

Molecular cloning and characterization of two novel transport proteins from rat kidney

Edgar Schömig, Folker Spitzenberger, Martin Engelhardt, Fátima Martel¹, Nicola Örding, Dirk Gründemann*

Department of Pharmacology, University of Heidelberg, Im Neuenheimer Feld 366, 69120 Heidelberg, Germany

Received 6 January 1998; revised version received 13 February 1998

Abstract The recent cloning of renal transport systems for organic anions and cations (OAT1, OCT1, and OCT2) opened the possibility to search, via polymerase chain reaction (PCR) homology screening, for novel transport proteins. Two integral membrane proteins, UST1 and UST2, were cloned from rat kidney. RT-PCR revealed that UST1 is confined to the kidney whereas UST2 mRNA was detected in all tested tissues. Sequence analyses suggest that UST1 and UST2, together with four related transporters, comprise, within the major facilitator superfamily, a so far unrecognized transporter family, termed amphiphilic solute facilitator (ASF) family. Characteristic signatures for the ASF family were identified.

© 1998 Federation of European Biochemical Societies.

Key words: Organic cation transport; Organic anion transport; Amphiphilic solute facilitator family; Phylogeny; Evolution

1. Introduction

Transport proteins in the kidney play a key role in various physiological functions such as fluid homeostasis as well as tubular reabsorption and excretion of endobiotics and xenobiotics. A detailed knowledge of the molecular structure of the membrane proteins involved is expected to provide new insights into kidney physiology and pathophysiology.

Several of the renal transport proteins that have recently been identified by molecular cloning seem to be members of the major facilitator superfamily (MFS) as proposed by Marger and Saier [1]. These new transporters from kidney tubule cells include OAT1, the basolateral renal organic anion/dicarboxylate exchanger [2,3], OCT2, the apical transporter for organic cations [4,5], and OCT1, the basolateral counterpart of OCT2 [6]. Whereas, on the basis of mRNA, OAT1 and OCT2 are expressed predominantly in the kidney [2–5], substantial amounts of OCT1 were also found in some non-renal tissues [5,6].

Here we report the primary structure and tissue distribution of two so far unknown putative transporters, UST1 and UST2, which are related to OAT1, OCT2, OCT1, and NLT, a hepatic transporter of unknown function [7]. Sequence anal-

yses suggest that these transporters comprise, within the major facilitator superfamily, a distinct transporter family.

2. Materials and methods

2.1. PCR amplification and DNA sequencing

For polymerase chain reaction (PCR) on cDNA from rat kidney, the following degenerate oligonucleotides were employed (5'-end first): CCT GAR WSY CCN MGN TGG (forward primer, PESPRW motif), CCT GAR WSY CCN MGN TWY (forward primer, PESPR(F/Y) motif), GAR CTN TAY CCN AC (forward primer, ELYPT motif), GAR TTR TAY CCN AC (forward primer, ELYPT motif), TTN GTY TCN GGN AG (reverse primer, LPETK motif), TTN GTY TCN GGY AA (reverse primer, LPETK motif). PCR was performed as described [5]. Amplification products were isolated by UV-protected agarose gel electrophoresis [8] and cloned into pUC19 [9]. Inserts were sequenced with Thermo Sequenase cycle sequencing kit US 78500 (Amersham, Braunschweig, Germany) and the LI-COR 4000 electrophoresis system (MWG, Ebersberg, Germany).

2.2. cDNA library construction and inverse PCR

Total RNA, mRNA, and cDNA were prepared as described [5]. Size fractionated cDNA was ligated with pBluescript II SK(–) (Stratagene, Heidelberg, Germany) or pcDNA3 (Clontech, Palo Alto, CA, USA) and electroporated into *E. coli* DH10B. Resulting libraries were screened as described [5].

For inverse PCR [10] of UST2r, double-stranded rat kidney cDNA was synthesized with a specific primer (reverse complementary to pos. 1534–1557 of the UST2r cDNA), blunt-ended with T4 DNA polymerase, and incubated with T4 DNA ligase to create circular cDNA. PCR (annealing temperature 60°C, 38 cycles) across the junction was performed with primers corresponding to pos. 1015–1043 (forward) and 320–343 (reverse) of the UST2r cDNA. The major product at 0.9 kbp was cloned and sequenced to reveal the missing 5'-end of UST2r.

2.3. RT-PCR

Total RNA from various rat tissues was processed as detailed before [5]. The following primers were used in PCR (annealing temperature 60°C, 34 cycles): GAC CAA CAA ACT CCA GAA (forward primer UST1r), CCT GGC ATA GGC CAT TCG CCT TTT (reverse primer UST1r), CC ACT ATC TTC GAT CCC AGT GAG TTA CAA (forward primer UST2r), GCA GGA AGC GAT CAT AGG CAC CAA (reverse primer UST2r). GAPDH primers have been described [5].

2.4. Northern blot analysis

Poly(A)⁺-RNA (selected twice, 2 µg per lane) was denatured and fractionated by agarose gel electrophoresis as described [11] except that ethidium bromide staining was avoided. After downward alkaline capillary transfer [12] to Hybond-N (Amersham), the membrane was prehybridized for 1 h at 42°C in solution A (6× SSPE, 50% formamide, 0.5% SDS) plus 5× Denhardt solution and 0.1 mg/ml denatured herring sperm DNA. For overnight hybridization, random-primed, ³⁵S-labeled, double- or single-stranded cDNA fragments from UST1r (pos. 950–1265) or UST2r (pos. 276–3007) were used in solution A plus 5 mM DTT. Membranes were washed to a final stringency of 0.2× SSPE, 0.1% SDS at 60°C. Integrity of mRNA was verified with a GAPDH probe. Radioactivity was detected and quantified by radioluminography with the BAS1000 system (Fuji Photo Film, Tokyo, Japan).

