

Six mRNAs with different 5' ends are encoded by a single γ -glutamyltransferase gene in mouse

(transcription/kidney/promoter/gene expression)

SRIDHARAN RAJAGOPALAN, DA-FANG WAN*, GEETHA M. HABIB, ANTONIA R. SEPULVEDA, MICHAEL R. MCLEOD, RUSSELL M. LEBOVITZ, AND MICHAEL W. LIEBERMAN†

Department of Pathology, Baylor College of Medicine, Houston, TX 77030

Communicated by Salih J. Wakil, March 8, 1993 (received for review October 28, 1992)

ABSTRACT The 5' region of the mouse γ -glutamyltransferase (γ GT; EC 2.3.2.2) gene has been cloned and analyzed. This analysis, combined with sequence information obtained from γ GT cDNA clones, indicates that in mouse a single γ GT gene codes for six different mRNAs that differ in their 5' sequences. Analysis of steady-state levels of γ GT RNA reveals different expression patterns for these RNAs in different organs. The six different 5' sequences are widely separated within a 10-kb region and three of them show 75–86% identity with the three known rat γ GT cDNAs. Although the heterogeneity of the 5' ends of γ GT RNAs may be explained in part by alternative splicing, it is likely that multiple promoters are involved in their generation.

γ -Glutamyltransferase (γ GT) catalyzes the transfer of the γ -glutamyl moiety of glutathione and other γ -glutamyl compounds to a variety of amino acids and peptides. The activity of γ GT is high in cells with secretory or absorptive functions, including the epithelial cells of the proximal convoluted tubules of the kidney, the ciliary and retinal pigment epithelium of the eye, pancreatic ducts and acini, epididymis, and choroid plexus (1–6). γ GT is also expressed by bile duct epithelium and fetal liver but not by normal adult liver (7). Its expression is elevated in some rat and human hepatocellular cancers (8–16).

Cloning experiments from several laboratories including our own have led to the identification of three types (I, II, and III) of γ GT mRNA in the rat that differ in their 5' untranslated regions but not in the coding sequences (17–19). In a previous study, we identified the genomic sequences in the rat from which type I mRNA is initiated (19). However, attempts by us and by other investigators to clone regions further upstream that might contain the transcription start sites for type II and type III mRNA in rats have proved unsuccessful.

In this paper we report the organization and sequence of the 5' flanking region of the mouse γ GT gene.‡ Our findings are of interest because we have identified six γ GT RNAs each containing unique 5' sequences encoded by a single gene and expressed in a tissue-specific manner.

EXPERIMENTAL PROCEDURES

Genomic Cloning and Sequence Analysis. A BALB/c mouse genomic library constructed in EMBL3 phage vector from a *Sau3A1* partial digest of DNA was obtained from Clontech. Phage growth, library screening, phage DNA isolation, and subcloning were performed by standard procedures (20).

RNA Isolation and Northern and Southern Blotting. Total RNA was isolated by a guanidinium isothiocyanate method (21). Poly(A)⁺ RNA was prepared by chromatography using

oligo(dT)-cellulose. Southern and Northern blot hybridizations using nitrocellulose membranes (Schleicher & Schuell) were done by standard procedures (20).

Construction of cDNA Library. cDNA was synthesized by using poly(A)⁺ mouse kidney RNA with a cDNA synthesis kit (Bethesda Research Laboratories). An oligonucleotide primer complementary to mouse γ GT coding sequences (5'-AACCACAGCCACCAGGCCAG-3'; Fig. 1, common region, P1) was used in the reverse transcription reaction. The reverse-transcribed product was oligo(dA)-tailed with deoxynucleotidyltransferase, and the oligo(dA)-extended products were amplified by PCR. The 5' primer in this reaction (5'-GAAGATCTGGATCCTTTT TTT TTT TTT TTT TTT-3') contain *Bam*HI and *Bgl* II restriction enzyme sites and anneals to the oligo(dA) tail. The 3' primer (5'-CTGAATTC-CACCAGAAACCGATTCTTCAT-3'; Fig. 1, common region, P2) contains an *Eco*RI restriction site and a sequence that lies internally with respect to the product of the reverse transcription reaction. PCR-amplified products were digested with *Eco*RI and *Bam*HI and ligated into pBluescript vector (Stratagene).

PCR Analysis for the Presence of the Different Types of γ GT mRNA. Total RNA isolated from organs was reverse-transcribed by using primer P1 (Fig. 1, described above). The product was amplified in a PCR using a 5' primer complementary to each of the unique sequences present in the six different types of γ GT mRNAs (Fig. 1, type I, P3; type II, P5; type III, P9; type IV, P12; type V, P14; type VI, P16) and the primer used in reverse transcription as the 3' primer (P1; Fig. 1, common region). The product was electrophoresed in a 4% agarose gel and the detection of a DNA fragment of appropriate size (based upon the cDNA sequence) after staining the gel with ethidium bromide was considered to denote the presence of the particular γ GT mRNAs in the organ analyzed.

