JPET Fast Forward. Published on July 28, 2004 as DOI: 10.1124/jpet.104.073098

JPET #73098

Genetic Variants of the human H^+/dipeptide transporter PEPT2: Analysis of Haplotype Functions

Julia Pinsonneault[1], Carsten Uhd Nielsen[2], Wolfgang Sadée [1]

Department of Pharmacology Program in Pharmacogenomics, College of Medicine and Public Health, Ohio State University, Columbus, Ohio. (JP, WS)

Molecular Biopharmaceutics, Department of Pharmaceutics, The Danish University of Pharmaceutical Sciences, Universitetsparken 2, DK-2100 Copenhagen, Denmark (CUN)

a) Running title: Genetic Variants of the human H^+/dipeptide transporter PEPT2

b) Author for correspondence:
Julia Pinsonneault
Department of Pharmacology
Ohio State University
333 West 10th Avenue
Columbus OH 43210-1239
Phone: 614-292 5165
FAX: 614-292-7232
Email: pinsonneault

c) Number of text pages: 20

Number of tables: 2

Number of figures: 5

Number of references: 31

Number of words

Abstract: 220

Introduction: 739

Discussion: 1146

d) Nonstandard abbreviations:

SNP	Single nucleotide polymorphism
PCR	Polymerase chain reaction
EM	estimation maximization
HPLC	High pressure liquid chromatography
HBSS	Hanks Balanced Salt Solution
MES	2-(N-Morpholino)-ethanesulfonic acid

e) Recommended section: Gastrointestinal, Hepatic, Pulmonary & Renal

Abstract

PEPT2 is a high-affinity H^+/dipeptide transporter expressed in kidney, brain, lung and mammary gland. The physiologic role of PEPT2 in kidney is to reabsorb small peptides generated by luminal peptidases. PEPT2 is also a transporter for peptide-like drugs such as penicillins and cephalosporins. We have conducted a haplotype analysis of 27 SNPs located in or near exons of the human gene encoding hPEPT2 (SLC15A2), using genotyping data from 247 genomic DNA samples from the Coriell collection. Our analysis reveals that h*PEPT2* has a >6 kb sequence block with at least 10 abundant polymorphisms in almost complete linkage disequilibrium. As a result, only two main *hPEPT2* variants exist (*hPEPT2*1* and **2*) with several phased amino acid substitutions, present in substantial frequencies in all ethnic groups tested. When expressed in CHO cells, hPEPT2*1 and *2 displayed similar Vmax values for GlySar, but differed significantly in their K_m values (83±16 μM and 233±38 μM, respectively). Moreover, hPEPT2*1 and *2 differed in their pH sensitivity for H^+/GlySar transport. In addition, hPEPT2*1 and *2 generated varying levels of mRNA in 9 heterozygous kidney tissue samples, including one allele expressing no detectable mRNA, suggesting the presence of *cis*-acting polymorphisms affecting transcription or mRNA processing. The results indicate that polymorphisms in the gene encoding hPEPT2 can alter substrate transport and therefore could affect drug disposition *in vivo*.

Introduction

The dipeptide transporter PEPT2 (SLC15A2) is a high affinity proton-coupled oligopeptide transporter functioning in the cellular uptake of small peptides and peptide-like drugs such as beta-lactam antibiotics. PEPT2 is a member of the proton-dependent oligopeptide transporter family (Liu et al., 1995), along with PEPT1 (SLC15A1) which is expressed in the small intestine. The two proteins share 50% sequence identity and are 70% similar (Graul and Sadée, 1997). Unlike PEPT1, PEPT2 is expressed more broadly in kidney, lung, brain and mammary gland. Both transporters are located in the apical membrane of proximal kidney tubule cells (Daniel et al., 1991). In the nephron, PEPT1 and PEPT2 are sequentially expressed with PEPT1 located in the proximal part and PEPT2 in the distal parts of the proximal tubule (Smith et al., 1998).

The physiologic role of PEPT2 in kidney is to reabsorb small peptides left behind by luminal peptidases. It has been suggested that PEPT2 may be involved in drug disposition in the kidneys via reabsorption of beta-lactam antibiotics (Inui et al., 2000a; Inui et al., 2000b). Reabsorption of some cephalosporins, Gly-Sar (glycyl-sarcosine) and Gly-Sar-Sar into the kidneys from the ultrafiltrate has been demonstrated *in vivo* (Arvidsson et al., 1979; Garrigues et al., 1991; Minami et al., 1992). Brain tissues expressing hPEPT2 include astrocytes, ependymal cells, and choroid plexus epithelial cells (Berger and Hediger, 1999). In the central nervous system, hPEPT2 may play a role in neuropeptide signaling. Neuropeptide signaling is terminated by extracellular peptidases that breakdown peptides into smaller fragments which are then removed by oligopeptide transporters. Based on PEPT2 substrate requirements and localization, it has been suggested that PEPT2 may be involved in the removal of neuropeptide fragments from the brain and in the regulation of neuropeptide levels in the cerebrospinal fluid (Nielsen et

al., 2002). PEPT2 is present at the apical membrane of choroidal plexus epithelial cells and may function in the efflux of peptides across the blood-brain barrier from cerebrospinal fluid to blood (Shu et al., 2002). In mammary gland epithelia, PEPT2 is thought to be involved in the reuptake of small peptides that accumulate from the hydrolysis of milk proteins (Groneberg et al., 2002). PEPT2 is located in alveolar type II pneumocytes, bronchial epithelium, and endothelium of small vessels (Groneberg et al., 2001), and thought to be responsible for uptake of di- and tripeptides in the lung (Groneberg et al., 2001). Moreover, PEPT2 transports a number of drugs such as penicillins, cephalosporins, and ACE inhibitors (Daniel and Adibi, 1993; Ganapathy et al., 1995; Ganapathy et al., 1997; Ganapathy et al., 1998; Zhu et al., 2000), and it may be responsible for the relatively low levels of these drugs detected in breast milk of lactating women who are taking them.

PEPT2 (SLC15A2, solute carrier 15A2) is 729 amino acids in size with 12 proposed transmembrane spanning regions, and a large extracellular loop between transmembrane domains 9 and 10 (Figure 1). Located on chromosome 3 at 3q21.1, *hPEPT2* contains 22 exons spread over 47 kilobases (Figure 2). When the gene was cloned in 1995 by screening a human cDNA kidney library with a rabbit PEPT2 cDNA (Liu et al., 1995), a single transcript was identified and accepted as the consensus sequence. Numerous single nucleotide polymorphisms (SNPs) have been detected in the *SLC15A2* gene (see PharmGKB, http://www.pharmgkb.org/views/ and http://pharmacogenetics.ucsf.edu/). The functional significance of these polymorphisms remains unknown, but recently, a rare non-synonymous polymorphism (R57H) has been shown to disrupt PEPT2 function (Terada et al., 2004).

A haplotype is the combination of several sequence variants on a single chromosome at a specific locus. Haplotypes may more accurately reflect genetic diversity in the human

population. Here, using haplotype analysis, we show that *hPEPT2* exists in multiple variants and that two of these variants *(hPEPT2*1* and **2*) are significantly present in all ethnic groups we have tested. *In vitro* analysis indicates that the *1 and *2 protein variants differ in functional properties.

Genetic variations can affect protein structure on the one hand and mRNA processing, mRNA stability, and *cis*-regulatory mechanisms on the other. The latter variations have one common outcome: they affect the amount and nature of the mRNA generated from one allele versus another. To determine whether such *cis*-acting polymorphisms exist in *hPEPT2*, we have measured allele-specific mRNA expression of *hPEPT2* in heterozygous samples from kidney tissues. Taken together, the results indicate significant functional variations are present in hPEPT2. This work provides a basis for exploring the pharmacological relevance of genetic variability.

