Expression of the Rat γ-Glutamyl Transpeptidase Gene from a Specific Promoter in the Small Intestine and in Hepatoma Cells†

Tomomitsu Okamoto,‡ Mojtaba Darbouy, Arthur Brouillet, Olivier Lahuna, Marie-Noëlle Chobert, and Yannick Laperche*

INSERM U.99, Hôpital Henri Mondor, F 94010 Creteil, France

Received May 3, 1994; Revised Manuscript Received June 28, 1994®

ABSTRACT: In the small intestine and in HTC hepatoma cells, the γ-glutamyl transpeptidase (GGT) single-copy gene is transcribed into a 2.5 kb and a 2.2 kb mRNA. Cloning of the GGT cDNA sequences from HTC cells demonstrates that the 2.5 kb mRNA ($mRNA_{IV-1}$) differs from the other rat GGT transcripts by a 371-base unique leader sequence which maps in the gene as 2 separate exons upstream of the 3 promoters which have been previously characterized. We established that the transcription of these two mRNAs is initiated on a new promoter (promoter IV) and occurs in the small intestine, in the epididymis, and in some hepatoma cells. The primary transcript initiated on GGT promoter IV is then alternatively spliced into the 2.5 kb $mRNA_{IV-1}$ or the 2.2 kb $mRNA_{IV-2}$ which is shorter in its 5′-untranslated sequence. The rat GGT gene exhibits a complex transcriptional organization leading to the transcription of five mRNAs from four independent promoters in a tissue-specific manner. The expression of the GGT promoter IV in the HTC hepatoma cells as well as in the small intestine could reveal that the HTC-transformed cells originate from liver precursor cells which still have the capacity to evolve toward different lineages. Thus, the GGT promoter IV will be valuable to isolate factors involved in the differentiation and carcinogenic processes.

The γ-glutamyl transpeptidase (GGT)[1] [(5-glutamyl)-peptide:amino-acid 5-glutamyltransferase, EC 2.3.2.2] initiates the degradation of glutathione (GSH) on the outer surface of the plasma membrane. Degradation of extracellular GSH allows the cellular recovery of cysteine, a limiting amino acid for intracellular GSH or protein synthesis (Hanigan & Rickett, 1993). This heterodimeric enzyme is expressed mainly on the brush border membranes of epithelial cells particularly active in detoxication or in protein synthesis such as the kidney proximal tubule, the small intestine, the pancreas, and the epididymis (Tate & Meister, 1981). In the adult hepatocyte, which is able to synthesize its own cysteine (Pallardo et al., 1991), the GGT activity is very low, but it is induced in early preneoplastic liver foci and primary hepatomas (Hanigan & Pitot, 1985; Beer et al., 1986). The induction of GGT activity is the most widely used marker to follow liver cell transformation in *in vivo* and *in vitro* models of hepatocarcinogenesis (Beer et al., 1986; Hampton et al., 1990).

GGT is synthesized as a single-chain precursor (Nash & Tate, 1982; Finidori et al., 1984; Laperche et al., 1986) which is encoded by four different mRNAs (I–IV), transcribed from a single-copy gene in the rat (Pawlak et al., 1988), and differing only in their 5′-untranslated region (Darbouy et al., 1991). The 2.2 kb $mRNA_I$ and $mRNA_{II}$ (Chobert et al., 1990) are transcribed in the kidney from two proximal promoters (Rajagopalan et al., 1990; Kurauchi et al., 1991; Lahuna et al., 1992); the 2.4 kb $mRNA_{III}$ is transcribed from a more distal promoter on the gene (Brouillet et al., 1994). A fourth GGT mRNA ($mRNA_{IV}$) of a larger size (2.5 kb) was reported in the epididymis (Palladino et al., 1994) in the HTC hepatoma cells and in the small intestine where its accumulation increases along the crypt villus axis (Darbouy et al., 1991). We have now determined the complete sequence of the 5′-untranslated region specific for this 2.5 kb mRNA, and we have mapped its transcription start site on the gene upstream of the three promoters already characterized. This fourth promoter specifically regulates GGT gene transcription in the small intestine, in the epididymis, and in some hepatoma cells, and its expression results in a primary transcript which is then alternatively spliced into the 2.5 kb $mRNA_{IV-1}$ or the 2.2 kb $mRNA_{IV-2}$.

MATERIALS AND METHODS

Biological Material. Rat tissues were obtained from female Wistar rats. Livers were excised from 18 day old fetuses and adults; kidney and small intestine were obtained from adult animals, and mammary glands were obtained from lactating rats. The epididymis RNA samples were a gift from Dr. Hinton (University of Virginia, Charlottesville, VA).

The HTC cell line is derived from the chemically induced Morris hepatoma 7288 (Thomson et al., 1966). Cells were grown in a Dulbecco's modified Eagle's medium (DMEM; Gibco, Cambridge, U.K.) supplemented with 5% calf serum, 200 units/mL penicillin, 50 μg/mL streptomycin, and 0.5 μg/mL fungizone. The other hepatoma cell lines (Fao, C2, C2Rev7, and H5) were derived from the clonal H4IIEC3 (H4II) cell line established from Reuber H-35 hepatoma

†This work was supported by the INSERM, the Association pour la Recherche contre le Cancer (ARC Grant 88791), and the University Paris-Val de Marne.

* To whom correspondence should be addressed. Telephone: 33 (1) 49 81 35 44. Fax: 33 (1) 48 98 09 08. Electronic mail: GUELLAEN@CITI 2. FR.

‡ Recipient of a fellowship from the Association pour la Recherche contre le Cancer.

® Abstract published in *Advance ACS Abstracts,* September 1, 1994.

[1] Abbreviations: CAT, chloramphenicol acetyltransferase; cRNA, complementary RNA; Denhardt's (50×), 1% Ficoll, 1% polyvinylpyrrolidone, and 1% BSA; GGT, γ-glutamyl transpeptidase; NPC, nonparenchymal cells; PCR, polymerase chain reaction; SSC, 0.15 M NaCl/15 mM citrate; SSPE, 0.18 M NaCl, 10 mM NaH_2PO_4, and 1 mM EDTA; UTR, untranslated region.

0006-2960/94/0433-11536$04.50/0

(Reuber, 1961; Deschatrette, 1980); they were cultured in modified Ham's F12 medium (Fd, Gibco).

