Transcription of the Human Folylpoly-γ-glutamate Synthetase Gene*

(Received for publication, March 7, 1997, and in revised form, June 25, 1997)

Sarah J. Freemantle‡ and Richard G. Moran§
From the Department of Pharmacology and Toxicology and the Massey Cancer Center, Medical College of Virginia, Virginia Commonwealth University, Richmond, Virginia 23298

In mammals, folylpoly-γ-glutamate synthetase (FPGS) activity is found in any cell undergoing sustained proliferative phases, but this enzyme also displays a tissue-specific pattern of expression in differentiated tissues. It is now reported that the steady state levels of FPGS mRNA in normal and neoplastic cells reflect these patterns, supporting the concept that the control mechanisms underlying this distribution are transcriptional. To initiate an understanding of these interacting levels of control, we have determined the position and properties of the minimal FPGS promoter controlling transcription of the FPGS gene in human CEM leukemia cells, a line which expresses high levels of this enzyme and its mRNA. The TATA-less region immediately upstream of the major transcriptional start site previously mapped in human tumor cells, which includes several GC- and Y-boxes, functioned as a remarkably efficient promoter when used to drive expression of a luciferase reporter in transient expression studies in CEM cells. The minimal region of the FPGS promoter required for maximal transcriptional activation in CEM cells included the 80 base pairs over which the multiple transcriptional start sites were located, and the 43 base pairs immediately upstream. DNase I footprint analysis detected the binding of Sp1 at all seven of the consensus sites within the probe used, two of which are contained within the minimal promoter region. The several Sp1 sites immediately upstream of the first major transcriptional start activated transcription in Drosophila cells when cotransfected with an Sp1 construct, including those in the region which functioned as a minimal promoter in CEM cells. An additional region of the minimal promoter, situated between the two translational start codons of the FPGS gene, was bound by protein(s) from HeLa cell nuclear extracts. We conclude that transcription of the FPGS gene in CEM cells involves transcriptional events over a limited upstream DNA sequence and that the FPGS promoter used in proliferating human leukemic cells has strong similarity to other TATA-less promoters that utilize tandem, closely spaced Sp1 sites to initiate transcription.

Folylpoly-γ-glutamate synthetase (FPGS) catalyzes the ATP-dependent formation of an amide bond between the γ-carboxyl group of the naturally occurring folates and the amino group of glutamic acid. The addition of glutamic acid moieties to folate compounds allows their intracellular retention and concentration for both the folate cofactors (1–3) as well as for all of the "classical" folate antimetabolites studied to date. As a result of its role in the retention of folate cofactors in the cytosol and mitochondria, FPGS is essential for folate homeostasis and the survival of proliferating mammalian cells. The metabolism of antimetabolites by this enzyme plays a major role in the cytotoxicity, and perhaps the selective cytotoxicity, of several folate antagonists used or under development for the treatment of human cancers.

Because any cell that is attempting to proliferate during exposure to classical antifolates is susceptible to cytotoxicity if it expresses FPGS, the tissue distribution of this enzyme and the factors controlling its expression are of critical importance. From early studies on the distribution of enzyme activity, it was known that the levels of FPGS are moderately high in some tumors, in normal gut and bone marrow stem cells, and in the two specialized organs of folate metabolism, liver and kidney, but enzyme levels in other mouse adult tissues were barely detectable (4–6). Studies relating methotrexate polyglutamylation and FPGS activity in individual childhood leukemias to the therapeutic activity of this drug have suggested that the sensitivity of some of these tumors to antifolate chemotherapy is causally related to the level of expression of FPGS (6–9). In a recent study of FPGS in human leukemias (9), B-lineage leukemias, which are more responsive to methotrexate, had higher levels of FPGS activity than did T-lineage leukemias. Cells from acute nonlymphoblastic leukemias, which are generally resistant to antifolates, had lower enzyme levels (9). This lineage-specific effect was also seen in cells selected from normal bone marrow; lymphoid progenitor cells (CD10/CD19 + ) had high levels of FPGS activity similar to those seen in blast cells from patients with B-lineage leukemias, whereas normal nonlymphoid progenitors (CD34 + ) had lower levels (9). On the other hand, a decrease in cellular FPGS activity in tumor cells can contribute to or cause both acquired (10, 11) and intrinsic (12) resistance to methotrexate.

