Organization of the 5′ End of the Rat γ-Glutamyl Transpeptidase Gene: Structure of a Promoter Active in the Kidney†,‡

Osamu Kurauchi,§ Olivier Lahuna, Mojtaba Darbouy, Martine Aggerbeck, Marie-Noëlle Chobert, and Yannick Laperche*

Institut National de la Santé et de la Recherche Médicale, Unité INSERM 99, Hôpital Henri Mondor, 94010 Créteil, France

Received August 9, 1990; Revised Manuscript Received October 18, 1990

ABSTRACT: Two γ-glutamyl transpeptidase mRNAs ($mRNA_I$ and $mRNA_{II}$), with alternate 5′-untranslated regions, are expressed in the rat kidney. Oligonucleotides were designed based upon these two alternate 5′ sequences and used as primers to amplify GGT genomic DNA sequences. The genomic organization of the $mRNA_I$ and $mRNA_{II}$ 5′-untranslated sequences reveals that the mRNAs are encoded from two separate promoters which are 2.1 kbp apart on the single GGT gene. A 2775 base pair genomic sequence, which contains the proximal GGT promoter, was cloned from two overlapping amplified fragments. S1 mapping analysis shows that the kidney GGT $mRNA_I$ is transcribed from several start sites on this promoter which displays neither a classical TATA box nor Sp_1 binding sites. Chimeric plasmids, including the GGT promoter region for $mRNA_I$, associated with the chloramphenicol acetyltransferase (CAT) reporter gene, were transiently expressed in a kidney (LLCPK) and in a hepatoma (HTC) cell line. A sequence extending 308 bases upstream from the major GGT $mRNA_I$ start site drives a promoter activity which is 5-fold higher in LLCPK than in HTC cells and is sufficient to confer cell specificity to the GGT proximal promoter.

γ-Glutamyl transpeptidase (GGT)[1] [(5-glutamyl)-peptide:amino-acid 5-glutamyltransferase, EC 2.3.2.2] is a heterodimeric glycoprotein (Tate & Meister, 1981); its two subunits, encoded by a single mRNA (Finidori et al., 1984; Nash & Tate, 1984), are located on the outer surface of the plasma membrane (Tate & Meister, 1981). GGT initiates the degradation of extracellular glutathione (GSH) by the cleavage of the γ-glutamyl bond, allowing the recovery of extracellular cysteine from GSH by the cell. Therefore, GGT plays a key role in maintaining intraorgan thiol levels since most tissues rely on exogenous cysteine to support intracellular GSH synthesis (Meister, 1984). In the rat and humans, GGT expression exhibits a strong tissue specificity. The highest activity is found in the kidney (Tate & Meister, 1981); in the liver, the activity is low but increases during chronic alcoholism (Rosalki & Rau, 1972), during cholestasis (Chobert et al., 1989), and in various chemically induced tumors (Cameron et al., 1978).

In the rat, GGT is encoded by a single-copy gene (Pawlak et al., 1988) from which different mRNAs, with alternate 5′-untranslated regions, are transcribed in a tissue-specific manner (Chobert et al., 1990). In the kidney, two GGT mRNAs ($mRNA_I$ and $mRNA_{II}$) are expressed and are coordinately regulated during ontogeny (Chobert et al., 1990; Laperche et al., 1986). These two mRNAs, which have identical coding sequences, differ at their 5′ ends by two alternate sequences of respectively 138 and 154 bases (Chobert et al., 1990). Previously, we cloned the 3′ end of the rat GGT gene, but in spite of extensive screening of several rat genomic libraries, we have not been successful in cloning the 5′ part of the rat GGT gene. In the present study, we have taken advantage of the recent development of the polymerase chain reaction (PCR) technique (Saiki et al., 1988) to achieve amplification and cloning of genomic sequences flanked by two GGT gene specific primers designed from the GGT cDNA sequences. Analysis of the genomic organization of the sequences coding for the 5′-untranslated regions of the $mRNA_I$ and $mRNA_{II}$ reveals that these two mRNAs are encoded from two separate promoters on the GGT gene. A 2775 bp[1] GGT sequence was cloned from two overlapping amplified DNA fragments; it includes one of the two GGT promoters which are functional in the kidney.

Experimental Procedures

Genomic DNA was prepared from the livers of Wistar rats (Pawlak et al., 1988). Restriction endonucleases, T4 DNA ligase, T4 polynucleotide kinase, and the plasmid pGEM-3Z were obtained from Promega Biotec; deoxyribonucleotides and S1 nuclease were from Boehringer Mannheim, and the *Taq*I–DNA polymerase was from Stratagene. Radiolabeled nucleotides, [^{14}C]chloramphenicol, nylon membrane, and hyperfilms MP for autoradiography were purchased from Amersham Radiochemical Centre.

The oligodeoxyribonucleotides, designed from the rat kidney cDNA sequences (Chobert et al., 1990) (Figure 1A), were synthesized on an Applied Biosystems Model 380B DNA synthesizer and purified on an OPC cartridge as recommended by the supplier. Five oligonucleotides were used as primers in polymerase chain reaction experiments (PCR): oligomers 1 (5′-TTGGGAGAAGTCATG-CAT-3′) and 2 (5-CGCCTCCTTGCTGCTGCC-3′) correspond to the sequences on the RNA strand; oligomers 3 (5′-AGAGGGGCAAG-AGGT-CAGC-3′), 4 (5′-TAGAGCAGGCTGGTACTC-3′),

† This work was supported by the INSERM and the University of Paris Val-de-Marne.