Enzymes and Other Material. AMV reverse transcriptase, T4 DNA ligase, Taq DNA polymerase, and S1 nuclease were purchased from Stratagene cloning systems, La Jolla, CA; Sp6 and T7 RNA polymerases, T_4 polynucleotide kinase, and pGEM-3Zf(+) plasmid were from Promega, Charbonnières, France; the terminal transferase and RNA molecular weight markers were from Life Technologies, Eragny, France. The oligodeoxyribonucleotides were synthesized on an Applied Biosystems Model 380 B DNA synthesizer and purified on an OPC cartridge as recommended by the supplier. The labeled nucleotides and the nylon membrane (Hybond-N) were purchased from Amersham, Les Ulis, France.

5′ End cDNA Cloning. The procedure was based on the rapid amplification of cDNA 5′ ends (5′-RACE) technique (Frohman et al., 1988) with the modifications described previously (Brouillet et al., 1994). Poly(A)$^+$ RNA samples purified on an oligo[d(T)] column from HTC and from small intestine total RNA were used as a template to synthesize the first cDNA strand. An oligo[d(T)] sequence was added on the 3′ end of the first strand, and second cDNA strand synthesis was initiated from an anchor oligomer complementary to the poly[d(T)] sequence and flanked on its 5′ end by an *Eco*RI site [5′-CCAACAGAATTC(A)$_{18}$-3′]. The sequence between the anchor oligomer and a GGT-specific oligomer (oligomer 1), complementary to residues +29 to +53 in the GGT coding sequence (Laperche et al., 1986) (Figure 1), was amplified and cloned into the *Sma*I site of pGEM-3-Zf(+). The recombinants were screened for GGT sequences using labeled oligomer 2 which contains residues 1–21 in the coding GGT sequence (Figure 1). The longest GGT recombinant was sequenced using the Sequenase kit from United States Biochemical Corp. (Cleveland, OH) and [α-^{35}S]dATP.

Analysis of Rat Genomic DNA. A GGT recombinant phage, which contains the transcription start site for GGT mRNA$_{III}$ (Brouillet et al., 1994), was previously isolated from a rat genomic library. It was digested with either *Bam*HI or *Sac*I and analyzed on a Southern blot for HTC GGT-specific sequences. A 244 bp *Eco*RI fragment, which corresponds to the 5′ end of the HTC GGT-specific cDNA sequence, was labeled using the megaprime DNA labeling kit from Amersham Radiochemical Centre. This probe (30×10^6 cpm) was hybridized to the phage DNA in 2× SSC,[1] 1× Denhardt,[1] 0.5% SDS, 2 mM EDTA, 25 mM phosphate buffer, pH 7.2, and 100 μg/mL denatured salmon sperm DNA. The filters were washed in 0.1× SSC/0.1% SDS at 68 °C. The 5.9 kb *Sac*I and the 3.7 kb *Bam*HI overlapping fragments, which hybridize to the probe, were subcloned into the corresponding sites of pGEM-3-Zf(+) to give the pGEM-G-III and pGEM-G-IV plasmids as diagrammed in Figure 2. These two plasmids were partially sequenced using oligomers designed from the HTC cDNA and genomic DNA sequences.

S1 Nuclease Mapping Analysis. A 310 bp genomic DNA which contains the putative start site for GGT mRNA$_{IV}$ was synthesized by the polymerase chain reaction (PCR) using oligomers 4 and 3 (Figure 2). The downstream oligomer (oligomer 3) was labeled at its 5′ end by T4 polynucleotide kinase and [γ-^{32}P]ATP prior to the amplification reaction which was carried out in 100 μL of 1.5 mM of each dNTP, 50 mM KCl, 10 mM Tris-HCl, pH 8.3, 1 mM $MgCl_2$, 0.01% gelatin, and 2.5 units of Taq DNA polymerase. The DNA (100 ng) and the oligomers (40 pmol) were denatured by heating the reaction mixture to 94 °C for 5 min, and a 30 cycle amplification was performed. Each cycle consisted of a 1 min denaturation step at 94 °C, a 1 min annealing step at 55 °C, and a 1.5 min elongation step at 72 °C. The resulting labeled DNA fragment was purified on a 1.5% low-melting agarose gel, and an aliquot (4×10^5 cpm) was mixed with 20 μg of poly(A)$^+$ RNA and digested with S1 nuclease as previously described (Kurauchi et al., 1991). The size of the protected DNA fragments was analyzed on a polyacrylamide denaturing gel along with the related sequence ladder initiated from oligomer 3 which is the downstream primer used to generate the 310 bp DNA probe. The transcription start sites were directly identified from the sequencing fragments which comigrate with the DNA-protected fragments.

Construction of Chloramphenicol Acetyltransferase (CAT) Plasmids. The genomic sequence which extends upstream from the start site for GGT mRNA$_{IV}$ was inserted into the pMW-2-CAT vector which is a promotorless chloramphenicol acetyltransferase expression vector (Brouillet et al., 1994). The longest CAT plasmid was obtained by cloning of the 942 bp *Apa*I–*Hin*cII fragment (–903 +38) from pGEM-G-IV (after filling of the 3′ ends by the Klenow fragment) into the *Sma*I site of the pMW-2-CAT plasmid; we obtained the recombinant plasmids P$_{IV}$(–903 +38) and P$_{IV}$(+38 –903) which contain the GGT sequence in both orientations. The plasmids P$_{IV}$(–264 +38) and P$_{IV}$(+38 –264) were obtained by cloning of the *Pst*I–*Hin*cII fragment (–264 +38) in both orientations into the *Pst*I–*Stu*I or *Pst*I–*Sma*I sites of the CAT vector. Two 5′-deletion mutants, P$_{IV}$(–164 +38) and P$_{IV}$(–68 +38), were obtained by amplification of the GGT sequences between oligomers 4–5 and 6–5 (Figure 2), respectively. Oligomer 4 (–170 to –149) was mutated at nucleotide –165 (G for T) to generate a *Pst*I site, and oligomer 6 corresponds to the gene sequence from –76 to –57 where three nucleotides have been changed to create a *Pst*I site (–72 to –67). The amplified products were digested with *Pst*I and *Hin*cII and cloned into the *Pst*I–*Sma*I sites of pMW-2-CAT to give the recombinant plasmids P$_{IV}$(–164 +38) and P$_{IV}$(–68 +38). The length and the orientation of these plasmids were verified by sequencing. The pRSV-CAT vector, which was used as a positive control in the transfection experiments, was described previously (Kurauchi et al., 1991).