A previous study in this laboratory (6) concluded that FPGS levels are controlled by at least two mechanisms, one of which is linked to proliferation and the other results in a tissue-specific pattern of expression in differentiated tissues. Despite this complexity, all available evidence (13) appears to rule out the existence of multiple genes for FPGS function, a fact that has focused attention on control of the known FPGS locus. With the publication of the human FPGS cDNA (14), recent studies have defined the structure of the FPGS genomic locus (13) and the use of multiple transcriptional start sites (15), and they have begun to address how the control of FPGS gene expression is mediated. Thus, when HL60 promyelocytic leukemia cells were chemically induced to differentiate, levels of FPGS enzyme activity decrease (6), and its mRNA declined coordinately (16). FPGS enzyme activity was not detectable in human pe-
rribular blood lymphocytes (6), but both enzyme activity and FPGS mRNA increased when the cells were stimulated to proliferate with phytohaemagglutinin (17). Conversely, in a cell line selected for resistance to the second generation dihydrofolate reductase inhibitor edatrexate, the decreased FPGS levels causative of resistance were demonstrated to be due to a decrease in the rate of transcription of the FPGS gene (18). Hence, the limited information available implicates the level of transcription of this gene as the major determinant of cellular FPGS activity.

On these bases, we initiated a study to define the processes involved in the transcriptional control of this gene. In this study, we report the promoter region of the FPGS gene in a human leukemic cell. We demonstrate that a well defined short sequence, which covers the major transcriptional start sites mapped previously in several human tumor cell lines (19), and an additional 43 bp upstream, are necessary and sufficient for basal promoter activity and that discrete binding sites for Sp1 and for a second nuclear factor define this region. These studies constitute the first description of the promoter controlling transcription of this gene and set the stage for studies defining the proliferation-linked and tissue-specific controls on the expression of the FPGS gene.

**MATERIALS AND METHODS**

**Northern Analysis of FPGS mRNA**—RNA from murine normal tissues and human tumor cell lines was extracted using the Trizol reagent (Life Technologies, Inc.), denatured with glyoxal, and separated by size on a 1.2% denaturing agarose gel in 10 mM sodium phosphate buffer, pH 7.0. The RNA was transferred onto nylon membranes (Biotrans, ICN, Irvine, CA) and hybridized either with the first 690 bp of the published human cytosolic FPGS-coding region (14) or with a 1.7-kb probe representing the downstream sequences of the mouse L1210 cell FPGS cDNA (19). Probes were random-labeled to a specific activity of about 2 × 10^9 cpm/μg and were used at a concentration of 1 × 10^6 cpm/ml of hybridization solution (0.5 mM sodium phosphate, pH 7.0, 7% SDS, 1% bovine serum albumin, and 1 mM EDTA). A Northern blot of poly(A)^+ RNA (2 μg/lane) from normal human tissues was purchased from CLONTECH Laboratories (Palo Alto, CA). This latter blot was blocked and hybridized in 75 mM NaCl, 50 mM phosphate, 5 mM EDTA, 2% SDS, 100 μg/ml sheared salmon sperm DNA, and 0.1 × Denhardt's solution for 16 h at 65 °C, and filters were washed to a stringency of 0.2 × SSC and 0.1% SDS at 65 °C. For both murine and human Northern blots, a human β-actin probe was used to normalize for RNA loading.

**5′-Rapid Amplitification of cDNA Ends (RACE)—**5′-RACE was carried out starting with 2 μg of poly(A)^+ -selected RNA from CEM cells using a gene-specific primer. Both these primers were assayed in an initial PCR; poly(C)-tailoring; subsequent polymerase chain reaction (PCR) amplification used a nested primer within exon 4 and an anchor primer. A published method (20) was followed using the primers and reaction conditions previously described (15). The PCR products resultant from this 5′-RACE, representing the 5′ ends of the FPGS mRNA, were separated on a 1.5% agarose gel. The products ranged over approximately 100 bp, and three individual ligations into the pCRII vector (In-Vitrogen, San Diego, CA) were carried out using DNA purified from the 5′-end product. All PCR fragments were cloned into the pCRII vector, released with AccI and EcoRV sites of the pBlueScript II to analyze the antisense strand. This probe was used to generate these constructs were RC1, 5′-GGTACCGACG-CTGCACCTGG-3′; RC2, 5′-GGTACCTTTGGGGCGGTGCTGATT-5′-GATTGAGCAT-3′; and Mut1, 5′-CACGTGCAGCTGATACCTGGTACTCCTTGC-3′. All PCR fragments were cloned into the pCRII vector, released with enzymes Acc65I and EcoI for ligation into the Acc65I and Smal sites of the pG33-basic vector. The cloned segments which were PCR-generated were sequenced to confirm a lack of PCR-generated artifacts.

**SP1 Expression Constructs—**The SP1 expression construct pPacSP1 containing the 2.1-kb SP1 cDNA downstream of the Drosophila actin promoter and the corresponding vector without insert, pPac-NdeI, were generously provided by Dr. Robert Tjian (University of California, Berkeley).

