Expression Patterns of the Multiple Transcripts from the Folylpolyglutamate Synthetase Gene in Human Leukemias and Normal Differentiated Tissues*

Fiona B. Turner‡, Shirley M. Taylor§, and Richard G. Moran¶¶

From the Departments of ¶Pharmacology and Toxicology and §Microbiology and Immunology and the ¶Massey Cancer Center, Medical College of Virginia, Virginia Commonwealth University, Richmond, Virginia 23298

Published, JBC Papers in Press, August 29, 2000, DOI 10.1074/jbc.M005228200

Received for publication, June 15, 2000, and in revised form, August 21, 2000

Folylpoly-γ-glutamate synthetase (FPGS) catalyzes the activation of folate antimetabolites in mammalian tissues and tumors. We have determined the sequence, abundance, and function of human FPGS transcripts and found some striking differences to transcription of the mouse gene that allow production of FPGS isoforms in mouse liver and dividing tissues. Multiple human transcripts were identified, including the homolog of the mouse transcripts that initiate at two upstream exons. However, the human FPGS upstream promoter is infrequently used, and transcripts from this promoter include sequences homologous with only one of the upstream exons found in the mouse. The downstream promoter generates an array of transcripts, some of which do not produce active enzyme, a phenomenon not seen in the mouse. Hence, the dual promoter mechanism directing expression of FPGS isoforms in mouse tissues is not conserved in humans, and, unlike the mouse downstream promoter, the human downstream promoter is active in both dividing and differentiated tissues. This study raises questions about the differences in function served by the two mouse FPGS isoforms and how, or if, human tissues fulfill these functions. How humans and mice produce FPGS in only a subset of tissues using such different promoter structures also becomes a central issue.

Folylpoly-γ-glutamate synthetase (FPGS) catalyzes the activation of folate antimetabolites to intracellular folate cofactors. The folylpolyglutamate products of this reaction are no longer substrates for cellular efflux mechanisms, so that folate cofactors accumulate in the cell to levels that drive metabolism. The folylpolyglutamates are also superior substrates for many folate-dependent enzymes (1), and the function of FPGS has been shown to be essential for the survival of dividing mammalian cells (2). FPGS is recognized to play a central role in the activity of cancer chemotherapeutic antifolates. Metabolism by FPGS increases their intracellular concentration, and their cytotoxicity is closely linked to this metabolism (3). FPGS is expressed in mouse liver and kidney and all tumors and dividing cell lines (4, 5); significant levels of mRNA for FPGS were not found in a scan of other mouse differentiated tissue (6). Recently, it was discovered that mouse liver and kidney express a different isozyme of FPGS than that found in normal and malignant dividing tissues (7). We do not yet understand the physiological significance of these FPGS isoforms, but the mouse has evolved the tissue-specific expression of two subtly different enzyme species.

The expression patterns from the two promoters have been well characterized in the mouse (7, 8). Two 5′ terminal exons, located 10 kb upstream from the body of the gene (8), are spliced to exon 2 in transcripts that are expressed only in a few differentiated tissues (7). In contrast, any normal or neoplastic dividing tissues expresses transcripts that initiate at the downstream exon 1 (7). Interestingly, kidney contains both classes of transcripts. Mouse FPGS transcripts initiating at the upstream or downstream promoters have been shown to translate to isoforms of FPGS with different kinetic properties (7). Clearly, a very tightly controlled mechanism has evolved to direct transcription from either the upstream or downstream promoters in the mouse, although the details of this mechanism are not yet clear.

The human FPGS gene is located on chromosome 9 (9) and was originally reported (10) to have 15 exons spanning 11.2 kb of genomic DNA in an organization nearly identical to that found later for the downstream region of the mouse gene (11). An initial characterization of the downstream promoter of the human gene (6) showed it to be very similar to that of the mouse (12). Others have reported that a complex set of transcriptional start sites exist in the vicinity of this downstream human promoter (9). An understanding of whether different FPGS isoforms are expressed in different human tissues is of paramount importance because of the ability of FPGS to activate drugs of potential therapeutic use in humans, particularly because the differences between mouse isoforms result in differences in substrate preference (7). Such a difference in activation of cytotoxic drugs by different FPGS isoforms could potentially be used for targeting of drugs or modification of drug toxicity profiles (7). Others have suggested that human ALL cells, which are sensitive to methotrexate, and AML cells, which are refractory, express species of FPGS that differ in the kinetics of activation of methotrexate (13).