Transient Transfection and CAT Assay. The HTC rat hepatoma cells were transfected, and the CAT activity was assayed as previously described (Kurauchi et al., 1991). Each fusion plasmid was transfected at least 4 times using at least two different plasmid preparations. In each experiment, the transfections were performed in duplicate.

Northern Blot Analysis. Poly(A)$^+$ RNA samples purified on an oligo[d(T)] column were separated on a 1.5% agarose denaturing gel along with an RNA ladder, and transferred onto a nylon membrane. The membranes were analyzed for GGT mRNA sequences by hybridization to three different probes. The cRNA-4–1 probe contains a sequence complementary to the 140 bases downstream of the major transcription start site (+1) (Figure 2), which corresponds to the entire leader exon of mRNA$_{IV-1}$ and mRNA$_{IV-2}$. This probe was transcribed from a plasmid (pGEM-7-HTC 3) which contains the mRNA$_{IV}$ leader exon inserted into the *Eco*RI site of pGEM-7Zf(+) in front of the T_7 promoter. A second hybridization employed a 238 bp cDNA probe (cDNA-4–2) which covers the second exon of mRNA$_{IV}$. This DNA sequence was synthesized by amplifying the genomic DNA between oligomers 7 and 8 (Figure 2) in the conditions described above for the S1 mapping analysis; it was ^{32}P labeled using the random kit from Amersham. After dehybridization, this blot was hybridized to the cRNA-139A probe. This

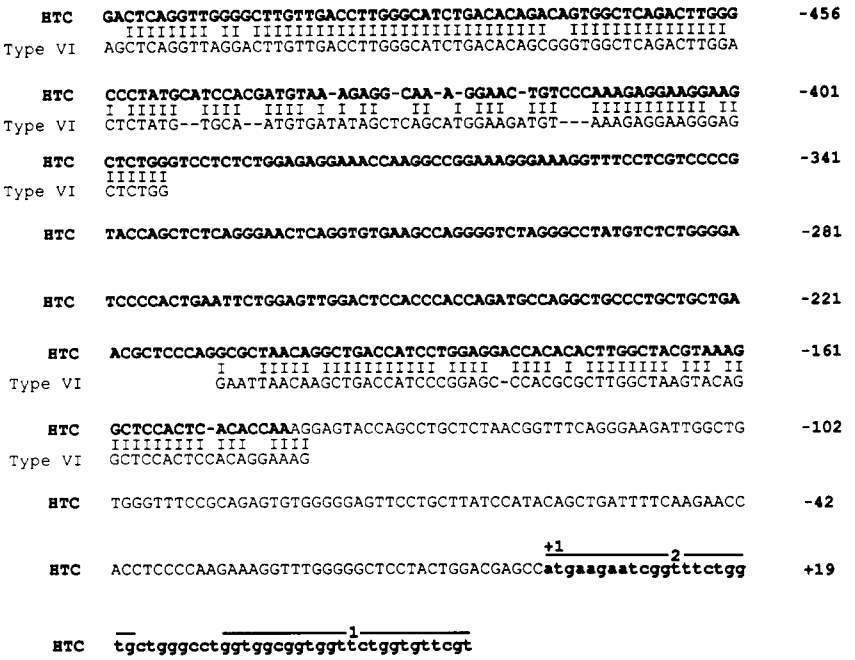

FIGURE 1: Nucleotide sequence of the cDNA cloned from HTC hepatoma cells. Nucleotides are numbered from the GGT initiation codon (+1). The sequence in boldface upper case letters is specific to HTC GGT cDNA; the HTC sequences in light face upper case letters and in boldface lower case letters (coding sequence) are 100% identical to the sequences reported for other rat GGT mRNA types. The unique HTC cDNA sequence was aligned to the specific sequence reported for mouse type VI cDNA. The sequences of oligomers 1 and 2 used in cDNA synthesis are overlined.

complementary probe covers the GGT mRNA sequence from nucleotide −72 in the 5′ UTR to nucleotide +1557 in the coding region (Laperche et al., 1986) which is common to the different GGT mRNAs. All the conditions for the *in vitro* transcription have been described elsewhere (Darbouy et al., 1991). The blots were washed in 0.1× SSPE at 68 °C (cDNA probe) or at 72 °C (cRNA probe).

RESULTS

Cloning the 5′ End of the GGT mRNA Expressed in the Small Intestine and in HTC Hepatoma Cells. The longest GGT cDNA recombinant plasmid obtained from the HTC mRNA contains a 568 bp sequence extending from nucleotide +53 in the coding region up to nucleotide −515 in the 5′-untranslated region (UTR) (Figure 1). The length of the 5′ UTR (515 nucleotides) is close to the size (525 nucleotides) previously estimated for that of the GGT $mRNA_{IV}$ expressed in the HTC and small intestine (Darbouy et al., 1991). The sequence downstream of nucleotide −144 is 100% homologous to the other cDNA sequences already reported for GGT $mRNA_{I}$, $mRNA_{II}$, and $mRNA_{III}$ (Chobert et al., 1990; Brouillet et al., 1994); however, this cDNA differs totally from the other rat mRNA sequences by a unique sequence of 371 nucleotides upstream of nucleotide −144. All the other GGT recombinant clones, obtained from HTC or the small intestine, exhibit the same sequence, which was shorter in its 5′-specific end.

Isolation and Characterization of the GGT Genomic Sequence around the GGT $mRNA_{IV}$ Transcription Start Site. A GGT positive recombinant phage, already known to contain the transcription start site for GGT $mRNA_{III}$ (Brouillet et al., 1994), was digested by either *Bam*HI or *Sac*I and analyzed on a Southern blot for HTC-specific cDNA sequences. These restriction sites, mapped in Figure 2 (upper part), delineate overlapping fragments of 3.7 kb (*Bam*HI) and 5.9 kb (*Sac*I), which both hybridize to the HTC GGT cDNA-specific sequence. These fragments were ligated into the homologous sites of pGEM-3 Zf(+), giving rise to the recombinant plasmids pGEM-G-III and pGEM-G-IV as diagrammed in Figure 2 (upper part). A partial sequence obtained from oligonucleotides designed from the cDNA, and the subsequent genomic DNA sequence, is reported in Figure 2 (lower part). Alignment of the rat HTC cDNA to the genomic sequence shows that the 371 nucleotides at the 5′ end of the cDNA map on the gene onto two separate exons separated by a 0.4 kb intron. The second exon for the GGT $mRNA_{IV}$ (247 nucleotides) lies on the gene 4 kb upstream of the 5′ end of the $mRNA_{III}$ leader exon as shown in Figure 2 (upper part). All the intron–exon junctions follow the GT/AG rule (Breatnach & Chambon, 1981).