*Corresponding author. Fax: (49) (6221) 548607.
E-mail: dirk.gruendemann@urz.uni-heidelberg.de

¹Present address: Institute of Pharmacology and Therapeutics, Faculty of Medicine, 4200 Porto, Portugal

Abbreviations: RT, reverse transcriptase; PCR, polymerase chain reaction; TM, transmembrane segment; kb, kilobase(s); GAPDH, glyceraldehyde-3-phosphate dehydrogenase

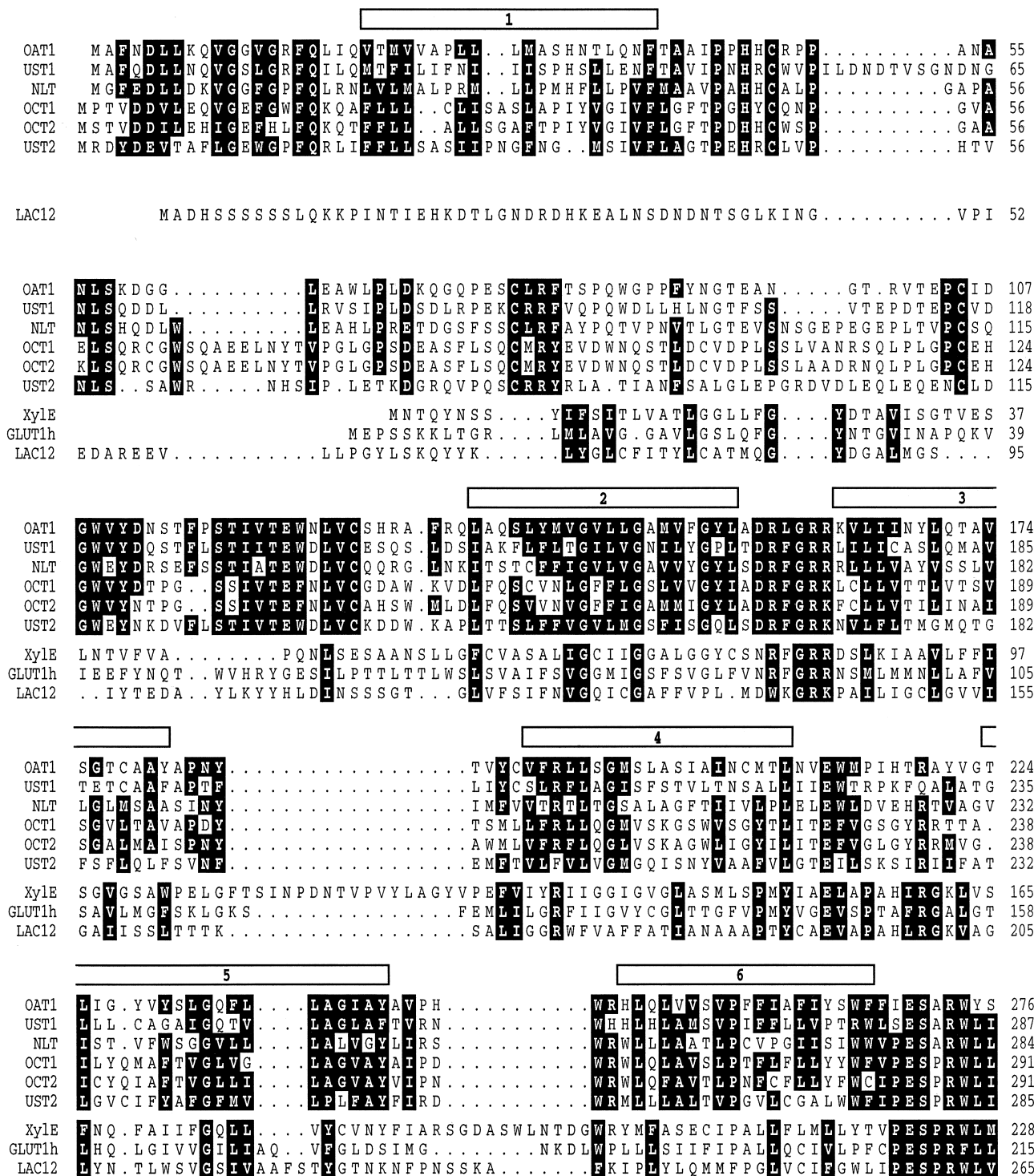


Fig. 1. Amino acid sequences of UST1r and UST2r in alignment to homologous transporters from rat. Boxed residues are conserved in at least four transporters as determined by program PRETTYPLOT with the default normalized Dayhoff substitution matrix and a threshold of 0.8. Open bars indicate putative transmembrane segments of the ASF family transporters. Periods indicate alignment gaps. For comparison, a separate alignment of three sugar transporters from family I was generated as above and aligned to the first set by program PROALIGN. In this second set, boxed residues are conserved in all three transporters. See Table 1 for description of transporters.

2.5. Sequence analysis

The local alignment program BLASTX was used to search databases for proteins most closely related to UST1r and UST2r [13]. Sequences were analyzed with the HUSAR package (DKFZ, Heidelberg, Germany), which is based on the GCG suite of programs. Multiple amino acid sequences were aligned by program PILEUP, with

default substitution matrix, a gap creation penalty of 3.0 and a gap extension penalty of 0.1.

For phylogenetic tree construction, output from program PILEUP, after conversion to PHYLIP format, was analyzed with program PUZZLE, version 3.1 [14].