In the work reported here, we sought to identify whether a mechanism similar to the one directing the expression of FPGS isoforms in different tissues of the mouse also exists in humans. We now report that the human FPGS gene contains two upstream exons similar to those of the mouse and that an
alternative, functional human FPGS species can be produced from a transcript that includes only one of these exons. We found no molecular evidence for differences in primary sequence between the enzymes expressed in AML and ALL leukemias. When we evaluated the expression and functional significance of the multiple FPGS transcripts found in humans, a tissue-specific pattern of isozyme expression was not found, suggesting an evolutionary divergence in the control of FPGS expression between these two species.

EXPERIMENTAL PROCEDURES

Materials—Human cell lines MOLT-3 (acute lymphocytic leukemia), K-562 (acute myelogenous leukemia), HL-60 (acute promyelocytic leukemia), and CEM (acute lymphocytic leukemia) were obtained from the American Type Culture Collection (Manassas, VA). AUXB1 cells were originally obtained from Dr. Gordon Whitmore of the Ontario Cancer Center. Poly(A)+ RNA from normal human liver pooled from two Caucasian males and from skeletal muscle from a pool of 10 males and females was purchased from CLONTECH. The α modification of minimum Eagle’s medium formulated with (α−) or without nucleosides (α−) was purchased from Life Technologies, Inc.

5'-RACE—Poly(A)+ RNA from human liver (CLONTECH) was reverse transcribed with Superscript II (Life Technologies, Inc.) or Thermoscript RT (Life Technologies, Inc.) using an antisense primer (α) (5′-ACTCGTGCGAGGAGTACGTCGAGGTTAGC) in exon 2 (see Fig. 1B) of the human FPGS gene. A polydeoxycytidylate tail was added with terminal deoxynucleotidyl transferase, and PCR was performed using an internally nested antisense primer (β) (5′-GGTACGTCGAGGTTAGCAGGTTAGCG) to clone the 5′-GAGGAAGTACGTCGAGGTTAGC) in exon 2 and the 5′ α anchor primer. The cDNAs were cloned into the pCR2.1-TOPO vector (Invitrogen) and transformed into Top10 One Shot cells (Invitrogen). Forty-five clones were manually sequenced using Sequenase 2.0 (Amersham Pharmacia Biotech). To amplify exon 1a a single primer was used and a second nested primer (c) (5′-AATGATGCAGAATTCGAC) in exon 2 and the 5′ anchor primer. The cDNAs were cloned into the pCR2.1-TOPO vector (Invitrogen) and transformed into Top10 One Shot cells (Invitrogen). Forty-five clones were manually sequenced using Sequenase 2.0 (Amersham Pharmacia Biotech). To amplify exon 1a a single primer was used and a second nested primer (c) (5′-AATGATGCAGAATTCGAC) in exon 2 and the 5′ anchor primer. The cDNAs were cloned into the pCR2.1-TOPO vector (Invitrogen) and transformed into Top10 One Shot cells (Invitrogen). Forty-five clones were manually sequenced using Sequenase 2.0 (Amersham Pharmacia Biotech). To amplify exon 1a a single primer was used and a second nested primer (c) (5′-AATGATGCAGAATTCGAC) in exon 2 and the 5′ anchor primer. The cDNAs were cloned into the pCR2.1-TOPO vector (Invitrogen) and transformed into Top10 One Shot cells (Invitrogen). Forty-five clones were manually sequenced using Sequenase 2.0 (Amersham Pharmacia Biotech). To amplify exon 1a a single primer was used and a second nested primer (c) (5′-AATGATGCAGAATTCGAC) in exon 1 (see Fig. 1B). The exon 1c+2 probe contains the sequence from the HindII site at genomic nt 448–1469 in exon 2. The exon 1a+2 probe initiates at the AvaII site at nt 1314–1449 in exon 2. The exon 1+2 probe was transcribed from an EcoRI/NcoI fragment (nt 104–1455) cloned into pGEM blue (Promega). RNA transcripts were generated from linearized cDNAs using α-32PUTP and either SP6 or T7 RNA polymerase. The full lengths of the undigested probes are as follows (in nt): exon 1a+2, 194; exon 1c+2, 220; exon 1c+2, 222; and exon 2a+2, 214. The sense complement of each probe was transcribed in vitro and used as a standard for the mobility of a single molecular species that was a perfect match to the probe. Total RNA was extracted with Trizol (Life Technologies, Inc.), and poly(A)+ RNA was selected using an Oligotex mRNA kit (Qiagen). Ribonuclease protection assay (RPA) analysis (Molecular Dynamics), and data were corrected by the U content of each fragment to allow comparison of the molar proportions of expressed transcripts.

Reverse Transcriptase-PCR—Human liver, skeletal muscle, and CEM poly(A)+ cDNAs (1 μg) were reverse transcribed with Superscript II (Life Technologies, Inc.) and random hexamers. Genomic DNA was prepared from CEM cells by phenol-chloroform extraction. Primers from the human sequence homologous to mouse exon A1a (see Fig. 2) were paired with primers in exons 4 and 6 for human liver. In other experiments, sense primers from exon 1A were paired with antisense primers in exon 2, using human liver, skeletal muscle, and CEM cDNAs as templates to detect trace levels of transcripts containing exon 1A.

