHLGTLHTIAMIGPILGFIMSSVFAKIYVDVGYV
rlst-1c   YKYAKENDIGSLGNST---LTCFINQMTSPTGPSPEIVEKGCEKGLKSHMWIYVLMGNMLRGIGETPIVPLGISYLDDFAKEGHTSMHL-------------------------------
rlst-1b   YKYAKENDIGSLGNST---LTCFINQMTSPTGPSPEIVEKGCEKGLKSHMWIYVLMGNMLRGIGETPIVPLGISYLDDFAKEGHTSMHLGTLHTIAMIGPILGFIMSSVFAKIYVDVGYV
hLST-1    YRYSKETNINSSENSTSTLSTCLINQILSLNKASPEIVGKGCLKESGSYMWIYVFMGNMLRGIGETPIVPLGLSYIDDFAKEGHSSLYLGILNAIAMIGPIIGFTLGSLFSKMYVDIGYV
          *:*:**.:*.*  ***     **:***: * . .***** *** *   *:*****:****************:**:********:*::*

rlst-1a   DLNSVRITPNDARWVGAWWLSFIVNGLLCITSSIPFFFLPKIPKRSQEERKNSVSLHAPKTDEEKKHMTNLTKQEEQDPSNMTGFLRSLRSILTNEIYVIFLILTLLQVSGFIGSFTYLF
rlst-1c   --DSVRITPNDARWVGAWWLSFIVNGLLCITSSIPFFFLPKIPKRSQEERKNSVSLHAPKTDEEKKHMTNLTKQEEQDPSNMTGFLRSLRSILTNEIYVIFLILTLLQVSGFIGSFTYLF
rlst-1b   DLNSVRITPNDARWVGAWWLSFIVNGLLCITSSIPFFFLPKIPKRSQEERKNSVSLHAPKTDEEKKHMTNLTKQEEQDPSNMTGFLRSLRSILTNEIYVIFLILTLLQVSGFIGSFTYLF
hLST-1    NLSTIRITPTDSRWVGAWWLNFLVSGLFSIISSIPFFFLPQTPNKPQKERKASLSLHVLETNDEKDQTANLTNQGKNITKNVTGFFQSFKSILTNPLYVMFVLLTLLQVSSYIGAFTYVF
            .::****.*:********.*:*.**:.* *********: *::.*:*** *:***. :*::**.: :***:* :: ..*:***::*::***** :**:*::*******.:**:***:*

rlst-1a   KFIEQQFGRTASQANFLLGIITIPTMATAMFLGGYIVKKFKLTSVGIAKFVFFTSSVAYAFQFLYFPLLCENKPFAGLTLTYDGMNPVDSHIDVPLSYCNSDCSCDKNQWEPICGENGVT
rlst-1c   KFIEQQFGRTASQANFLLGIITIPTMATAMFLGGYIVKKFKLTSVGIAKFVFFTSSVAYAFQFLYFPLLCENKPFAGLTLTYDGMNPVDSHIDVPLSYCNSDCSCDKNQWEPICGENGVT
rlst-1b   KFIEQQFGRTASQANFLLGIITIPTMATAMFLGGYIVKKFKLTSVGIAK-------------------------------------MNPVDSHIDVPLSYCNSDCSCDKNQWEPICGENGVT
hLST-1    KYVEQQYGQPSSKANILLGVITIPIFASGMFLGGYIIKKFKLNTVGIAKFSCFTAVMSLSFYLLYFFILCENKSVAGLTMTYDGNNPVTSHRDVPLSYCNSDCNCDESQWEPVCGNNGIT
          *::***:*:.:*:**:***:**** :*:.*******:*****.:*****                                      *** ** ***********.**:.****:**:**:*

rlst-1a   YISPCLAGCKSFRGDKKPNNTEFYDCSCISNSG----NNSAHLGECPR-YKCKTNYYFYIILQVTVSFFTAMGSPSLILILMKSVQPELKSLAMGFHSLIIRALGGILAPIYYGAFIDRT
rlst-1c   YISPCLAGCKSFRGDKKPNNTEFYDCSCISNSG----NNSAHLGECPR-YKCKTNYYFYIILQVTVSFFTAMGSPSLILILMKSVQPELKSLAMGFHSLIIRALGGILAPIYYGAFIDRT
rlst-1b   YISPCLAGCKSFRGDKKPNNTEFYDCSCISNSG----NNSAHLGECPR-YKCKTNYYFYIILQVTVSFFTAMGSPSLILILMKSVQPELKSLAMGFHSLIIRALGGILAPIYYGAFIDRT
hLST-1    YISPCLAGCKSSSGNKKP--IVFYNCSCLEVTGLQNRNYSAHLGECPRDDACTRKFYFFVAIQVLNLFFSALGGTSHVMLIVKIVQPELKSLALGFHSMVIRALGGILAPIYFGALIDTT
          *********** *:***    **:***:. :*    * *********   *. ::**:: :**   **:*:*..* :::::* ********:****::***********:**:** *

rlst-1a   CIKWSVTSCGKRGACRLYNSRLFGFSYLGLNLALKTPPLFLYVVLIYFTKRKYKRNDNKTLENGRQFTDEGNPDSVNKNGYYCVPYDEQSNETPL
rlst-1c   CIKWSVTSCGKRGACRLYNSRLFGFSYLGLNLALKTPPLFLYVVLIYFTKRKYKRNDNKTLENGRQFTDEGNPDSVNKNGYYCVPYDEQSNETPL
rlst-1b   CIKWSVTSCGKRGACRLYNSRLFGFSYLGLNLALKTPPLFLYVVLIYFTKRKYKRNDNKTLENGRQFTDEGNPDSVNKNGYYCVPYDEQSNETPL
hLST-1    CIKWSTNNCGTRGSCRTYNSTSFSRVYLGLSSMLRVSSLVLYIILIYAMKKKYQEKDINASENG-SVMDEANLESLNKN-KHFVPSAGADSETHC
          *****...**.**:** *** *.  ****. *:...*.**::*** *:**:.:* :: *** .. **.* :*:***  : **    ..**
```

FIG. 1. Comparison of the amino acid sequence of human LST-1 (hLST-1) and all forms of rat lst-1, i.e., rlst-1a, rlst-1c and rlst-1b. The missing amino acids in rlst-1c and rlst-1b (compared to rlst-1a) are indicated by dashed lines. Asterisks indicate the conserved amino acids between all the sequences.

