Organization and Alternate Splicing of the Murine Folylpolyglutamate Synthetase Gene  
DIFFERENT SPlice VARIANTS IN L1210 CELLS ENCODE MITOCHONDRIAL OR CYTOSOLIC FORMS OF THE ENZYME*

(Received for publication, February 28, 1996, and in revised form, June 18, 1996)

Krishnendu Roy‡‡, Kenji Mitsugi‡‡‡, and F. M. Sirotnak‡**

From the ‡‡Program in Molecular Pharmacology and Therapeutics, Memorial Sloan-Kettering Cancer Center and the ‡‡‡Graduate School of Medical Sciences, Cornell University, New York, New York 10021

The organization of the murine folylpolyglutamate synthetase (FPGS) gene has been determined by sequence analysis that also revealed an interesting complexity in the case of exon 1. The entire nucleotide sequence of the L1210 FPGS cDNA, the 3′- and 5′-untranslated regions, the mitochondrial leader sequence, and the coding region were found to be distributed on 15 exons with an overall length of 10.358 kilobases. Two splice variants of exon 1 were identified by screening of an L1210 cell cDNA library. Variant I (exons 1a + 1b plus 2–15) incorporates all of the sequence homologous to the recently reported (Taylor, S. M., Freemantle, S. J., and Moran, R. G. (1995) Cancer Res. 55, 6030–6034) human exon 1, including two ATG start codons at positions +1 and +126, and encodes both mitochondrial and cytosolic form of FPGS. The most prevalent variant, Variant II (exons 1b plus 2 to 15), incorporates only a portion (92 nucleotides at the 3′ end) of this sequence, incorporates only one ATG start codon at position +126, and encodes only cytosolic FPGS. The existence of this variant is consistent with the identification of an appropriately situated internal donor/acceptor site in what was believed to be exon 1. A third related variant (Variant III) with a novel 5′ termini was originally identified by screening of a mouse liver cDNA library. This variant, which occurs at moderately low frequency in the L1210 cell cDNA library, incorporates an alternate to exon 1a (exon 1c) spliced to exon 1b plus exons 2–15 and encodes a different mitochondrial leader peptide than Variant I. The identification of these variants suggests another possible mechanism, i.e. at the level of precursor mRNA splicing, for regulating synthesis of mitochondrial versus cytosolic forms of FPGS in the cell. Exon 1c is positioned in the gene upstream of exon 1a separated by an intron of 56 nucleotides within a region of DNA sequence that like the homologous human sequence is distinctly promoter-like. However, the sequence of this region differs from the human sequence in terms of the number, position, and type of putative regulatory elements, particularly in regard to the number of SP-1 binding sites and the position of multiple transcription start sites as determined by enzymatic primer extension.

Folates exist in mammalian cells primarily as γ-polyglutamate peptides (1–5) of varying chain length. The anabolism of folates and their analogues to γ-polyglutamates is mediated (1–5) by the enzyme folylpolyglutamate synthetase (FPGS), and metabolic turnover of these anabolics appears to be modulated by folylpolyglutamate hydrolase after their mediated entry (reviewed in Ref. 6) into lysosomes. The process of folylpolyglutamylation in mammalian cells is important to the conservation and efficient utility of folate coenzymes required for macromolecular biosynthesis (1–5). It has been suggested in the context of earlier reports (7–10) that one way in which tumor cells may control their macromolecular synthesis is through regulation of intracellular folate homeostasis. In addition to the metabolic conversion (7–10) of folate compounds, this could occur at the level of mediated entry of exogenous folates (9) and/or through the biosynthesis of folylpolyglutamates (1–5). Because classical folate analogues are also substrates (1–5) for FPGS in tumors and normal proliferative tissues of animals and man, the process of their polyglutamylation has pharmacologic and therapeutic relevance. Intracellular levels of FPGS and its preferences among folate analogues as substrate appear to at least partially determine (11, 12) the extent of cytotoxic action of these analogues in these proliferative tissues as well as their therapeutic utility (13–15).

Earlier studies by others (16) and our own laboratory as well (17, 18) suggest that expression of FPGS activity is under stringent regulation. Levels of this enzyme, which are high in liver, vary among nonhepatic tissues differing in their proliferative capacity (16) and during terminal maturation (16–18). The recent cloning by Shane and co-workers (19) of a human FPGS cDNA now make it possible to address these issues at the level of gene expression. Recent additional studies by Freemantle et al. (20) have provided information pertaining to the 5′ region of the human FPGS gene, which included characterization of a putative promoter like region and partial identification of the proximal exon/intron sequence. These workers have also provided structural and functional evidence to suggest that both cytosolic and mitochondrial forms of the enzyme are generated by alternate use of different translational ATG.

* This work was supported in part by Center Core Grant CA08748 and Grant CA56517 from the National Cancer Institute and the Elsa U. Pardee Foundation. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked “advertisement” in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

‡‡‡ These authors contributed equally to the studies presented in this publication.