RESULTS

Mapping Transcriptional Start Sites of the FPGS Gene in Human Liver by 5'-RACE—Because mouse FPGS shows a tightly controlled, tissue-specific pattern of FPGS isoform expression, we set out to define the distribution of FPGS transcripts in human tissues, initially focusing on liver. Using an antisense primer in exon 6 (21), we reverse transcribed liver poly(A)+ RNA and amplified the resulting cDNA by PCR with an internally nested antisense primer specific to exon 2 (Fig. 1B) and a 5′ anchor primer. These RACE products were ligated into pCRII, and 47 clones were sequenced (Table 1). Among these clones, we detected a RACE product whose initial sequence was remarkably similar to that of mouse exon 1B (7, 8)
We previously demonstrated that the mouse exon A1b has mitochondrial and cytosolic translational start sites (7) similar to what was reported for the human exon 1 (14). Mouse proteins initiating at the more downstream ATG in exon A1b encode cytosolic FPGS, whereas those initiating at the upstream ATG encode mitochondrial FPGS. Two triplets encoding methionine were conserved in the human exon A1b homolog at the identical position to the translational start codons encoded in mouse exon A1b (Fig. 1A).

Two striking differences were noted from our experience with mouse liver FPGS gene expression. In contrast to the mouse, the human exon A1b homolog was infrequently represented in the RACE clones sequenced; 2 of 47 clones contained a human exon A1b homolog (Table I), whereas 10 of 11 mouse liver RACE clones contained exons A1a and A1b (7). Secondly, human clones containing a sequence corresponding to mouse exon A1b did not extend into the sequence corresponding to mouse exon A1a (Fig. 1B). We repeated the 5' RACE on human liver poly(A)1 RNA using an antisense primer in exon A1b (Fig. 1B). For all 35 of the longest clone inserts sequenced, we again found only exon A1b and not the mouse exon A1a-homologous sequence. A search of the NCBI database of expressed sequence tags found an exon A1b 5' sequence extending to nt 2102 (Fig. 1A), but the adjacent sequence homologous to exon A1a was again not represented. Hence, RACE data suggested that upstream transcriptional initiation was infrequent in human hepatic tissue, and even then transcription began at a different start site than that used in the mouse.

Several other 5' termini were represented in the human liver FPGS RACE clones. As previously reported for human HepG2 hepatoma cells (9), these included transcripts containing the initial exons 1, 1b, and 1c and a new transcriptional start site, denoted exon 2a, all individually linked to exon 2 (Fig. 1B). In contrast to the human A1b- and exon 1-containing RACE clones, which have translational start sites in-frame with exon

| Initial exon | Experiment 1a | Experiment 2b | % of total |
|-------------|---------------|---------------|------------|
| Exon A1b    | 1             | 1             | 4.3        |
| Exon 1      | 8             | 4             | 25.5       |
| Exon 1b     | 1             | 0             | 2.1        |
| Exon 1c     | 14            | 8             | 46.8       |
| Exon 2a     | 0             | 10            | 21.3       |

a Sequences of 24 clones were analyzed.

b Sequences of 23 clones were analyzed.

(Fig. 1A). We previously demonstrated that the mouse exon A1b has mitochondrial and cytosolic translational start sites (7) similar to what was reported for the human exon 1 (14). Mouse proteins initiating at the more downstream ATG in exon A1b encode cytosolic FPGS, whereas those initiating at the upstream ATG encode mitochondrial FPGS. Two triplets encoding methionine were conserved in the human exon A1b homolog at the identical position to the translational start codons encoded in mouse exon A1b (Fig. 1A).

Two striking differences were noted from our experience with mouse liver FPGS gene expression. In contrast to the mouse, the human exon A1b homolog was infrequently repre-
2, the exon 1b and 1c variants did not, suggesting that they either used a translational start site in exon 2 or did not allow translation of active enzyme. The initial sequence identified as exon 2a was contiguous with exon 2 in the genomic sequence (Fig. 1C) and could represent either an alternative transcriptional start site or unprocessed pre-mRNA, but it was similar to exon 1c in having stops in all three reading frames and no consensus translational start sites. Differences from the mouse FPGS gene (7) were apparent; exon 1-containing human RACE products were common, even in liver, and transcripts without a clear translational start site were abundant in these RACE products.