‡ The nucleic acid sequence in this paper has been submitted to GenBank under Accession Number J05310.

* Correspondence should be addressed to this author.

§ Recipient of a fellowship from the Fondation pour la Recherche Médicale and from the Association pour la Recherche contre le Cancer (ARC). Present adress: Department of Obstetrics and Gynecology, Nagoya University School of Medicine, Nagoya, Japan.

[1] Abbreviations: GGT, γ-glutamyl transpeptidase; 1×SSC, 150 mM NaCl/15 mM sodium citrate; PIPES, piperazine-*N*,*N*′-bis(2-ethanesulfonic acid); bp, base pair(s); kbp (kb in figures), kilobase pair(s).

0006-2960/91/0430-1618$02.50/0 © 1991 American Chemical Society

and 5 (5′-CCAGCACC-AGAAACCGATTCTTCAT-3′) are complementary to the RNA strand. Two oligonucleotides, 6 (5′-GAGTTCAGCGGGAGACAGA-3′) and 7 (5′-CTTCTGAGTGGGATGTGGGATA-3′), were used as probes to identify amplified genomic sequences corresponding to cDNA-17 ($mRNA_I$) and cDNA-12 ($mRNA_{II}$), respectively.

DNA Amplification. Amplification of genomic DNA fragments was carried out by using the conditions of Saiki et al. (1988) in a DNA thermal cycler from Perkin Elmer—Cetus Instruments, except that we did not use gelatin and that the $MgCl_2$ concentration was brought to 1.5 mM. Genomic DNA samples (1 μg, digested by *Bam*HI) were subjected to 30 cycles of amplification with 100 ng of each set of two primers in a 100-μL volume. Samples were denatured at 94 °C for 1 min, cooled to 55 °C over 1 min for primer annealing, and heated to 72 °C for a 2.5-min elongation. After the last cycle, cell samples were incubated for an additional 5 min at 72 °C to ensure completion of the final extension step.

Analysis of the Amplified DNA. An aliquot (20 μL) of the amplified DNA was analyzed on a 1% agarose gel and transferred onto a Hybond N membrane. The 19-mer and 22-mer oligonucleotides 6 and 7, labeled at their 5′ end by T4 polynucleotide kinase and [γ-^{32}P]ATP, were hybridized to the amplified DNA. Prehybridization (6 h) and hybridization (18 h) were performed in 6×SSC,[1] 0.1% SDS, 5 mM EDTA, 1× Denhardt, and 200 μg/mL denatured salmon sperm DNA at 37 °C; the blots were washed 3 times for 1 h at 45 °C in 6×SSC/0.1% SDS.

Cloning and Sequencing of the Amplified DNA. The DNA fragments, amplified from oligonucleotides 2 and 4 (fragment A), digested by *Hin*dIII, and from oligonucleotides 1 and 5 (fragment B), digested by *Sac*I, were subcloned in the *Hin*dIII–*Sma*I sites or *Sac*I–*Sma*I sites of pGEM-3 Z, respectively, using standard recombinant DNA procedures (Figure 2). The resulting plasmids (pGEM-3A and pGEM-3B) were used as templates and hybridized to synthetic oligonucleotides for DNA sequencing. Sequencing reactions were performed by using the Sequenase kit from United States Biochemical Corp. in the presence of [α-^{35}S]dATP, or the Taq Track DNA sequencing kit from Promega Biotec in reactions primed by 5′-^{32}P-labeled oligonucleotides.

S1 Nuclease Mapping Analysis. Mapping of the 5′ end of the kidney GGT mRNA onto the gene was performed as described previously (Berk & Sharp, 1977). The plasmid pGEM-3A was digested by *Sac*I, 5′ end labeled by T4 polynucleotide kinase and [γ-^{32}P]ATP, and further digested by *Kpn*I. The resulting *Kpn*I–*Sac*I fragment labeled only at the 5′ end of the RNA complementary strand was purified on a 1% low-melt agarose gel. The 584 bp *Kpn*I–*Sac*I probe (5 × 10^6 cpm) was hybridized to 20 μg of kidney poly(A+) RNA for 16 h at 50 °C in 30 μL of 80% formamide, 1 mM EDTA, 0.4 M NaCl, and 40 mM PIPES.[1] The mixture, completed to 300 μL with 280 mM NaCl, 50 mM sodium acetate, pH 4.5, and 4.5 mM $ZnSO_4$, was digested by 200 units of S1 nuclease for 1 h at 37 °C. The sample was then extracted with phenol and precipitated with ethanol, and the S1-protected fragments were resolved by electrophoresis on a 6% acrylamide/7 M urea gel. The size of the fragments was determined by comparison to a sequence ladder run on the same gel.