‡‡‡‡ To whom correspondence should be addressed: Laboratory for Molecular Therapeutics, Memorial Sloan-Kettering Cancer Center, 1275 York Ave., New York, NY 10021. Tel.: 212-639-7952; Fax: 212-794-4342.

1 The abbreviations used are: FPGS, folylpolyglutamate synthetase; PCR, polymerase chain reaction; kb, kilobase(s); RT, reverse transcriptase.
start codons. As suggested by Moran and co-workers (20), this appears to be an interesting example of a phenomenon only rarely documented in mammals (20) of the alternate use during translation of two ATG codons resulting in the synthesis of proteins differing only by the presence or the absence of an amino-terminal leader peptide.

Recently, Moran and colleagues (21) provided information on the structural organization of the human FPGS gene. Data were presented in that report that showed that the entire nucleotide sequence of the FPGS cDNA was represented in 15 exons spanning approximately 11 kb of genomic DNA. Additional information on the structure of the human gene and evidence for the existence of alternates to exon 1 in the form of splice variants were more recently provided by Shane and co-workers (22). In the studies presented here, we describe the cloning and organization of the murine FPGS gene, which appears to be smaller in size but like the human gene (21) is also organized into 15 exons. We also reveal an unexpected complexity pertaining to exon 1 in the form of splice variants found in L1210 cells and mouse liver identified during cDNA cloning that encode either different mitochondrial leader peptides or only the cytosolic form of FPGS. The existence of these splice variants raises the possibility of an additional mechanism regulating at the level of precursor mRNA splicing a differential in synthesis within the cell of mitochondrial and cytosolic forms of FPGS.

**Experimental Procedures**

**Cloning of a Murine FPGS cDNA—**Full-length (2.247 kb) murine FPGS cDNA clones were obtained by hybridization screening (23) of an L1210 cDNA library in λgt11 (ZAP, Stratagene, La Jolla, CA) using a human FPGS cDNA (19) as a probe. The hybridization screening of the cDNA library was carried out under conditions described earlier (24). The radioactive FPGS cDNA probe was prepared by random priming (Boehringer Mannheim) with (α-32P)dCTP and the human FPGS cDNA as a template.

**Isolation of Genomic Clones by Hybridization Screening—**A Balb/c mouse liver genomic library in EMBL 3 SP6/T7 (Clontech, Palo Alto, CA) was screened with radioactive (α-32P)dCTP murine FPGS cDNA that was prepared by random priming (Boehringer Mannheim). The nucleotide sequence of this FPGS cDNA was submitted to GenBank (GenBank®/EMBL accession number U33557). After screening of 106 plaques and purification, several positive clones were obtained. The DNA insert from these clones was purified and characterized by restriction mapping and Southern hybridization using fragments encompassing sequences from the most 3' and 5' end of the FPGS cDNA. Two nonidentical clones designated mFPGS-3 and -5 were selected by the above procedure for further analysis. Restriction fragments of these clones generated with SacI or EcoRI were selected for further study on the basis of hybridization with the region-specific probes. The fragments of interest after extraction were subcloned into Bluescript SK+ (Stratagene, Menask, WI) for sequencing.

**Quantitative PCR—**The relative amount of L1210 cell FPGS mRNA transcript representing different splice variants was determined by RT-PCR. L1210 cell poly(A)+ RNA was prepared by a published procedure (26), and 2 μg was utilized in the synthesis of cDNA using a standard protocol (27) with a reaction volume of 50 μl, and oligo(dT) was used as the primer. Prior to cDNA synthesis, poly(A)+ RNA was digested with RNase free DNase I, treated with phenolchloroform, and precipitated with ethanol followed by washing twice with 70% ethanol before dissolving in H2O. All poly(A)+ RNA was treated with methylmercuric hydroxide to reduce secondary structure. 5 μl of the prepared cDNA was utilized in a PCR reaction using three different exon 1-specific sense primers and an exon 5 antisense primer. 10 pmol of an 18-mer oligonucleotide encompassing the ATG start site were used for each PCR reaction. 5'-TAAGACTATGTCGCTGCCC-3' was used as the 1a-specific sense primer, 5'-GAGGCCGGCATGGAGAT-3' was used as the 1b-specific sense primer, and 5'-ACAAAGATTGGCGT-3' was used as the 1c-specific sense primer. 5'-TTCCAGGCCCTATTCCGGC-3' was used as the 5-exon 5-specific antisense primer. Initially, reactions were run with different amounts of cDNA and different PCR cycles to obtain conditions that insured that amplification was in the linear range at the time the amount of product of each primer-specific reaction was compared. As a control, the cDNA reaction mixture without reverse transcriptase was utilized to initiate a PCR reaction to make sure that there is no genomic contamination (data not shown). As an internal standard, sense (5'-CTCGGTGTCCTGCGTCAC-3') and antisense (5'-CTTATGGATGAAAACCTCC-3') primers of FPGS representing exons 9 and 14, respectively, were used in a parallel reaction. The experimental reaction was run for 25 cycles using a standard protocol (28), and the relative amount of product generated with each set of primers were determined by 1.5% agarose gel electrophoresis and staining with ethidium bromide (29).