Positioning the Upstream Exon A1b in the Human FPGS Genomic Locus—Because we had detected a sequence homologous with mouse exon A1b in FPGS transcripts from human liver (Fig. 1A), we sought to place this alternative initial exon on the map of the human genomic locus. We had previously mapped the human FPGS gene using the overlapping λ clones diagrammed in Fig. 2A (10). Because exons A1a and A1b were located 10 kb upstream of exon 1 in the mouse FPGS locus (8), we looked for the human A1b homolog on BL, which extended upstream of exon 1 by ~15 kb. When we initially probed Southern blots of BL with a mouse A1a/A1b-containing RACE clone, a 4.0-kb HindIII fragment hybridized, which was 6.5 kb upstream of human exon 1 (Fig. 2A). We subcloned the HindIII 4.0-kb fragment and probed a Southern blot of this subclone with an end-labeled 24-nt primer from the human A1b sequence; a 1.1-kb NcoI fragment hybridized (Fig. 2B). Direct

FIG. 2. Placement of the human exon A1b on the FPGS genomic map and sequence homology between mouse exon A1a and the boxed region of clone ABL. A, precise location of the human FPGS exon A1b, located 8 kb upstream of the remaining body of the gene. The overlapping λ clones containing FPGS exons 1–15 are described (9). The positions of human A1b and A1a homologs were determined by Southern blotting with a human-specific exon A1b oligonucleotide (B) and by sequencing, respectively. FPGS Exons 1 and A1b each have two translational start sites (A) that are in-frame with the downstream exons. B, Southern blot of the 4.0-kb HindIII subclone probed with a human exon A1b-specific γ-32P end-labeled oligonucleotide. C, comparison of the human genomic sequence upstream of exon A1b and the mouse exon A1a cDNA. Sequence identity is indicated by dots. The PCR primers from this region used in attempts to detect A1a-related sequences in mRNA are noted.
The human FPGS exon A1b transfected into FPGS-null AUXB1 cells supports colony growth, but the exon 1c and 2a FPGS constructs do not. Each initial exon (1, A1b, 1c, and 2a) was linked to the downstream exons 2–15 and cloned into pcDNA3. Constructs were transfected into AUXB1 cells as calcium phosphate precipitates. Selection of transfectants with G418 began after 48 h in one of the following media: a+ medium contained G418, thymidine, purines, and glycine and selects only for the integration of pcDNA3; α− medium did not contain nucleosides and tested the complementation of the cytosolic purine and thymidine auxotrophy. The exon 2a and pcDNA3 plates are from a separate experiment from the others, but identical controls were run in each experiment.

We sought to define the intracellular role of any proteins encoded by each of the several FPGS transcripts from the Human FPGS Gene. We concluded that the human homolog to the mouse exon A1a is absent in mRNA from CEM leukemic cells or from normal human liver or skeletal muscle. Mouse exon A1a contains only a 5′ untranslated sequence (7); therefore, the proteins encoded by transcripts initiating at upstream exon(s) in the mouse and the human would not be expected to differ in length at the amino-terminal end. The omission of exon A1a in human FPGS transcripts suggests that there are major variations between the human and mouse upstream promoters, resulting in different transcriptional start sites.

Testing the Function of FPGS Produced from the Several Human Transcripts in Vivo—We sought to define the intracellular role of any proteins encoded by each of the several FPGS transcripts detected in human liver. The Chinese hamster ovary-derived cell line AUXB1 contains an inactivating point mutation in the FPGS gene and, therefore, cannot survive in the absence of thymidine, purines, and glycine (2). cDNAs containing human FPGS exons 1, A1b, 1c, or 2a individually linked to exons 2–15 were cloned into pcDNA3 and transfected into AUXB1 cells. Transfection of the exon 1 construct served as a positive control (14), and transfection of pcDNA3 alone was the negative control. As seen in Fig. 3, only the exon 1 and exon A1b constructs complemented the FPGS-null phenotype in the absence of nucleosides (α− medium). Constructs containing exon 1c linked to exons 2–15 did not complement the AUXB1 phenotype. In a separate but identically designed set of experiments, transfection of exon 2a constructs into AUXB1 cells also did not yield colony growth in α− medium (Fig. 3). These results ruled out the possibility of initiation of translation of active FPGS from a cryptic translational start site in exons 1c or 2a or translation of active enzyme from a methionine in exon 2.