The $mRNA_{IV}$ transcription start site on the gene was determined by a protection assay of a 310 bp genomic sequence around the putative start site (Berk & Sharp, 1977). This 310 bp DNA probe, amplified from oligonucleotides 4 and 3 (Figure 2), was labeled at the 5′ end of its complementary strand as indicated under Materials and Methods. Hybridization of this DNA probe to poly(A)$^+$ RNA samples from the small intestine or HTC hepatoma cells prior to S1 nuclease digestion gave similar DNA protection patterns (Figure 3). The size of the four labeled protected DNA fragments (117, 121, 130, and 140 nucleotides) was determined by comparison to the sequence ladder initiated on the genomic DNA from oligomer 3 and run on the same gel. The 5′ end of the protected DNA fragments identifies 4 transcription start sites for the GGT $mRNA_{IV}$ within 24 bases and, therefore, a fourth promoter region (promoter IV) in the GGT gene. The longest protected fragment (140 nucleotides) corresponds to the major start site in the small intestine and HTC cells; its 5′ end maps on the nucleotide which is numbered +1 (Figure 2). The initiation of transcription at +26 occurs with a higher efficiency in the HTC cells as compared to the small intestine. No protection against the nuclease occured when the DNA was hybridized to fetal liver RNA; the upper band (310 bp) represents the undigested probe.

FIGURE 2: Structure of the 5′ region of the rat GGT gene. Upper part: Schematic representation of the phage containing the 5′ region of the GGT gene. The recombinant phage containing the 16 kb GGT genomic sequence is diagrammed: the leader exon for GGT $mRNA_{III}$ is represented by the hatched box and the two leader exons for the GGT $mRNA_{IV}$ sequences by the quadrilled boxes. The arrows map the major transcription start site for GGT $mRNA_{IV}$, and P_{III} stands for promoter III. Restriction sites used for subcloning are mapped. Subcloning of the *Sac*I and *Bam*HI overlapping fragments into pGEM-3Zf(+) results in the pGEM-G-III and pGEM-G-IV recombinant plasmids. The striped boxes represent vector DNA. Lower part: Nucleotide sequence of the rat GGT gene around the mRNAIV transcription start sites. The arrows map the rat $mRNA_{IV}$ transcription start sites, and the major one is numbered +1. The sequence of the oligomers used to synthesize the DNA fragments for S1 mapping experiments (3 and 4), the subcloning in the CAT vector (4, 5, and 6), and the cDNA-4–2 probe (7 and 8) used for Northern blot analysis are overlined. Nucleotides in boldface upper case letters correspond to sequences present in the mature rat $mRNA_{IV}$ or mouse type VI mRNA which map into two separate exons spaced by a 0.4 kb intron. Nucleotides in lower case letters are for sequences in the promoter region and in the introns. The asterisk maps to the 5′ end of the cDNA cloned from HTC. Consensus sequences for the binding of C/EBP-like proteins (CCAAT), the peroxisome proliferator activated receptor (AGGTCA), and the AP-1 (GACTCAG) and PEA3 (AGGAAA) factors are boxed.

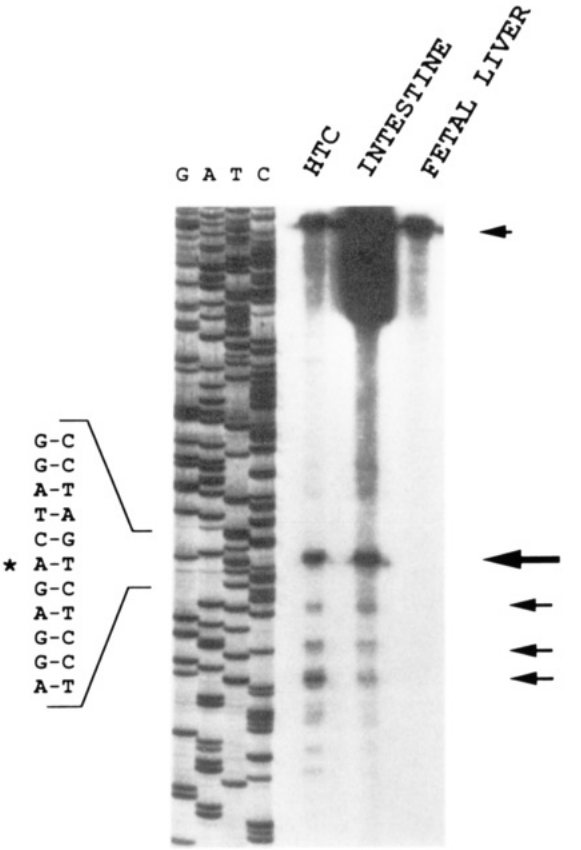

FIGURE 3: S1 nuclease mapping of the GGT $mRNA_{IV}$ transcription start site. A 310 bp genomic DNA fragment labeled at the 5′ end of its complementary strand, and which encompasses the putative $mRNA_{IV}$ start site, was used as a probe hybridized to 20 μg of poly-$(A)^+$ RNA and subjected to S1 nuclease digestion. The labeled DNA probe was obtained by amplification of the genomic sequence between oligomer 4 and oligomer 3 which had been labeled on its 5′ end by T_4 polynucleotide kinase and [γ-^{32}P]ATP prior to the amplification reaction. The protected DNA fragments (arrows) were resolved on a sequencing gel. The large arrow corresponds to the longest protected fragment; its 5′ end maps at the nucleotide (dot) numbered +1 in Figure 2. The upper band represents the 310 bp undigested probe. The position of the start sites on the gene was identified from sequencing reaction products generated from oligomer 3 run in adjacent lanes (GATC).

The GGT promoter IV is transcribed from multiple start sites like many other promoters which do not exhibit a TATA-like motif (Breatnach & Chambon, 1981) or an initiator element which acts alone or in concert to direct a specific transcription from a single site (Jahavery et al., 1994). All the start sites appear to consist of A residues, and the two major ones are preceded by pyrimidine residues. This consensus is generally reported for the start sites of genes transcribed with RNA polymerase II (Breatnach & Chambon, 1981; Jahavery et al., 1994). Several consensus sequences for transcription factors are found in the vicinity of the mRNA IV initiation sites: an AP-1 binding site (Schule et al., 1990) at +17, a CCAAT element at –79 (Benoist et al., 1980), a sequence at –94 which could bind receptors of the peroxisome proliferator activated receptor (PPAR) family (Isseman et al., 1990), and two PEA_3 sites (–225 and –174) which could bind transcription factors activated by oncogenes (Roth et al., 1990) (Figure 2).