OAT1	.SSGRDLDTLRALQRVARINGKQEEGAKLSIE.VLRTSLQKELTSLSKGQASAMELLRCPTLR.HLFLC	341
UST1	.MTNKLQKGLKELIKVAHINGMKNSTDVLTIE.VVRTIMKEELEASQTKSSLWDLFRTPNLR.KRICL	352
NLT	.TQGRVEEAKKYLSCAKLNGRPVGEGLSQE.ALNNVV.TMERALQRPSYLDLFRTSQRLR.HISLC	347
OCT1	.SQKRITTRAVRIMEQIAQKNGKVPPADLKML.....CLEEDASEKRSPTS..FADLFRTPNLR.KHTVI	350
OCT2	.SQNKIVKAMKIHKHIAKNGKSVVPSLQNL.....TPDEDAGKKLNPS..FLDLVTRTPQIR.KHTLI	350
UST2	.SQGRVKEAEVIRKAIAKFNIGIVAPSTIFDP.....SELQDLNSKKPQSHHLYDLVTRNIR.IITIM	346
XylE	.SRGKQEQAEGLRKKIMGNLTATQAVQEI.....HS...LDHGRKTGGRLLMFGVG...VIVIG	281
GLUT1h	.INRNEENRAKSVLKKLRGTADVTHDLQEMK.....EESRQMMREKKVTILELFRSPAYRQPILIA	275
LAC12	.GVGREEEAREFLIKYHLNGDRTHPLLDMEAEIIESFPHGTDLSNPLEMLDVRSLFRTRSDRYRAMLV	332
<hr/>		
7		
OAT1	LSMLWFATSFAYYGLVMDLQGFVS.....MYLTIQVIFGAVDLPKAFVCFELVINSMGRPPAQMAS	401
UST1	LSLVRFVWVLSVIGLLINEQHLRIN.....VFLQLQLLGIITIPANLVGIFLVNHLGRRISQLFI	412
NLT	CMVWFGVNFSSYYGLTLDVSGLGLN.....VYQTQLLEGAVELPSKIMVYFLVRRLGRRLTEAGM	407
OCT1	LMYLWFSCAVLYQGLIMHVGTGAN.....LYLDFEFYSSSLVEFPAAFIILVTIDRIGRIYPIAAS	410
OCT2	LMYNWFTSSVLYQGLIMHMGLAGDN.....IYLDFFYSALVEFPAAFIILVTIDRIGRIYPIAAS	410
UST2	SIILWLITISVGYFGLSLDTPNLHG.....IYVNCFLLAAVEVPAYVLAWLLQLHLPRRYSISAA	406
XylE	VMLSIFQQFVGINVVLYAPEVFKTLG.ASTDIALLQTTIIVGVINLTFTVLAIMTVDKFGRKPLQIIG	348
GLUT1h	VVLQLSQQSLGINAVFYYSSTISFEK...AGVQQPVYATIGSGIVNTAFTVVSFLFVVERAGRRTLHLIG	340
LAC12	ILMAWFGQFSGNVNCSSVYLPMTLRNVGMKSVSLNVLMNGVYSIVTWISSICGAFFIDKIGRREGFLGS	400
<hr/>		
9		
OAT1	LLLAGICILVNGIIPKSHTIIR.....TSLAVLKGCLASSFNCIFLYTGELYPTVIRQTGLGM	460
UST1	ISLFGISILAIIFVFPQEMQILR.....MVLATFGGVFSFVSVSALVHANELLPTTIIRATALGV	471
NLT	LLGAALTFGTSLLVSLLETKSWI.....TALVVVGKAFSEAAFTTAYLFTSELYPTVLRQTGLGL	466
OCT1	NLVGTGAACLLMIFIPHEHHLN.....VTLACLGRMGATIVLQMVCLVNABELYPTFIRNLGMMV	469
OCT2	NMVAGAACLASVFIIPDDDLQWLK.....ITLACLGRMGITMAYEMVCLVNABELYPTYIRNLGVLV	469
UST2	LFLGGSVLLFIQLVPSSELYLS.....TALVMVGKFGITSAYSMVYVYTAELYPTVVRNMGVGV	465
XylE	ALGMAIG..MFSLGTAFYTQAP..GI..VALLSMLFYVAAFAMSWGPVVCWVLLSEIFPNAIRGKALAI	410
GLUT1h	LAGMAACAILMTIALALLEQLPWSY..LSIVAFIFGEVAFFEVGPPIPIPFIVAEELFSQGPRAAIAV	406
LAC12	ISGAALALTLGLSICTARYEKTKKKSASNGALVFIYLEGGIFSFAPTPMQSMYSTEVSTNLTRSKAQLL	468
<hr/>		
10		
OAT1	GSTMARVGSIVSPLVSM.....TAEFYPSMPLFIIFGAVPVVASAVT.ALLPETLGPQLPDDTVQDL	519
UST1	IGIAGSTGAALSPLFMI.....LRTYSDSLPWIIFGVLSFLGGLVV.LLLPETKNQPLPDISIQDV	530
NLT	TALMGRGLGASLARLAL.....LDGVWLLLPKVAYGGIALVAACATA.LLLPETKKAQLPETIQDV	525
OCT1	CSALCDLGGIFTFFMVER.....LMEVWQALPLILFGLGLTAGAMT.LLLPETKGVALPETIEEA	529
OCT2	CSSMCDIGIITPFLVYR.....LTDIWMFELPVFAVVGVLVAGALV.LLLPETKGAALPETIEDA	529
UST2	SSTASRLGSILSPYFVY.....LGAYDRFLPVYILMGLTILTAIILT.LFFPESFGAPLPDITDQM	524
XylE	AVAAQWLANYFVSWTFPMMDKNSWLVAHFHNGFSYWIYGCIMGVLAALFMWKFPVETKKGKLEELE.AL	477
GLUT1h	AGFSNWTSNFIVGMCFCQ.....YVEQLCGPYVFIIFETVLLLVLFIFTYFKVPEBKGRTFDEIASGF	467
LAC12	NFVVSQVAQFVNQFATPKAMKNIKY.....WFYVYFVFEDIFEFIVIVYFFFVETKGRSLEELEVVVF	529
<hr/>		
11		
OAT1	KSRSRGKQNNQQQEQQKQMMPLQASTQEKNGL	551
UST1	ENEGRASRQGGKQNDTLIKVTQF	552
NLT	ERKSTQEEDV	535
OCT1	ENLGRKSKAKENTIYLVQVQTKSSST	556
OCT2	ENMQRPKRKKERKENL.PPSQASRPSAKLKRKGI IAAGADFALSEARDGASLSPPPPTQTNLTYP	593
UST2	LRVKGIKQWQIQSQTRTQKGGESPTVLKSTAF	557
XylE	WEPETKKTKQTATL	491
GLUT1h	RQGGASQSDKTPEELFHPLGADSQV	492
LAC12	EAPNPRKASVDQAFLAQVRATLVQRNDVRVANAQNLKEQEPLKSDADHVEKLSAEESV	587