**Enzymatic Primer Extension—**In separate reactions, 10 pmol of an antisense oligonucleotide primer (5'-ATACCTCAGTGGCGCCTC-3' and 5'-CGATCCGCTTCTATTTTCTCT-3') corresponding to nucleotide sequence in the 3' end of exon 1b and 1c, respectively, of the murine FPGS gene were end-labeled (Promega, Madison, WI) at the 5' terminus using T4 polynucleotide kinase and [γ-32P]ATP. The labeled primer was hy-

---

**Fig. 1.** Organization of the murine FPGS gene showing a restriction map, exon/intron structure, and different splice variant formation. The two genomic clones shown had inserts of 17 kb (clone 3) and 12 kb (clone 5), respectively. The lengths of the exons and introns identified are shown to scale.
FIG. 2. Nucleotide sequence of a murine FPGS cDNA. The nucleotide sequence and deduced amino acid composition is given in panel A and is numbered from the initial nucleotide. The mitochondrial and cytosolic ATG and polyadenylation signal in the 3' of the sequence are in boldface type. The derived amino acid sequence is compared with that reported (19, 20) for the human FPGS cDNA in panel B. The amino acid sequence is numbered from the most upstream ATG start codon.
bridized to 5 μg of poly(A)^+ RNA from L1210 cells, which was first treated with methyl mercurichloride to reduce secondary structure, and the primer extension reaction was carried out as well with a standard kit as specified (Promega, Madison, WI). An mRNA preparation derived from a 1.2-kb kanamycin-resistant plasmid cDNA with an antisense primer was used as a control. Polyacrylamide gel electrophoresis (6 or 8%) was carried out by a standard protocol (29).

**Materials**—All radioactive isotopes used for the above studies were obtained from NEN (Waltham, MA). Specific activities for [α-^32^P]dCTP, [α-^35^S]dATP, and [γ-^32^P]ATP were 3000, 1000, and 3000 Ci/mmol, respectively. DNA restriction enzymes were purchased from Boehringer Mannheim. Amplitag DNA polymerase was obtained from Perkin-Elmer. Nitrocellulose was purchased from Schleicher & Schuell, and oligonucleotide primers were synthesized by Operon Inc. All other materials were reagent grade.

**RESULTS AND DISCUSSION**

**Organization of the Murine FPGS Gene**—Several relevant clones were derived following screening of the mouse liver genomic library. Two of these clones designated λ mFPGS-3 and -5 were found to be nonidentical following characterization by restriction enzyme mapping. The partial restriction maps for these clones are given in Fig. 1. This map was based upon the correspondence of the restriction sites for the genomic clone.