The genomic sequence flanking the transcription start sites for GGT $mRNA_{IV}$ was examined for its ability to drive the expression of the CAT reporter gene in transient transfection assays in two hepatoma cell lines (Figure 4). In Fao cell, the P_{IV}(–68 +38) GGT plasmid displays a low promoter activity which represents only 6% of the activity driven by the RSV promoter in control experiments. Extension of the GGT sequence up to nucleotide –164 results in a slight increase (×1.7) in the promoter activity which is not altered by further extension of the GGT sequence. In HTC cells, the shorter GGT-CAT plasmid displays a promoter activity (6% of the control RSV value) which is comparable to that measured in Fao cells. However, extension of the GGT sequence to base –164 gives a sharp increase in the promoter activity (5-fold), thus revealing the presence of a strong positive cis-acting element between –164 and –68 in the GGT sequence. A further increase in the size of the GGT sequence in transfected plasmid results in a decrease of the activity to the level found in Fao cells, indicating the presence of a negative element between bases –264 and –164. No CAT activity was detected in the cells transfected by the promoterless pMW-2-CAT vector or by the constructs bearing the GGT sequence oriented in the opposite direction.

In summary, we have identified a fourth promoter in the rat GGT gene which lies 4.7 kb upstream of promoter III. The positions on the gene of the 5′ UTR sequence specific to the

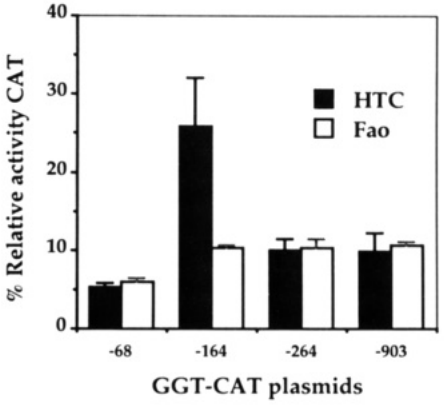

FIGURE 4: Transient expression of CAT activity in Fao and HTC hepatoma cells transfected by GGT-CAT plasmids. The GGT-CAT plasmids, described under Materials and Methods, are identified by the position of the most distal nucleotide relative to the transcription start site. The CAT activity driven by these plasmids was corrected for the protein amount and normalized to the values obtained with the RSV-CAT vector. Each value (±SEM) is the average of five (for Fao cells) and eight (for HTC cells) independent experiments, each performed with duplicate dishes.

four rat GGT mRNAs are reported in Figure 5 together with their transcription start sites. They map onto a 12 kb fragment upstream from the initiation codon (ATG).

Northern Blot Analysis. RNA hybridization to the leader exon of the mRNA (cRNA-4–1) transcribed from GGT promoter IV reveals a major band at 2.5 kb ($mRNA_{IV-1}$) in the small intestine, HTC, C2Rev7, and epididymis (Figure 6A) and a second band which migrates at 2.2 kb. The 2.2 kb mRNA ($mRNA_{IV-2}$) is different from the 2.2 kb $mRNA_I$ and $mRNA_{II}$ since there is no hybridization signal in the kidney sample (Figure 6A) where these two mRNAs are expressed at high levels (Chobert et al., 1990). Hybridization to the cDNA-4–2 probe, which covers most of the second exon of $mRNA_{IV}$, reveals the 2.5 kb $mRNA_{IV-1}$, also detected by the cRNA-4–1 probe, but not the 2.2 kb $mRNA_{IV-2}$ species (Figure 6B). Therefore, the absence of the second exon (251 bp) of $mRNA_{IV-1}$ in $mRNA_{IV-2}$ fully accounts for the size difference between these two transcripts. Both transcripts are initiated on the same promoter, and the primary transcript is alternatively spliced to the mature $mRNA_{IV-1}$ or $mRNA_{IV-2}$ transcripts. Hybridization of the RNA samples to the cRNA 139A probe, which recognizes the GGT coding sequence, reveals all the GGT mRNA species (2.2–2.5 kb) which, in addition to the $mRNA_{IV}$ species, includes the 2.2 kb $mRNA_I$ and $mRNA_{II}$ (Chobert et al., 1990) and the 2.4 kb $mRNA_{III}$ (Brouillet et al., 1994) (Figure 6C).

Comparison of the Rat and Mouse GGT Gene and mRNAs. The alignment of the rat $mRNA_{IV}$-specific sequence to the partial sequences reported for several GGT transcripts cloned from mouse kidney (Rajagopalan et al., 1993) reveals two highly conserved areas in the 5′ UTR of the mouse type VI GGT mRNA which, therefore, appears to be the mouse equivalent of the rat $mRNA_{IV}$ (Figure 1). The 138 nucleotide leader sequence of the rat $mRNA_{IV}$, which corresponds to its first exon, can be aligned with 77% identity to the 123 nucleotides at the 5′ end of the mouse type VI mRNA. The 68 nucleotide sequence at the 3′ end of the second exon of rat $mRNA_{IV}$ is 83% identical to the 3′ end of the mouse type VI mRNA-specific sequence. Therefore, the rat $mRNA_{IV}$ and the mouse type VI GGT mRNA share a similar leader exon but are differentially spliced. In the mouse, the splice acceptor site for the second exon is 184 bases downstream on the gene, resulting in a shorter 5′-untranslated sequence for the mouse as compared to the rat transcript (Figure 5).