Probe and Primers. Probes were labeled by random priming (Boehringer Mannheim kit) or labeled by using [α -³²P]dATP during PCR synthesis. Primers for PCR and sequencing were synthesized with an Applied Biosystems Synthesizer.

DNA Sequence Analysis. DNA sequencing was performed by the Sanger dideoxy method with minor modifications (20) and primers described in the text. The sequence data reported in this paper were obtained by sequencing in both directions. DNA sequences were analyzed for consensus sequences and homology by using the EUGENE computer programs provided by the Molecular Biology Information Resource at the Baylor College of Medicine.

Abbreviation: γ GT, γ -glutamyltransferase.

*On leave from the Department of Molecular Biology, Shanghai Cancer Institute, Shanghai, China.

†To whom reprint requests should be addressed.

‡The sequences reported in this paper have been deposited in the GenBank data base (accession nos. L17331–L17336).

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Type I									
Mouse	P3	TTGGGAGAAG	TCATGCATAC	GTGGAACCGC	TGGCAGCTTC	T**GAGGCTC	*AGGAGTTCA	60	
Rat		XXXXXXXXXX	XXXXXXXXXX	AXXXXGGXXG	XCXXXXCXX	XTTX**XX	CXX*XXXXXXXX		
Mouse		ATGGGAGACA	GGAAGTCAGT	GAGAGCCAGA	GGAACGTCGC	CACATCAGGC	TCTCCAGAAA	120	
Rat		GXXXXXXX	*X*****G	XXXXXXXXXX	***X**XAT	XXXXXXXXXX	AXCXXXXXXG		
P4									
Mouse		AGTTCTGGGC	CTGCCTCACA	CACGTTTGAC	TTTGAGATT	TTGAC		180	
Rat		XXXXXXXXXX	XXXXTXXXX*	*****AXX	XXXXTX**X	XXXXX			
Type II									
Mouse		CGGCTC*TTG	CGGCTGTGAC	GTCAGCGTTG	ACCATCCTCC	CAGCCTATCT	GAACC*TCTT	60	
Rat		XCKXXCKXX	XTXXXCKXX	**XXXXTXXXX	XXXXXXXXXX	XXXXXXXXGX	XGXXCKXXXX		
P7									
Mouse		ACAGAGGACT	CCCTACCTCA	CACCCCACTC	AGAAGCCTTC	AGATGCCCTT	CT*GGGATGC	120	
Rat		GXXXXXXXX	XTXXTXCXX	XTXXXXXX	XXXXXXXXTX	XCXXXXXX	XXAX*XXXX		
Mouse		GGGCACCCCT	TAGCTGACCT	CTTGCCCCC	TGGGATTAG			180	
Rat		XXXXXXXXAXT	XXXXXXXXXX	XXXXXXXXTX	XXXXXXXXXX				
P8									
Type III									
Mouse		P9	CACCACACTC	CAGAATAACG	GGGCAGCAAG	CTCCTGTGTC	CTGGTTCCCC	TG*GTTTACA	60
Rat			XCGXX*GXAX	XXXXXXXXXX	X*XXXXXXXX	XXXXXXXXXX	XXXXXXXXXX	XXGXXXXXXXX	
Mouse		P10						120	
Rat		CAG							
		CAG							
Type IV									
Mouse		P12	GCAAGTGTGG	AGCTGAGAGG	GAACGCAGCA	GCACGTGTCT	GTGTGCCCA	GCTGCTGAAT	60
			GACACAAA	GCCCTCCTCA	TCGTCTCTGA	GCTGCTGGAC	ATC		
							P13	120	
Type V									
Mouse		P14	GACTTGTCCG	AAGCTGCCTC	TTCAGGAAGT	CTCCAGCTCC	AGCCTCTAAG	CTTACAG	60
								P15	
Type VI									
Mouse		P16	GCTTAGCTCA	GGTTAGGACT	TGTTGACCTT	GGCATCTGA	CACAGCGGT	GGTCAGACT	60
			TGGACTCTAT	GTGCAATGTG	ATATAGCTCA	GCATGGAAGA	TGTAAGAGG	AAGGAGCTC	120
			TGGGAATTAA	CAAGCTGACC	ATCCGGGAGC	CCACGCGCTT	GGCTAAGTAC	AGGCTCCACT	180
			CCACAGGAAA	GG					240
				P17					
Common Region									
Mouse		P11	GAGTGCCAGG	AGGTAGCCTG	CTCTAGCAGT*	TCAGGGAAGA	TCGGTCTCTG	CAGAGTGTGG	60
Rat			XXXXAXXX**	*****XXXXX	XXXXXAXGX*T	XXXXXXXXXX	XTXX*****	*****XTXCC	
Mouse			AGAGAGT***	*****TCCTG	CCCATCCATA	CAGCTGTTT	TCAAAGCCA	CCTCCCAAG	120
Rat			CXXXXXXGGG	GGAGTXXXX	XTTXXXXXX	XXXXXXXXAXX	XXXXGXAXX	XXXXXXXXXXXX	
Mouse			AAAGTTTGG	GGGCTCTAC	TGGAAGAGCC	ATGAAGATC	GGTTTCTGGT	GCTGGGCTG	180
Rat			XXXXXXXXXX	XXXXXXXXXX	XXXXCXXXX	XXXXXXXXXX	XXXXXXXXXX	XXXXXXXXXX	
Mouse			GTGGCTGTGG	TTCTGGTATT	TGTCATC				240
Rat			XXXXXGXXXX	XXXXXXXXGX	CXXXXXX				