In the present communication, we propose that the full-length rat lst be called rat lst-1a (rlst-1a), and all others be assigned names based on their order of discovery. Hence, the splice variant reported by Kakyo *et al.* (10) be called rat lst-1b (rlst-1b), and the splice variant reported in this communication be called rat lst-1c (rlst-1c). In the text that follows, this nomenclature has been followed.

MATERIALS AND METHODS

Materials. Restriction enzymes were obtained from Gibco-BRL (Gaithersburg, MD). Terminal transferase was purchased from Roche Molecular (Indianapolis, IN). [α-^{32}P]dATP and Ready-To-Go DNA ligase were from Amersham Pharmacia Biotech (Piscataway, NJ). Zeta probe membrane was obtained from Bio-Rad (Hercules, CA).

cDNA cloning. cDNA cloning of rat lst-1a and lst-1c was performed by 5′- and 3′-RACE (Rapid Amplification of cDNA Ends) reactions using Sprague–Dawley (SD) rat liver-derived adaptor-ligated Marathon-Ready cDNA library (Clontech, Palo Alto, CA). 5′-RACE reactions using the primer 5′-CTGAAACGCATAGGCCACTGAGGAGGTGAA-3′ yielded a ~1.4 kb fragment, and 3′-RACE reactions using the primer 5′-ATGGGGAACATGCTTCGTGGGATAGGGGAA-3′ yielded a ~1.8 kb fragment. These fragments had a 745-bp overlap. A touchdown PCR was performed for RACE reactions with the following conditions: 1 cycle of 94°C × 30 s; 6 cycles of 94°C × 5 s, 72°C × 4 min; 6 cycles of 94°C × 5 s, 70°C × 4 min; 26 cycles of 94°C × 5 s, 68°C × 4 min; and a final extension of 72°C × 7 min. The RACE products were subcloned into pT-Adv TA cloning vector (Clontech, Palo Alto, CA), and sequenced with an ABI Prism 377 DNA sequencer (Perkin–Elmer, Foster, CA). To confirm the accuracy of the sequence, a total of three independent reactions were performed. The initial 5′- and 3′-RACE reactions yielded two composite sequences: a 2299-bp product that contains an open reading frame (ORF) of rat lst-1a cDNA encoding a protein of 687 amino acids, 100 bp of the 5′-untranslated region (5′-UTR) and 135 bp of the 3′-untranslated region (3′-UTR); and a 2200-bp product (rat lst-1c cDNA) that has a 99 bp deletion from nt 620 to nt 718 (relative to the first base of the start codon) compared to the rat lst-1a cDNA. Additional 5′- and 3′-RACE reactions were performed in an effort to obtain more regions from the 5′ and the 3′ end. Sequencing of the cDNAs was done by primer walking, and both strands were sequenced. Synthesis of all oligonucleotide primers and sequencing were carried out by the Biotech Support Facility, University of Kansas Medical Center.