Fig. 1 (continued).

3. Results

3.1. Cloning of two putative transporters from rat kidney

Degenerate consensus primers were derived from the amino acid sequences of NLT, OCT1r, OCT2p, and OCT2r and used in PCR on cDNA from rat kidney. Resulting fragments were cloned into pUC19 and sequenced. Two fragments showed some degree of similarity to the afore mentioned

transporters. The corresponding putative transporters are referred to as UST1r and UST2r.

The UST1r fragment, when used as a probe in Northern blot analysis of rat kidney mRNA, detected a single band with a length of approximately 2.9 kbp (data not shown). Subsequently, an appropriately size-fractionated primary cDNA library was generated from rat kidney mRNA and screened by colony hybridization, with the same probe as above. A clone,

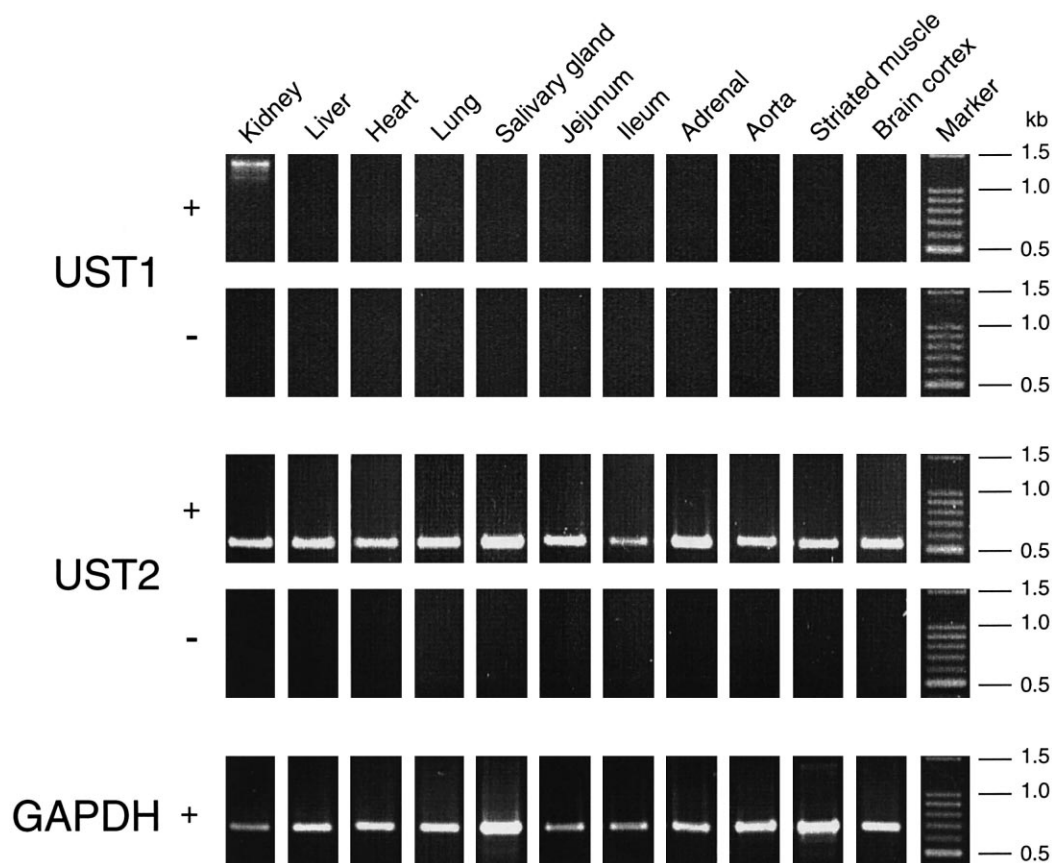


Fig. 2. Tissue distribution of UST1r and UST2r mRNA. RT-PCR products were separated by agarose gel electrophoresis, followed by staining with ethidium bromide, as shown. The mRNA-derived signal is the difference of band intensities between samples generated in the presence (+) or in the absence (–) of RT during reverse transcription. As control for the integrity of mRNA, GAPDH was detected with specific primers.

2.6 kbp long, with a complete open reading frame was isolated and sequenced.

The UST2r fragment, in Northern analysis of rat kidney mRNA, produced a single band of about 3.0 kbp (Fig. 3). A clone with a length of 2.7 kbp (pos. 276–3007 of the published sequence) was isolated from a rat kidney cDNA library as described above. This clone turned out, however, to lack

some coding sequence at the 5'-end. The missing information was gathered by inverse PCR.