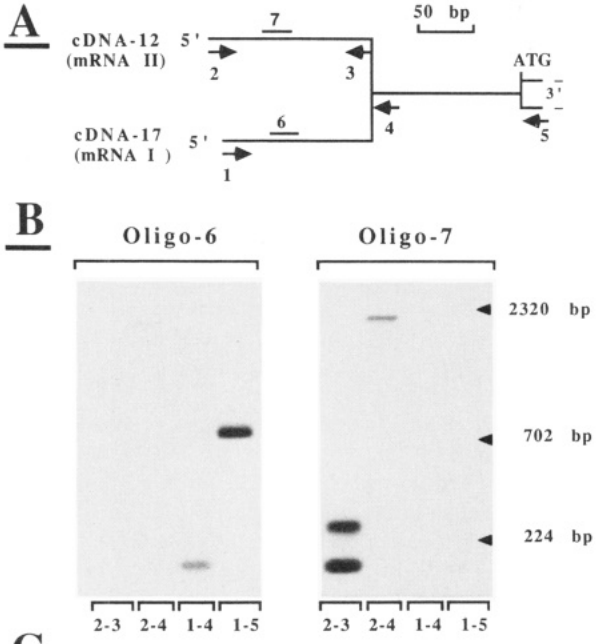
GGT-CAT Plasmid Constructions. Several GGT-CAT (chloramphenicol acetyltransferase) plasmids were constructed by subcloning GGT genomic sequences, from −1960 bases upstream to +158 bases downstream of the major transcription initiation site, into pMW1-CAT, which carries the CAT reporter gene (Gorman et al., 1982). *Hin*dIII–*Eco*RI and *Kpn*I–*Kpn*I fragments from pGEM-3A were cloned in the corresponding sites in pMW_I-CAT. The resulting clones contain GGT sequences from −1960 to +158 bases (pK1 −1960 +158) and from −508 to +158 bases in both orientations (pK1 −508 +158 and pK1 +158 −508). Two other CAT constructs were obtained by amplification of GGT sequences in pGEM-3A between a set of two 18-mer synthetic primers and their cloning into the *Sma*I site of the polylinker of pMW1-CAT. The resulting plasmids, pK1 −308 +158 and pK1 +158 −308, contain GGT sequences cloned in both orientations. The pRSV-CAT and the pRSV-β-galactosidase plasmids have been previously described (Gorman et al., 1982; Edlund et al., 1985).

Cell Culture and Transfection. The rat hepatoma cell line HTC (Richards et al., 1982) and the pig kidney cell line LLCPK, derived from the proximal convolution of the renal tubule (Stevens et al., 1986), were cultured in Dulbecco's-modified Eagle's medium (Gibco) supplemented with 5% fetal calf serum, 2 mM glutamine, and 1 mM pyruvate. Twenty-four hours prior to transfection, subconfluent culture dishes were trypsinized, and 5 × 10^5 cells were plated per 10-cm-diameter tissue culture dish. These cells were cotransfected by 10 μg of a CAT construct and 10 μg of the plasmid RSV-β-galactosidase per dish using the calcium phosphate technique followed by a glycerol shock (Wigler et al., 1979). The cells were cultured for 40 h, and cell extracts were prepared by three cycles of freezing and thawing. CAT assays were performed at 37 °C for 2 h in the presence of 4 mM acetyl-CoA and 62.5 nCi of [^{14}C]chloramphenicol as described by Gorman et al. (1982). The β-galactosidase activity was determined in every extract to correct for differences in transfection efficiency as described by An et al. (1982). Proteins were assayed according to the method of Bradford (1976).

Results

Amplification of GGT Genomic Sequences. Two GGT cDNA sequences (cDNA-17 and cDNA-12) were cloned previously from a rat kidney cDNA library (Chobert et al., 1990). These two clones contain the complete 5′-untranslated sequence of two mRNAs ($mRNA_I$ and $mRNA_{II}$) which exhibit alternate 5′ ends starting from 144 bp upstream of the initiation codon. Oligonucleotides 1–5 (Figure 1A) were used to amplify the corresponding genomic DNA sequences.

Amplification of genomic DNA sequences between primers 1 and 4 yields a fragment of 0.15 kbp which hybridizes to oligonucleotide 6 and not to oligonucleotide 7 (Figure 1B,C); this indicates that, in the GGT gene, the sequence delineated by oligonucleotides 1 and 4 is not interrupted by an intervening sequence and that the 5′ end of cDNA-17 ($mRNA_I$) lies within the same exon as the 5′ end of the sequence common to both mRNAs. A fragment of 0.8 kbp, which hybridizes to oligonucleotide 6, was amplified from primers 1 and 5 (Figure 1B,C); it is about 500 bases longer than the cDNA sequence and, thus, includes an intervening sequence in the gene between these two primers. A 2.2 kbp fragment was obtained by using oligonucleotides 2 and 4; it hybridizes to oligonucleotides 6 and 7 (Figure 1B,C), showing that the 5′ end of the cDNA-12 ($mRNA_{II}$) maps onto the gene approximately 2.1 kbp upstream from the initiation site for $mRNA_I$. A 0.15 kbp DNA fragment was obtained by using oligonucleotides 2 and 3 (Figure 1B,C); therefore, the 5′-end sequence specific for $mRNA_{II}$ is encoded by a single exon. In this case, a fragment of 350 bases was also obtained. It results most probably from the annealing of primer 3 on two sites on the gene; in fact, a



C

OLIGOS	SIZE OF THE FRAGMENTS (base pair)	HYBRIDIZATION TO	
		OLIGO-6	OLIGO-7
2 - 3	150-350	-	+
2 - 4	2200	+	+
1 - 4	150	+	-
1 - 5	800	+	-

FIGURE 1: Amplification of GGT genomic sequences. (A) Amplification strategy: the positions of the oligonucleotides used as primers (→) or specific probes (–) are shown on the cDNA sequences. (B) Southern blot analysis of the amplified products. Four sets of two primers (2-3, 2-4, 1-4, and 1-5) were used. The amplified genomic sequences were run on a 1% agarose gel, transferred to a Hybond membrane, and hybridized to oligonucleotides 6 or 7. Arrows correspond to the positions of DNA markers run on the same gel. (C) Properties of the amplified DNA fragments.

sequence highly homologous to this primer is found 180 bases downstream of primer 3 on the gene (Figure 3, nucleotides −1714 to −1696).