---

**ATGCAAAATTATTTTATGAAAAAAA**

2247

**Mur. Var. I**

MNSAWRSRCSTLSLAAVSARGATTBPPRMGSGAPFAPQFPEGMVQDAVTLTML 55

Human MNSAWRSRCSTLSLAAVSARGATTBPPRMGSGAPFAPQFPEGMVQDAVTLTML 55

**Mur. Var. I**

QTNQ.YLQQKQRSQGRQLEMQMRKLALRQGLVEDNRNLH.QVHTGKGS 110

QTNQ.YLQQKQRSQGRQLEMQMRKLALRQGLVEDNRNLH.QVHTGKGS 110

**Mur. Var. I**

CAPTER1LSYGLTQFennQPLQVRQIPEFREPLTTKMFCLNQLES 165

CAPTER1LSYGLTQFennQPLQVRQIPEFREPLTTKMFCLNQLES 165

**Mur. Var. I**

PKDDNHSMLAFYFPFLTLMPLAFMLQVHGDNLALPPLFRTLYPQPAECNQLIPVCQ 220

Human TKQCVSFLMPYPFLTLMPLAFMLQVHGDNLALPPLFRTLYPQPAECNQLIPVCQ 220

**Mur. Var. I**

VSSGLIDHTSHLDGTVERIQAQGRSPQGPAVPQTVQGPILGMAYAQTDNS11REPVIC 275

Human VSSGLIDHTSHLDGTVERIQAQGRSPQGPAVPQTVQGPILGMAYAQTDNS11REPVIC 275

**Mur. Var. I**

PLYLCFSLELSELVPLSPGLGGLRQHRSWALGLQNLQKLEQHDQIQLKVE 330

Human PLYLCFSLELSELVPLSPGLGGLRQHRSWALGLQNLQKLEQHDQIQLKVE 330

**Mur. Var. I**

SPRS1R0Q1PLAPVFRPTTRMRGLRTDTCRGTOLIQGPKLTVGDLGATTVSS 385

Human SPCFRMLTAPLVPQTSRMLGLRTDTCRGTOLIQGPKLTVGDLGATTVSS 385

**Mur. Var. I**

QACVRHVRQQLERMRTSEQDGGEVLHVEVNEUTMDRNASSNLKQLQCPFDYVFPC 440

Human QACVRHVRQQLERMRTSEQDGGEVLHVEVNEUTMDRNASSNLKQLQCPFDYVFPC 440

**Mur. Var. I**

NVTVESISIGNAQNNFTVTLQQLRCLMQRVWNEGMLREQANSNLSGCCDOPPA 495

Human NVTVESISIGNAQNNFTVTLQQLRCLMQRVWNEGMLREQANSNLSGCCDOPPA 495

**Mur. Var. I**

GPQSLIIAPPHQPQPRSTVLSVFICSHALLWISQSQGDIPIQPSLPEHNLNPTPTA 550

Human GQASLLAPPHQTCSASLIVFICSHALIWIQISQSQGDIPIQPSLPEHNLNPTPTA 550

**Mur. Var. I**

NSGAIHEAAAIIHVLTSGLHLQVYVKKLDPSMQ 587

Human NSGAIHEAAAIIHVLTSGLHLQVYVKKLDPSMQ 587

---
to the hybridization pattern obtained when the murine FPGS cDNA was used as a probe in a Southern blot of the restricted mouse liver genomic DNA. Clone 3, which was 17 kb in length, contained the region corresponding to the 5’ end of the murine FPGS cDNA (Fig. 2) approximately 1 kb of sequence beyond that corresponding to the 5’-untranslated region and an additional 10 kb of nucleotide sequence extending to the 3’ end of the cDNA. Clone 5 (12 kb in length) lacked the region containing the 5’ end of the cDNA and overlapped with 2.5 kb of clone 3 at the region corresponding to the 3’ end of the open reading frame of the cDNA. A number of restriction fragments of clones 3 and 5 were subcloned and sequenced in each direction, and the entire sequence of the gene was obtained with the exception of intron 4. In this case, partial sequencing of the relevant subclones (data not shown) using the appropriate exon 3 and 4 primers identified the intron/exon boundaries of the cDNA (Fig. 2).

The organization of the murine FPGS gene shown in Fig. 1, with the exception of exon 1, is very similar to that reported (21, 22) for the human gene. However, the mouse gene is smaller in size spanning slightly more than 10.3 kb in length but also consists of 15 exons and 14 introns. Similar to that recently reported (22) for the human FPGS gene, exon 1 in the murine gene appears to exhibit substantial complexity. Most all intron/exon splice junctions at the 3’ and 5’ termini of the specific cDNA sequences were found (Table I) to correspond to the consensus sequences (29) for such junctions, the so-called GT-AG rule. The 5’-untranslated region, two ATG start codons, and the intervening sequence coding for a mitochondrial leader peptide (Fig. 2), and the first 164 nucleotides of coding region (Fig. 2) were situated in exon 1. Exons 2–14, ranging in size from 54 to 151 nucleotides, incorporate only coding regions, and exon 15 included 865 nucleotides of coding region and the 3’-untranslated region.

Nucleotide Sequence of a Murine FPGS cDNA—The unpublished (GeneBank\textsuperscript{\textregistered}/EMBL accession number U33557) nucleotide sequence of our initial full-length cDNA clone (Variant I) is given in Fig. 2. The overall identity between the sequence of this cDNA clone and the published human FPGS cDNA sequence (19) was 77.3%. This murine cDNA includes 3’- and 5’-untranslated regions of 460 nucleotides (64.55% identity), including the stop codon and 26 nucleotides (61.54% identity), respectively, and an open reading frame of 1761 nucleotides (81.83% identity), which like the human FPGS cDNA (19, 20) appears to code for both a mitochondrial leader peptide as well as the cytosolic protein. Consistent with this sequence data, the 5’-untranslated region of this cDNA includes an ATG start codon 126 nucleotides upstream and in-frame with a second start codon at position 152 (+126 nucleotides) with properly positioned Kozak’s consensus sequences (30). The identity (Fig. 2) between the murine and human mitochondrial leader peptides was predicted to be 61.9%. The open reading frame for murine cytosolic FPGS downstream from the 3’ most ATG encodes a protein with 80.24% identity to human cytosolic FPGS.