Similarly to human exon 1 and mouse exon A1b, human exon A1b also encodes an RNA transcript that contains two in-frame potential start methionines (Fig. 1A). Freemantle et al. (14) demonstrated that the species of human FPGS translated from the downstream start codon in human exon 1 (Fig. 2A) complemented the thymidine and purine auxotrophy of AUXB1 cells but not the glycine requirement and that FPGS translated from the upstream codon complemented the glycine auxotrophy of these cells. These functions were equated to the cytosolic and mitochondrial components of folate metabolism, and the 42 codons between the two AUGs encode a mitochondrial leader sequence. We sought to determine whether human exon A1b also encodes cytosolic and mitochondrial isoforms of FPGS. AUXB1 cells were transfected with constructs initiating at either of the ATGs in exon A1b and continuing through exon 15. Several independent clones of cells stably transfected with the exon 1-, A1b-, and 1c-containing constructs were selected in α+ medium containing G418, mass-cultured in this non-selective medium, and used to determine the sub-cellular biochemical function of any species of FPGS produced by plating 200 cells in either α+ medium, α− medium, or α+ medium formulated without glycine. The resultant colony growth (Table II) demonstrated that constructs initiating with the exon 1c sequence supported neither mitochondrial nor cytosolic folate metabolism, whereas the constructs initiating at the first ATG in exons A1b and 1 supplied the glycine requirement of AUXB1, and those starting at the second ATG of both exons allow growth in the absence of nucleosides but not in the absence of glycine. Hence, we have established in vivo that the human exon A1b-initiated FPGS transcripts encode two functional forms of FPGS, as do those initiating with exon 1 (14).

Quantification of FPGS Transcripts by RPA—To define the significance of RNAs encoding FPGS isoforms in a number of human tissues, we turned to ribonuclease protection assays, which do not have the inherent biases associated with PCR-based methods such as RACE or with cDNA library screening that may over- or under-represent RNA species. The RPA probes were constructed from cDNA clones containing exons A1b, 1, 1c, or 2a linked to exon 2. We quantitated the percent of transcripts containing the same initial exon as the probe and compared it directly to other FPGS species that contained exon 2 but had a sequence from any other initial exon than that represented in the probe. The RPA experimental conditions, such as the hybridization temperature, salt concentration, and RNase digestion temperature, were individually optimized for each probe. The high GC content of exon 1c made it particularly troublesome, resulting in a background of spurious protected fragments from each of several probes. Because the greatest differences in mouse tissues were found between liver and leukemic cells (7), poly(A)+ RNA from human liver and the ALL cell line CEM were initially analyzed. Additional human ALL and AML leukemic cell lines, MOLT-3 (ALL), K-562, and HL-60 (AML), were studied, as well as skeletal muscle, which was previously shown (6) to have a much higher level of FPGS expression in human than in mouse.

The exon 1+2 probe detects the FPGS species containing exon 1 linked to exon 2, represented by a signal at 145 nt (Fig. 4A); other species of FPGS, containing a different exon up-
stream of exon 2, are represented by a signal at 100–110 nt. In human liver, the predominant FPGS species contained exon 1 linked to exon 2, representing the majority of total FPGS transcripts (Table III). The remaining transcripts contain a different upstream exon. In skeletal muscle and ALL and AML cell lines, transcripts containing the exon 1 sequence represented a larger proportion (nine-tenths) of total FPGS transcripts than detected in liver (two-thirds); transcripts with a sequence other

![Fig. 4.](image)

**Fig. 4. Quantitation of FPGS transcripts containing alternate initial exons in human normal and tumor tissues.** Antisense riboprobes were generated by in vitro transcription using either SP6 or T7 RNA polymerase. A single-stranded RNA standard complementary to each probe was generated to mark the gel migration of a protected product resulting from a perfect match. A, the predominant species of FPGS mRNA in the human tissues studied contains exon 1 linked to exon 2. The liver and skeletal muscle (1 \( \mu \)g of poly(A) mRNA each) and CEM (2.5 \( \mu \)g of poly(A) mRNA) lanes were exposed to film for 46 h. The other lanes (4 \( \mu \)g of poly(A) mRNA each) were exposed to film for 18 h. B, human exon A1b-containing transcripts were detected only in liver. The standard, liver, and skeletal muscle lanes were exposed to film for 136 h; a faint signal at 140 nt is visible in the liver lane. The tumor cell lines and yeast lanes shown were exposed to film for 72 h, but even at 136 h, exon A1b transcripts were not detectable. Two \( \mu \)g of poly(A)+ RNA were used for each tissue. C, exon 1c-containing species represent a minor proportion of FPGS transcripts. Protected fragments also found in yeast were ignored. The liver and skeletal muscle (1 \( \mu \)g of poly(A)+ RNA each) and CEM (2.5 \( \mu \)g of poly(A)+ RNA) lanes were exposed to film for 46 h. The MOLT-3 and HL-60 (4 \( \mu \)g of poly(A)+ RNA each) and K-562 (1 \( \mu \)g of poly(A)+ RNA) lanes were exposed to film for 18 h. D, FPGS transcripts containing exon 2a are a minor population in the human tissues studied. The liver and skeletal muscle (1 \( \mu \)g of poly(A)+ RNA each) and CEM (2.5 \( \mu \)g of poly(A)+ RNA) lanes were exposed to film for 46 h. The other lanes (4 \( \mu \)g of poly(A)+ RNA each) were exposed to film for 18 h.