From these results, we deduced the organization of the 5′-untranslated sequences of $mRNA_I$ and $mRNA_{II}$ on the GGT gene (Figure 2); they map onto a 2.83 kbp fragment and are interrupted by two intervening sequences as confirmed by the sequence (see below).

Cloning and Sequencing of the Amplified DNA. The two overlapping amplified fragments A and B were subcloned as pGEM-3A and pGEM-3B (Figure 2), and a 2775 bp GGT genomic sequence was obtained from these two clones. The precise intron–exon structure was obtained by its alignment to the two full-length 5′-untranslated sequences of cDNA-12 and cDNA-17 (Chobert et al., 1990). This genomic sequence extends 1960 bases upstream from the major initiation site for $mRNA_I$ (+1), determined as reported below, and stops 25 bases downstream of the GGT initiation codon (Figure 3). The predominant cap site (+1) of $mRNA_I$ maps onto the gene exactly at the first base of cDNA-17; the leader exon for this mRNA extends to base +278 and precedes a 508 bp long intervening sequence; the second exon for $mRNA_I$ starts 4 bases before the initiation codon. The first 66 bases in the 2775 bp GGT genomic sequence align to the 3′ end of the leader exon for $mRNA_{II}$. The first intervening sequence for the $mRNA_{II}$ primary transcript ends within the leader exon for $mRNA_I$ at position +139 where there is the consensus

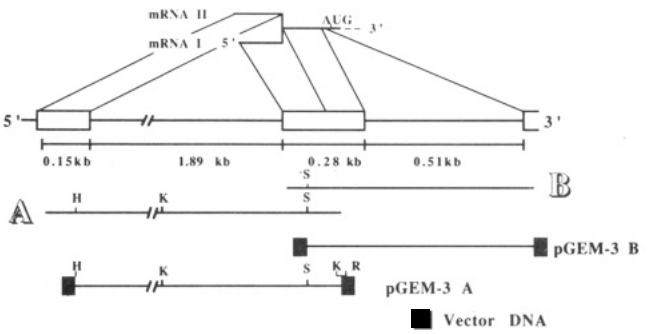


FIGURE 2: Organization of the 5′-untranslated sequences of GGT $mRNA_I$ and $mRNA_{II}$ in the gene, and cloning of the GGT genomic sequences. The structural organization of the gene is deduced from the results presented in Figure 1. Exons are shown as open boxes. The strategy for cloning GGT genomic sequences from the amplified fragments A (2200 bp) and B (800 bp) in the plasmid pGEM is illustrated. Abbreviations: H, *Hin*dIII; K, *Kpn*I; S, *Sac*I; R, *Eco*RI. The symbol (–//–) indicates an interruption in the scale.

splice site sequence CAG. Thus, the second exon for $mRNA_{II}$ corresponds to the 3′ part of the $mRNA_I$ leader exon (Figure 2). All the intron–exon junctions follow the GT/AG rule (Breatnach & Chambon, 1981). Upstream of the predominant initiation site for $mRNA_I$, there is no TATA box, but there are two classical CCAAT boxes (−136; −605) and a related CCAAAT sequence at position −73. The sequence from −13 to −6 fits exactly to the consensus motif 5′-CC(G/C)-C(G/A)-GGC-3′ which has been identified as the binding site for the cellular enhancer binding factor AP-2 (Mitchel et al., 1987). We also note the presence of several stretches of GAGA or complementary CTCT motifs which could be targets for the GAGA transcription factor (Biggin & Tjian, 1988). A short GT motif (TGTGGTTG) is found between nucleotides −468 and −462 which corresponds to the core sequence common to many cellular and viral enhancers (Weiher & Botchan, 1984).

S1 Nuclease Mapping of the Initiation Site of the mRNA onto the GGT Gene. A nuclease protection assay was performed to map precisely the initiation site for $mRNA_I$ onto the gene. The *Kpn*I–*Sac*I genomic fragment, which extends upstream from the putative start site and ends in the leader exon for $mRNA_I$, was hybridized to kidney poly(A+) RNA and digested with S1 nuclease. Three groups of genomic DNA fragments which are 60–73 (A), 76–78 (B), and 97–102 (C) nucleotides long were protected from S1 digestion (Figure 4). The presence of a cluster of bands around each start site is generally attributed to the instability of the RNA–DNA hybrid around the mRNA cap, leading to a slight heterogeneity in the DNA protection (Weydert et al., 1983; Kopito et al., 1987). The 5′ terminus of the predominant fragments in each cluster corresponds to three mRNA start sites. Thus, three initiation sites for $mRNA_I$ synthesis are present on the gene. The major site in B (numbered +1 in Figure 3) coincides with the end of cDNA-17 and also coincides with the extremity of the longest extension product on kidney $mRNA_I$ (Chobert et al., 1990; Figure 2). The minor start site, 21 bases upstream, was not detected in primer extension experiments, but it coincides with the 5′ end of a cDNA which is 21 bases longer than cDNA-17; this cDNA[2] was obtained from a kidney library in λgt 10.