Northern blot analysis of rat lst-1 mRNA. Total RNA was isolated from pooled liver, kidney, brain, heart, lungs and intestine from 11-week-old SD rats (3 rats for each tissue), using Trizol reagent following the manufacturer's direction. We used 11-week-old rats because their age matches the age of the SD rats (10–12 weeks old) from which the Marathon-Ready cDNA library was prepared. Poly(A)$^+$ mRNA from each tissue was purified from total RNA using poly(A)$^+$ pure mRNA isolation kit from Ambion (Austin, TX). 3 μg of poly(A)$^+$-enriched RNA was electrophoretically resolved in 1% formaldehyde-agarose gel, capillary-blotted overnight onto Zeta-probe nylon membrane, UV cross-linked and hybridized at 46°C overnight in 1× hybridization solution (Sigma, St. Louis, MO) containing 25% formamide. Antisense oligonucleotide probes were 3′-end labeled using terminal transferase and the average specific activity was about 3–4 × 10^8 cpm/μg DNA. More than one oligonucleotide probe was used to increase the sensitivity of detection. The probe sequences were as follows: Probe 1b$^-$ #1, 5′-ATAGGCCACTGAGGAGGTGAAAAATA-3′; Probe 1b$^-$ #2, 5′-AGCAAATGGCTTGTTTTCACAGAG-3′; Probe 1c$^-$ #1, 5′-GGTCCAATCATTGCTAT-

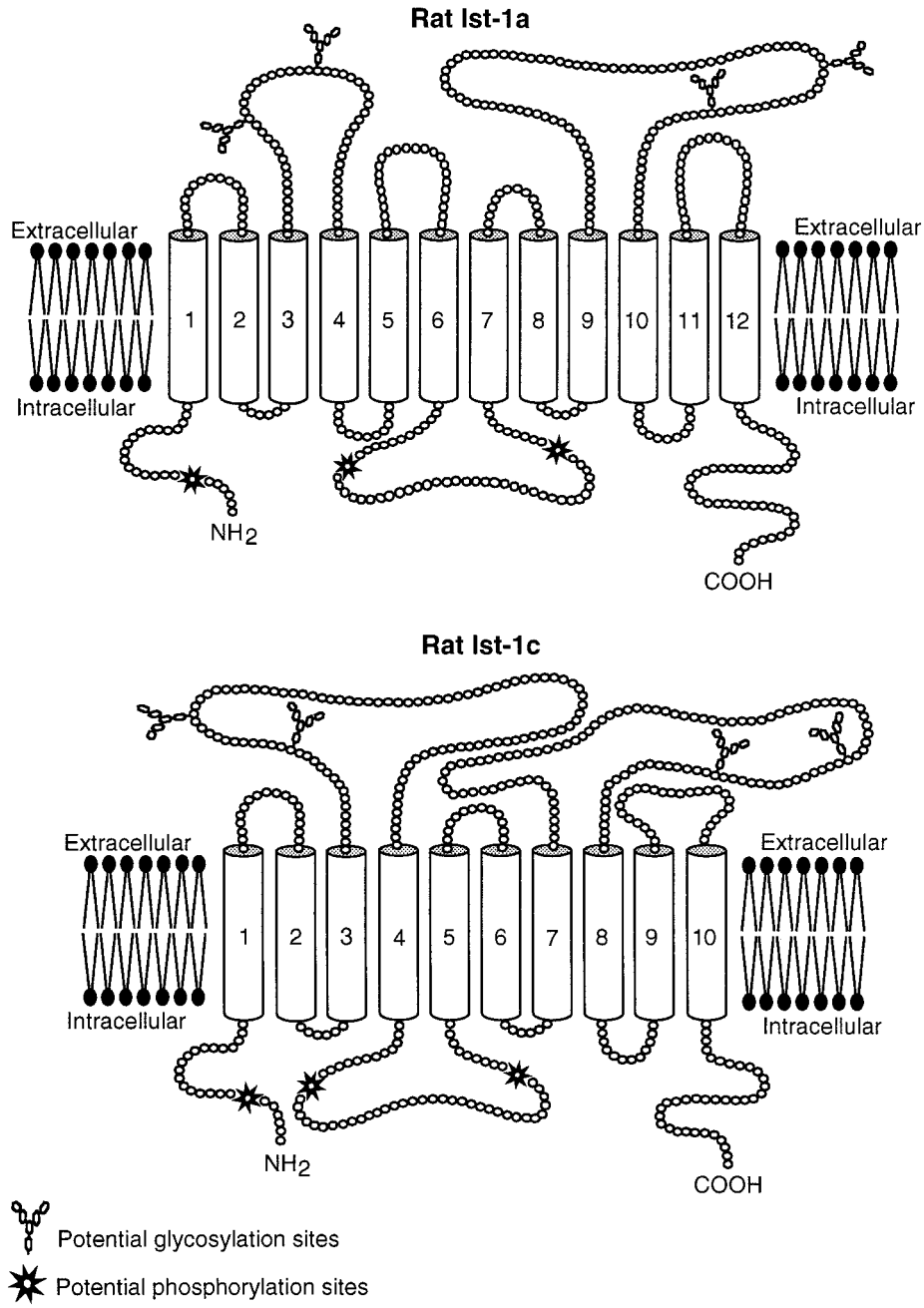


FIG. 2. Topology of rat lst-1a and -1c proteins. The transmembrane domains are indicated by cylinders. Each circle in the extracellular and intracellular loop represents one amino acid. Because of the loss of 2 transmembrane domains, rlst-1c appears to have an additional large extracellular loop domain compared to rlst-1a.

AGTATGCAGAGTAC-3′; Probe $1c^-$ #2, 5′-CAGGTCTACATATCCA-ACGTCTACATATATC-3′.

The above antisense oligonucleotide probes were designed from the regions of rat lst-1a that are missing from rat lst-1b and rat lst-1c, so that the probes hybridize with rat lst-1a mRNA and also with either rat lst-1b or rat lst-1c mRNA. Thus, $1b^-$ (#1 and #2) probes hybridize with rat lst-1a and -1c, but not with rat lst-1b because they have been chosen from the region that is missing from rat lst-1b. Similarly, $1c^-$ (#1 and #2) probes hybridize with rat lst-1a and rat lst-1b, but not with rat lst-1c.