![Table II](image)

**Table II**

| DNA transfected | Colonies formed per plate |
|-----------------|---------------------------|
|                 | \( a^+ \) | \( a^- \) | \( a^+ \) minus glycine |
| Controls        |                |          |                        |
| AUXB1 None      | 181 ± 18       | 0 ± 0    | 0 ± 0                 |
| CHO None        | 192 ± 19       | 193 ± 5  | 178                   |
| FC2/2D Human genomic* | 142 ± 13       | 134 ± 9  | 134 ± 22              |
| 10–1 pcDNA3     | 162 ± 16       | 0 ± 0    | 0 ± 0                 |
| 10–2 pcDNA3     | 226 ± 10       | 0 ± 0    | 0 ± 0                 |
| cDNA transfectants |            |          |                        |
| Clone 5–2 Exon A1b upstream | 177 ± 11       | 172 ± 11 | 186 ± 5               |
| Clone 5–4 Exon A1b upstream | 172 ± 13       | 65 ± 6   | 31 ± 1                |
| Clone 6–3 Exon A1b downstream | 196 ± 7        | 208 ± 11 | 0 ± 0                 |
| Clone 6–12 Exon A1b downstream | 251 ± 11       | 254 ± 13 | 0 ± 0                 |
| Clone 8–1 Exon 1 upstream | 193 ± 14       | 185 ± 5  | 173 ± 5               |
| Clone 8–2 Exon 1 upstream | 131 ± 10       | 140 ± 9  | 139 ± 8               |
| Clone 9–3 Exon 1 downstream | 102 ± 4        | 111 ± 3  | 0 ± 0                 |
| Clone 9–4 Exon 1 downstream | 194 ± 19       | 189 ± 9  | 0 ± 0                 |
| Clone 7–1 Exon 1c | 137 ± 6        | 0 ± 0    | 0 ± 0                 |
| Clone 7–2 Exon 1c | 216 ± 13       | 0 ± 0    | 0 ± 0                 |

The indicated cDNAs were transfected into 5000 AUXB1 cells/100-mm dish and selected in \( a^+ \) medium supplemented with 1.2 mg/ml G418, and multiple well isolated colonies were picked and mass-cultured. Cloned transfectants that expressed FPGS were plated at 250 cells/100-mm dish in \( a^+ \) medium, and the indicated media were applied after 16 h. Colonies were fixed and stained 10 days later.

* See Ref. 14.
than exon 1 represented 6% in skeletal muscle and 7–11% in the leukemic cell lines. Hence, unlike what has been found in the mouse, both differentiated and dividing human tissues express primarily exon 1-containing transcripts as the major FPGS RNA species.

Exon A1b spliced to exon 2 is the main FPGS transcript (≥95%) expressed in mouse liver (7), but in human liver, the exon A1b homolog is expressed only in trace amounts (1.2%) (Fig. 4B and Table III). Exon A1b transcripts were not detectable by RPA in mRNA from skeletal muscle and ALL and AML cell lines. Each human tissue contained predominantly FPGS transcripts with a sequence other than exon A1b upstream of exon 2, as indicated by the signal at 109 nt in Fig. 4B, in agreement with the reciprocal experiment shown in Fig. 4A. In a parallel experiment, PCR primer pairs in exon A1b and exon 2 were applied to human liver, skeletal muscle, and CEM 89% 11% (FPGS RNA species. Exon A1b represent a very small proportion of the FPGS transcripts from human liver and was not detectable by RPA in skeletal muscle or ALL and AML cell lines. Unlike what has been found in mouse tissues, and the downstream promoter was not active in liver; in humans, the downstream promoter was active in all tissues tested.

**DISCUSSION**

A very discrete choice is made between two strong promoters in the mouse FPGS gene in different tissues, generating one enzyme isoform in liver and kidney and a second in dividing mouse tissues (7, 12). In the majority of mouse and human differentiated tissues, neither promoter is active (6). Metabolism of antifolates by FPGS is central to the action of these drugs, and the production of two proteins with different catalytic properties (7) suggested the use of these isozyme patterns to alter the cytotoxicity of antifolates in *vivo*. Others have reported what appears to be significant differences in the kinetics of FPGS in human leukemia cells sensitive (ALL) and refractory (AML) to antifolates (13), results that underscore the therapeutic importance of this gene in humans. Against this background, we set out to define the distribution of transcripts from the FPGS gene in human tissues and tumors and the functional consequences of any transcripts found. To our surprise, the precedent studies on mouse tissues did not predict the behavior of the human FPGS gene (Fig. 5).