Fig. 1. Sequence of types I–VI γ GT mRNA. The sequence is shown for the unique 5' region for each of six mouse γ GT RNAs and the common region. Corresponding sequences for rat types I, II, and III mRNA and the common region are also shown (17–19). x, Identity between mouse and rat sequence; *, absence of the corresponding nucleotide in that position in one species. P1–P17 are primers used in sequencing and PCRs described in *Experimental Procedures* and other figure legends. The translated sequences starting with ATG are shown in a box.

RESULTS

Isolation of Mouse γ GT Genomic Clones Containing the Unique Sequences of Types I, II, and III mRNA. Approximately 1.2×10^6 recombinant phages of a mouse genomic library were screened with rat γ GT cDNA that included part of the 5' untranslated exon (Fig. 2A; see ref. 19). Three independent phages (Fig. 2B, nos. 1703, 1705, and 1707) that hybridized to this probe were obtained. To identify mouse recombinant phages containing the region coding for 5' ends of different γ GT RNAs we employed the PCR. In this assay sense and antisense primers (P7 and P8, respectively; Fig. 1, type II) within the unique sequence of rat type II mRNA were used to amplify the DNA from the three recombinant phages (Fig. 2B, nos. 1703, 1705, and 1707; see ref. 17). These primers were chosen because they represent sequences at the 5'-most end of type II mRNA in rat and because our previ-

ously isolated rat genomic clones (19) extend up to that region. With phage 1703 DNA we detected reaction product of about 140 bp in length, a size predicted from the rat sequence (data not shown) (17). Digestion of 1703 DNA with *Xho* I released fragments of 8.0, 6.0 and 0.6 kb (Fig. 2B). Upon PCR with the same primers only the 6.0-kb DNA fragment yielded a 140-bp fragment (data not shown). Data from Southern blot analysis of mouse genomic DNA cut with appropriate restriction enzymes were consistent with a single-copy gene (data not shown).

The 6.0-kb fragment was cloned into KS M13⁻ Bluescript vector (Stratagene) and was sequenced by using primers P7 and P8 (Fig. 1) and T3 and T7 primers adjacent to the cloning site. Further sequencing was done with custom primers made on the basis of the deduced sequence. Sequences highly homologous to the unique sequences present at the 5' termini of rat types I, II, and III γ GT mRNA were present in the

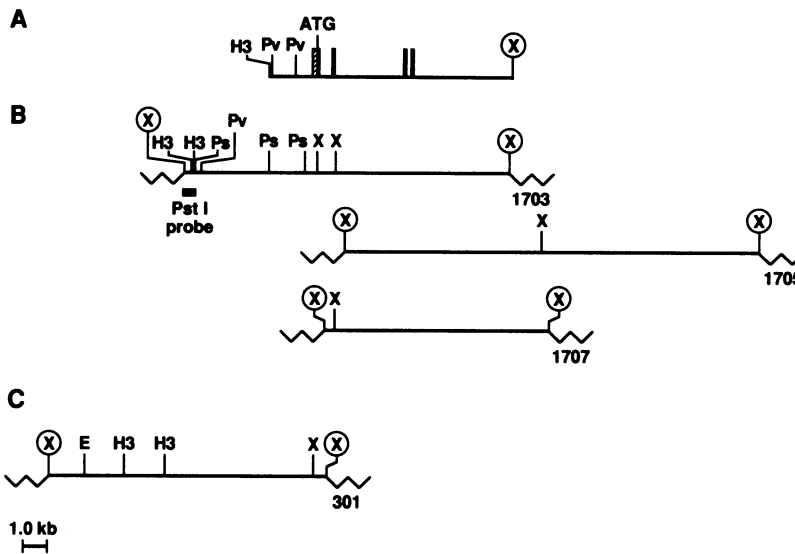


FIG. 2. Phage clones containing the mouse γ GT gene. (A) The 5' region of the rat γ GT gene previously described (19). Exons are shown as boxes. Filled boxes represent coding sequences. The DNA sequences of the first four exons were used as a probe to screen a mouse library to obtain phages shown in B. (B) Recombinant phages containing the mouse γ GT gene. Sites used in subcloning the fragments are shown. A filled box (*Pst* I probe) represents the region used as a probe in screening the mouse library to obtain the phage shown in C. (C) The 5' mouse γ GT clone obtained by using "*Pst* I probe." Inserts are shown as straight lines; wavy lines represent λ phage sequences. Phage DNAs are aligned with respect to the position of the homologous coding sequences present in rat (A). Restriction enzyme sites: E, *Eco*RI; H, *Hind*III; Ps, *Pst* I; Pv, *Pvu* II; X, *Xho* I in the genomic DNA; circled X, *Xho* I sites in the bacteriophage linker.