PCR for determining the relative abundance of different forms of rat lst-1 (rlst-1a, -1b and -1c). To determine which form of rat lst-1 (rlst-1a, -1b or -1c) has the highest expression, PCRs were performed with the following conditions: 1 cycle of 94°C × 30 s; 35 cycles of 94°C × 10 s, 58°C × 1 min, 72°C × 1.5 min; and a final cycle of 72°C × 7 min. The primers used were as follows: Primer FP-1, 5′-GCACCTAGGTACTCTGCATACTATAGCAATGATTGG-3′; Primer FP-2, 5′-GCTCAGGGTAGACGACCATTTCAAAAGAAA-3′; Primer AP-1, 5′-CCATCCTAATACGACTCACTATAGGGC-3′; Primer RP-1, 5′-GGTACATCTATGTGAGAGTCCACTGGGTTCA-3′; Primer RP-2,

5′-CAGCAAATGGCTTGTTTTCACAGAGTAGAGG-3′; Primer RP-3, 5′-CAACCCAACGAGCATCATTAGGAGTTATTC-3′.

Characterization of rlst-1 gene. An SD rat genomic BAC library (BAC rat library) was screened with a 362-bp *Bam* H1 fragment containing the 5′-untranslated region (5′-UTR) of the rat lst-1 cDNA. The hybridization screening was performed by Genome Systems (St. Louis, MO). A positive BAC clone containing the entire rat lst gene was used for sequencing the exon-intron boundaries. Sequencing was done by genomic walking using primers designed from each newly obtained genomic sequence.

Sequence analysis. DNA sequence analysis, alignments, contig assembly, amino acid sequence prediction and construction of phylogenetic tree were done by a combination of DNAsis, Autoassembler and Clutal W software. Prediction of putative membrane topology was done with the aid of hidden Markov model for predicting transmembrane helices (http://www.cbs.dtu.dk/services/TMHMM-1.0).

RESULTS AND DISCUSSION

Full-Length Coding Region of Rat lst-1a Bears High Sequence Identity with Human and Mouse lst-1

The full-length form (rat lst-1a) encodes a protein containing 687 amino acids with a predicted molecular mass of approximately 77 kDa. Figure 1 shows amino acid sequence comparisons between human LST-1, and all three forms of rat lst-1 (rlst-1a, rlst-1b and rlst-1c). Gaps in rat lst-1b and -1c indicate the missing amino acids compared to rlst-1a. Interestingly, the Asn 240 in rlst-1a and -1b has been replaced by Asp (at the position immediately after the deletion) in rlst-1c. At the nucleotide level, the coding sequence of the full-length rat lst-1 (rlst-1a) is 73.5% and 88% identical to that of human LST-1 (hLST-1) (9) and mouse lst-1 (mlst-1), respectively (13).

Both Splice Variants (rlst-1c and rlst-1b) Have Different Predicted Topology Compared to the Full-Length Form (rlst-1a)

Prediction of possible topology of rat lst-1a and -1c was performed by hidden Markov model for predicting transmembrane helices (TMHMM program), and the results are presented in Fig. 2. Rat lst-1a (rlst-1a) has 12-putative transmembrane domains (TMDs). This is similar to other Na^+-independent organic anion transporters which all have 12 TMDs. The splice variant rlst-1c has a 99 bp deletion compared to rlst-1a (nt #620 through nt #718, relative to the first base of the start codon as #1). This results in a deletion of 33 amino acids (aa #207 through aa #239 compared to rlst-1a), and the loss of two TMDs (TMD #4 and #5, compared to rlst-1a). Because of the loss of two TMDs, rlst-1c possesses one additional large extracellular domain after TMD #3, compared to rlst-1a (Fig. 2). Rat lst-1b cDNA (10) has a 105 bp deletion compared to rlst-1a (nt #1218 through nt #1322, relative to the first base of the start codon as #1). This results in a deletion of 35 amino acids (aa #407 through aa #441 compared to rlst-1a), and the loss of one TMD (TMD #9 compared to rlst-1a).