3.2. Primary structures of UST1r and UST2r

The amino acid sequences of UST1r and UST2r are shown in Fig. 1, in alignment to the most similar transporters known so far, i.e. OAT1, NLT, OCT1, and OCT2. With an identity

Table 1
Membrane proteins analyzed in this study, in the order of output from the alignment program

Designation ^a	Substrate ^b	Accession number ^c	Reference
OAT1m ^d	?	U52842	[2]
OAT1r	PAH, glutarate	AB004559	[3]
UST1r	?	Y09945	This study
NLT1r	?	L27651	[7]
OCT2p	TEA, MPP ⁺	Y09400	[5]
OCT2h	TEA, MPP ⁺	X98333	[21]
OCT2r	TEA, MPP ⁺	D83044	[4]
OCT1r	TEA, MPP ⁺	X78855	[6]
OCT1m ^e	?	U38652	[20]
OCT1h	TEA, MPP ⁺	U77086	[19]
UST2r	?	AJ001933	This study
XylE	Xylose	J02812	[23]
GLUT1h	Glucose	K03195	[24]
LAC12	Lactose	X06997	[25]

^ah, m, p, and r designate a protein from human, mouse, pig, and rat, respectively. XylE is from *E. coli*, and LAC12 from *Kluyveromyces lactis*.

^bPAH, *p*-aminohippurate; TEA, tetraethylammonium; MPP⁺, 1-methyl-4-phenylpyridinium.

^cOf the respective GenBank/EBI Data Bank nucleotide sequence.

^dOriginally termed NKT (novel kidney transporter).

^eOriginally termed LX1.

score of 28%, UST1 and UST2 are not closely related. From hydropathy analysis, UST1r (552 amino acids) and UST2r (557 amino acids) are both predicted to contain 12 transmembrane segments (TMs), with an overall topology like OAT1, NLT, OCT1, and OCT2 (Fig. 1). Both proteins carry three potential N-glycosylation sites (pos. 56, 66, and 102 for UST1r, pos. 57, 64, and 91 for UST2r) in the proposed large extracellular loop between TM 1 and TM 2. Consensus sites from PROSITE pattern search were found for phosphorylation by protein kinase C (pos. 167, 282, and 289 for UST1r, pos. 164, 280, 323, and 541 for UST2r), and cAMP-dependent protein kinase (pos. 405 and 399 for UST1r and UST2r, respectively).

3.3. Tissue distribution

RT-PCR revealed that UST1r and UST2r differ markedly in terms of tissue distribution (Fig. 2). The UST1r mRNA is strictly confined to the kidney. The UST2r mRNA, on the

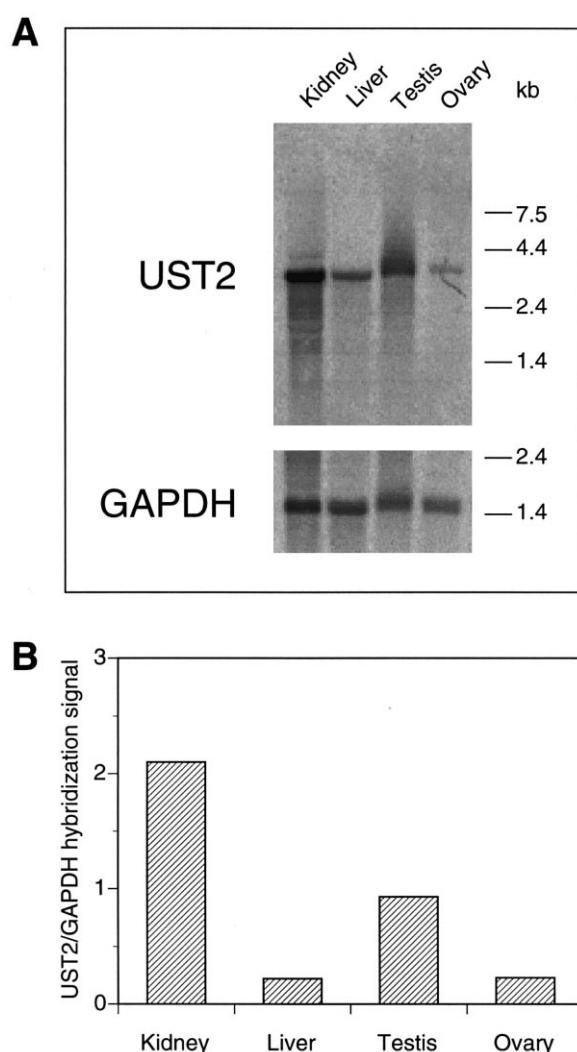


Fig. 3. Northern blot analysis with a UST2r probe of various rat tissues. A: The blot was first hybridized with a 2.7-kbp double-stranded, 35 S-labeled UST2r probe. Loading of mRNA was then checked with a GAPDH probe. B: Comparison of UST2r mRNA tissue contents. Radioactivity in the UST2r bands relative to GAPDH bands was measured by radioluminography.

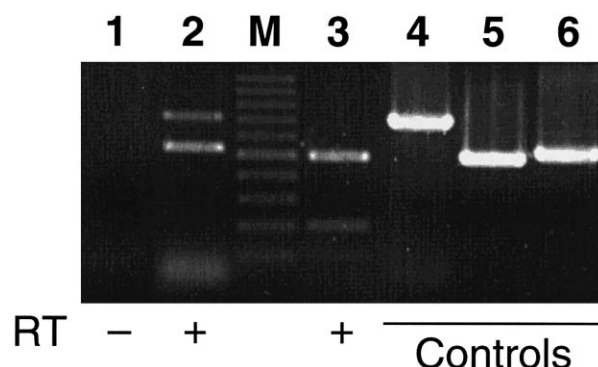


Fig. 4. RT-PCR with UST1r specific primers on rat kidney mRNA. An ethidium-stained agarose gel is shown with PCR products obtained, in lanes 1, 2, 4, and 6, with forward primer pos. 950–967 of UST1r and reverse primer pos. 1641–1658. For lanes 3 and 5, reverse primer pos. 1425–1442, i.e. within the deletion, was used. PCR parameters were 50°C annealing temperature, 30 cycles. For controls, PCR was performed on linearized plasmids with the cDNA of UST1r (lanes 4 and 5) or UST1rD (lane 6).

other hand, was easily detectable in all tested tissues. Northern analysis with subsequent quantification of the hybridization signals revealed that kidney and testis contain considerably higher levels of UST2r mRNA than liver and ovary (Fig. 3).