6.0-kb fragment obtained from mouse (Fig. 1; see refs. 17–19). The identity between sequences found in mouse and rat types I, II, and III mRNA is 75%, 86%, and 84%, respectively (Fig. 1). In rats, the three different types of mRNA contain 143 nt that are common and are located at the 3' end of the untranslated sequence (17–19). DNA sequence that is 77% identical to these 143 nt was found in the 6.0-kb fragment (Fig. 2B) and are presented in Fig. 1 (common region). Also shown in Fig. 1 in the common region is the sequence from the 0.6-kb fragment that contains the last 4 nt of the untranslated sequence and the first 19 codons of the γ GT gene. This sequence is 95% identical to the rat sequence and predicts the same amino acid sequence.

Cloning of New Mouse γ GT cDNAs. We made a cDNA library enriched for γ GT mRNA from mouse kidney by using a primer complementary to the coding sequences of the mouse γ GT gene for reverse transcription followed by poly(A) tailing of the product (see *Experimental Procedures*). It was screened with a PCR-synthesized probe made by using P11 and P2 as primers (Fig. 1, common region) and the 6.0-kb DNA fragment as the template; thus the product would contain sequences that are common to the various γ GT mRNA species. Twelve clones were obtained and the insert fragments were sequenced. One clone contained sequence (26 nt) unique to type I mRNA. Five clones of varying length containing type II sequences were obtained. One type III clone with 85 nt 5' of the common splice site was obtained. Also present were three other types of clones referred to as type IV (three clones), type V (one clone) and type VI (one clone) (Fig. 1).

We analyzed the expression of γ GT in adult mouse kidney by Northern blot hybridization. Radioactive probes specific for types I–VI γ GT mRNA were synthesized in separate PCRs with 5' and 3' primers whose sequences were based upon the unique sequences present at the 5' ends of the six different γ GT mRNAs. All six probes hybridized to kidney mRNA, and a band of the same size was detected in all six blots (Fig. 3A). This band corresponded in size to the same RNA that was detected by hybridizing kidney RNA and rat γ GT coding sequences (Fig. 2A) as probe (Fig. 3). No hybridization was observed in the RNA isolated from mouse spleen for any of the probes used in this experiment. Type II mRNA appeared to be present at higher levels than the other five different types of mRNA. However, in this experiment, even though the same amount of radioactivity was used in all hybridization reactions, the probe lengths varied in size (127, 83, 61, 103, 56, and 192, nt for types I–VI, respectively) precluding any attempt to quantitate the relative abundance of the various types of mRNAs.