Transient Expression of the GGT-CAT Fusion Gene in LLCPK and HTC Cells. The promoter activity of the GGT genomic sequence was assessed by transient transfections of

[2] R. Hughey, personal communication.

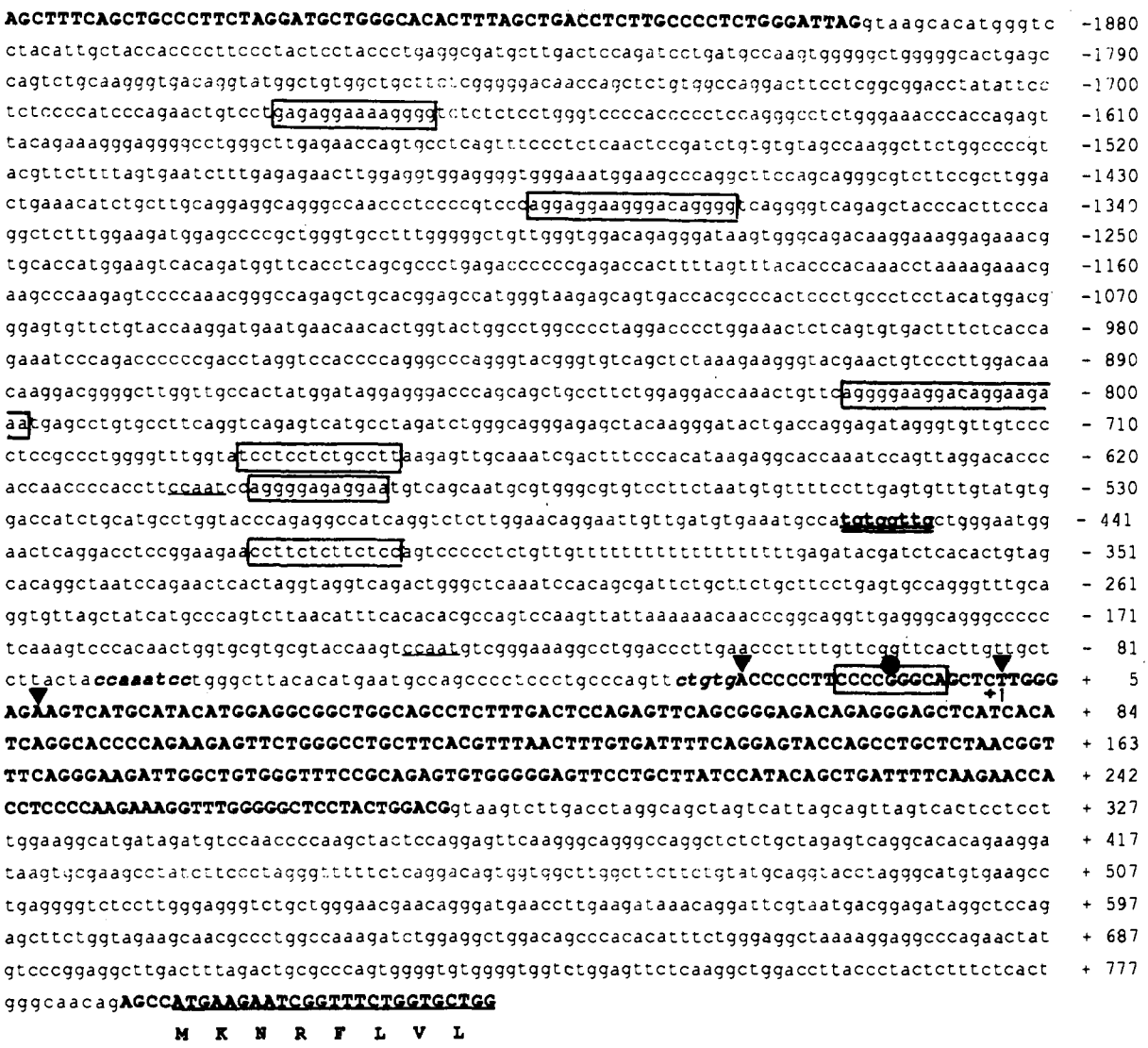


FIGURE 3: Nucleotide sequence of the promoter region of the GGT gene. Nucleotides are numbered from the major start site for $mRNA_I$ (+1); start sites are indicated by arrows. Nucleotides present in the mature mRNA are printed in boldface upper case letters; nucleotides in introns are in lower case letters. Nucleotides in underlined boldface upper case letters code for the amino acids printed below the sequence using single-letter abbreviations. The related CCAAATCC sequence at −72 and the CTGTG box at −26 are printed in boldface lower case italics. The core enhancer consensus sequence TGTGGTTG is underlined twice and the CCAAT box sequences once. Repeated GAG motifs are boxed, and the box with the symbol (●) indicates homology to the AP-2 responsive element.