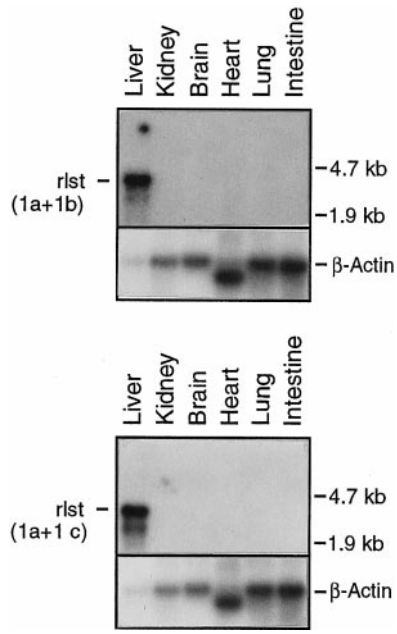
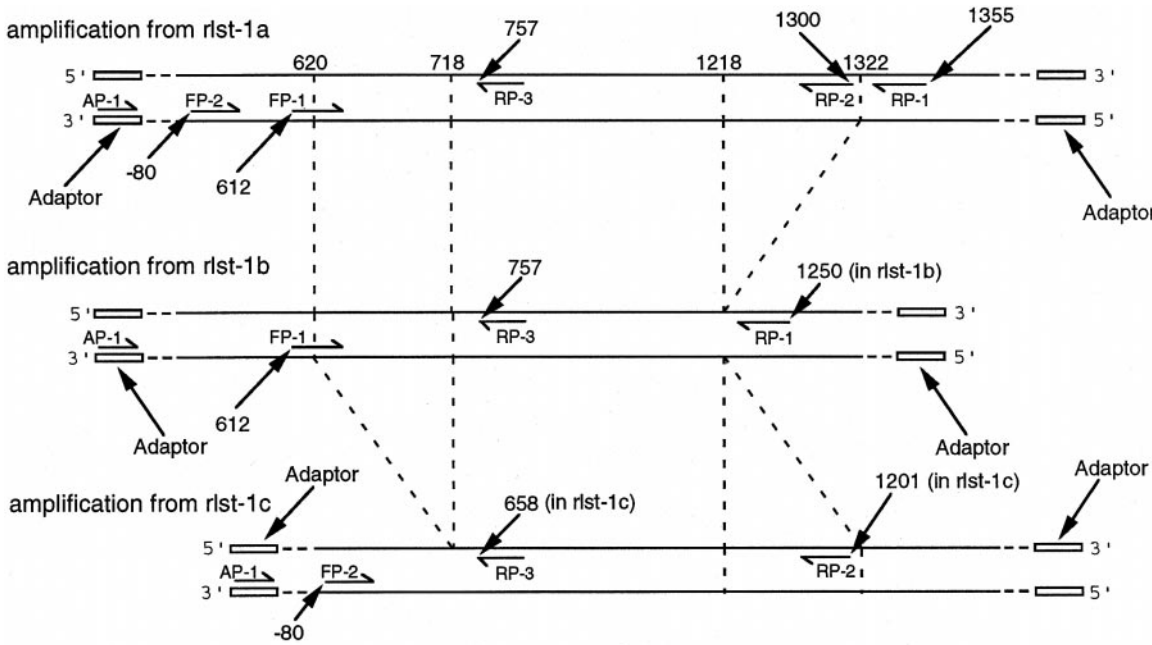


FIG. 3. Expression of rat lst-1a, -1b, and -1c in different tissues in rat. Upper panel shows Northern blot done with probe $1c^-$ (#1 and #2), and lower panel shows Northern blot done with probe $1b^-$ (#1 and #2) (see Materials and Methods for explanation). Each lane contains 3 μg of poly$(A)^+$ enriched RNA. Each blot was probed with antisense β-actin probe as control. The expression of rat lst-1a, -1b and -1c was liver specific.

Analysis of the putative amino acid modification sites (http://maple.bioc.columbia.edu/predictprotein/ and http://www.cbs.dtu.dk/services/NetOGly) indicated that rat lst-1a protein possesses three potential *N*-linked glycosylation sites on Asn 134, 496, 511 and a potential *O*-linked glycosylation site on Ser 146. Likewise, rat lst-1c protein possesses three potential extracellular *N*-linked glycosylation sites on Asn 134, 463, 478 and a potential *O*-linked glycosylation site on Ser 146. Both rlst-1a and -1c also possess a potential protein-kinase A (PKA) phosphorylation site on Ser 290 and Ser 257, respectively, and two potential protein-kinase C (PKC) phosphorylation sites each; Ser 7 and Ser 325 in rlst-1a, and Ser 7 and Ser 292 in rlst-1c.

All the Three Forms of Rat lst-1 (rlst-1a, -1b, and -1c) Are Expressed Only in Liver and rlst-1a Is the Most Highly Expressed Form

Expression of rat lst-1 in various tissues was determined by Northern-blot as previously described (14) (Fig. 3). The full-length form (rlst-1a) as well as the splice variants (rlst-1b and rlst-1c) are all expressed only in liver. The size of rlst mRNA is approximately 3.5 kb. However, the Northern-blot data



Primer Combination	Expected Amplicon Size
FP-1 + RP-1	744 bp from rlst-1a template 639 bp from rlst-1b template
FP-2 + RP-2	1380 bp from rlst-1a template 1281 bp from rlst-1c template
AP-1 + RP-3	906 bp from both rlst-1a and 1b templates 807 bp from rlst-1c template

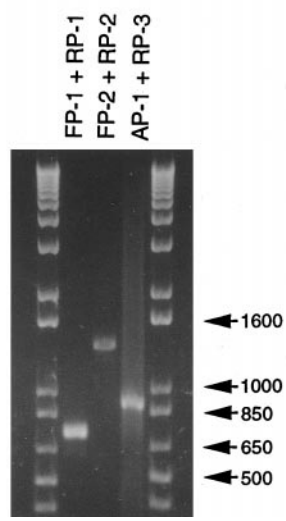


FIG. 4. PCR strategy to determine the relative abundance of the three forms, i.e., rlst-1a, -1b, -1c. Converging dashed lines indicate the regions that are missing in rlst-1b and rlst-1c. The numbers above the dashed lines indicate the first and the last nucleotide position of the regions that are missing. This numbering is based on the first nucleotide position of the start codon (i.e., A) as #1 in all three cDNAs. The designation −80 for FP-2 indicates a location 80 nucleotides upstream from the start codon. Arrows pointing at the primers indicate the nucleotide position of the template that corresponds to the first 5′-end nucleotide of the primer. Note that the position numbering for RP-1, RP-2 and RP-3 changed in rlst-1b and rlst-1c compared to rlst-1a. This is due to the loss of regions in rlst-1b and rlst-1c.