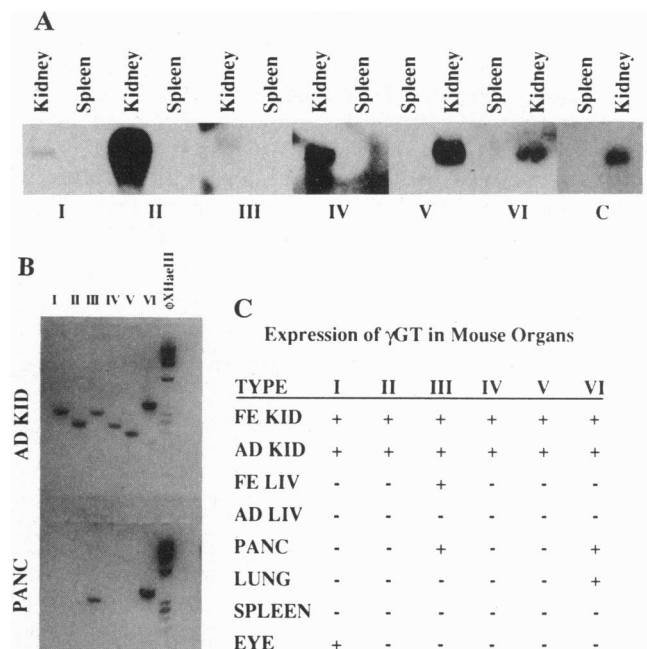


FIG. 3. (A) Expression of six γ GT mRNAs in kidney shown by Northern blot hybridization. Total RNA (20 μ g per lane) from mouse adult kidney and spleen was electrophoresed, transferred, and hybridized to probes specific for type I, II, III, IV, V or VI mRNA. Probes were synthesized using primers (see Fig. 1) specific for types I (P3 and P4), II (P5 and P6), III (P9 and P10), IV (P12 and P13), V (P14 and P15), or VI (P16 and P17) in PCRs with [α - 32 P]dATP and plasmid DNA as substrate. A fragment containing coding sequences (Fig. 2A) was labeled by a random-priming reaction and used in the blot labeled C. Filters labeled I, II, IV, V, VI, and C were exposed for 2 days; the filter labeled III was exposed for 4 days. (B) Analysis of γ GT expression by reverse transcription-PCR. Total RNA isolated from mouse organs was reverse-transcribed by using primer P1 (Fig. 1). The product was amplified in a PCR with the common 3' primer P1 and 5' primers specific for unique sequences present in the six different types of mRNA: P3, P5, P9, P12, P14, and P16 for types I–VI, respectively. Equal amounts of the PCR products were electrophoresed in a 4% agarose gel (SeaKem, FMC) and stained with ethidium bromide. Upper and lower panels show products obtained from adult kidney and pancreas, respectively. *Hae* III digest of ϕ X174 phage DNA provided size markers. (C) Survey of γ GT expression by reverse transcription-PCR in different tissues. The reactions were done as described for B. Presence (+) or absence (-) of a detectable fragment of appropriate size based upon the sequences of the cDNA clones is indicated. FE KID, fetal kidney; AD KID, adult kidney; FE LIV, fetal liver; AD LIV, adult liver; PANC, pancreas.

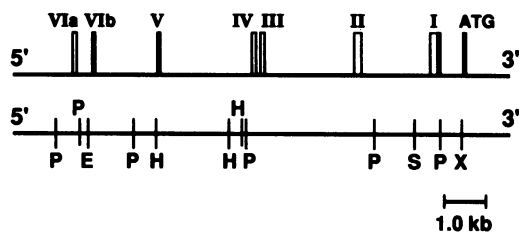


Fig. 4. Structure of the 5' region of the mouse γ GT gene. The unique exons of the six different types of γ GT mRNAs are shown by vertical bars labeled I–VI. Type VI mRNA contains unique sequences that are interrupted, resulting in two exons, VIa and VIb. The filled portion of bar I and the filled bar below ATG represent the only exons that are common to all six types of γ GT mRNAs. ATG denotes the translation initiation codon. Only the distances between the exons, and not the sizes of the exons, are drawn to scale. The bottom line shows some of the restriction sites present in this region of the gene. P, *Pst* I; E, *Eco*RI; H, *Hind*III; S, *Sst* I; X, *Xho* I.

We surveyed eight tissues for the expression of these six γ GT RNAs with the reverse transcriptase–PCR technique (Fig. 3C). Each of the six PCR mixtures contained a common coding-region 3' primer (P2; Fig. 1, common region) and a 5' primer complementary to only one of the six γ GT mRNAs (Fig. 1, type I, P3; type II, P7; type III, P9; type IV, P12; type V, P14; type VI, P16). The presence of a band of the appropriate size based upon the sequence of the cDNA clones in an agarose gel was considered as evidence of expression (Fig. 3B). Adult and fetal kidney express all six types of RNAs (Fig. 3). Fetal liver expresses only type III mRNA, whereas adult liver does not express γ GT (22).

5' -GGGCAAAATGGGGCCCTTTGAAGGGTTTCCAGCGGGAGAGGAAAATGTTCCATTAAGGTAACAGCAGC
CCTTGGCAAACCACTGGTAGCAGGCTACACAGGTTGATAGGAGTGGTGGGGGAGGGTACTAGGTGGTGGCT
TGGAGGAGGTCACAGATAACACCAATTTGAAGGGCCAGTCCGAGGCGAGGGGTCCGTTCTGAAGGGCTG
GACTCTGGCTTCTGTGGCTGGTAGAGGTTTCTGGAAGTGAGTTGCTTAGCTCAGGTTAGGACTTGTGTGAC

Vla
CTTGGGCATCTGACACAGCGGGTGGCTCAGACTTGGACTCTATGTGCAATGTGATATAGCTCAGCATGGA
AGATGTAAGAGGAAGGGAGCTCTGGgttaagctgag-----0.6kb.,-----atgctcccagGAATT

Vlb
AACAAAGCTGACCATCCCGGAGCCACCGGCTTGGCTAAGTACAGGCTCCACTCACACCAAGGcgagtac
tgg-----1.3kb.,-----TCACAGCTGCCCTCTGCCAGTTCGCCAGTCAAACCAAGGAGCAAAG
GACACCCGCAAGGGTCTTAATGAGCAAGCGGGTAAATGGGTTACCCAGCATGTGATATATTCACACGC
CTCACAGGAAACACAAAGTCCAGACTAGCCGAACCTTCAAGGCTTAGCCAGCTAGGCTTTCACCCCTG
ACACCTTAAGAGTTCAAGGGTGGAGCAAGGGAAGTCAAGAGACCTCCTCCACACCACCCATGACTTGTGTC