METHODS:

Genotyping

Genomic DNA samples were obtained from the Coriell Institute (Camden, NJ). Primers for Polymerase chain reaction (PCR) were synthesized to amplify each exon and a minimum of 35 bases of 5' and 3' flanking sequence. Sequencing and genotyping of the Coriell samples was performed in the genetics core facility at the University of California at San Francisco (UCSF). Primers for exons and adjoining intron regions (ca. 50 basepairs) for hPEPT2 (National Center for Biotechnology Information (NCBI) reference sequence: NM 021082.1) were designed using the Virtual Genome center website at http://alces.med.umn.edu/VGC.html and ordered from Operon (Alameda, CA). A collection of 247 ethnically identified genomic DNA samples was used to screen for hPEPT2 variants. PCR was performed in split 20 μL reactions using TaqGold (Perkin-Elmer, Boston, MA) on the GeneAmp 9700 thermocycler from Perkin-Elmer. One of three PCR protocols was chosen based on an optimization performed prior to amplifying all 247 samples. Occasionally, it was necessary to add glycerol or dimethyl sulfoxide or both to a reaction to optimize it. Agarose gel electrophoresis (2%, 150V, 35 minutes) was performed to verify the PCR products. Samples were pooled three-fold, and denaturing high pressure liquid chromatography (dHPLC) was performed on Varian HPLC machines with Varian Helix columns (Varian, Palo Alto, CA) to detect the presence of sequence variants. The dHPLC melt program described at http://insertion.stanford.edu/melt.html was used to approximate optimal temperature conditions for each run. Typically, the highest recommended temperature was run along with at least one lower temperature depending on the complexity of the melt profile. The results were scored manually. If a well scored positive, the three PCRs corresponding to that well were cleaned with 2 units of shrimp alkaline phosphatase and Exonuclease I from USB (Cleveland,

OH) and sequenced using ABI BigDye v2 (Applied Biosystems (ABI), Foster City, CA), cleaned with 96-well gel filtration blocks from Edge BioSystems (Gaithersburg, MD), and run on an ABI 3700 DNA Analyzer. Sequences were scored in Sequencher v4. All singleton SNPs were verified with a fresh PCR reaction and sequence.

Haplotype analysis:

Genotype data of the 14 most common SNPs from all 247 samples were converted into A or B. AA is the homozygous major allele sequence, AB is heterozygous, and BB represents the homozygous minor allele. The data were formatted for Fallin and Schork's SNPEM.2001 program, (Fallin and Schork, 2000; Fallin et al., 2001) which was compiled on a LINUX operating system. Up to 10 SNPs were run at a time over several runs, for example SNPs 5a, 9, 12c, 13b, 13c, 15, 17b, 20, 21a, and 22. The numbers in the SNP names correspond to exon position, and the lower case letter corresponds to SNP number in the exon (See table 1). In exons with only a single SNP, no letter is specified. For further information, refer to the SNP description of *hPEPT2* genotyping data at http://pharmacogenetics.ucsf.edu/

Construction of hPEPT2∗2 variant:

A 1.2 Kb fragment of *hPEPT2* was amplified by polymerase chain reaction from human cerebellum QUICK-Clone cDNA (BD Biosciences Clontech, Palo Alto, CA)(forward primer E7F: GAGAAGACTGCTATGCATTGG; reverse primer E20R: ACCTGTGACAGAGAACATGACCTC). The Quick-Clone cDNA was made from the cerebella of 24 pooled male and female Caucasians. The collection of PCR fragments, representing multiple hPEPT2 variants was ligated into the SmaI site of pBluescript (Stratagene, La Jolla,

CA). Several DNA minipreps were screened by sequencing to obtain a 1.2 Kb fragment of the hPEPT2∗2 variant. The NsiI/ XbaI fragment of hPEPT2∗2 was cut out and ligated into the pTLN2 vector containing the hPEPT2∗1 variant, obtained from M. Hediger (Harvard University), with its NsiI and XbaI fragment removed. The final construct was verified by sequence analysis.

Identification of partial chimpanzee sequence:

Primers used to amplify a 1.2 Kb fragment from human cDNA (same as above) were used to amplify the same fragment in a first strand synthesis cDNA library made from chimpanzee (*pan troglodyte*) mRNA (Lena 10-77). Chimpanzee RNA was kindly provided by Daniel J. Birmingham (Ohio State University). The amplified fragment was sequenced using internal human gene specific primers.

Sequence alignments

Alignments were conducted online using the CLUSTAL W algorithm. Input sequences were in FASTA format. Accession number for sequences: Reference sequence (human *hPEPT2*1*) NM_021082; *hPEPT2*2* XM00292; Rhesus monkey (*Macaca mulatto*) AAQ54588; Chimpanzee (*Pan troglodytes hPEPT2*) genomic sequence from Washington University sequencing center.

Cell culture, transfection, and uptake experiments

CHO cells were cultured under standard conditions. The cells were seeded onto 12 well plates, and transfected once they reached approximately 50 % confluency. pcDNA 3.1 containing wild-

JPET Fast Forward. Published on July 28, 2004 as DOI: 10.1124/jpet.104.073098
This article has not been copyedited and formatted. The final version may differ from this version.

type *hPEPT2* and its variant were transfected into CHO cells using Jet-PEI (Qbiogene, Carlsbab, CA). 24h after transfection, the cell monolayers were used for experiments. *Uptake experiments.* Cells were grown on TC-treated 12 well plates (Corning, Acton, MA). Uptake of [^{14}C]-Gly-Sar ([^{14}C]-Gly-Sar (Perkin-Elmer) with a specific activity 49.9 mCi · mmol^{-1}) was measured in Hanks Balanced Salt solution (HBSS) supplemented with 0.05% bovine serum albumin and buffered with 10 mM 2-[*N*-morpholino] ethanesulfonic acid (MES) with pH adjusted to 6.0. Cells were placed on a shaking plate, pre-heated to 37°C, and allowed to equilibrate for 10 min in the buffer solution. The experiment was started by adding 0.5 ml fresh buffer per well containing Gly-Sar (total concentration from 40-1000 μM) including 2 μCi/ml [^{14}C]-Gly-Sar, and 1 μCi/ml [^{3}H]-mannitol (specific activity 51.5 mCi · mmol^{-1} (Perkin-Elmer)) as a marker of extracellular space. The uptake of Gly-Sar in transfected cells was linear for more than 30 min, tested at extracellular Gly-Sar concentrations of 40 and 1000 μM. Uptake experiments were terminated after 30 minutes by removing the uptake medium, followed by three washes of the monolayers with ice-cold HBSS. The cells were detached from the well with 0.1% TritonX (polyethylene glycol tert-octylphenyl ether) in phosphate buffered saline. Radioactivity was determined in a liquid scintillation counter. Experiments were performed using 4-9 individual transfections. *pH-dependent uptake.* We investigated the pH-dependence of the Gly-Sar transport via the two PEPT2 variants. To ensure that the intracellular pH of the CHO cells was not determined by the extracellular pH, and to eliminate the proton-motive force across the cell membrane, we cultured the CHO cells on permeable polycarbonate filter support with a cell growth area of 1.1 cm^2. This allowed us to incubate the cells with a pH 7.4 buffer before the experiments and to change pH on the donor side, while maintaining the pH of 7.4 on the trans-side and thus, minimizing the rate of intracellular pH change caused by the buffer. The lower

JPET Fast Forward. Published on July 28, 2004 as DOI: 10.1124/jpet.104.073098
This article has not been copyedited and formatted. The final version may differ from this version.

chamber contained 1 ml HEPES buffer [N-(2-Hydroxyethyl)piperazine-N'-(2-ethanesulfonic acid); 4-(2-Hydroxyethyl)piperazine-1-ethanesulfonic acid] (pH 7.4), and the upper chamber 0.5 ml 40 μM [^{14}C]-Gly-Sar solution in 10 mM MES buffer of varying pH values. Uptake into the cells was measured over a 5 min period. Experiments were performed with 6 individual transfections, all occurring after seeding. To validate the method we took samples from the trans-side, and here we did not detect any [^{14}C]-Gly-Sar counts, indicating that the donor and trans solutions did not mix during the experiment. *Affinity studies.* The affinity experiment was initiated by adding 0.5 mL MES-buffer (pH 6.0) containing 2.0 μCi/ml [^{14}C]-Gly-Sar (40 μM) and various amounts of unlabeled test compounds to the transfected CHO cells. From this point on, the affinity experiment was similar to the Gly-Sar uptake experiment.

Kinetic Analysis

Concentration dependent Gly-Sar uptake. Uptake of Gly-Sar in transfected CHO cells was corrected for non-cellular uptake using mannitol as previously described, (Bravo et al., 2004), and the cellular uptake as a function of apical Gly-Sar concentration was fitted to a Michaelis-Menten type equation (Equation 1),

$$V = \frac{V_{\max} \times [S]^n}{K_m{}^n + [S]^n}$$

Equation 1

where V = uptake rate (pmol × cm^{-2} × min^{-1}), V_{max} = maximum uptake rate (pmol × cm^{-2} × min^{-1}), K_m = the Michaelis-Menten constant (μM), [S] = Gly-Sar concentration (μM), and n = Hill-factor. The Hill-factor was between 0.94 and 0.96.