3.4. Detection of an UST1r mRNA isoform

While screening the respective rat kidney cDNA library for UST1r, some of the clones, by DNA sequencing, were found to carry an internal deletion of 184 bp, with the first deleted nucleotide somewhere at pos. 1370–1372 of the UST1r sequence. To examine whether this truncated UST1r form, termed UST1rD, represents a cDNA synthesis artefact, RT-PCR was performed on rat kidney mRNA (Fig. 4). Interestingly, a pair of primers designed to flank the presumptive deletion gave rise to two bands of the expected sizes for UST1r and UST1rD (Fig. 4, lane 2). In support of this assignment, PCR with one of the primers placed within the deletion produced a single major band of the expected size for UST1r (lane 3). Thus, results both from cDNA cloning and RT-PCR suggest the UST1rD mRNA to be authentic. It appears from Fig. 4 that, on the mRNA level, UST1rD may even be more abundant than UST1r.

3.5. Evolutionary analysis

The amino acid sequences of UST1r and UST2r, together with 9 known homologous proteins, were aligned by progressive multiple alignment [15]. For reference, three representative members from the family I of sugar transporters as defined by Griffith et al. [16] were included in the analysis (Table 1). Based on this alignment, identity and similarity scores were computed (Table 2). Also, an evolutionary tree was reconstructed (Fig. 5) by the maximum likelihood method of quartet puzzling [14]. The tree topology, the branch lengths – which correspond to evolutionary distances – and the scores for identity and similarity suggest that UST1r and UST2r, together with NLT and the organic cation and organic anion transporters, constitute a family of homologous proteins, termed amphiphilic solute facilitator (ASF) family. The ASF transporters are clearly set off against the family I sugar transporters.

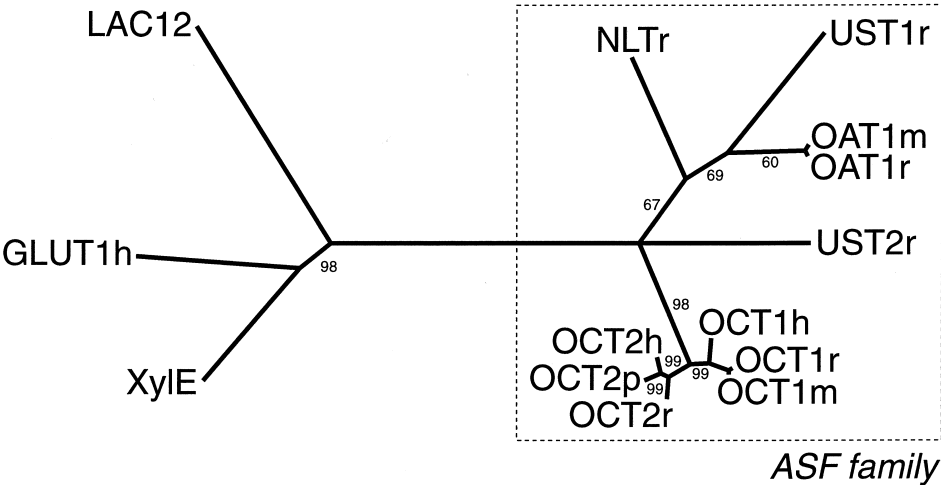


Fig. 5. Evolutionary relationships among the proteins from Table 1. Shown is the unrooted majority rule consensus tree from quartet puzzling (1000 puzzling steps, default options). The low percentage of unresolved quartets, 2.0%, indicates that the data are well suited for phylogenetic analysis. Reliability values, in percent, for each internal branch are given if smaller than 100%. The reliability value tells how often the corresponding cluster was found among the intermediate trees. Excluded bipartitions cluster the six organic cation transporters together with the non-ASF transporters (reliability value 48%), and OAT1, NLTr, UST1r together with the non-ASF transporters (reliability value 43%). All other dismissed bipartitions have a support below 30%. Similar results were generated using parsimony or distance methods available in PHYLIP [26] with a posteriori bootstrap resampling to evaluate support for the branches.

4. Discussion

In our search for novel transport proteins – by RT-PCR analysis and subsequent molecular cloning – we identified two 2.6 kbp and 3.0 kbp long cDNAs from rat kidney which code for integral membrane proteins, UST1r and UST2r. Both proteins are new, i.e. closely related homologs or orthologs are not known. UST1r and UST2r, which consist of 552 and 557 amino acid residues, respectively, are predicted to share a common transporter topology of 12 transmembrane segments (TMs).

RT-PCR analysis failed to detect UST1 mRNA in any of the tested tissues other than kidney which suggests that UST1

plays a role in kidney-specific physiological processes such as tubular reabsorption, excretion of endo- and xenobiotics, or body fluid homeostasis. During the present study, we identified a second species of UST1r mRNA which carries a 184 bp deletion (UST1rD). UST1rD represents an authentic mRNA, since it was identified not only by molecular cloning from a primary kidney cDNA library, but also demonstrated by RT-PCR to exist in the kidney. The deletion could be due to alternate splicing and results in a truncated protein which lacks TMs 10–12.