V
CGAAGCTGCCCTCTCAGGAAGTCTCCAGCTCCAGCCTCTAAGCTTACAGgttttcttcta-----2.7kb
.,-----GGCTCCGTAAGACCCCTTAGGCTTATGAGCCTAAATCATCCTCTCTCTTTGATATACAGCTTCTTTCA
CCGCTTAGACTGTCTCGGCTCGGGTTTCTCTCGGGGACGCTGACGCTCTGAGTTCTCCGATGCTGTT
ACAGGTTACCAATGCTACCTGAGTCTGTTCTGTACCTTTAAACCTTGCGCCACCCCAAGCCAGCTAGGGG
AGGGAAGGGAGGAGCTGGCTGCTGCGGGAAGTGGACTTAGCCTTCCGCAAGTGTGGAGCTGAGAGGGAA

IV
CGCAGCAGCAGTGTCTGTGTCGCCAGCTGCTGAATGACACCAAAAGCCCTCCATCCTGCTCTGAGCT
GCTGGACATCgtagggtggggGCAATAGACTTGGGGGACCTAAGAATGTGTGGAGACCTGGTTCTGTGTT
CTCAGTTCTGTCCTCCATTCACATTCACAGAGGCTCTGGTGATCTCCAGAGGCTCTCTGTGTAGCCACC

III
TGCGCAACAGGAGTGCACCACACTCCAGGATAACGGGGCGGCGAGGCTCCTGTGCCCTGGTCCCCCTGGTT
CACACAGgttaggcagag-----2.2kb.,-----CTAGTGCTGGGGTACCCTACTTTATTGCCCTCA
GGAACATTTTGAAGACATCATATGCTGTCTCTAAATTCACAGAGGACTTAGCAGGGGCTGTCTGGAA
GGGTTACCCGTTGGCCTCTGCCCTTAGAGGGAAACCAAAATCTGGAAATCCCCAGCCCTTCCAGGCT
CTTAGTACCCTCCCTCCACAGACTTCCAGACTTCCAGGCTGCAATGAACTTTAGTCACTGTGTGCAAGCC
TGACCGTGGGCTCTTTCGGGCTGTGACGTCAGGTTGACCATCTCCAGCCTATCTGAACCTTTACAG

II
AGGACTCCCTACCTCACAGCCACTCAGAAGCCTTCAGATGCCCTTCTGGGATGCTGGGCACCCCTTAGC
TGACCTCTTGGCCCTGGGATTTAGgttaagcgtg-----1.6kb.,-----CGAGTGTAAAGGTTTAC
AGGTGTTAGCTTCCATGCTCAGTTATAGCATTTACACATCCAAAGTTATTCAAAACACCCCGGAGGCA
AGTCCGAGGCTCCGCAAAAGTCCACCCACAGGTCATACCAACAGCTCTCAGGAAGGGCCCTGGACC
CTTGAACCGTTTTATACGGTTTACATGCTGCTATTACTACCAATCTGGGTTTACGCAATGAAATGCCAG
TTCTGTGACCCCTTCCCGGAGCTCTTGGGAGAAGTCAATGCACTGGAACCCCTGGCAGCTTCTGTA

I
GGCTCAGGAGTTCAATGGGAGACAGGAAGTCACTGAGAGCCAGAGGAACGTTCCGCACATCAGGCTCTCCA
GAAAGTTCTGGGCTGCTCACACAGTTTGTACTTTGAGATTTTTAGGAGTCCAGGAGGTAGCCTGTC

Com
TCTAGCAGTTTCAGGGAAGTCCGCTCTGTCAGAGTGTGGAGAGTTCCTGCCATCCATACAGCTGTTT
TCAAAGCCACCTCCCAAGGTTTGGGGCTCTTACTGGAAGgtaagctctg-----0.5kb.,--
-----tgggcagcagAGCCATGAAGAAATCGGTTCTG-3'

Pancreas expresses type III and VI, while in lung only type VI is expressed. Type I RNA is expressed in the eye, and no expression of any of the six types of γ GT RNA is found in spleen. These results indicate that the RNAs can be detected by PCR. The actual abundance of these RNAs must in the future be assessed by more quantitative methods such as Northern blot analysis or nuclease protection.

Isolation and Localization of Sequences Unique to Types IV, V, and VI mRNA in Mouse Genomic DNA. A computer search for sequences unique to types IV, V, and VI mRNA within the 6.0-kb DNA of phage 1703 (Fig. 2B) was carried out. The 104 nt that are unique to type IV RNA were found in the 5' region of the 6.0-kb fragment and are present about 120 nt 5' of the type III region.