Differential Splicing and Alternate Exon Usage—Further screening of the murine tumor cDNA library identified (Fig. 1) a large number of clones (Variant II) that had a shorter insert size and incorporated an exon at the most 3’ end spliced to exon 2 that was different from the murine homologue (Variant I) of human exon 1 described above. These clones were derived from mature mRNA transcripts with an appropriate poly(A) tail and properly positioned polyadenylation signal sequence at the 3’ end. The analysis of these Variant II clones showed that the insert was a truncated version (Figs. 1 and 3) 92 nucleotides in length of what was considered to be exon 1 beginning at position +47 and ending at position +138, which incorporates only the downstream ATG start site (Figs. 4 and 5). These results suggested that the murine homologue of human exon 1 exists (Fig. 1) as two contiguous exons (exons 1a and 1b). This suggestion is supported by the results of other analyses that identify (Table I) what is believed to be an internal donor/acceptor splice site (29) within exon 1 at the putative junction (positions +46 and +47) between the proposed exons 1a and 1b.

More conclusive evidence for the complexity pertaining to exon 1 was subsequently obtained in the form of a unique variant (Variant III) originally obtained as a cDNA clone in a mouse liver cDNA library. This variant (Figs. 1 and 3) bears a novel 5’ end resulting from the incorporation of an alternate to exon 1 (exon 1c) spliced precisely at nucleotide +47 (Fig. 5) to exon 1b plus exons 2–15. An appropriate poly(A) tail and potential polyadenylation signal sequence is also found at the 3’ terminus of this variant. This variant has differently spaced ATG start codons and encodes a different (31) somewhat larger mitochondrial leader peptide (MAVIRRCFSVIRENLETR) than Variant I. The position of this exon in the murine FPGS gene is upstream of exon 1a with both exons separated by an intron (1c) 56 nucleotides in length. Consistent with this in-

---

**Table I**

| Exon number | Exon length\(^a\) | 5’ junction | 3’ junction | Intron length\(^{b, c}\) |
|-------------|-------------------|-------------|-------------|------------------|
| 1c          | 66                | aacctcccc  | GGATG      | 56               |
| 1a          | 72                | cctgcaccc  | TGCCAG      | 0                |
| 1b          | 92                | gttgtccag  | TATCAG      | 120              |
| 2           | 129               | GTGCT      | CTGCAAG     | 87               |
| 3           | 54                | cttgttagc  | GTAAGC      | 1100             |
| 4           | 65                | GCTTC      | CCTTAG      | 2300\(^c\)       |
| 5           | 115               | CCTTACCC   | TCTTACCC    | 81               |
| 6           | 78                | GACGCCAG   | GAGAAC      | 93               |
| 7           | 62                | GCCGTACC   | CATTACAG    | 82               |
| 8           | 103               | AAACCCAG   | TTAAAG      | 190              |
| 9           | 78                | CTGCTCAG   | ATTTGAG     | 250              |
| 10          | 148               | CGCCCGAG   | ACAACAG     | 94               |
| 11          | 90                | ACACCCAG   | GCGGCTG     | 650              |
| 12          | 151               | CTTCGACC   | CGATGG      | 328              |
| 13          | 76                | AGGGTC     | CTGCAAG     | 560              |
| 14          | 67                | CTTCCACAG  | ATCGCAAG    | 900              |
| 15          | 860               | ACCCAG     | GATGTA      | 920              |

\(^a\) Measured in nucleotides.  
\(^b\) Intron length determined by sequencing except for intron.  
\(^c\) Intron length determined by PCR from subclones.
tron/exon organization in this region of the gene, data obtained by primer extension analysis (see following section) showed that although there was extension of an exon 1c-specific primer, there was no extension (data not shown) of an intron 1c-specific primer during the enzymatic reaction.

The frequencies of Variants I, II, and III in the L1210 cDNA library were in the proportion of 1:14:5 (Variant I:Variant II:Variant III) as determined by sequencing using the appropriate exon-specific primers of the various cDNA clones. The results of quantitative RT-PCR subsequently carried out were consistent with this observed frequency. They showed (Fig. 4) that substantially more cDNA product was generated with L1210 cell mRNA with an exon 1b-specific primer (Fig. 4, lane 3) than either an exon 1a- (Fig. 4, lane 2) or exon 1c-specific (Fig. 4, lane 1) primer. These primers were paired with an exon 5 antisense primer to initiate the reaction. The PCR products generated with each primer were the size expected if only exon 1b or both exons 1a plus exon 1b or exons 1c plus 1b were incorporated in each transcript. The same results were obtained (data not shown) when different exon 1a-, 1b-, or 1c-specific primers were utilized in the PCR reaction, making it less likely that these results only reflected the efficiency of each primer in the PCR reaction.