JPET Fast Forward. Published on July 28, 2004 as DOI: 10.1124/jpet.104.073098
This article has not been copyedited and formatted. The final version may differ from this version.

Affinity experiments. Affinity for hPEPT2 in transfected CHO cells was determined from inhibition curves of 40 μM [^{14}C]-Gly-Sar uptake in the presence of varying concentrations of test compound. The degree of inhibition was fitted to a Michaelis-Menten type equation (Equation 2),

$$1-(U/U_0) = \frac{(1-(U/U_0))_{\max} \times [I]^n}{IC_{50}{}^n + [I]^n}$$

Equation 2

where U = uptake of [^{14}C]-Gly-Sar, U_0 = uptake of [^{14}C]-Gly-Sar at zero inhibitor concentration, IC_{50} = the concentration required to inhibit the maximum uptake by 50% (μM), [I] = compound concentration (μM), and n = Hill-factor. The Hill factor was between 0,96 and 1.14. K_i-values were calculated as described by Cheng and Prusoff (Cheng and Prusoff, 1973) using the K_m-values for Gly-Sar obtained for the two variants.

Kidney tissue preparation for allele-specific assays.

Genomic DNA and RNA were prepared from frozen kidney and tissue autopsy samples (obtained from the Ohio State University tissue bank). Frozen tissue samples were pulverized under liquid nitrogen. The frozen powder was portioned into aliquots for DNA and RNA extractions. DNA was prepared by digestion of the pellet or frozen powder with sodium dodecyl sulfate and proteinase K followed by sodium chloride precipitation of proteins (Miller et al., 1988). The DNA in the supernatant was further purified and recovered by ethanol precipitation. For RNA preparation, the starting material was homogenized in TRIzol ™ reagent (monophasic solution of phenol and guanidine isothiocyanate) (Invitrogen, Carlesbad, CA), and extracted with chloroform. The RNA was recovered by precipitation with isopropanol followed by

centrifugation. For additional purification, the RNA precipitate was dissolved in RNase free water or Qiagen buffer (Qiagen, Valenia, CA), then extracted using Qiagen RNeasy™ columns according to the manufacturers instructions. Complimentary DNA (cDNA) was generated from the mRNA by Superscript II reverse transcriptase (Invitrogen) employing oligo dT as a primer, to selectively bind to the polyA tail of mRNA. This ensured that all of the gene transcripts were represented at a rate that reflected the original abundance of each gene product.

SNaPshot assay and quantitative analysis of allelic ratios in genomic DNA and mRNA

This is a primer extension method originally intended for use with the ABI 3100 sequencer. We have successfully adapted this genotyping assay for the ABI 3730, with a few modifications that were needed to accommodate data handling. A stretch of genomic DNA (~70 bp's) or cDNA was amplified by PCR, and the allelic ratio was measured by primer extension using fluorescently labeled terminator nucleotides. We assume that this amplifies both alleles equally. Genomic DNA serves as an internal control compared with amplification of cDNA of the same samples. Standard PCR conditions were used on 15 μl reactions. Amplification conditions consisted of 35 cycles of denaturation at 95^{o}C for 30 seconds, then primer annealing at 60^{o} C for 1 minute, followed by extension at 72^{o} C for 1 minute. (*hPEPT2* exon 14 amplification primer pairs: forward primer: AGGAAAATGGCTGTTGGTATGATC, reverse primer CGCAACTGCAAATGCCAG) After amplification, the reactions were treated with exonuclease I and bacterial Antarctic alkaline phosphatase (New England Biolabs, Beverly, MA). For the primer extension, a gene-specific primer was designed with its 3'-end one base from the SNP position (*hPEPT2* exon 14 extension primer: GCTGTTGGTATGATCCTAGC). Primers used to analyze the SNP in exon 15 were: forward primer GAAATGGCCCCAGCCC; reverse primer

CATCTGCCAGATTCAAGACTTGTAG; extension primer AACCTCCTGGGGACCTG. SnaPshot™ reagent from Applied Biosystems was used to incorporate a single fluorescently-labeled dideoxynucleotide to the 3' end of the primer in a template dependent manner. The final primer extension reactions were analyzed using an ABI 3730 capillary electrophoresis DNA instrument, and calculated with Gene Mapper ™ 3.0 (ABI) software. The data for each incorporated fluorescently labeled nucleotide was measured as a peak area, which is proportional to the amount of amplified allele. Heterozygous samples have two differently labeled peaks at approximately the same position. Since differing fluorophores may influence nucleotide incorporation and migration rates, the peak areas are not identical between two alleles present in equal abundance. Therefore, peak area ratios of genomic DNAs assumed to be present in equal amounts (ratio=1) were used to normalize average genomic and cDNA ratios of heterozygous samples. The mRNA results in Figure 5b are averages of 4 separate experiments, normalized to genomic DNA. Ratios are derived by dividing peak areas of each of the two alleles of *hPEPT2* (B/A) where **1* is the B allele and **2* is the A allele. Deviations from unity in the normalized peak area ratios between alleles in the cDNA are attributed to differences in allele-specific mRNA levels.

RESULTS:

Genotype analysis and identification of variants:

247 human genomic DNA samples from different ethnic groups were genotyped. The samples include 100 Caucasian, 100 African American, 30 Asian, 10 Mexican, and 7 Pacific Islander individuals. Samples were genotyped over all 22 exons of *hPEPT2*, and 27 single nucleotide polymorphisms (SNPs) were detected. The genotyping results are summarized in Table 1. Of the 27 SNPs, 8 were in coding regions with 4 SNPs being synonymous and 4 non-synonymous.

SNP frequencies. 17 SNPs occurred with frequencies greater than 1%, 14 SNPs were detected with allele frequencies greater than 7%, and 10 of these had frequencies of 48-49%. Of the latter, 3 were non-synonymous with a frequency of 49%. There are some notable ethnic differences with respect to genotype frequency of some SNPs. For example, the synonymous SNP in exon 9 ranges from 36.7% in Asians up to a frequency of 71.4% in Pacific Islanders, but all of the abundant SNPs are present in all ethnic groups we have tested. Frequencies and occurrence of the rare SNPs varied between ethnic groups. For example the second SNP in exon 21 is present in 3% of Asians and 1.8% of Caucasians, but is missing in the African American, Mexican and Pacific Islander populations tested.

The abundant non-synonymous SNPs affect the following amino acids: F350L (exon 13, rs2257212), located in transmembrane domain 8 near the cytoplasmic side of the membrane; S409P (exon 15, rs1143671) and K509R (exon 17, rs1143672), both located in the large extracellular domain. All three of these amino acids are indicated by black shading in Figure 1.

Haplotype:

Haplotypes were inferred using the estimation maximization (EM) algorithm (SNPEM.2001) (Fallin and Schork, 2000; Fallin et al., 2001). 14 SNPs with frequencies over 7% were included with the analysis. For simplicity, all of the genotype data for each SNP in every sample were converted to "A" or "B", where A usually represented the major allele, but in some cases the designation was arbitrary. Figure 3 depicts predicted haplotypes with frequencies over 1%. The haplotypes are grouped by similarity. The haplotype with the highest frequency (number 1) has the allele designation of "A" in all 14 SNPs. Since its frequency (38%) is much larger than that of other haplotypes, we considered haplotype 1 to be the reference**.** However, haplotype 1 does not represent the *hPEPT2* sequence cloned by Liu et al. (Liu et al., 1995), which we have termed here *hPEPT2*1*, because it was published first while not representing the most abundant haplotype variant.

Several patterns emerge in Figure 3. The 10 SNPs with frequencies over 49% are in near complete linkage disequilibrium and form the core of several related haplotypes. The core haplotype block begins at exon 9 and continues through the middle of exon 17, spanning at least 6500 base pairs of genomic DNA. These haplotypes include the 3 most abundant non-synonymous SNPs –each allele present in nearly equal abundance - that affect amino acid positions 350, 409 and 509; two synonymous SNPs at amino acid positions 284 (exon 9, rs2293616) and 387 (exon 14, rs1143670); and at least 5 additional SNPs located in introns. Haplotypes 7 through 12 contain B alleles in all 10 SNPs within the 6.5 kilobase (kb) block. Together they constitute one variant: *hPEPT2∗1.* The hPEPT2∗1 protein contains the amino acids L350, P409 and R509, as cloned by Liu et al (Liu et al., 1995). Haplotypes 1 through 5 are closely related, containing the reference genotype in all 10 SNPs within the 6.5 kb block and comprising the second variant *hPEPT2∗2.* The hPEPT2∗2 protein contains amino acids F350,

JPET Fast Forward. Published on July 28, 2004 as DOI: 10.1124/jpet.104.073098
This article has not been copyedited and formatted. The final version may differ from this version.