The rat $mRNA_{III}$ leader exon, which has been characterized previously (Brouillet et al., 1994), also displays two areas of

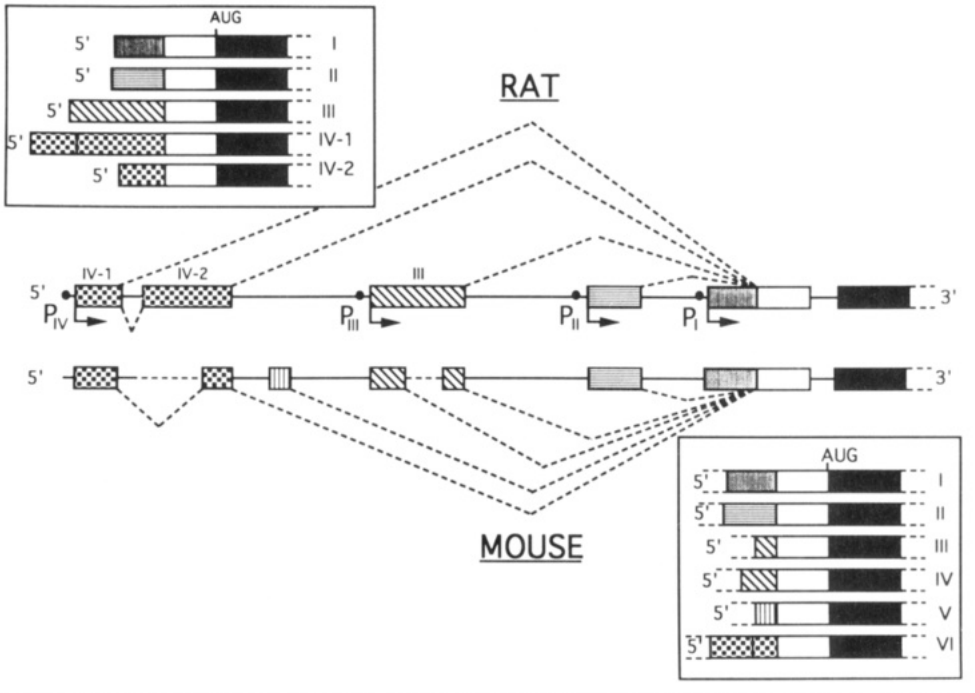

FIGURE 5: Comparison of the organization of the rat and the mouse GGT gene upstream of the translation start site. The 5′ end of the rat gene (upstream of ATG) is diagrammed at the top: the exons specific for $mRNA_I$, $mRNA_{II}$, $mRNA_{III}$, and $mRNA_{IV}$ are represented by stippled, striped, hatched, and quadrilled boxes, respectively. The open box stands for the 5′-untranslated sequence which is common to the four mRNAs. The black box represents the first translated exon. The transcription start sites are indicated by P_I, P_{II}, P_{III}, and P_{IV}, and the splicing patterns initiated on the four promoters are represented. The 5′ regions of the five GGT mRNA transcripts are represented in the left inset. The mouse gene sequence (Rajagopalan et al., 1993) is represented in the lower part. The mouse exons homologous to the rat sequences are represented by the same symbols; the vertically hatched box stands for an exon which is found only in mouse mRNA (type V). The splicing patterns and the 5′ regions of the mouse transcripts (right inset) are also represented.

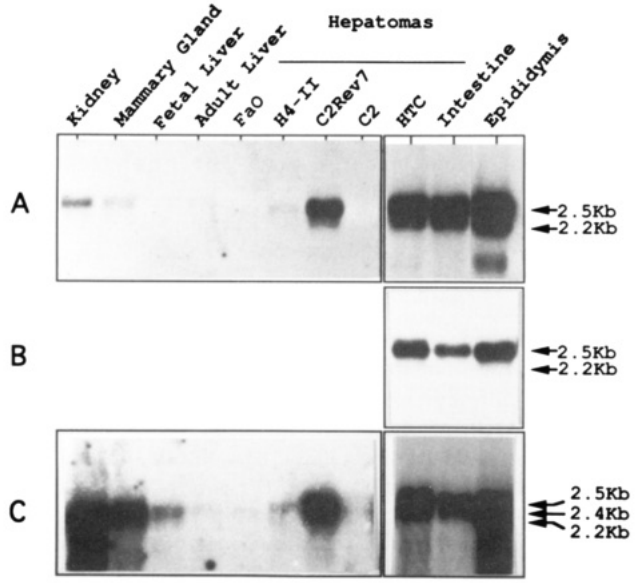

FIGURE 6: Northern blot analysis of GGT mRNAs. Poly(A)+ RNA (5 μg) samples were separated on two agarose denaturing gels and blotted onto a nylon membrane. These blots were hybridized to the cRNA-4–1 probe, which is specific for the HTC cDNA (A), and to the cRNA-139A probe, which hybridizes to the coding region common to all the GGT mRNA species (C). The blot on the right of the figure which includes HTC, small intestine, and epididymis samples was also hybridized to the cDNA-4–2 probe, which corresponds to the second exon of the HTC-specific sequence (B). Autoradiography was performed with two intensifying screens at –80 °C for 48 h.

extensive homology to mouse mRNA sequences: the 106 nucleotide leader sequence of the rat $mRNA_{III}$ is 78% identical to the sequence at the 5′ end of the mouse type IV mRNA, and the 49 nucleotide sequence at the 3′ border of the $mRNA_{III}$ leader exon can be aligned (94% identity) to the mouse type III mRNA leader sequence (Figure 7). The strong identity of the GGT gene sequence in these two different species extends beyond the $mRNA_{IV}$ and $mRNA_{III}$ transcription start sites in the promoter regions. The 181 bp sequence upstream of the major transcription start site for rat $mRNA_{IV}$ and the 233 nucleotide sequence upstream of the $mRNA_{III}$ start site are 82% and 90% identical to the mouse genomic sequences flanking the 5′ end of type VI and type IV mRNA, respectively (Figure 7). A comparison of mouse type I and type II leader sequences to rat $mRNA_I$ and $mRNA_{II}$ has also revealed a strong identity between the mouse and rat transcripts (Rajagopalan et al., 1993). The areas of identity between the rat and mouse exons upstream from the initiation codon are diagrammed in Figure 5.

DISCUSSION

In HTC hepatoma cells, small intestine, and epididymis, the GGT gene is transcribed into a 2.5 kb and a 2.2 kb mRNA. Cloning of the cDNA demonstrates that the 2.5 kb mRNA ($mRNA_{IV-1}$) differs from the other GGT transcripts ($mRNAs_{I,II,III}$) by a 371 base unique leader sequence which, in the gene, maps into 2 separate exons 4 kb upstream of the most distal promoter previously identified (Brouillet et al., 1994). The genomic sequence flanking the $mRNA_{IV}$ start site exhibits a promoter activity when transfected into HTC hepatoma cells in front of the CAT gene. The $mRNA_{IV-1}$ and the $mRNA_{IV-2}$ are transcribed from a fourth promoter (pIV) on the GGT gene which is distal to the three transcription units previously reported (Rajagopalan et al., 1990; Kurauchi et al., 1991; Lahuna et al., 1992; Brouillet et al., 1994). Their size difference results from the alternate splicing of the second exon. The complex transcriptional organization of the rat GGT gene gives rise to the synthesis of five mRNAs with an