Several studies have demonstrated two levels of heterogeneity in transcripts from the FPGS gene, both of which were due to differences at the 5′ end. In both mouse and human, longer and shorter FPGS transcripts are derived from the multiple transcriptional start sites of a TATA-less promoter, now recognized as the downstream promoter (6, 7, 14). The shorter transcripts encoded a cytosolic form of the enzyme; the longer forms contained an additional upstream AUG codon that added a mitochondrial leader sequence to the protein and supplied FPGS for mitochondrial folate metabolism. Initial genomic organization studies defined this downstream start site as exon 1, and transcriptional start site mapping in CEM cells indicated that all FPGS transcripts contained exon 1 spliced to exon 2 and to a series of exons further downstream (10, 14). However, Chen et al. (9), applying 5′ RACE to RNA from HepG2 cells, found an array of FPGS transcripts containing either the exon 1 originally defined by Freemantle et al. (14) or

---

**Table III**

Quantitation of FPGS transcripts containing alternate initial exons by RPA

Percentages were obtained by PhosphorImager analysis of a representative experiment (*n* ≥ 2). The sum of % for each transcript amounted to 108–111% in the various tissues, indicating that some of the species were somewhat overestimated by this analysis. Inspection of the x-ray films for the various probes indicated that these errors were not associated with the exon 1+2 or A1b+2 probes but were more likely to be caused by the higher background on the exon 1c films, together with some degree of overestimation of exon 2a+2 transcript due to unspliced precursors.

| Tissue | Exon 1+2 probe | Exon A1b+2 probe | Exon 1c+2 probe | Exon 2a+2 probe |
|--------|----------------|-----------------|-----------------|----------------|
|        | Exon 1          | Other than 1    | Exon A1b        | Other than A1b |
|        | Exon 1c         | Other than 1c   | Exon 2a         | Other than 2a  |
| Liver  | 66%            | 34%            | 1.2%           | 98.8%          |
|        |                |                | 35%            | 65%            |
| Skeletal muscle | 94% | 6%   | <0.2%          | >99%           |
|        |                |                | 15.2%          | 85%            |
| CEM    | 89%            | 11%            | <0.2%          | >99%           |
|        |                |                | 6.3%           | 94%            |
| MOLT-3 | 89% | 11% | <0.2%          | >99%           |
|        |                |                | 8.0%           | 93%            |
| K-562  | 91%            | 9.3%           | <0.2%          | >99%           |
|        |                |                | 9.1%           | 91%            |
| HL-60  | 93%            | 7.3%           | <0.2%          | >99%           |
|        |                |                | 6.5%           | 93%            |

* The limits of detection with the four probes were as follows: exon 1+2, 5%; exon A1b+2, 0.2%; exon 1c+2, 5%; and exon 2a+2, 2%, based on the background in the region of the quantitated protected fragments.
one of three additional exons entitled 1a, 1b, and 1c, all of which were spliced to exon 2 in mature transcripts. In the human genomic locus of FPGS, these alternate exons are clustered within a few hundred nucleotides of exon 1 (9). cDNA cloning (8), 5' RACE, and ribonuclease protection assays (7) also demonstrated another level of heterogeneity at the 5' region of mouse FPGS; mouse liver and kidney contain transcripts that initiate at two small exons (denoted exons A1a and A1b) located 10 kb upstream from exon 1, whereas all dividing mouse tissues initiate in exon 1 (7) (Fig. 5). This latter phenomenon had not been reported for the human FPGS gene prior to this study.

A region of human genomic DNA about 8 kb upstream of FPGS exon 1 contained close homologs of mouse exons A1a and A1b (Figs. 1 and 2), but the A1a homolog did not appear in any FPGS transcripts in human liver, skeletal muscle, or CEM cells (Fig. 4B and Table III). The A1b sequence was only found by RPA in FPGS transcripts from human liver, and then infrequently so. Although the transcripts initiated at human exon A1b would produce a functional FPGS species (Fig. 3 and Table II) with a different amino-terminal peptide (and different kinetic characteristics), this alternative FPGS isoform is unlikely to contribute to the landscape of human folate metabolism in any major way, even in human liver. In addition, FPGS transcripts were detected that did not translate to active FPGS species, due to a somewhat promiscuous choice of transcriptional start sites from the downstream promoter; such species are not detected in any mouse tissue. Furthermore, FPGS transcripts from the downstream promoter represent the exclusive forms of mRNA for FPGS in several dividing tissues of the mouse but are only rarely represented in mouse liver RNA; the human equivalent exon 1 transcripts are the predominant forms of FPGS mRNA in human tumors, liver, and skeletal muscle. Hence, it is clear that the mechanism determining the tissue-specific expression of two distinct isozymes of FPGS seen in adult mouse tissues has not been conserved in man.