do not yield information about the relative expression levels of rat lst-1a, -1b and -1c. Therefore, in an attempt to determine the relative expression levels of all three lst forms (rlst-1a, -1b, -1c), PCR was done with specific primer sets for different forms of rat lst-1 (rlst-1a, -1b, -1c).

PCR conditions and the primer sequences have been described under Materials and Methods. The strategy

TABLE 1

Exon–Intron Content, Exon Size, and Intron Phases of Rat lst-1 Gene

Exons		Introns	
No.	Size (bp)	No.	Phase
1	26 + 5′-UTR	1	
2	158	2	0
3	142	3	1
4	133	4	2
5	113	5	1
6	147	6	1
7	99	7	1
8	243	8	1
9	165	9	1
10	196	10	2
11	172	11	0
12	170	12	2
13	65	13	1
14	118	14	2
15	217 until STOP + 3′-UTR		

Note. Start codon begins at nt# 75 of Exon 2.

for primer design and the results of the PCR are shown in Fig. 4. More than one primer set was used to amplify different rat lst-1 forms (1a, 1b, 1c). The results of PCR suggest that the expression of rlst-1a is the highest among the three forms. Primers FP-1+RP-1 were used so that they can amplify from both rlst-1a and rlst-1b cDNAs. The result shows that the major amplicon (744 bp) was derived from rlst-1a. A minor amplicon (very faint band) of expected size (639 bp) is also visible, indicating the existence of rlst-1b template in the cDNA library. However, rlst-1a is the predominant form. Similarly, use of primer sets FP-2 + RP-2 and AP-1 + RP-3 also demonstrated that rlst-1a is the major form compared to rlst-1c. Thus, of the three forms of rat lst (1a, 1b and 1c), the full-length form (rlst-1a) is the most highly expressed.

Comparison of Rat lst-1 cDNAs and the Gene Reveals That rlst-1c Is the Product of One Complete Exon Deletion, but the Previously Reported rlst-1b Is a Product of Alternative Usage of the Splice Donor Site from within the Exon

Table 1 shows the size of each exon and the intron phases, while Fig. 5 shows the exon boundaries. The gene contains 15 exons and 14 introns. The last exon is the longest exon. There was no difference in the nucleotide sequence between the cDNA and the gene, indicating the absence of posttranscriptional mRNA editing. The rat lst-1 gene contains two phase 0, seven phase 1 and four phase 2 introns. Four exons (exons 6 through 9) are symmetrical exons (class 1, 1), indicating that legitimate alternative splicing is possible involving these exons. The fact that the number of phase 0 introns is minimum (2 out of 14) in rat lst-1 gene suggests that it is probably not an ancestral gene but evolved later (15). During preparation of this manuscript, the genomic structure of human OATPs (OATP-1, OATP-2 [LST-1] and OATP8) has been reported (16). Human LST-1 gene organization is very similar to its rat ortholog, and the reported intron phases of human LST-1 gene are identical to those of rat and mouse lst-1 genes.

Comparison of the rlst-1 gene and the cDNA sequences (rlst-1a, -1b and -1c) showed that rat lst-1c is a product of alternative splicing involving deletion of the entire exon 7. However, rat lst-1b is not a typical splice variant. It is produced by the alternative usage of splice donor site (GT) from within exon 10. Hence, rlst-1b retains 91 bp out of a total of 196 bp of exon 10. Thus, rat lst-1b represents an interesting but rare deviation from the normal phenomenon of alternative splicing. Because exon 10 is an asymmetrical exon (class 1, 2), a legitimate alternative splicing involving exon 10 will result in frameshift mutation downstream in the open reading frame. In contrast, alternative usage of the splice donor site from within exon 10 can produce a truncated protein without causing frameshift mutation. Kakyo *et al.* (10) reported that they had obtained a number of rlst-1b clones during cloning of rat lst-1 cDNA, and we were also able to PCR amplify rlst-1b using specific primers. Therefore, it appears that rlst-1b is consistently expressed in liver.

The generation of rat lst-1b by circumventing the phase limitation on exon shuffling, is an interesting finding in the present study and its significance remains to be explained.