Sequence Analysis of the 5’ Region of the Murine FPGS Gene—The region of this genomic DNA immediately 5’ of exon 1c contains (Fig. 5) nucleotide sequence that like the human gene (20) is distinctly promoter-like. This region is GC-rich and contains putative binding sites for various transcription factors but lacks TATA-like and CCAAT sequence motifs. However, a number of differences were evident from the homologous region of the human DNA were evident. Only one Y box was found but at a position (−318 nucleotides) different from that found within the human sequence. Also, putative binding sites for ETS, Aml-1, and EKLF were found in this region but not in the human 5’ region. Three putative Sp-1 binding sites in addition to other putative binding sites for transcription factors were found downstream of exon 1c. Enzymatic primer extension utilizing an antisense primer beginning at +135 nucleotides within the 3’ end of exon 1b identified (Fig. 6) a number of potential transcription start sites within exon 1a and 1b and immediately 5’ of exon 1a. Interestingly, the major transcripts identified (Figs. 5 and 6) were initiated at +83 and +84 nucleotides within exon 1b. Additional transcripts appeared to be initiated at positions +15, +17, +31, +48, +52, +57, and +60 nucleotides within exons 1a and 1b and at −3, −11, −21, −26, and −47 in the region 5’ of exon 1a. All of these start sites in the region 5’ of exon 1a were preceded by putative SP-1 binding sites. Enzymatic primer extension utilizing an antisense primer beginning at −88 at the 3’ end of exon 1c suggested that transcription initiated (Figs. 5 and 6) primarily at −260, to a lesser extent at −147, and to a much lesser extent at −146, −175, and −176.

From the results provided here it appears that the organization of the FPGS gene is highly conserved between murine and human tissues with, perhaps, the exception of exon 1. The nucleotide sequence incorporated in exons 2–15 of the murine gene was approximately 82% homologous to the same exons in...
the human gene (16, 17). However, as expected differences were found in length among the introns in the gene of each species. The identification of different splice variants at the level of their DNA product incorporating a sequence analogous to either a portion (exon 1b) or all (exons 1a plus 1b) of the human exon 1 was a surprising finding and would appear to suggest that this region of the gene in the mouse is represented as two distinct exons of 72 (exon 1a) and 92 (exon 1b) nucleotides in length. This conclusion is consistent with the identification of an internal donor/acceptor site at the appropriate position within this region of the gene. More conclusive evidence for the divisibility of exon 1 by internal splicing was obtained in the form of a third cDNA variant, which incorporated a novel 5' end (exon 1c) spliced to exon 1b. The position of exon 1c is upstream of exon 1a in a region of the FPGS gene, which is promoter-like. The organization of this portion of the gene as suggested by the results may offer alternative options for control of transcription of different splice variants by specific trans-acting factors binding at different sites within this region. To what extent this might actually happen will require additional studies. It should also be pointed out that the complexity pertaining to exon 1 in the murine FPGS gene is different from that recently reported (22) for the human gene. However, the murine exon 1c and the human exon 1a as reported earlier (22) by Shane and co-workers appear to be located at approximately the same position within the highly GC-rich 5' region. However, only the murine exon 1c incorporates an ATG start codon.

The above results bear on the question as to the manner in which synthesis of mitochondrial and cytosolic forms of FPGS are regulated. Earlier work on the human FPGS gene by others (20, 21) suggests that this could occur at the level of different usage of transcriptional start sites that generate transcripts with either one or both ATG start codons. Alternatively, such regulation could occur at the level of translation by differential usage of each ATG start codon. In addition, we now have evidence that suggests that regulation of the synthesis of each form of FPGS could also occur as a consequence of alternate exon selection during processing of immature FPGS transcripts. Specifically, transcripts were isolated in the form of cDNA that contained either both ATG start codons (exon 1a plus 1b in Variant I and exon 1c plus 1b in Variant III) or only the downstream ATG start codon (exon 1b in Variant II). Further work will be required to determine which of these three possible mechanisms, if any, are predominant in situ. It should also be noted that alternate usage of exons 1a and 1c will result in the synthesis of mitochondrial forms of FPGS with leader peptides of different length and composition. The significance of this finding in terms of the mitochondrial targeting of FPGS and its functional properties is unknown and will require

Fig. 5. Nucleotide sequence of the 5' region of the murine FPGS gene. The DNA sequence is numbered from the ATG start codon for the mitochondrial leader sequence in Variant I. The boundaries for exon 1a, 1b, and 1c are designated in the figure. Major putative transcriptional start sites identified by primer extension are indicated by vertical arrows, and other potential regulatory elements are in boldface type and identified by the appropriate symbols.

CAGAAATCGT CCGCTCTCTG CTCCTCCAGT GCTGGATTAA AAGGTTGCGC GACACAGCGC -1028
CAGCCAATAA TTCTTTATCT CTTTTGGCCT TTTTTGACT TGGTTTGGAG -988
GCCGAATTCC GCTCATACCC CCAGTTGAGA GCTCAATGG TGGACCGCC TGGGGTGGTC -928
CTAAAGCTTG CTACCTGGGA GACCACTTCC TCCAAA CCTAACTTCTACGC -868
GGGTCAAGG TTCTCGTCGA GGAAAGCACC GAGCTCTTCC AAAAAAGGAG -808
CCTTCGCCAC CTTCTCCCTT TTCTCCCTTT TCTGTTGCAA TCCCCCGGTT CAACACCCACA -748
GATA-1/2/3 C-RTs
GCTGCTGAGGG GTGCTAT CTTCCTGAGAACGC TTCTCTA GCTTCTAGT AGATGTATT -688
AAAGCTCTGC AGAAGCCCGGC GGTGCTGGCA CAGCTCTTAA AATCGGGAAG GCCTCAACAA -628
GAGGAACGCG CATTCTTGAG TCTACGCGCA GCCTGACCTA CAGATATACG TACCTGTACA -566
GGGCCTGACA GAGAAAGCCT GCTCTGACTT GAGGAGCCA AAAAAAAA GCCTGAGGAG -508
ACCGGCTGT ATCTCTGACC GAAAAAGCTG CACCCCATGT CTGCGGTGTA GCGAGCAT -448