S409 and K509. Both the ∗1 and ∗2 variants are distributed roughly equally in the population at about 44 - 47%. There is a 3rd protein sequence variant represented with lower abundance (1.6%) (haplotype 6) termed hPEPT2∗3. It is the only haplotype where the 3 abundant non-synonymous SNPs are not in linkage disequilibrium.

Conservation of human coding region polymorphisms across primate sequences

We found five common SNPs that differ in the cDNA of the three human *PEPT2* variants. The approximately even distribution of the *hPEPT2∗1* and *hPEPT2∗2* alleles in the population suggests that the two alleles and their corresponding core haplotypes are ancient, raising the question how these two variants have evolved, carrying multiple phased SNPs. We compared the cDNA of the three human variants with cDNA from chimpanzee (*Pan troglodytes*) and rhesus monkey (*Macaca mulatto*) (Table 2). There were 8 single base pair differences in the coding region between the chimpanzee and human sequences: the 5 SNPs that differ between the human variants and 3 additional chimpanzee-specific SNPs. Two out of the three in common non-synonymous SNPs match the *hPEPT2∗2* variant (position 409 and 509) and the other SNP (350) matches the *hPEPT2∗1* variant. Hence, the chimpanzee sequence resembles the human *hPEPT2∗3* variant.

The cDNA from rhesus monkey had 38 base pair changes that were not found in either human or chimpanzee sequence, but overall the cDNA was 98.3% identical. Using the 8 nucleotides that differ between chimpanzee and human (Table 2), a comparison of the 5 primate sequences allowed the construction of an ancestral primate *PEPT2* allele, which resembles neither the **1* or **2* variant. SNPs in exon 14 (S409P) and 15 (K509R) most likely diverged from the ancestral allele before chimpanzees and humans diverged, generating a haplotype that

JPET Fast Forward. Published on July 28, 2004 as DOI: 10.1124/jpet.104.073098
This article has not been copyedited and formatted. The final version may differ from this version.

encodes *3. The exon 13 SNP, found only in the *2 variant, originated later, as did the SNPs in exon 9 and exon 17 which are specific to the **1* variant. Overall, the **1* variant is more similar to the ancestral primate allele, while the **2* variant is more similar to the chimpanzee sequence.

Functional analysis of the hPEPT2*1 and *2 variants

To test for the presence of functional differences between the two main protein variants, hPEPT2∗1 and ∗2, the cDNAs for both variants were subcloned into pcDNA3.1 expression vector and transiently transfected into CHO cells. Analysis of Gly-Sar uptake in transfected CHO cells (Figure 4a) showed that the two variants had similar V_{max} values (hPEPT2*1: 0.96±0.07 pmol·cm^{-2}·min^{-1}; hPEPT2*2: 1.11±0.07 pmol·cm^{-2}·min^{-}1. The *1 variant had a K_m of 83±16 μM, whereas the *2 variant had a significantly higher K_m value 233±38 μM ($p<0.05$, n=4-9). Since the transport of substrates by hPEPT2 is partly driven by the co-transport of protons, we investigated the effect of extracellular pH on the transport of Gly-Sar. For hPEPT2*1 the highest uptake was seen at pH 6.0 (Figure 4b), whereas for hPEPT2*2 uptake at pH 6.0 was significantly lower ($p<0.01$, n=6-8). At the other extracellular pH values investigated no difference in uptake was observed. The affinities of cephalexin and kyotorphin for hPEPT2 were also investigated in CHO cells transfected with *hPEPT2*1* and **2* (Figure 4c). The K_i values obtained for *1 and *2 were similar for both cephalexin (98±23 μM and 102±3 μM, respectively) and kyotorphin (5.1 ±0.3 μM and 6.0 ±1.2 μM, respectively).

Allele specific mRNA analysis in kidney tissue

In order to test whether the two variants of *hPEPT2* are associated with differences in mRNA expression in their native tissue, we measured the quantity of each allele in heterozygous

kidney samples using SNaPshot, and compared them to each other. Figure 5a is an example of allelic peaks obtained for one DNA and two RNA samples. Allelic ratios of the synonymous SNP in exon 14 (rs1143670) and the non-synonymous SNP in exon 15 (rs1143671) of *hPEPT2* were tested separately in DNA and mRNA from heterozygous kidney samples. Out of 20 kidney samples analyzed, 9 were heterozygous for the *1 and *2 SNPs. Allele-specific measurements were precise and highly reproducible for the exon 14 SNP in genomic DNA, deviating from unity with a S.D. of only 4% (Figure 5b). In contrast, significant deviations from unity were observed for allelic mRNA ratios (Figure 5b). We also measured allelic ratios in DNA and mRNA using the SNP in exon 15 (data not shown). While the variability of the results was somewhat greater than with the exon 14 SNP, the results were comparable, confirming the validity of the allele-specific analysis. In particular, only the B allele (representing variant*2*) was detectable in mRNA of the heterozygous sample K018; this was confirmed with both SNPs located in exons 14 and 15, in independent experiments. Complete absence of one allele suggest the presence of a mutation in the regulatory region, or one that is affecting mRNA processing (e.g., splicing), or gene silencing by epigenetic effects. We found 3 samples: K004 (B/A ratio 1.18±0.03), K005 (1.28±0.03) and K009 (1.52±0.22) where the *1 allele mRNA was significantly higher. The other 5 kidney samples did not exhibit significant allele-specific differences in gene expression

DISCUSSION:

A genetic and haplotype analysis of the gene encoding hPEPT2, *SLC15A2*, revealed the presence of 2 distinct, common variants that are present at a high frequency in all of the ethnic groups tested. The gene contains a block of 10 polymorphisms in a 6 kb region, including 3 non-synonymous SNPs. All of the SNPs in this block have the same allele frequency and are in near complete linkage disequilibrium with each other. There are additional genomic SNPs outside of this block, which increases the number of haplotypes. The two common haplotype groups encode two protein variants, termed *hPEPT2*1* and *hPEPT2*2*, that differ by 3 amino acids. *hPEPT2*3* is a less frequent variant with a slightly different pattern of non-synonymous SNPs.

The presence of several linked SNPs with similar allele frequencies that form a haplotype block with only two major variants, suggested that the haplotype block is ancient and highly resistant to cross-over events. It also suggested the possibility that all of these polymorphisms arose together. In order to examine the evolution of the haplotypes, we compared the cDNA sequence of the three human *PEPT2* variants with cDNA from chimpanzee and rhesus monkey. Humans and chimpanzees diverged from each other approximately 3.5 to 5.5 million years ago (Sarich and Cronin, 1976)and the genomes are 95% identical (Britten, 2002). In contrast, rhesus monkeys diverged from other primates approximately 50 million years ago (Martin, 1993), consistent with 38 additional base pair differences in the rhesus sequence absent in the chimpanzee or human sequences. We inferred an ancestral primate *PEPT2* allele from sequence comparisons showing that the **1* variant more closely resembles the ancestral primate allele, while the **2* variant is more similar to the chimpanzee sequence. The comparison also indicated that the block of polymorphisms did not arise at the same time. However, if the SNPs arose independently, we would have expected to find multiple transitional haplotypes in the human

JPET Fast Forward. Published on July 28, 2004 as DOI: 10.1124/jpet.104.073098
This article has not been copyedited and formatted. The final version may differ from this version.

population, which is not the case. The rare **3* variant encoded by haplotype 6 does fit the description of a transitional haplotype. Considering that the history of human evolution is full of bottle necks and dead ends, perhaps there are no additional human haplotypes in the *PEPT2* block because only families carrying the 3 current *hPEPT2* alleles survived.