Rat lst-1, Mouse lst-1 and Human LST-1 Are Orthologs and Are Significantly Different from the Organic Anion Transporting Polypeptide Family (oatps)

In an attempt to understand the evolutionary relationship/divergence of the three families of hepatic Na^+-independent organic anion transporters (oatps, PGT, lst), nucleotide sequences were aligned and a phylogenetic tree was constructed (Fig. 6). The nucleotide sequence of the coding region of rat lst-1a shows about 50% identity with that of rat oatps (oatp 1, 2, 3 and 5), and about 44% identity with rat PGT. The similarity in amino acid sequence is about 41% between rat lst-1a and rat oatps (oatp 1, 2, 3, 5), and 32% between rat lst-1a and rat PGT. In contrast, the identity between rat, human and mouse lst-1 is higher at both the nucleotide and the amino acid level. At the nucleotide level, the identity is 74% (rat and human) and 88% (rat and mouse). At the amino acid level, the identity is 63% (rat and human) and 81% (rat and mouse). It is, therefore, evident that human, rat and

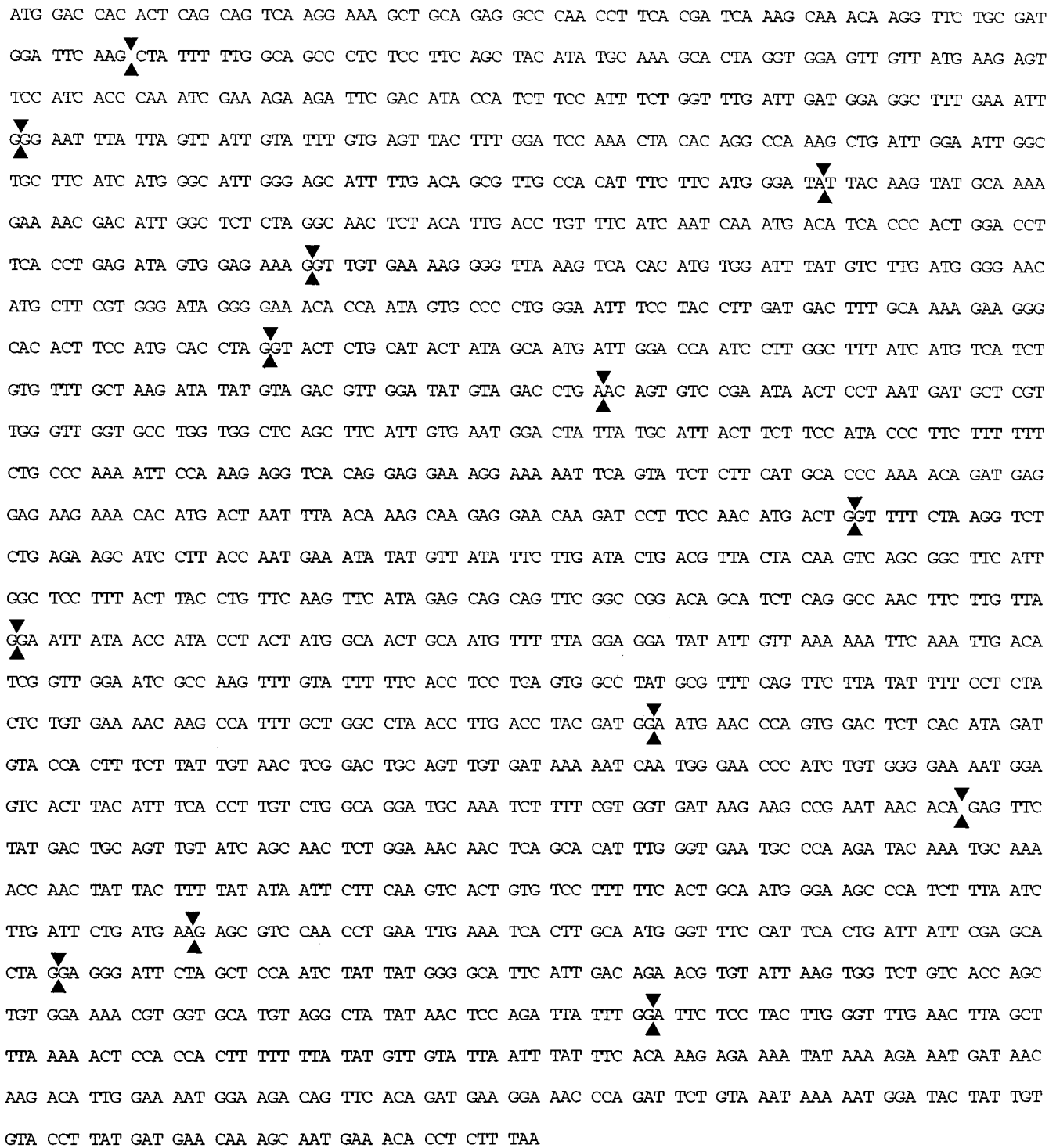


FIG. 5. Exon boundaries of rat lst-1 gene. Only the coding sequence (sense strand) is shown. The arrows indicate the exon breakpoints. From the breakpoints it is evident that the gene contains two phase 0, seven phase 1 and four phase 2 introns that interrupt the coding region.

mouse lst-1 are orthologous genes and the lst family appears to have diverged from the oatp family from a common root. This probably explains a poor homology of rat lst-1 with rat oatps and rat PGT, but higher homology of rat lst-1 with mouse lst-1 and human LST-1. Phylogenetic analysis thus strongly suggests that lsts, oatps, and PGT belong to different transporter families.