UST2, on the other hand, appears to be ubiquitous as it was detected by RT-PCR analysis in all tested tissues. Hence, UST2 is expected to be involved in more general cellular

Table 2
Pairwise comparisons of amino acid sequences of proteins from Table 1

	OAT1m	OAT1r	UST1r	NLTr	OCT2p	OCT2h	OCT2r	OCT1r	OCT1m	OCT1h	UST2r	XylE	GLUT1h	LAC12
OAT1m		95	39	38	32	32	32	34	33	32	29	17	16	13
OAT1r	98		39	38	31	32	31	33	33	31	29	17	15	13
UST1r	63	62		33	27	26	26	28	27	28	28	16	18	14
NLTr	62	60	53		29	28	31	30	30	29	31	17	17	15
OCT2p	54	52	51	51		85	81	67	68	69	32	18	17	15
OCT2h	54	53	52	50	92		81	68	68	69	32	18	17	16
OCT2r	53	52	52	53	90	88		68	67	67	32	17	18	15
OCT1r	54	52	52	52	81	81	83		95	78	33	17	19	15
OCT1m	53	52	52	52	82	82	83	98		78	32	17	19	16
OCT1h	53	52	52	51	83	84	82	89	90		33	17	19	16
UST2r	54	53	49	55	55	56	55	56	56	57		15	17	13
XylE	42	42	44	41	43	43	43	43	43	42	41		30	23
GLUT1h	39	40	41	39	42	43	44	41	41	43	41	60		22
LAC12	35	35	39	36	37	38	36	37	37	37	35	52	45	

Percentage identity (above diagonal) and percentage similarity (below diagonal) were computed from the multiple sequence alignment file created by PILEUP. Similarity was scored with the default GAP matrix nwsgappep.cmp, which is a normalized Dayhoff matrix, and a threshold value of 0.5.

processes. Quantitation by Northern analysis revealed, however, that mRNA levels, by contrast to house-keeping proteins such as GAPDH, vary among tissues. The kidney and the testis should be especially rich in UST2.

Sequence comparison with the recently cloned organic cation transporters OCT1 [6,19,20] and OCT2 [4,5,21], the organic anion transporter OAT1 [2,3], and NLT [7], a putative transporter expressed predominantly in liver, revealed that both UST1 and UST2 are about 30% identical with these transporters (Table 2). Although it is no sensitive measure of homology, the percentage identity between amino acid sequences has the appeal of being easily understood and wholly unbiased. As a rule of thumb, for two protein sequences of sufficient length, i.e. 50 residues or more, a percentage identity of more than 25% indicates significant structural homology [17]. A group of proteins with mutual identity scores equal to or greater than 30% is generally considered to constitute a family [18], which implies a clear evolutionary relationship. Thus, on the basis of identity scores, OCT1, OCT2, OAT1, NLT, UST1, and UST2 constitute a new protein family. Since the transporters for which substrates are known, i.e. OCT1, OCT2, and OAT1, in essence facilitate the transmembrane movement of organic cations or organic anions, i.e. amphiphilic solutes, this family of transporters was tentatively termed amphiphilic solute facilitator (ASF) family.

Based on statistical sequence analysis, Griffith et al. have sorted a large collection of transporters for various sugars, antibiotics, antiseptics, and carboxylates into four distinct, but homologous transporter families [16]. Out of these, family I (equivalent to cluster 2 in Ref. [1]), a group consisting almost exclusively of sugar transporters, is most closely related to the ASF family. This is readily apparent from conserved motifs [3,7,20] (see below) and from a prominent intracellular loop of 60–65 hydrophilic residues between TMs 6 and 7 which is present in both family I and ASF family transporters.

In the present study, representative members from family I, i.e. XylE, GLUT1h, and LAC12, were only 13–19% identical to the ASF transporters, compared with the lower limit of 26% identity within the ASF family. The analysis of similarity scores reveals essentially the same picture (Table 2). Note, in particular, that most highly conserved residues within the TMs of the sugar transporters [16,27] are absent from the ASF family (Fig. 1). Consequently, the detailed analysis of evolutionary relations places the family I members in some distance to the ASF transporters (Fig. 5). Thus, in contrast to previous assignments [2,7,20], transporters from the ASF family do not belong to the family I of sugar transporters.

A low identity score, i.e. below 20%, does not disprove the existence of homology. In this 'twilight zone', protein families may still be grouped into a superfamily – provided that structural or functional features suggest a common evolutionary origin. Indeed, the following common sequence motifs suggest that the ASF transporters and the family I transporters belong to the same superfamily, termed major facilitator superfamily (MFS) by Marger and Saier [1]: GX₃GX₄GX₃DRFGRR through and after TM 2, EX₆R after TM 4, and PESPRWL after TM 6 (allow here and below for conservative replacements and slight deviations). All three motifs are also discernible in the second half both of the family I and the ASF family transporters, suggestive of internal duplication of a primordial gene [1,16]: (D/N)(R/

H)LGRR after TM 8, EX₆R after TM 10, and PETKG after TM 12.

Are there, apart from isolated residues, any sequence features that positively distinguish the ASF transporters? We suggest the following motifs (Fig. 1), which are not present in the family I transporters: STIVTEW(D/N)LVC before TM 2, ELYPT after TM 10, which is a special case of the EX₆R motif described above, and LP(D/E)TI after TM 12. Another motif, DLFRT, before TM 7, is also present in the GLUT1 and GLUT3 transporters from family I [22].