Our *in vitro* experiments suggest that the two main variants differ significantly in biochemical properties. The Gly-Sar K_m value for the ∗1 variant (83±16 μM) is similar to the K_m values previously obtained for hPEPT2 and rPEPT2 (from rat). The K_m for Gly-Sar uptake via hPEPT2 has previously been reported to be 74 and 110 μM, (Ramamoorthy et al., 1995; Fujita et al., 2004) and 90 μM for rPEPT2. (Bravo et al., 2004) However, the K_m value for the *2 variant (233±38 μM) is significantly higher than for the *1 variant. Yet, the inhibitory potencies of cephalexin and kyotorphin on Gly-Sar uptake were similar for both variants. The inhibitory potency of cephalexin for both variants (98±23 μM for *1 and 102±3 μM for *2) are similar to reported values for rPEPT2 (63 and 73 μM) (Ganapathy et al., 1995; Brandsch et al., 1997). On the other hand, the affinity of kyotorphin for rPEPT2 has been reported to be 29 μM, which is higher than the values obtained in this study for hPEPT2 (5.1 ±0.3 μM for *1 and 6.0 ±1.2 μM for *2). The reasons for this discrepancy remain to be clarified. Because the two variants differed in their pH dependence, yielding disparate Gly-Sar Km values, while the affinities for cephalexin and kyotorphin did not differ, it appears that the functional difference between the two variants stems from the proton-dependent part of the translocation cycle. Alternatively, cephalexin and kyotorphin may bind differently from Gly-Sar to hPEPT2. Further studies are needed to determine kinetic differences for drug substrates between the hPEPT2 variants.

The sequencing analysis had focused on coding regions of *hPEPT2* and only portions of intronic splice regions and regulatory domains. Functional diversity is likely also to arise from

polymorphisms that affect transcription or mRNA processing. However, responsible functional polymorphisms are difficult to identify. Therefore, we have developed an experimental approach to determine whether *hPEPT2* harbors any such *cis*-acting polymorphisms. Allele-specific mRNA analysis in relevant target tissues is a potentially powerful tool for detecting *cis*-acting functional polymorphisms. To reflect the native regulatory environment of tissues where hPEPT2 is functionally expressed, this type of analysis must be performed on DNA and mRNA extracted from the target tissues. We have developed a primer extension assay capable of precisely measuring allelic differences in hPEPT2 mRNA levels in kidney tissues, using two indicator SNPS present in both genomic DNA and mRNA. Even though our sample was small (n=9 kidneys heterozygous for the marker SNPs (out of 20 kidney samples analyzed), significant allelic differences were observed in three samples. Moreover, only the *hPEPT2*2* allele was expressed in a fourth heterozygous kidney sample, indicating the presence of a mutation in *hPEPT2*1* that prevented mRNA accumulation. However, we cannot rule out gene imprinting or other epigenetic events. *hPEPT2* allelic mRNA differences were not associated solely with one of the two main variants, suggesting that *cis*-acting polymorphisms affecting mRNA levels are present, but are likely to be located in a separate haplotype block. Using allelic differences in mRNA expression, we can now begin to identify the mechanisms and functional polymorphisms underlying this observation. Given the magnitude of the expression differences, it is possible that a combination of allele-specific biochemical differences at the protein level in concert with *cis*-regulatory polymorphisms could exert strong effects on drug disposition of various hPEPT2 substrates in some individuals. Dose-dependent renal clearance of cefadroxil in healthy human males has been demonstrated, and the renal clearance was significantly increased in the presence of high doses of cephalexin (Garrigues et al., 1991). Although the contribution of hPEPT2 to the

clearance was not specifically investigated, both cephadroxil and cephalexin are known substrates for hPEPT2. The cefadroxil plasma concentrations in this study were in μM range, which indicate that the transporter involved in the reabsorption of the compound is more likely to be the high affinity transporter, hPEPT2, rather than hPEPT1. Increased urinary excretion of Gly-Sar-Sar in the presence of high amounts of the peptide has also been shown in perfused rat kidneys (Minami et al., 1992). In this study, the initial concentration of Gly-Sar-Sar in the perfusate was 100 μM, which again falls into the range of recognition by PEPT2 rather than PEPT1.

In summary, *SLC15A22* encodes two main variants of the human di/tri-peptide transporter hPEPT2 (*1 and *2) that differ in biochemical properties; moreover, *cis*-acting polymorphisms appear to affect expression levels in kidneys. As both protein variants are highly abundant in all populations, it is likely that these variants have not been subject to selection pressures in evolution. However, the impact of *hPEPT2* sequence variations on drug disposition could be substantial and needs to be studied further.

Acknowledgements

We are grateful to Danielle Fallin (Johns Hopkins University) for providing the SNPEM.2001 software, to Audrey Papp for SNaPshot analysis and helpful discussions about the project, Gloria Smith for kidney sample DNA/RNA preparation, and Daniel J. Birmingham (Ohio State University) for providing chimpanzee RNA. Maria Pedersen and Dr. Jan Amstrup (University of Copenhagen) have contributed to the functional analysis of the hPEPT2 variants.

REFERENCES

Arvidsson A, Borgå O and Alván G (1979) Renal excretion of cephapirin and cephaloridine:evidence for saturable tubular reabsorption. *Clin Pharmacol Ther* **25**:870-876.

Berger U and Hediger M (1999) Distribution of peptide transporter PEPT2 mRNA in the rat nervous system. *Anat Embryol* **199**:439-449.

Brandsch M, Brandsch C, Ganapathy M, Chew C, Ganapathy V and Leibach F (1997) Influence of proton and essential histidyl residues on the transport kinetics of the H+/peptide cotransport systems in intestine (PEPT 1) and kidney (PEPT 2). *Biochim Biophys Acta* **1324**:251-262.

Bravo S, Nielsen C, Amstrup J, Frokjaer S and Brodin B (2004) Epidermal growth factor decreases PEPT2 transport capacity and expression in the rat kidney proximal tubule cell line SKPT0193 cl.2. *Am J Physiol Renal Physiol* **286**:F385-F393.

Britten R (2002) Divergence between samples of chimpanzee and human DNA sequences is 5%, counting indels. *Proc Natl Acad Sci U S A* **99**:13633-13635.

Cheng YC and Prusoff WH (1973) Relationship between the inhibition constant (Ki) and the concentration of inhibitor which causes 50 per cent inhibition (I50) of an enzymatic reaction. *Biochem Pharmacol* **22**:3099-3108.

Daniel H and Adibi S (1993) Transport of beta-lactam antibiotics in kidney brush border membrane. Determinants of their affinity for the oligopeptide/H+ symporter. *J Clin Invest* **92**:2215-2223.

Daniel H, Morse E and Adibi S (1991) The high and low affinity transport systems for dipeptides in kidney brush border membrane respond differently to alterations in pH gradient and membrane potential. *J Biol Chem* **266**:19917-19924.

Fallin D, Cohen A, Essioux L, Chumakov I, Blumenfeld M, Cohen D and Schork N (2001) Genetic analysis of case/control data using estimated haplotype frequencies: application to APOE locus variation and Alzheimer's disease. *Genome Res* **11**:143-151.

Fallin D and Schork NJ (2000) Accuracy of haplotype frequency estimation for biallelic loci, via the expectation-Maximization algorithm for unphased diploid genotype data. *Am J Hum Genet* **67**:947-959.

Fujita T, Kishida T, Wada M, Okada N, Yamamoto A, Leibach F and Ganapathy V (2004) Functional characterization of brain peptide transporter in rat cerebral cortex: identification of the high-affinity type H+/peptide transporter PEPT2. *Brain Res* **997**:52-61.

Ganapathy M, Brandsch M, Prasad P, Ganapathy V and Leibach F (1995) Differential recognition of beta -lactam antibiotics by intestinal and renal peptide transporters, PEPT 1 and PEPT 2. *J Biol Chem* **270**:25672-25677.

Ganapathy M, Huang W, Wang H, Ganapathy V and Leibach F (1998) Valacyclovir: a substrate for the intestinal and renal peptide transporters PEPT1 and PEPT2. *Biochem Biophy Res Commun* **246**:470-475.

Ganapathy M, Prasad P, Mackenzie B, Ganapathy V and Leibach FH (1997) Interaction of anionic cephalosporins with the intestinal and renal peptide transporters PEPT 1 and PEPT 2. *Biochim Biophys Acta* **1324**.

Garrigues T, Martin U, Peris-Ribera J and Prescott L (1991) Dose-dependent absorption and elimination of cefadroxil in man. *Eur J Clin Pharmacol.* **41**.

Graul RC and Sadée W (1997) Sequence Alignments of the H+-Dependent Oligopeptide Transporter Family. Inferences on Structure and Function of the Intestinal PET1 Transporter. *Pharm Res* **14**:388-400.