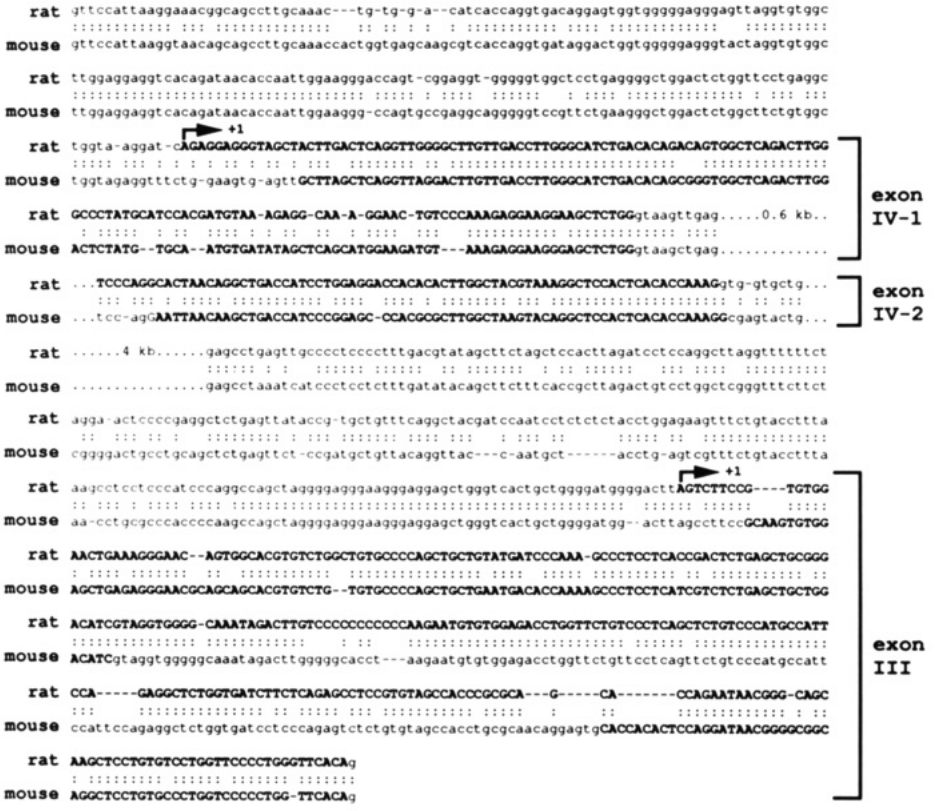

FIGURE 7: Comparison of the rat and mouse GGT sequences. The rat genomic sequence (upper row) around the $mRNA_{III}$ and $mRNA_{IV}$ transcription start sites is aligned to the mouse genomic one (lower row). Exons are in boldface upper case letters and the major transcription initiation sites for $mRNA_{III}$ and $mRNA_{IV}$ are indicated by arrows.

alternate 5′ end (Figure 5), which encode the same GGT polypeptidic chain.

Recently, six mRNAs displaying heterogeneous 5′ ends have been cloned from the mouse kidney (Rajagopalan et al., 1993); however, their transcription start sites on the gene are still unknown. Comparison of the rat and mouse cDNAs to the genomic sequences reveals that the mouse type I, II, IV, and VI mRNA-specific sequences exhibit a strong identity to the leader sequences of rat type I, II, III, and IV mRNA, respectively, and that the identity between the rat and mouse genes extends upstream from the transcription start sites that we have identified in the rat gene. From this comparison, it can be concluded that the mouse type I, II, III, and IV mRNAs are also transcribed from four independent promoters on the mouse gene, thus indicating that the GGT gene exhibits a similar transcriptional organization in both species. However, the origin of the mouse type III and type V mRNAs is still unclear. They could result from alternative splicing of a primary transcript initiated on promoter III or on promoter IV.

Three other genes, coding for IGF-II (Dijk et al., 1991), for cytochrome P-450 aromatase (Mahendroo et al., 1993), and for UDP glucuronyltransferase (UDPGT) (Ritter et al., 1992), are also transcribed from four promoters into mRNA with alternate 5′ ends. The IGF-II and aromatase transcripts are translated into the same polypeptide chain whereas the sequence differences observed for the UDPGT mRNAs extend into the reading frame and generate different isoforms of the enzyme with alternate amino-terminal sequences (Ritter et al., 1992).

Analysis of GGT gene expression by Northern blotting shows that the two mRNAs IV (IV-1 and IV-2) are transcribed from promoter IV which exhibits a strong tissue or cell specificity since a significant accumulation of these transcripts was observed only in the small intestine, in the epididymis, and in some hepatoma cells. Transient transfection assays in hepatoma cells indicate that this specificity is mediated by the interplay of positive and negative cis-acting elements. In the small intestine, $mRNA_{IV}$ accumulation accounts for all the GGT expression, which increases from the undifferentiated cells at the basis of the villi to the differentiated enterocytes at the top of the villi (Marathe et al., 1979). In fact, the GGT promoters I and II are not expressed in this tissue, and the GGT $mRNA_{III}$ is expressed only at an extremely low level only detectable after a 30 cycle PCR experiment (Darbouy et al., 1991). The GGT promoter IV activity, which follows that of the fatty acid binding protein (FABP) (Iseki & Kondo, 1990) or the sucrase isomaltase gene (Traber, 1990) along the villi, can also be viewed as a hallmark of the differentiated enterocyte phenotype.

Unexpectedly, GGT promoter IV, which otherwise is silent in fetal and adult liver, is expressed in HTC and C2Rev7 cells at a level comparable to that found in differentiated enterocytes. This could be indicative of the common cellular origin of intestinal and liver cells (Balinsky, 1981). Treatment of rats with carcinogenic or noncarcinogenic toxic compounds induces the proliferation of non parenchymal cells (NPC), also known as oval cells, in the periportal area of the liver (Lemire et al., 1991; Dabeva & Shafritz, 1993). These cells are often recognized as hepatic precursor cells since they have the capacity to differentiate into hepatocytes or biliary cells (Germain et al., 1988). In some circumstances, they can undergo differentiation toward a phenotype that closely ressembles the epithelial cells covering the villi of the small intestine (Tatematsu et al., 1985), and they form ductlike structures lined by a tall columnar epithelium with scattered goblet cells near the biliary ducts. Hepatomas could originate from these pluripotent cells, which still have the capacity to evolve toward different lineages. Expression of the GGT promoter IV could be a marker revealing that these hepatoma cells originate from precursor cells whose differentiation program has been altered by the carcinogenic process. In contrast, expression of the 2.4 kb GGT $mRNA_{III}$ in primary tumors induced by diethylnitrosamine treatment (Darbouy et al., 1991) could reveal that these tumors originate from more differentiated cells.