Future work in the field of the ASF family faces two challenges: first, to define the yet unknown substrate specificities of NLT, UST1 and UST2. Hints may come from analysis of transporter localization. Second, to clarify which residues determine substrate specificity. In this respect, the growing sequence dataset might help to identify key residues. One might speculate, for example, that charge selectivity filters have evolved to attract substrates and to repel solutes of opposite charge. The organic cation transporters, then, should contain negative charges that line, by a simple model, the substrate passage way. Candidate residues, which are conserved in all the organic cation transporters, but which are absent in the organic anion transporters and the other ASF members, are located at pos. 150, 379, and 475 of the OCT1 sequence, i.e. within or adjacent to the transmembrane segments shown in Fig. 1. Oppositely charged residues, such as those in pos. 368 and 487, could help to orient the organic ions in the solute recognition site. From this simple analysis of conserved charged residues, one might speculate that UST2 and NLT would accept anionic solutes as substrates whereas UST1 would transport uncharged compounds.

The present study describes two novel putative transporters from rat kidney, UST1 and UST2, which were identified by molecular cloning. The tissue distribution suggests that UST1 has a kidney-specific function, whereas UST2 might be involved in more general cellular processes. Finally, it was possible to define, within the major facilitator superfamily, a novel family of transporters (ASF family). Currently the ASF family comprises, besides UST1 and UST2, the organic cation transporters OCT1 and OCT2, the organic anion transporter OAT1, and a putative transporter from liver, NLT.

Acknowledgements: Supported by the Deutsche Forschungsgemeinschaft (Grants Scho373/1-2, Un34/19-1/B2, and SFB601/A4) and the Humboldt Foundation. We thank Anke Ripperger and Barbara Wallenwein for skillful technical assistance.

References

- [1] Marger, M.D. and Saier Jr., M.H. (1993) Trends Biochem. Sci. 18, 13–20.
- [2] Lopez-Nieto, C.E., You, G., Bush, K.T., Barros, E.J.G., Beier, D.R. and Nigam, S.K. (1997) J. Biol. Chem. 272, 6471–6478.
- [3] Sekine, T., Watanabe, N., Hosoyamada, M., Kanai, Y. and Endou, H. (1997) J. Biol. Chem. 272, 18526–18529.
- [4] Okuda, M., Saito, H., Urakami, Y., Takano, M. and Inui, K. (1996) Biochem. Biophys. Res. Commun. 224, 500–507.
- [5] Gründemann, D., Babin-Ebell, J., Martel, F., Örding, N., Schmidt, A. and Schömig, E. (1997) J. Biol. Chem. 272, 10408–10413.
- [6] Gründemann, D., Gorboulev, V., Gambaryan, S., Veyhl, M. and Koepsell, H. (1994) Nature 372, 549–552.
- [7] Simonson, G.D., Vincent, A.C., Roberg, K.J., Huang, Y. and Iwanij, V. (1994) J. Cell Science 107, 1065–1072.

- [8] Gründemann, D. and Schömig, E. (1996) *BioTechniques* 21, 898–903.
- [9] Liu, Z. and Schwartz, L.M. (1992) *BioTechniques* 12, 28–30.
- [10] Huang, S., Hu, Y., Wu, C. and Holcenberg, J. (1990) *Nucleic Acids Res.* 18, 1922.
- [11] Gründemann, D. and Koepsell, H. (1994) *Anal. Biochem.* 216, 459–461.
- [12] Chomczynski, P. (1992) *Anal. Biochem.* 201, 134–139.
- [13] Altschul, S.F., Gish, W., Miller, W., Myers, E.W. and Lipman, D.J. (1990) *J. Mol. Biol.* 215, 403–410.
- [14] Strimmer, K. and von Haeseler, A. (1996) *Mol. Biol. Evol.* 13, 964–969.
- [15] Feng, D.-F. and Doolittle, R.F. (1987) *J. Mol. Evol.* 25, 351–360.
- [16] Griffith, J.K., Baker, M.K., Rouch, D.A., Page, M.G.P., Skurray, R.A., Paulsen, I.T., Chater, K.F., Baldwin, S.A. and Henderson, P.J.F. (1992) *Curr. Opin. Cell Biol.* 4, 684–695.
- [17] Jones, D.T., Taylor, W.R. and Thornton, J.M. (1992) *Comput. Applic. Biosci.* 8, 275–282.
- [18] Murzin, A.G., Brenner, S.E., Hubbard, T. and Chothia, C. (1995) *J. Mol. Biol.* 247, 536–540.
- [19] Zhang, L., Dresser, M.J., Gray, A.T., Yost, S.C., Terashita, S. and Giacomini, K.M. (1997) *Mol. Pharmacol.* 51, 913–921.
- [20] Schweifer, N. and Barlow, D.P. (1996) *Mamm. Genome* 7, 735–740.
- [21] Gorboulev, V., Ulzheimer, J.C., Akhoundova, A., Ulzheimer-Teuber, I., Karbach, U., Quester, S., Baumann, C., Lang, F., Busch, A.E. and Koepsell, H. (1997) *DNA Cell Biol.* 16, 871–881.
- [22] Goswitz, V.C. and Brooker, R.J. (1995) *Protein Science* 4, 534–537.
- [23] Davis, E.O. and Henderson, P.J.F. (1987) *J. Biol. Chem.* 262, 13928–13932.
- [24] Mueckler, M., Caruso, C., Baldwin, S.A., Panico, M., Blench, I., Morris, H.R., Allard, W.J., Lienhard, G.E. and Lodish, H.F. (1985) *Science* 229, 941–945.
- [25] Chang, Y.D. and Dickson, R.C. (1988) *J. Biol. Chem.* 263, 16696–16703.
- [26] Felsenstein, J. (1988) *Annu. Rev. Genet.* 22, 521–565.
- [27] Gould, G.W. and Holman, G.D. (1993) *Biochem. J.* 295, 329–341.