Groneberg D, F. Doring, S.Theis, M. Nickolaus, Fischer A and Daniel. H (2002) Peptide transport in the mammary gland: expression and distribution of PEPT2 mRNA and protein. *Am J Physiol Endocrinol Metab* **282**:E1172-E1179.

Groneberg D, Nickolaus M, Springer J, Doring F, Daniel H and Fischer A (2001) Localization of the peptide transporter PEPT2 in the lung: implications for pulmonary oligopeptide uptake. *Am J Pathol* **158**:707-714.

Inui K, Masuda S and Saito H (2000a) Cellular and molecular aspects of drug transport in the kidney. *Kidney Int* **58**:944-958.

Inui K, Terada T, Masuda S and Saito H (2000b) Physiological and pharmacological implications of peptide transporters, PEPT1 and PEPT2. *Nephrol Dial Transplant.* **15**:11-13.

Liu W, Liang R, Ramamoorthy S, Fei Y-J, Ganapathy ME, Hediger MA, Ganapathy V and Leibach FH (1995) Molecular cloning of PEPT2, a new member of the H(+)/peptide cotransporter family, from human kidney. *Biochim. Biophys. Acta.* **1235**:461-466.

Martin R (1993) Primate origins: plugging the gaps. *Nature* **363**:223-234.

Miller S, Dykes D and Polesky H (1988) A simple salting out procedure for extracting DNA from human nucleated cells. *Nucleic Acids Res* **16**:1215.

Minami H, Daniel H, Morse E and Adibi S (1992) Oligopeptides: mechanism of renal clearance depends on molecular structure. *Am J Physiol* **263**:F109-F115.

Nielsen C, Brodin B, Jorgensen F, Frokjaer S and Steffansen B (2002) Human Peptide Transporters: therapeutic applications. *Exp Opin Ther Patents* **12**:1329-1350.

Ramamoorthy S, Liu W, Ma Y, Yang-Feng T, Ganapathy V and Leibach F (1995) Proton/peptide cotransporter (PEPT 2) from human kidney: functional characterization and chromosomal localization. *Biochim Biophys Acta* **1240**:1-4.

Sarich V and Cronin J (1976) Molecular systematics of the primates, in *Molecular Anthropology* (Tashian R ed) pp 141-170, Plenum Press, New York.

Shu C, Shen H, Teuscher N, Lorenzi P, Keep R and Smith DE (2002) Role of PEPT2 in peptide/mimetic trafficking at the blood-cerebrospinal fluid barrier: Studies in rat choroid plexus epithelial cells in primary culture. *J Pharmacol Exp Ther* **301**:820-829.

Smith D, Pavlova A, Berger U, Hediger M, Yang T, Huang Y and Schnermann J (1998) Tubular localization and tissue distribution of peptide transporters in rat kidney. *Pharm Res* **15**:1244-1249.

Terada T, Irie M, Okuda M and Inui K (2004) Genetic variant Arg57His in human H+/peptide cotransporter 2 causes a complete loss of transport function. *Biochem Biophys Res Commun* **316**:416-420.

Zhu T, Chen X, Steel A, Hediger M and Smith D (2000) Differential recognition of ACE inhibitors in Xenopus laevis oocytes expressing rat PEPT1 and PEPT2. *Pharm Res* **17**:526-532.

Footnotes:

a) Financial support: This work was in part supported by the UCSF Plasma Membrane Transporter Group (Grant GM61390 from General Medical Sciences, NIH), by funds form the Ohio State University, and The Carlsberg Foundation (CUN).

b) Reprint requests:
Wolfgang Sadée
Department of Pharmacology
Ohio State University
333 West 10th Avenue
Columbus OH 43210-1239
Phone: 614-292 5593
FAX: 614-292-7232
Email: sadee-1@medctr.osu.edu

c) [1]Department of Pharmacology Program in Pharmacogenomics, College of Medicine and Public Health, Ohio State University, Columbus, Ohio.

[2]Molecular Biopharmaceutics, Department of Pharmaceutics, The Danish University of Pharmaceutical Sciences, Universitetsparken 2, DK-2100 Copenhagen, Denmark

Figure Legends:

Fig 1: Predicted hPEPT2 secondary structure. There are 12 transmembrane domains and a large extracellular domain located between transmembrane domains 8 and 9. This image was modified from an image generated using TOPO, a transmembrane protein display program, by S.J. Johns of UCSF and R.C. Speth of WSU. The three most abundant non-synonymous amino acid changes are shaded black.

Fig. 2: The genomic structure of *SLC15A2*, the gene encoding the H^+/peptide transporter PEPT2.

Fig. 3: A summary of the haplotype data for the 14 most abundant SNPs with frequencies above 7%. Each column, beginning with the one labeled 5a represents a SNP. The top row specifies the exon (number) and position of the SNP (lower case letter if more than one per exon was identified. See Table 1). The second row specifies the base pair change for each SNP. The letter on the left is the reference sequence and the letter on the right is the variant. The third row from the top indicates whether the SNP is intronic or if coding and non-synonymous, the amino acid change caused by the SNP. The row at the bottom lists the genotype frequencies obtained for each SNP. The first column on the left identifies each haplotype with a number, 1 through 12. The second column from the left indicates which protein variant: ∗1, ∗2 or ∗3 the haplotype encodes. The column on the far right lists the haplotype frequencies obtained from the SNPEM.2001 algorithm. Frequencies in bold text indicate haplotypes that are high frequency. The middle blocks represent each SNP. White blocks labeled “A” are the reference sequence. Gray blocks labeled “B” are variants.

JPET Fast Forward. Published on July 28, 2004 as DOI: 10.1124/jpet.104.073098
This article has not been copyedited and formatted. The final version may differ from this version.

Figure 4: **a**) Gly-Sar concentration dependent effect of apical uptake of [^{14}C]-Gly-Sar in CHO cells transfected with *hPEPT2*1* (•) or *hPEPT2*2* (o). The apical uptake of Gly-Sar was measured over a 30 min incubation period. The lines describe the best fit to the experimentally obtained points using the Michaelis-Menten equation. Each data point represents the mean ± SE of 4-9 individual transfections. hPEPT2*1: Km=82.9±16, Vmax=0.96±0.07 pmol/(cm2*min), n=1.01±0.16. hPEPT2*2: Km=233±38, Vmax=1.11±0.07 pmol/(cm2*min), n=0.94±0.07. **b**) pH dependence of Gly-Sar uptake. CHO cells were grown on filter support. The lower chamber contained a HEPES buffer (pH 7.4) solution, and the upper chamber had a 40 μM Gly-Sar solution in MES buffer with varying pH values. Uptake was measured over a 5 min period. Each point represents average ± SE of 3-9 individual transfections. **c**) Inhibition of Gly-Sar uptake. The uptake of 40μM Gly-Sar in MES-buffer at an extracellular pH of 6.0 was measured in the presence of varying amount of kyotorphin and cephalexin in transfected CHO cells. The lines describe the best fit to the experimentally obtained points using Equation 2. Each data point represents the mean ± SE of 6 individual transfections.

Figure 5: Allele-specific mRNA analysis of *Pept2* mRNA from heterozygous kidney samples. **a**) Allele-specific quantification of DNA and mRNA from heterozygous samples. Genomic DNA is shown for one sample K009, but it is representative of all of the genomic DNA samples. Allele-specific levels in cDNA in sample K020 are not significantly different from DNA, however levels in K009 deviate significantly. **b**) Allelic ratios showing *1/*2 (or B/A) are calculated for genomic DNA and RNA. A ratio of 1 indicates that both alleles are expressed equally. Ratios greater than 1 indicate that the *1 allele is expressed more than *2. Ratios under 1 indicate that the *2 allele is expressed more. RNA samples with significant allelic differences are marked ($^{\#}$).

JPET Fast Forward. Published on July 28, 2004 as DOI: 10.1124/jpet.104.073098
This article has not been copyedited and formatted. The final version may differ from this version.