GGT-CAT constructs into HTC cells and LLCPK cells. LLCPK cells transfected by the constructs pK1 −308 +158 and pK1 −508 +158 display a significant CAT activity which represents 20% and 10% of the strong promoter activity driven by the RSV promoter associated with its enhancer (Figure 5A). No CAT activity was detected in the cells transfected by the construct bearing the same GGT sequence in the opposite orientation (Figure 5A) or the promoterless pMW_I-CAT (not shown). Addition to the CAT plasmid of GGT sequences extending up to position −508 and to −1960 decreases the promoter activity in LLCPK cells up to 5-fold. In HTC cells, the GGT promoter activity of pK1 308 +158 is 5 times lower than in LLCPK cells (Figure 5B), whereas RSV-CAT drives a similar CAT activity in both cell lines. Addition of sequences up to position −1960 did not alter significantly the GGT promoter activity in HTC cells. These experiments reveal a tissue specificity in the expression of GGT promoter activity.

Discussion

We previously isolated a clone from a rat genomic DNA library which contained the 3′ end of the rat GGT gene (Pawlak et al., 1988). The 5′ end of this clone maps in the cDNA between bases 159 and 170 downstream from the initiation codon (unpublished data). Screening of the same rat genomic library as well as three others, using the 5′-untranslated sequences of $mRNA_I$ and $mRNA_{II}$ as probes, did not result in the isolation of any further upstream sequences. This may indicate that the genomic sequence of interest is underrepresented in the libraries or that it has been selectively lost due to the presence of unstable sequences as suggested earlier for the cloning of two human genes (Heisterkamp et al., 1989; Soininen et al., 1989). Therefore, we used the polymerase chain reaction technique to amplify and clone the 5′ part of the rat GGT gene. Analysis of the GGT genomic sequence for homologous sequences in Genbank reveals that the 41 nucleotides from −461 to −421 exhibits 90–95% homology to a sequence located within Alu type 2 repeats in several rat or mouse DNA sequences like the rat chymotrypsin B gene (Bell et al., 1984) and the mouse c-abl cDNA (Ben-Neriah et al., 1986). Repetitive DNA sequences are often associated with genomic instability and participate in recombination events (Calabretta et al., 1982) which could account for the difficulties encountered in the isolation of the 5′ part of the GGT gene from rat genomic libraries.

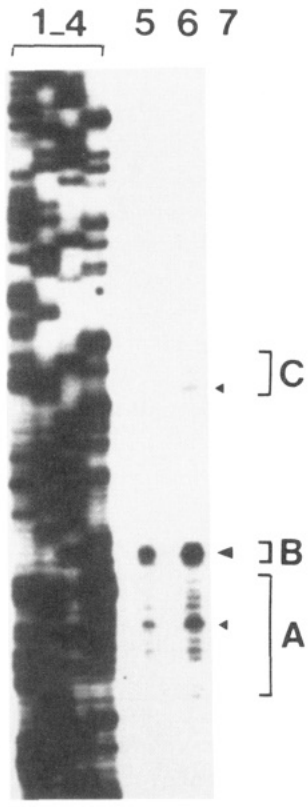


FIGURE 4: S1 mapping of the GGT mRNA start sites. The *Kpn*I–*Sac*I labeled genomic fragment was annealed to kidney poly(A+) RNA and subjected to S1 nuclease digestion as described under Experimental Procedures. A sequencing reaction was also performed from pGEM-3A using a 25-mer oligonucleotide whose 5′ end maps 20 bases downstream from the *Sac*I site on the complementary strand. The reaction products were loaded on a 6% acrylamide denaturing gel: lanes 1–4, sequencing reactions (GACT); lanes 5 and 6; DNA fragments resistant to S1 digestion after hybridization to 10 μg (lane 5) and 20 μg (lane 6) of poly(A+) RNA; lane 7, *Kpn*I–*Sac*I probe digested by S1 (control). The arrow in each group of bands (A–C) indicates the positions of the mRNA I start sites.

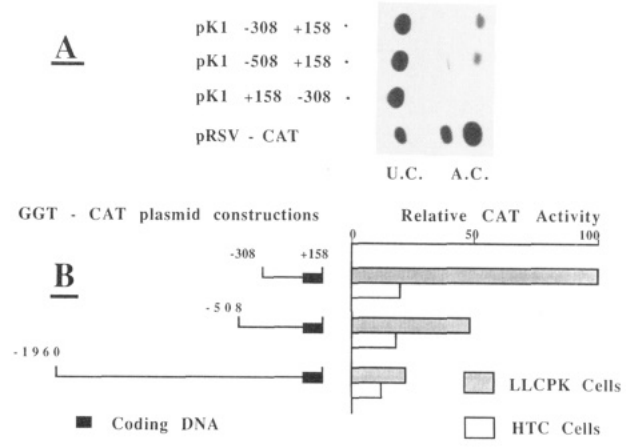


FIGURE 5: Transient expression of CAT activity in LLCPK and HTC cells transfected by GGT-CAT constructs. (A) CAT assay in LLCPK cells transfected by pK1 −308 +158 and pK1 −508 +158 which contain, respectively, a 308 and 508 bp of sequence upstream from the major initiation site for mRNAI. The vectors pK1 +158 −308 containing the GGT sequence in the reverse orientation and pRSV-CAT containing the Rous sarcoma virus promoter and enhancer sequences were used as negative and positive controls. UC, unreacted chloramphenicol. AC, acetylated chloramphenicol. (B) On the left side, the GGT sequences used to drive the CAT activity in the plasmid pMW1-CAT are shown, numbered from the major initiation site. The relative CAT activity driven by the constructs in LLCPK and HTC cells is shown on the right. Each value, corrected for the transformation efficiency as reported under Experimental Procedures, is normalized to that obtained with pK1 −308 +158 in LLCPK cells and is the average of three values obtained from three separate experiments.