To obtain genomic DNA corresponding to regions spanning type V and type VI sequences, we rescreened the mouse genomic library with a *Pst* I fragment (450 nt) obtained from the 5'-most region of the 6.0-kb *Xho* I fragment of phage 1703 (Fig. 2B) and obtained several phages, including phage 301 (Fig. 2C). We identified two DNA fragments, each of which contained one of the sequences: a 1.4-kb *Eco*RI–*Hind*III fragment and a 1.4-kb *Xho* I–*Eco*RI fragment (phage 301, Fig. 2C) that hybridized to probes for types V and VI mRNA, respectively. These two fragments were subcloned into a Bluescript vector and sequenced. The type V sequences are 2.7 kb 5' of type IV sequences. The unique sequences of type VI RNA are split into two exons that are interrupted by about 0.5 kb of DNA. The sequence data combined with our restriction enzyme digest data indicate that type VI sequences are about 10.0-kb 5' from the start of translation (ATG). The

FIG. 5. Partial sequence of the 5' region of the mouse γ GT gene. The unique sequences present at the 5' end of the six γ GT mRNAs are underlined. Type VI RNA contains two exons coding for unique sequences, denoted as VIa and VIb. Shown italicized and heavily underlined are the untranslated and the coding sequences that are common (com) to all six types of mRNAs. Only the first six codons, starting with the initiation codon ATG, are shown (in bold-face type). For each of the RNAs (10 nt that constitute the intron region of the splice donor junction are shown in lowercase letters. For each type of mRNA 250 nt immediately 5' of the longest stretch of cDNA are presented in uppercase letters. Interspersing regions between the different exons whose sequences are not shown are represented by dashed lines and the distances between exons are shown in kilobases.

organization of this region along with the location of the six unique sequences of the γ GT mRNAs are shown in Fig. 4.

For each of the six unique RNAs, genomic sequences 250 nt 5' of the longest cDNA stretch and the splice donor sequences are presented in Fig. 5. Also shown are splice acceptor sequences and the initial coding region of γ GT. All of the splice donor junctions contained the consensus GT dinucleotide sequence at the 5' end of the intron except for type VI mRNA. In this case at the VIb exon/intron junction the GT nucleotides were replaced by CG. Such variation in splice junction sequences has been observed for several genes (23).

DISCUSSION

To our knowledge this is the first report of a single gene coding for six different mRNAs with alternative 5' sequences. Because some organs express all six RNAs (kidney) whereas others express only one (fetal liver, lung, and eye), it is clear that the γ GT gene is regulated in a complex manner. In addition, the gene shows developmental regulation; in mouse and rat, type III RNA is made in fetal liver but not in adult liver, and in kidney types I–III are known to be expressed at lower levels during development (22).

In theory the six γ GT mRNAs might arise from initiation at independent promoters 5' of the common region, alternative splicing in the 5' flanking region, or both. Recent evidence, based on transfection of chloramphenicol acetyltransferase constructs and RNA mapping studies, indicates that there are at least five and probably six promoters in the upstream region of the mouse γ GT gene (A.R.S., B. W. Carter, G.M.H., R.M.L., and M.W.L., unpublished work). These data and the data in this paper are also consistent with our transgenic mouse data which show that transgenes driven by sequences from the type I region (Figs. 4 and 5) are expressed only in the kidney and eye (Fig. 3) (24, 25). It is unclear why so many promoters have evolved for this gene. It is likely, given the wide tissue distribution of γ GT and its changes during development, that multiple regulatory strategies must be necessary.

Multiple tissue-specific differences in the 5' untranslated regions of mRNA as observed in this study have also been observed for genes encoding aldolase A (26), porphobilinogen deaminase (27), acetyl-coenzyme A carboxylase (28), and α_1 -antitrypsin (29). The human aminopeptidase N gene uses two separate promoters to control transcription in myeloid and intestinal epithelial cells (30). Such alternative use of tissue-specific promoters has also been found in the carbonic anhydrase I gene (31), the human *lck* gene (32), and the phosphofructokinase gene (33). The human gene coding for insulin-like growth factor II contains four promoters that are subject to tissue-specific and development-dependent regulation (34). Much remains to be learned about the control of alternative splicing and promoter usage in the regulation of gene expression. These analyses should be facilitated by the study of genes such as the γ GT gene which code for multiple RNAs and show varied tissue expression.

This work was supported by National Institutes of Health Grants R37 CA39392 (M.W.L.) and RO1-CA49845 (R.M.L.).