Table 1: Genotyping of exons and adjoining intronic regions of the hPEPT2 transporter. * When more than one SNP per exon: a=1, b=2, c=3, d=4. [a]AA: African American, [b]CA: Caucasian, [c]AS: Asians (Japanese, Chinese, SE Asians), [d]ME: Mexicans, [e]PI: Pacific Islander

Exon	SNP*	CDS Pos	Exon Pos	Nucleotide Change	Amino Acid Position	Amino Acid Change	Total	AA[a]	CA[b]	AS[c]	ME[d]	PI[e]
1			(+13)	A → G	-	-	0.002	0.005	0	0	0	0
2			(-41)	A → T	-	-	0.004	0	0	0	0.1	0
5	a		(-27)	G → A	-	-	0.248	0.148	0.323	0.217	0.3	0.643
5	b		(+26)	T → G	-	-	0.033	0.005	0.035	0.05	0.25	0
9		852	72	A → G	284	syn	0.49	0.459	0.52	0.367	0.7	0.714
10		895	28	C → T	299	syn	0.002	0.005	0	0	0	0
11			(-21)	T → A	-	-	0.492	0.465	0.52	0.362	0.7	0.714
12	a		(-28)	T → A	-	-	0.49	0.465	0.51	0.383	0.7	0.714
12	b		(+69)	C → A	-	-	0.002	0	0	0.017	0	0
12	c		(+73)	A → T	-	-	0.494	0.465	0.52	0.383	0.7	0.714
12	d		(+107)	A → G	-	-	0.494	0.465	0.52	0.383	0.7	0.714
13	a		(-40)	A → G	-	-	0.025	0.045	0.01	0	0.05	0
13	b	1048	13	T → C	350	Phe → Leu	0.498	0.475	0.525	0.375	0.65	0.714
13	c		(+13)	C → T	-	-	0.498	0.48	0.52	0.375	0.65	0.714
14		1161	37	G → A	387	syn	0.504	0.505	0.505	0.383	0.7	0.714
15		1225	19	T → C	409	Ser → Pro	0.488	0.465	0.51	0.367	0.7	0.714
16	a	1436	95	A → C	479	His → Pro	0.002	0	0.005	0	0	0
16	b	1443	102	G → T	481	syn	0.004	0.01	0	0	0	0
17	a		(-9)	C → T	-	-	0.002	0	0.005	0	0	0
17	b	1526	20	A → G	509	Lys → Arg	0.492	0.464	0.52	0.367	0.7	0.714
17	c		(+39)	T → C	-	-	0.004	0.01	0	0	0	0
19	a		(+29)	C → A	-	-	0.002	0	0	0.017	0	0

19	b	(+41)	A → C	-	-	0.002	0	0.005	0	0	0
20		(+34)	A → G	-	-	0.073	0.06	0.07	0.133	0	0.143
21	a	(-34)	A → G	-	-	0.231	0.275	0.19	0.25	0.125	0.25
21	b	(-29)	G → A	-	-	0.014	0	0.03	0.018	0	0
22		(+6)	T → C	-	-	0.103	0.065	0.13	0.133	0.2	0

Table 2: The genotype of 8 SNPs from human, chimpanzee and rhesus monkey cDNA. [#]Rhesus monkey cDNA contains 38 additional sequence variations not present in human and chimpanzee sequences and is 98.3% identical. The chimpanzee and human cDNA sequences are 99.7% identical. The Ancestral primate allele was determined by comparing the human sequences with sequences from chimpanzee and rhesus monkey, giving more weight to the rhesus monkey genotypes in cases where a SNP is found in multiple sequences (example: SNPs in exons 14 and 15).

		Exon9	Exon13	Exon14	Exon15		Exon17	
		Syn	F350L	Syn	S409P		K509R	
Relative cDNA Position	826	889	1085	1198	1262	1384	1563	1586
Chimpanzee	G	A	C	G	T	G	A	T
Human *3	T	A	C	G	T	A	A	C
Human *2	T	A	T	G	T	A	A	C
Human *1	T	G	C	A	C	A	G	C
Rhesus monkey[#]	T	A	C	A	C	A	A	C
Ancestral Primate	T	A	C	A	C	A	A	C

JPET Fast Forward. Published on July 28, 2004 as DOI: 10.1124/jpet.104.073098
This article has not been copyedited and formatted. The final version may differ from this version.

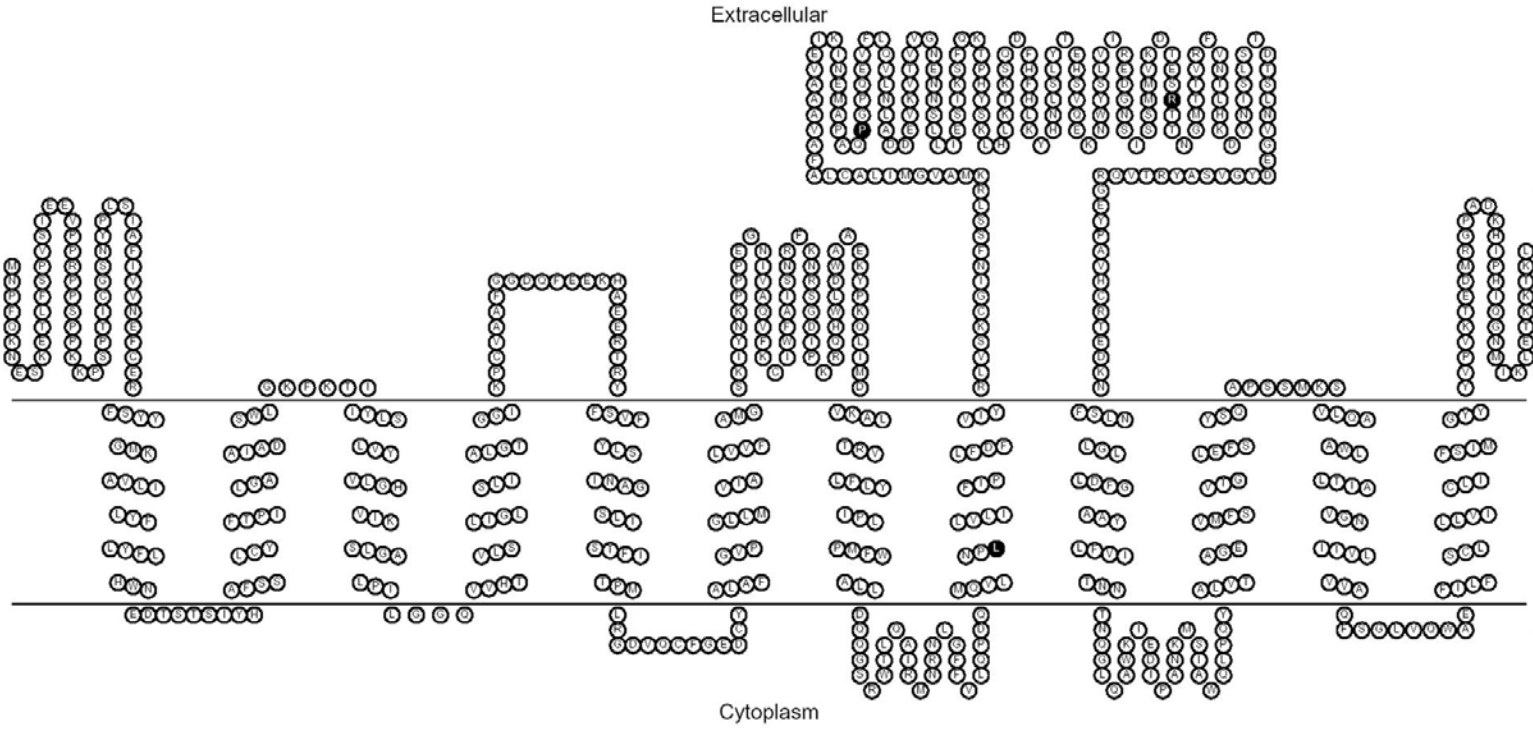


Figure 1

JPET Fast Forward. Published on July 28, 2004 as DOI: 10.1124/jpet.104.073098
This article has not been copyedited and formatted. The final version may differ from this version.