The transcriptional organization of the 5′-untranslated sequences of the two mRNAs characterized from the kidney reveals that these two transcripts are initiated on two separate promoters 2.1 kbp apart on the gene. A third GGT mRNA species is transcribed from ethoxyquin-treated rat liver which differs from kidney $mRNA_I$ and $mRNA_{II}$ in its 5′-untranslated region; interestingly, the three GGT mRNA sequences diverge from the same point at nucleotide 144 upstream from the AUG (Griffiths & Manson, 1989). The 59 bp sequence immediately upstream from the divergent point in the liver mRNA is not present in the cloned genomic sequence, thus revealing that the GGT mRNA expressed in the liver is initiated on a third promoter upstream on the GGT gene. Transcription of two mRNAs species from two separate promoters on a single gene has been reported for several eukaryotic genes (Young et al., 1981; Benyajati et al., 1983); however, the presence of three independent transcription units has been described only for the IGF-II gene in the rat (Ueno et al., 1988). Thus, the GGT gene appears to provide another example of a gene with three active transcription units.

The different GGT mRNA species vary only at the end of their 5′-untranslated regions and are potentially active in producing the same GGT protein. However, the use of independent promoters allows several mechanisms for the modulation of GGT activity in various tissues and/or under different physiological situations; for example, the stability or the translational capacity of the mRNA can be specifically modulated by the alternative 5′-leader exons, or each promoter may contain target sequences specific for different trans-acting regulatory elements (Rouault et al., 1988; Lowe et al., 1987).

The kidney $mRNA_I$ is transcribed from the proximal promoter. This GGT promoter contains two CCAAT boxes at −136 and −605 but has neither a TATA box nor an Sp_1 binding site. The structure of this proximal GGT promoter was recently reported, and a 1061-base region extending 836 bases upstream the initiation site was sequenced (Rajagopalan et al., 1990). The major transcription initiation site was mapped at the same position on the gene; however, some sequence discrepancies were found in the 836 bases upstream of this site. We resequenced this region and confirmed our data. Moreover, the sequence of this promoter region determined by Dr. D. Goodspeed was identical with ours.[3] In particular, the two Sp1 binding sites reported by Rajagopalan et al. (1990) at −101 and −746 do not appear in our sequence.

In spite of the absence of a classical TATA box and the presence of several initiation sites for mRNA transcription, this GGT promoter does not fall in the class of housekeeping genes (Dynan, 1986; Tsuzuki et al., 1987); its expression is under the control of tissue-specific factors since the encoded $mRNA_I$ transcript has been detected only in the kidney, so far (Chobert et al., 1990). Moreover, at positions −26 and −73, the usual positions for the consensus TATA and CCAAT boxes in polymerase II promoters (Myers et al., 1986), we find the CTGTG and CCAAATCC motifs which are also present, at similar positions, in the mouse transducin gene. The latter cannot be viewed as a housekeeping gene since its expression is restricted to photoreceptor cells (Raport et al., 1989). Transfection studies show that the 308-nucleotide sequence immediately upstream from the major start site for the GGT $mRNA_I$ has a promoter activity which is 5 times larger in

[3] D. Goodspeed, personal communication.

LLCPK cells than in HTC cells. Addition of upstream sequences up to −1960 bases decreases the promoter activity in LLCPK cells approximately to the level found in HTC cells; this reveals the presence of putative negative cis-acting elements which could function as binding sites for nuclear proteins in LLCPK cells but not in HTC cells.

The proximal GGT promoter coding for $mRNA_I$ is under the control of cell-specific factors. In fact, this mRNA is not transcribed in liver cells, and in the kidney, its expression is restricted to the proximal tubule cells (Chobert et al., 1990). Our results indicate that the 308 bp sequence upstream from the major initiation site for GGT $mRNA_I$ is sufficient to confer such cell specificity. This proximal sequence will now be analyzed for its interaction with kidney and liver nuclear factors in order to identify the sequence motifs which could be involved in the repression of the GGT proximal promoter in the liver or in its activation in the kidney.

Acknowledgments

We thank Drs. L. Aggerbeck, R. Barouki, G. Guellaën, J. Hanoune, and D. Stengel for critical reading of the manuscript and L. Rosario and E. Grandvilliers for skillful secretarial assistance. We are indebted to Drs. M. Walker (Rehovot) and W. J. Rutter (San Francisco) for providing the relevant plasmids. We are very grateful to Drs. R. Hughey, D. Goodspped, and H. Pitot for sharing their unpublished data.

Registry No. GGT, 9013-62-1.

References

An, G., Hidaka, K., & Siminovitch, L. (1982) *Mol. Cell. Biol. 2*, 1628–1632.

Bell, G. I., Quinto, C., Quiroga, M., Valenzuela, P., Craik, C. S., & Rutter, W. J. (1984) *J. Biol. Chem. 259*, 14265–14270.

Ben-Neriah, Y., Bernards, A., Paskind, M., Daley, G. Q., & Baltimore, D. (1986) *Cell 44*, 577–586.

Benyajati, C., Spoerel, N., Haymerle, H., & Ashburner, M. (1983) *Cell 33*, 125–133.

Berk, A. J., & Sharp, P. A. (1977) *Cell 12*, 721–732.