- Goldberg, J. A., Friedman, O. M., Pineda, E. P., Smith, E. E., Chatterji, R., Stein, E. H. & Rutenberg, A. M. (1960) *Arch. Biochem. Biophys.* **91**, 61–70.
- Albert, A. Z., Orlowski, J., Orlowski, M. & Szweczuk, A. (1964) *Acta Histochem.* **18**, 78–89.
- Rutenberg, A. M., Kim, H., Fischbein, J. W., Hanker, J. S., Wasserkrug, H. L. & Seligman, A. M. (1969) *J. Histochem. Cytochem.* **17**, 517–526.
- Tate, S. S. (1980) in *Enzymatic Basis of Detoxification*, ed. Jakoby, W. (Academic, Orlando, FL), pp. 95–120.
- Hanigan, M. H. & Pitot, H. C. (1985) *Carcinogenesis (London)* **6**, 165–172.
- Rathbun, W. B. & Wicken, K. (1973) *Exp. Eye Res.* **15**, 161–171.
- Tateishi, N., Higashi, T., Nomura, T., Naruse, A., Nakashima, K., Shiozaki, H. & Sakamoto, Y. (1976) *Gann* **67**, 215–222.
- Fiala, S., Fiala, A. E. & Dixon, B. (1972) *J. Natl. Cancer Inst.* **48**, 1393–1401.
- Jaken, S. & Mason, M. (1978) *Proc. Natl. Acad. Sci. USA* **75**, 1750–1753.
- Cameron, R., Kellen, J., Kolin, A., Malkin, A. & Farber, E. (1978) *Cancer Res.* **38**, 823–829.
- Williams, G., Ohmori, T., Katayama, S. & Rie, J. (1980) *Carcinogenesis (London)* **1**, 813–818.
- Dempo, K., Elliott, K. A., Desmond, W. & Fishman, W. H. (1981) *Oncodev. Biol. Med.* **2**, 21–37.
- Solt, D. B. (1982) *J. Natl. Cancer Inst.* **67**, 193–197.
- Peraïno, C., Richards, W. L. & Stevens, F. J. (1983) *Environ. Health Perspect.* **50**, 1–53.
- Yaniguchi, N., Iizuka, S., Zhe, N. A., House, S., Yokosawa, N., Ono, M., Kinoshita, K., Makita, A. & Sekiya, C. (1985) *Cancer Res.* **45**, 5835–5839.
- Li, Y., Seyama, T., Godwin, A. K., Winokur, T., Lebovitz, R. M. & Lieberman, M. W. (1988) *Proc. Natl. Acad. Sci. USA* **85**, 344–348.
- Chobert, M. N., Lahuna, O., Lebargy, F., Kurauchi, O., Darybouy, M., Bernaudin, J. F., Guellaen, G., Barouki, R. & Laperche, Y. (1990) *J. Biol. Chem.* **265**, 2352–2357.
- Griffiths, S. A. & Manson, M. M. (1989) *Cancer Lett.* **46**, 69–74.
- Rajagopalan, S., Park, J. H., Patel, P. D., Lebovitz, R. M. & Lieberman, M. W. (1990) *J. Biol. Chem.* **265**, 11721–11725.
- Maniatis, T., Fritsch, E. F. & Sambrook, J. (1982) *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Lab., Plainview, NY).
- Chirgwin, J. W., Przybyla, A. E., MacDonald, R. J. & Rutter, W. J. (1979) *Biochemistry* **18**, 5294–5299.
- Habib, G. M., Rajagopalan, S., Godwin, A. K., Lebovitz, R. M. & Lieberman, M. W. (1992) *Mol. Carcinog.* **5**, 75–80.
- Shapiro, M. B. & Senapathy, P. (1987) *Nucleic Acids Res.* **15**, 7156–7174.
- Schaffner, D. L., Barrios, R., Massey, C., Bañez, E. I., Ou, C.-N., Rajagopalan, S., Aguilar-Cordova, E., Lebovitz, R. M., Overbeek, P. A. & Lieberman, M. W. (1993) *Am. J. Pathol.* **142**, 1051–1060.
- Chévez-Barrios, P., Schaffner, D. L., Barrios, R., Overbeek, P. A., Lebovitz, R. M. & Lieberman, M. W. (1993) *Am. J. Pathol.*, in press.
- Maire, P., Gautron, S., Hakim, V., Gregori, C., Menecier, F. & Kahn, A. (1987) *J. Mol. Biol.* **197**, 425–438.
- Chretien, S., Dubart, A., Beaupain, D., Raich, N., Grandchamp, B., Rosa, J., Goossens, M. & Romeo, P. H. (1988) *Proc. Natl. Acad. Sci. USA* **85**, 6–10.
- López-Casillas, F. & Kim, K.-H. (1989) *J. Biol. Chem.* **264**, 7176–7184.
- Perlino, E., Cortese, R. & Ciliberto, G. (1987) *EMBO J.* **6**, 2767–2771.
- Shapiro, L. H., Ashmun, R. A., Roberts, W. M. & Look, A. T. (1991) *J. Biol. Chem.* **266**, 11999–12007.
- Fraser, R., Cummings, P. & Curtis, P. (1989) *Mol. Cell. Biol.* **9**, 3308–3313.
- Takadera, T., Leung, S., Gernone, A., Koga, Y., Takihara, Y., Miyamoto, N. G. & Mak, T. W. (1989) *Mol. Cell. Biol.* **9**, 2173–2180.
- Yamasaki, T., Nakajima, H., Kono, N., Hotta, K., Yamada, K., Imai, E., Kuwajima, M., Noguchi, T., Tanaka, T. & Tarui, S. (1991) *Gene* **104**, 277–282.
- Dijk, M. A., Van Bootsma, H. J., Van Schaik, F. M. A., Holthuisen, P. & Sussenbach, J. S. (1991) *Mol. Cell. Endocrinol.* **81**, 81–94.