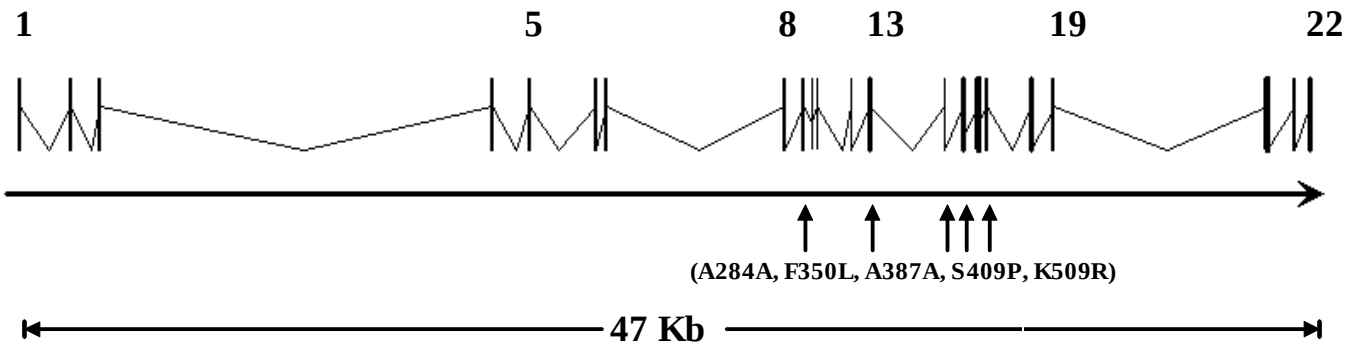


Figure 2

JPET Fast Forward. Published on July 28, 2004 as DOI: 10.1124/jpet.104.073098
This article has not been copyedited and formatted. The final version may differ from this version.

	Exon:	5a	9	11	12a	12c	12d	13b	13c	14	15	17b	20	21a	22	
		G/A	A/G	T/A	T/A	A/T	A/G	T/C	C/T	G/A	T/C	A/G	A/G	A/G	T/C	
Haplotype Number	Protein Variant	Intron	A284A	Intron	Intron	Intron	Intron	F350L	Intron	A387A	S409P	K509R	Intron	Intron	Intron	Haplotype Frequency
1	*2	A	A	A	A	A	A	A	A	A	A	A	A	A	A	**0.379**
2	*2	A	A	A	A	A	A	A	A	A	A	A	A	A	B	0.046
3	*2	A	A	A	A	A	A	A	A	A	A	A	A	B	A	0.016
4	*2	A	A	A	A	A	A	A	A	A	A	A	B	A	A	0.026
5	*2	A	A	A	A	A	A	A	A	A	A	A	B	A	B	0.013
6	*3	A	A	A	A	A	A	B	B	A	A	A	A	A	A	0.016
7	*1	B	B	B	B	B	B	B	B	B	B	B	A	A	A	**0.191**
8	*1	B	B	B	B	B	B	B	B	B	B	B	B	A	A	0.015
9	*1	B	B	B	B	B	B	B	B	B	B	B	A	A	B	0.021
10	*1	A	B	B	B	B	B	B	B	B	B	B	A	B	A	**0.175**
11	*1	A	B	B	B	B	B	B	B	B	B	B	A	B	B	0.013
12	*1	A	B	B	B	B	B	B	B	B	B	B	A	A	A	0.026
	Genotype Frequency	0.248	0.490	0.492	0.490	0.494	0.494	0.498	0.498	0.504	0.488	0.492	0.073	0.231	0.103	total: 0.935

A	Reference Sequence
B	Variant

Figure 3

JPET Fast Forward. Published on July 28, 2004 as DOI: 10.1124/jpet.104.073098
This article has not been copyedited and formatted. The final version may differ from this version.

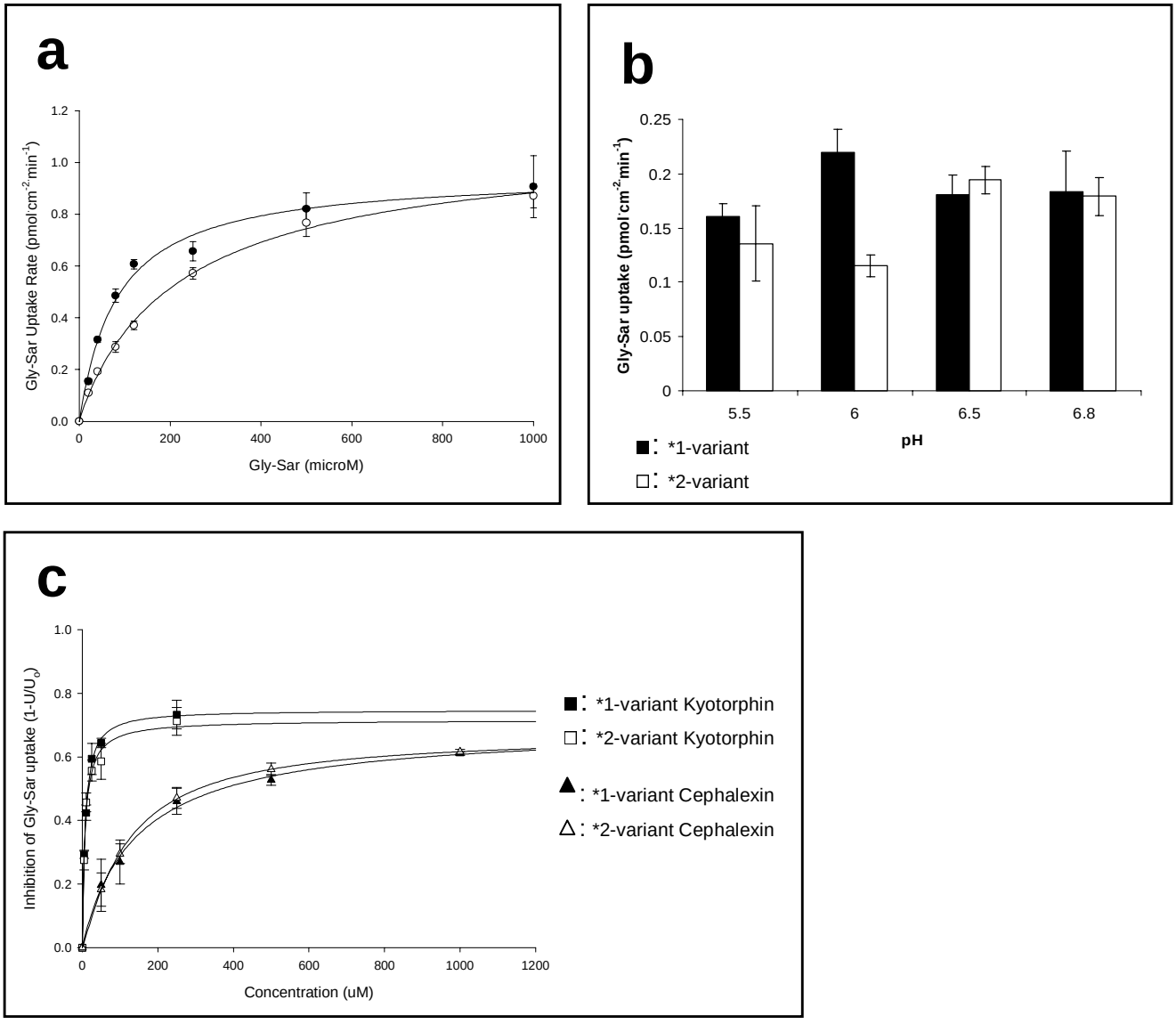
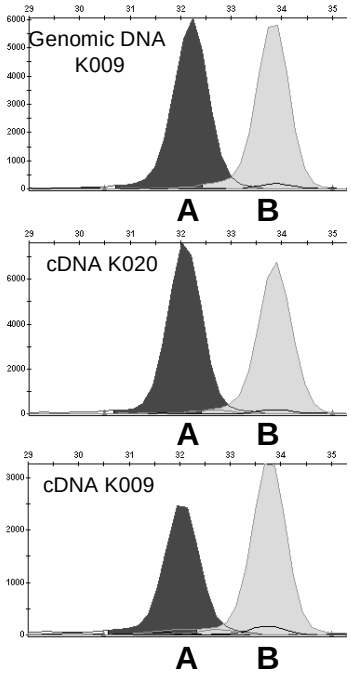


Figure 4

JPET Fast Forward. Published on July 28, 2004 as DOI: 10.1124/jpet.104.073098
This article has not been copyedited and formatted. The final version may differ from this version.

a



b

Kidney Samples	B/A DNA	B/A RNA	Count	Average St. Dev
K004	0.96	1.18[#]	4	0.03
K005	1.05	1.28[#]	4	0.03
K006	1.03	0.91	4	0.05
K007	1.02	1.03	4	0.07
K009	0.98	1.52[#]	4	0.22
K011	0.94	0.95	3	0.02
K018	1.00	0[# $]	4	0
K019	1.00	0.96	4	0.06
K020	1.05	1.03	4	0.02
Average	1.00	0.98		
St. Dev	0.04	0.42		
Maximum	1.05	1.52		
Minimum	0.94	0		

[#]Ratios deviated significantly from 1.00
[$]Allele B was undetectable in this sample.

Figure 5