In conclusion, the tissue specificity observed for the expression of the γ-glutamyl transpeptidase gene is based on the use of multiple promoters. A further study of the factors that govern GGT promoter IV expression in small intestinal cells will contribute to knowledge of the molecular mechanisms that control cell differentiation in the enterocyte lineage.

ACKNOWLEDGMENT

We thank Edith Grandvilliers for her skillfull secretarial assistance, and Drs. Aggerbeck, Barouki, Guellaën, and Pawlak for critical reading of the manuscript.

REFERENCES

Balinsky, B. I. (1981) *An introduction to embryology*, pp 524–525, Holt Saunders, Philadelphia PA.

Beer, D. G., Schwartz, M., Sawada, N., & Pitot, H. C. (1986) *Cancer Res. 46*, 2435–2441.

Benoist, C., O'Hare, K., Breathnach, R., & Chambon, P. (1980) *Nucleic Acids Res. 8*, 127–142.

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

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

Brouillet, A., Darbouy, M., Okamoto, T., Chobert, M. N., Lahuna, O., Garlatti, M., Goodspeed, D., & Laperche, Y. (1994) *J. Biol. Chem. 269*, 14878–14884.

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.

Dabeva, M. D., & Shafritz, D. A. (1993) *Am. J. Pathol. 143 (6)*, 1606–1620.

Darbouy, M., Chobert, M. N., Lahuna, O., Okamoto, T., Bonvalet, J. P., Farman, N., & Laperche, Y. (1991) *Am. J. Physiol. 261*, C1130–C1137.

Deschatrette, J., Moore, E. E., Dubois, M., & Weiss, M. C. (1980) *Cell 19*, 1043–1051.

Dijk, M. A., Van Bootsma, H. J., Van Schaik, F. M., Holthuizen, P., & Sussenbach, J. S. (1991) *Mol. Cell. Endocrinol. 81*, 81–94.

Finidori, J., Laperche, Y., Haguenauer-Tsapis, R., Barouki, R., Guellaën, G., & Hanoune, J. (1984) *J. Biol. Chem. 259*, 4695–4698.

Frohman, M. A., Dush, M. K., & Martin, G. R. (1988) *Proc. Natl. Acad. Sci. U.S.A. 85*, 8998–9002.

Germain, L., Blouin, M. J., & Marceau, N. (1988) *Cancer Res. 48*, 4909–4918.

Hampton, L. L., Worland, P. J., Yu, B., Thorgeisson, S. S., & Huggett, A. (1990) *Cancer Res. 50*, 7460–7467.

Hanigan, M. H., & Pitot, H. C. (1985) *Carcinogenesis 6*, 165–172.

Hanigan, M., & Rickett, W. A. (1993) *Biochemistry 32*, 6302–6306.

Iseki, S., & Kodo, H. (1990) *J. Histochem. Cytochem. 38*, 111–115.

Isseman, I., & Green, S. (1990) *Nature 347*, 645–651.

Jahavery, R., Khachi, A., Lo, K., Zenzie-Gregory, B., & Smalet, T. S. (1994) *Mol. Cell. Biol. 14*, 116–127.

Kurauchi, O., Lahuna, O., Darbouy, M., Aggerbeck, M., Chobert, M. N., & Laperche, Y. (1991) *Biochemistry 30*, 1618–1623.
Lahuna, O., Brouillet, A., Chobert, M. N., Darbouy, M., Okamoto, T., & Laperche, Y. (1992) *Biochemistry 31*, 9190–9196.
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.
Lemire, J. M., Shiojiri, N., & Fausto, N. (1991) *Am. J. Pathol. 139*, 535–552.
Mahendroo, M. S., Mendelson, C. R., & Simpson, E. R. (1993) *J. Biol. Chem. 268*, 19463–19470.
Marathe, G. V., Nash, B., Haschemeyer, H., & Tate, S. S. (1979) *FEBS Lett. 107*, 436–440.
Nash, B., & Tate, S. S. (1982) *J. Biol. Chem. 257*, 585–588.
Pallardo, F. Y., Sastre, J., Asensi, M., Rodrigo, F., Estrela, J. M., & Vina, J. (1991) *Biochem J. 274*, 891–894.
Palladino, M., Laperche, Y., & Hinton, B. (1994) *Biol. Reprod. 50*, 320–328.
Pawlak, A., Lahuna, O., Bulle, F., Suzuki, A., Ferry, N., Siegrist, S., Chikhi, N., Chobert, M. N., Guellaën, G., & Laperche, Y. (1988) *J. Biol. Chem. 268*, 9913–9916.
Rajagopalan, S., Park, J., Patel, H., Lebovitz, R. M., & Lieberman, M. W. (1990) *J. Biol. Chem. 265*, 11721–11725.
Rajagopalan, S., Wan, D. F., Habib, G. M., Sepulveda, A. R., McLeod, M. R., Lebovitz, R. M., & Lieberman, M. W. (1993) *Proc. Natl. Acad. Sci. U.S.A. 90*, 6179–6183.
Reuber, M. D. (1961) *Cancer Int. 26*, 891–899.
Ritter, J. K., Chen, F., Sheen, Y. Y., Tran, H. M., Kimura, S., Yeatman, M. T., & Owens, I. S. (1992) *J. Biol. Chem. 267*, 3257–3261.
Roth, P., Nerlov, C., Blasi, F., & Johnsen, M. (1990) *Nucleic Acids Res. 18*, 5009–5017.
Schüle, R., Umesono, K., Mangelsdorf, D. J., Bolado, J., Pike, J. W., & Evans, R. M. (1990) *Cell 61*, 497–504.
Tate, S. S., & Meister, A. (1981) *Mol. Cell. Physiol. 39*, 357–368.
Tatematsu, M., Kaku, T., Medline, A., & Farber, E. (1985) *Lab. Invest. 52*, 354–362.
Thomson, E. B., Tomkins, G. M., & Currare, J. F. (1966) *Proc. Natl. Acad. Sci. U.S.A. 56*, 296–303.
Traber, P. G. (1990) *Biochem. Biophys. Res. Commun. 173*, 765–773.