Yet, the degree of conservation in the sequence of upstream exons A1b and A1a between man and mouse is noteworthy. Exon A1a, which is not used in man, is 48–65% homologous with the corresponding mouse exon at the nt level (depending on where one chooses to begin the comparison), whereas exon A1b demonstrates a 61% identity between the two species. This compares with a 69% degree of conservation between exons 2–15 in man and mouse. Also conserved are the upstream and downstream ATG codons in exon A1b, identically placed within the sequence in man and mouse (Fig. 1A), and the functional (Table II) mitochondrial leader sequence between these codons. As in the mouse, the sequence of human A1b encodes enzymatically active but distinct species of FPGS to those encoded by transcripts initiating in exon 1. This degree of conservation is very surprising in exons that are apparently minimally utilized in the adult organism. We suspect that a developmental role for the upstream promoter is involved in this preservation of these structures and functions.

We have attempted to bring the literature on the transcripts produced by the FPGS gene together in this and the preceding study (7), which addressed transcription from the mouse gene. RACE can overestimate transcripts, presumably because it relies on reverse transcription and PCR, and the quantitative nature of these processes can be modified by RNA secondary structure and experimental conditions respectively; cDNA library screening shares some of these limitations. As a result, we have relied on RPA (Fig. 4 and Ref. 7) to estimate the abundance of species within a complex mixture of RNAs, using a series of probes extending across exonic borders. The conclusions we drew were that 1) human exon 1c occurs in several human tissues but is overestimated by RACE, 2) exon A1b is measurable only in liver, and 3) transcripts containing exon 1 represent the vast majority of FPGS mRNA species and are the only major species capable of encoding functional FPGS (Table II) in all human tissues studied.

The question of why tissue-specific isozymes of FPGS are carefully produced in mouse and not in man is perplexing, as is why mouse tissues have evolved the need for a different enzyme in liver and kidney than that which is sufficient for other, dividing tissues. However, the related question of how the human and mouse FPGS genes behave so differently is also of interest. The downstream promoters of mouse and human FPGS genes are rather similar; both are TATA-less and are driven by a set of concatameric Sp1 sites spaced within 60 bp of the major transcriptional start site (Fig. 5) (6, 12). There is an E-box motif in the human downstream promoter that is altered in the mouse, and we have previously speculated that this element is involved in the expression of FPGS in human but not mouse cardiac and skeletal muscle (6). Much less is known about the upstream FPGS promoter in either species. We are currently studying how transcription in liver is activated at the upstream promoter in the mouse but not in man and the related question of how the downstream promoter is activated (or not silenced) in liver of man but not of the mouse.

Acknowledgments—We are very grateful to Jennifer Ferguson for her technical expertise and contributions to this work and to Drs. Geoffrey Krystal and Peter Melera for their meticulous critiques of the manuscript.

REFERENCES
1. Shane, B. (1989) Vitam. Horm. 45, 263–335
2. McBurney, M. W., and Whitmore, G. F. (1974) Cell 2, 173–182
3. McGuire, J. J., Haieh, P., Franco, C. T., and Piper, J. R. (1986) Biochem. Pharmacol. 35, 2607–2613
4. Moran, R. G., and Colman, P. D. (1984) Anal. Biochem. 140, 326–342
5. Sirotnik, F. M., Johnson, T. B., Samuels, L. L., and Galivan, J. (1988) Biochem. Pharmacol. 37, 4239–4241
6. Freemantle, S. J., and Moran, R. G. (1997) J. Biol. Chem. 272, 25373–25379
7. Turner, F., Andreassi, J., Ferguson, J., Titus, S., Tse, A., Taylor, S. M., and Moran, R. G. (1999) Cancer Res. 59, 6074–6079
8. Roy, K., Mitsugi, K., and Sirotnak, F. M. (1997) J. Biol. Chem. 272, 5587–5593
9. Chen, L., Qi, H., Korenberg, J., Garrow, T. A., Choi, Y., and Shane, B. (1996) J. Biol. Chem. 271, 13077–13087
10. Taylor, S. M., Freemantle, S. J., and Moran, R. G. (1995) Cancer Res. 55, 6030–6034
11. Roy, K., Mitsugi, K., and Sirotnak, F. M. (1996) J. Biol. Chem. 271, 23820–23827
12. Chen, J., Hayes, P., Roy, K., and Sirotnak, F. M. (2000) Gene 242, 257–264
13. Longo, G., Gorlick, R., Tong, W., Ercikan, E., and Bertino, J. (1997) Blood 90, 1241–1245
14. Freemantle, S. J., Taylor, S. M., Krystal, G., and Moran, R. G. (1995) J. Biol. Chem. 270, 8579–8584
15. Graham, F. L., and Van der Eb, A. J. (1973) Virology 52, 456–457