| Exon 1c | GATA |
|--------|------|
| CAGAAATCGT | C-RTs |
| GCCGAATTCC | GCTCATACCC |
| CCAGTTGAGA | GCTCAATGG |
| GAGCTCTTCC | AAAAAAGGAG |

** FIG. 5.** Nucleotide sequence of the 5' region of the murine FPGS gene. The DNA sequence is numbered from the ATG start codon for the mitochondrial leader sequence in Variant I. The boundaries for exon 1a, 1b, and 1c are designated in the figure. Major putative transcriptional start sites identified by primer extension are indicated by vertical arrows, and other potential regulatory elements are in boldface type and identified by the appropriate symbols.
Murine Folylpolyglutamate Synthetase Gene

Fig. 6. Enzymatic primer extension analysis of the 5′ region of the murine FPGS gene defining potential transcription start sites. The oligonucleotide primer used (5′-ATACTCCATGGCCCCGCTC-3′) initiates transcription (panel A) at position +135 (Fig. 5). The oligonucleotide primer used (5′-CTCTATGGATGAACTC-3′) initiates transcription (panel B) at position +98 (Fig. 5). The source of poly(A)′′ mRNA was wild-type L1210 cells. Additional details are given in the text. A, 6% polyacrylamide electrophoresis gel. Lane 1, complete primer extension reaction of L1210 cell poly(A)′′ RNA plus primer; lane 2, the same poly(A)′′ RNA with primer but no RT; lane 3, complete primer extension reaction of control poly(A)′′ RNA with primer; lane 4, dephosphorylated dX174 HindIII DNA markers which were labeled and added to the gel according to Promega AMV RT Primer Extension Protocol. B, 8% polyacrylamide electrophoresis gel. Lane 1, complete primer extension reaction of L1210 cell poly(A)′′ RNA plus primer; lane 2, the same poly(A)′′ RNA with primer but no RT; lane 3, complete primer extension reaction of control poly(A)′′ RNA with primer; lane 4, DNA markers as in A. The units of length are given in terms of the extended size of the transcript (total primer length). The arrows indicate potential transcription start sites. Additional experimental details are given in the text.

further work.

Analysis of the 5′ region of the murine FPGS gene revealed interesting similarities and differences from the same region (20) of the human FPGS gene. This GC-rich region also appears to be a member of the TATA-less group of promoters controlling genes associated with cellular metabolism that nevertheless are able to respond to a variety of transcriptional regulatory signals. The results of enzymatic primer extension revealed the presence in the murine FPGS gene of multiple potential transcriptional start sites as was shown (20) earlier for the human gene. This appears to be a property (30) of many TATA-less promoters and may offer opportunities for diversity in expression of this gene in terms of tissue specificity during embryonic development and differentiation. Otherwise, the significance of these findings is unknown. Of considerable interest was the finding of a major pair of transcriptional start sites (+83 and +84) in exon 1b just 5′ of the ATG start codon for cytosolic FPGS. The role of these start sites is unknown but they could be involved in the transcription of additional splice variants excluding exon 1 or incorporating other alternatives to this exon within the first intron. The screening of cDNA clones we have carried out so far focused on the selection of clones with insert sizes similar in length to the published (19, 20) human FPGS cDNA. Finally, the mouse gene showed differences from the human gene in the content of putative binding sites for various transcriptional factors, particularly with regard to the number of SP-1 sites. The functional significance of these differences is unknown, and thus further work will be required on this aspect of the FPGS gene and its relevance for transcriptional regulation of this enzymic property.