Biggin, M. D., & Tjian, R. (1988) *Cell 53*, 699–711.

Bradford, M. M. (1976) *Anal. Biochem. 72*, 248–254.

Breatnach, R., & Chambon, P. (1981) *Annu. Rev. Biochem. 50*, 349–383.

Calabretta, B., Robberson, D. L., Barrera-Saldana, H. A., Lambrou, T. P., & Saunders, G. F. (1982) *Nature (London) 296*, 219–225.

Cameron, R., Kellen, J., Kolin, A., Malkin, A., & Farber, E. (1978) *Cancer Res. 38*, 823–829.

Chobert, M. N., Bernard, O., Bulle, F., Lemonnier, A., Guellaën, G., & Alagille, D. (1989) *J. Hepatol. 8*, 22–25.

Chobert, M. N., Lahuna, O., Lebargy, F., Kurauchi, O., Darbouy, M., Bernaudin, J. F., Guellaën, G., Barouki, R., & Laperche, Y. (1990) *J. Biol. Chem. 265*, 2352–2357.

Dynan, W. S. (1986) *Trends Genet. 2*, 196–197.

Edlund, T., Walker, M. D., Barr, P. J., & Rutter, W. J. (1985) *Science 230*, 912–916.

Finidori, J., Laperche, Y., Tsapis, R., Barouki, R., Guellaën, G., & Hanoune, J. (1984) *J. Biol. Chem. 259*, 4687–4690.

Gorman, C. M., Merlino, G. T., Willingham, M. C., Pastan, I., & Howard, B. H. (1982) *Proc. Natl. Acad. Sci. U.S.A. 79*, 6777–6781.

Griffiths, S. A., & Manson, M. (1989) *Cancer Lett. 46*, 69–74.

Heisterkamp, N., Morris, C., & Groffen, J. (1989) *Nucleic Acids Res. 17*, 8821–8831.

Kopito, R. R., Andersson, M. A., & Lodish, H. F. (1987) *Proc. Natl. Acad. Sci. U.S.A. 84*, 7149–7153.

Laperche, Y., Bulle, F., Aissani, T., Chobert, M. N., Aggerbeck, M., Hanoune, J., & Guellaën, G. (1986) *Proc Natl. Acad. Sci. U.S.A. 83*, 937–941.

Lowe, W. L., Roberts, C. T., Lasky, S. R., & Leroith, D. (1987) *Proc. Natl. Acad. Sci. U.S.A. 84*, 8946–8950.

Meister, A. (1984) *Hepatology 4*, 739–742.

Mitchel, P. J., Wang, C., & Tjian, R. (1987) *Cell 50*, 847–861.

Myers, R. M., Tilly, K., & Maniatis, T. (1986) *Science 232*, 613–618.

Nash, R., & Tate, S. S. (1984) *J. Biol. Chem. 259*, 678–685.

Pawlak, A., Lahuna, O., Bulle, F., Suzuki, A., Ferry, N., Siegrist, Chikhi, N., Chobert, M. N., Guellaën, G., & Laperche, Y. (1988) *J. Biol. Chem. 263*, 9913–9916.

Rajagopalan, S., Park, J. H., Patel, P. D., Lebovitz, R. M., & Lieberman, M. W. (1990) *J. Biol. Chem. 265*, 11721–11725.

Raport, C. J., Dere, B., & Hurley, J. B. (1989) *J. Biol. Chem. 264*, 7122–7128.

Richards, W. L., Tsukada, Y., & Potter, V. R. (1982) *Cancer Res. 42*, 1374–1383.

Rosalki, S. B., & Rau, D. (1972) *Clin. Chim. Acta 39*, 41–47.

Rouault, T. A., Hentze, M. W., Caughman, S. W., Harford, J. B., & Klausner, R. D. (1988) *Science 241*, 1207–1210.

Saiki, R. K., Gelfand, D. H., Stoffel, S., Scharf, S. J., Higuchi, R., Horn, G. T., Mullis, K. B., & Erlich, H. A. (1988) *Science 239*, 487–491.

Soininen, R., Huotari, M., Ganguly, A., Prockop, D. J., & Tryggvason, K. (1989) *J. Biol. Chem. 264*, 13565–13571.

Stevens, J., Hayden, P., & Taylor, G. (1986) *J. Biol. Chem. 261*, 3325–3332.

Tate, S. S., & Meister, A. (1981) *Mol. Cell. Biochem. 39*, 357–368.

Tsuzuki, T., Obaru, K., Setoyama, C., & Shimada, K. (1987) *J. Mol. Biol. 198*, 21–31.

Ueno, T., Takahashi, K., Matsuguchi, T., Endo, H., & Yamamoto, M. (1988) *Biochim. Biophys. Acta. 950*, 411–419.

Weiher, H., & Botchan, M. R. (1984) *Nucleic Acids Res. 12*, 2901–2916.

Weydert, A., Daubas, P., Caravatti, M., Minty, A., Bugaisky, G., Cohen, A., Robert, B., & Buckingham, M. (1983) *J. Biol. Chem. 258*, 13867–13874.

Wigler, M., Pellicer, A., Silverstein, S., Axel, R., Wilaub, G., & Chasun, L. (1979) *Proc. Natl. Acad. Sci. U.S.A. 76*, 1373–1376.

Young, R. A., Hagenbuchle, O., & Schibler, U. (1981) *Cell 23*, 451–458.