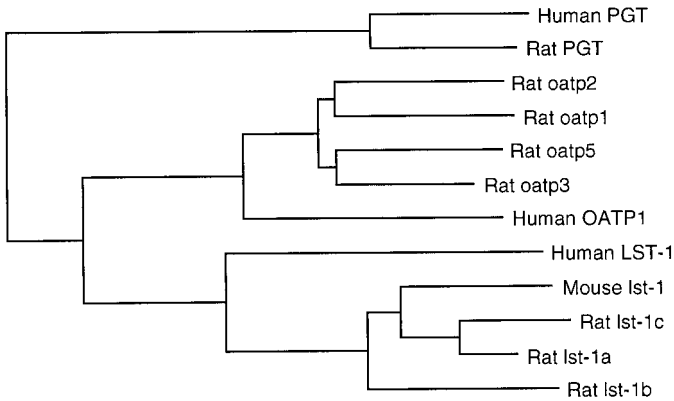


FIG. 6. Phylogenetic divergence analysis of the three major hepatic Na^+-independent organic anion transporters (oatps, PGT, lsts). The divergence was determined at the nucleotide level. All lst's can be grouped under the same family and the lst family appears to have diverged from the oatp family from a common root.

REFERENCES

1. Hagenbuch, B., Stieger, B., Foguet, M., Lubbert, H., and Meier, P. (1991) Functional expression cloning and characterization of the hepatocyte Na^+/bile acid cotransport system. *Proc. Natl. Acad. Sci. USA* **88,** 10629–10633.
2. Tiribelli, C., Lunazzi, G. C., and Sottocasa, G. L. (1990) Biological and molecular aspects of the hepatic uptake of organic anions. *Biochim. Biophys. Acta* **1031,** 261–275.
3. Muller, M., and Jansen, P. L. M. (1997) Molecular aspects of hepatobiliary transport. *Am. J. Physiol.* **272,** G1285-G1303.
4. Jacquemin, E., Hagenbuch, B., Stieger, B., Wolkoff, A. W., and Meier, P. J. (1994) Expression cloning of a rat liver Na^+-independent organic anion transporter. *Proc. Natl. Acad. Sci. USA* **91,** 133–137.
5. Noe, B., Hagenbuch, B., Stieger, B., and Meier, P. J. (1997) Isolation of a multispecific organic anion and cardiac glycoside transporter from rat brain. *Proc. Natl. Acad. Sci. USA* **94,** 10346–10350.
6. Abe, T., Kakyo, M., Sakagami, H., Tokui, T., Nishio, T., Tanemoto, M., Nomura, H., Hebert, S. C., Matsuno, S., Kondo, H., and Yawo, H. (1998) Molecular characterization and tissue distribution of a new organic anion transporter subtype (oatp3) that transports thyroid hormone and taurocholate and comparison with oatp2. *J. Biol. Chem.* **273,** 22395–22401.
7. Kanai, N., Lu, R., Satriano, J. A., Bao, Y., Wolkoff, A. W., and Schuster, V. L. (1995) Identification and characterization of a prostaglandin transporter. *Science* **268,** 866–869.
8. Lu, R., Kanai, N., Bao, Y., and Schuster, V. L. (1996) Cloning, *in vitro* expression, and tissue distribution of a human prostaglandin transporter cDNA (hPGT). *J. Clin. Invest.* **98,** 1142–1149.
9. Abe, T., Kakyo, M., Tokui, T., Nakagomi, R., Nishio, T., Nakai, D., Nomura, H., Uno, M., Suzuki, M., Naitoh, T., Matsuno, S., Kondo, H., and Yawo, H. (1999) Identification of a novel gene family encoding human liver-specific organic anion transporter LST-1. *J. Biol. Chem.* **274,** 17159–17163.
10. Kakyo, M., Unno, M., Tokui, T., Nakagomi, R., Nishio, T., Iwasashi, H., Nakai, D., Seki, M., Suzuki, M., Naitoh, T., Matsuno, S., Yawo, H., and Abe, T. (1999) Molecular characterization and functional regulation of a novel rat liver-specific organic anion transporter rlst-1. *Gastroenterology* **117,** 770–775.
11. Hsiang, B., Zhu, Y., Wang, Z., Wu, Y., Sasseville, V., Yang, W.-P., and Kirchgessner, T. G. (1999) A novel human organic anion transporting polypeptide (OATP2). *J. Biol. Chem.* **274,** 37161–37168.
12. Konig, J., Cui, Y., Nies, A. T., and Keppler, D. (2000) A novel human organic anion transporting polypeptide localized to the basolateral hepatocyte membrane. *Am. J. Physiol. Gastrointest. Liver Physiol.* **278,** G156–G164.
13. Ogura, K., Choudhuri, S., and Klaassen, C. D. (2000) Full length cDNA cloning and genomic organization of the mouse liver-specific organic anion transporter-1 (lst-1). *Biochem. Biophys. Res. Commun.* **272,** 563–570.
14. Choudhuri, S., Zhang, X. J., Waskiewicz, M. J., and Thomas, P. E. (1995) Differential regulation of cytochrome P450 3A1 and P450 3A2 in rat liver following dexamethasone treatment. *J. Biochem. Toxicol.* **10,** 299–307.
15. Strachan, T., and Read, A. P. (1996) Human Molecular Genetics. Wiley-Liss, New York.
16. Konig, J., Cui, Y., Nies, A. T., and Keppler, D. (2000) Localization and genomic organization of a new hepatocellular organic anion transporting polypeptide. *J. Biol. Chem.*, in press.