REFERENCES
1. McBurney, M. W., and Whitmore, G. F. (1974) Cell 2, 173–182
2. McGuire, J. J., Hsieh, P., Coward, J. K., and Bertino, J. R. (1980) J. Biol. Chem. 255, 5776–5788
3. Kislik, R. L. (1981) Mol. Cell Biochem. 39, 331–346
4. McGuire, J. J., and Coward (1984) in Folate and Pterins (Blakely, R. L., and Benkovic, S. J., eds) Vol. 1, pp. 135–190, Wiley-Interscience, New York
5. Moran, R. G., and Colman, P. D. (1984) Biochemistry 23, 4580–4589
6. Barrueco, J. R., O’Leary, D. F., and Sirotnak, F. M. (1992) J. Biol. Chem. 267, 15356–15361
7. Kislik, R. L. (1984) in Folate Antagonists as Therapeutic Agents (Sirotnak, F. M., Burchall, J. J., Ensminger, W. D., and Montgomery, J. A., eds) Vol. 1, pp. 2–55, Academic Press, New York
8. Jackson, R. C., and Grindey, G. B. (1984) in Folate Antagonists as Therapeutic Agents (Sirotnak, F. M., Burchall, J. J., Ensminger, W. D., and Montgomery, J. A., eds) Vol. 1, pp. 290–311, Academic Press, New York
9. Sirotnak, F. M. (1985) Cancer Res. 45, 3992–4000
10. Shane, B. (1989) Vitam. Horm. 45, 263–335
11. Samuels, L. L., Moccio, D. M., and Sirotnak, F. M. (1985) Cancer Res. 45, 1488–1495
12. Sirotnak, F. M., and DeGrave, J. L. (1984) in Folate Antagonists as Therapeutic Agents (Sirotnak, F. M., Burchall, J. J., Ensminger, W. D., and Montgomery, J. A., eds) Vol. 2, pp. 43–91, Academic Press, New York
13. Fabre, I., Fabre, G., and Goldman, I. D. (1984) Cancer Res. 44, 3190–3195
14. Moran, R. G., Colman, P. D., Rosowski, A., Forsch, R. A., and Chang, K. K. (1985) Mol. Pharmacol. 27, 156–166
15. Rumberger, B. G., Barrueco, J. R., and Sirotnak, F. M. (1990) Cancer Res. 50, 4639–4643
16. Barredo, J., and Moran, R. G. (1992) Mol. Pharmacol. 42, 687–692
17. Sirotnak, F. M., Johnson, T. B., Samuels, L. L., and Galivan, J. A. (1988) Biochem. Pharmacol. 37, 4229–4241
18. Egan, M. G., Sirlin, S., Rumberger, B. G., Garrow, T. A., Shane, B., and Sirotnak, F. M. (1990) J. Biol. Chem. 265, 5462–5468
19. Garrow, T. A., Adman, A., and Shane, B. (1992) Proc. Natl. Acad. Sci. U. S. A. 89, 9151–9155
20. Freemantle, S. J., Taylor, S. M., Krystal, G., and Moran, R. G. (1995) J. Biol. Chem. 270, 9579–9584
21. Taylor, S. M., Freemantle, S. J., and Moran, R. G. (1995) Cancer Res. 55, 6030–6034
22. Chen, L., Qi, H., Korenberg, J., Garrow, T. A., Choi, Y.-J., and Shane, B. (1996) J. Biol. Chem. 271, 13077–13087
23. Roy, K., Matsugi, K., Sirlin, S., Shane, B., and Sirotnak, F. M. (1995) J. Biol. Chem. 270, 26918–26922
24. Hendrickx, S. M., Cohen, D. I., Nielsen, E. A., and Davis, M. M. (1984) Nature 308, 149–153
25. Sanger, F., Nickleson, S., and Coulson, A. R. (1977) Proc. Natl. Acad. Sci. U. S. A. 74, 5463–5467
26. Badley, J. E., Bishop, G. A., St. John, T., andFrelinger, J. A. (1988) BioTechniques 6, 114–116
27. Gubler, V., and Hoffman, B. J. (1983) Gene (Amst.) 25, 269–275
28. Gilleland, G., Perrin, S., Blanchard, K., and Bunn, H. F. (1990) Proc. Natl. Acad. Sci. U. S. A. 87, 2725–2730
29. Padgett, R. A., Grabowski, P. J., Kanarska, M. M., Seiler, S., and Sharp, R. A. (1986) Annu. Rev. Biochem. 55, 1119–1150
30. Kozak, M. (1987) Nucleic Acids Res. 15, 8125–8148
31. Von Heijne, G. (1986) EMBO J. 5, 1335–1342

by guest on July 23, 2018http://www.jbc.org/Downloaded from http://www.jbc.org/
Organization and Alternate Splicing of the Murine Folylpolyglutamate Synthetase Gene: DIFFERENT SPLICE VARIANTS IN L1210 CELLS ENCODE MITOCHONDRIAL OR CYTOSOLIC FORMS OF THE ENZYME

Krishnendu Roy, Kenji Mitsugi and F. M. Sirotnak

J. Biol. Chem. 1996, 271:23820-23827.
doi: 10.1074/jbc.271.39.23820

Access the most updated version of this article at http://www.jbc.org/content/271/39/23820

Alerts:
- When this article is cited
- When a correction for this article is posted

Click here to choose from all of JBC's e-mail alerts

This article cites 31 references, 15 of which can be accessed free at http://www.jbc.org/content/271/39/23820.full.html#ref-list-1