**Transfection Assays—**CEM cells were cultured in RPMI 1640 medium supplemented with 10% fetal calf serum. Exponentially growing cells were washed once with phosphate-buffered saline and resuspended in serum-free medium at a density of 1.2 × 10^6 cells/ml. Aliquots (0.8 ml) of cells were mixed with 10 μg of the luciferase construct and 30 μg of the internal control vector CMV-β-galactosidase, and were transfected using a BTX transfectionator 300 electroporator (Biotechnologies and Experimental Research, Inc., San Diego, CA) at 250 V and 300 microfarads. As the size of the deletion constructs decreased, the amount of plasmid solution used to generate these constructs was decreased proportionally to maintain the construct molar constant between transfections; pG33-basic vector was used to keep the total DNA content per transfection sample at 10 μg. After electroporation, cells were cultured in 12 ml of complete medium for 48 h. The cells were then washed once in phosphate-buffered saline and lysed in 300 μl of reporter lysis buffer (Promega). The lysed cells were centrifuged at 14,000 rpm for 1 min in a Microfuge, and the supernatant was assayed for luciferase activity and β-galactosidase activity (Promega). The β-galactosidase activities of the extracts were determined and used to normalize luciferase levels for transfection efficiency. Each experiment was performed at least three times with triplicate determinations in each experiment. Results were first normalized for β-galactosidase activity, then expressed either relative to the luciferase expression driven by the SV40 promoter of the pG33-control vector in the same experiment or as a percentage of the activity produced by the most active FPGS promoter construct in a series.

**Drosophila Schneider 2 cells (1 × 10^6 cells) were plated onto 60-mm dishes in Schneider’s Drosophila medium (Life Technologies, Inc.) plus 10% fetal calf serum and used the next day for transfection. Each luciferase construct (10 μg) and 1 μg of either pPacSP1 or pPac-NdeI were transfected into the cells as a calcium phosphate precipitate. After 48 h, the cells were scraped from the plates, washed twice with phosphate-buffered saline, and lysed in 300 μl of reporter lysis buffer, as above. Supernatants were assayed for luciferase activity, which was then normalized to protein content measured using an adaptation of the Bradford technique (Bio-Rad).

** DNase I Footprinting**—DNase I footprints of individual probes were generated from the pStu-Eae construct (see Fig. 3) by cleaving with Acc65I to analyze the sense strand or with BglII to analyze the antisense strand. This probe stretched from position −180 bp to +124 bp where +1 represents the most 5′ major transcriptional start site. 5′-End-labeling was performed using Superscript reverse transcriptase (Life Technologies, Inc.) and high specific activity 32P-labeled dATP and dCTP (DuPont NEN). After labeling, the fragment was released using the opposite restriction en-
enzyme site, and the labeled fragment was gel-purified. Protein binding and DNase I reactions were carried out using 3 footprinting units of recombinant Sp1 or 25 μg of HeLa nuclear extract (Promega) in the buffers supplied with the Hotfoot footprinting system (Stratagene). Nuclear extracts or purified proteins were incubated for 10 min on ice with 100 μg/ml poly(dI-dC) before the addition of 15,000 cpm 32P-PAT. RNA from normal human tissues (19); total RNA (10 μg) from murine normal tissues from female DBA/2 mice was probed with a 1.7-kb downstream murine FPGS cDNA (19); a human normal tissue poly(A)+ mRNA (2 μg) was probed with a fragment of human FPGS cDNA which included the first 690 bp of coding region sequence; and c, total RNA (10 μg) from the indicated human tumor cell lines was probed with a 690-bp human FPGS cDNA fragment.

**RESULTS**

**FPGS mRNA Levels in Normal and Neoplastic Tissues**—The size and abundance of transcripts from the FPGS gene were determined by Northern blot analysis in normal tissues and neoplastic cell lines. Previous studies (4, 6) of the levels of FPGS activity in normal mouse tissues had indicated high enzyme activity in murine liver and kidney, with lower levels in spleen, and very low to undetectable levels in several other tissues. All mouse tumors surveyed in those studies expressed substantial levels of FPGS activity (4, 6). The hybridization of a near full-length mouse cDNA probe to Northern blots mirrored this pattern (Fig. 1a), with FPGS mRNA levels highest in the liver and kidney and equivalent levels in the L1210 leukemia cell line. Mouse brain, heart, skeletal muscle, lung, and spleen had low to undetectable levels of message. A single 2.4-kb band was detected in all mouse cells expressing FPGS. Northern analysis of poly(A)+ RNA from normal human tissues, using a 0.7-kb probe corresponding to the 5′-coding region of the FPGS gene, showed some interesting differences to the pattern in mouse tissues (Figs. 1, a and b). As in the mouse tissues, human liver showed high levels of FPGS mRNA, and kidney also expressed relatively high levels. Human pancreas, a tissue which was not studied in the mouse, had levels intermediate between that of the liver and kidney. A low level of FPGS message was detected in the brain as was seen with the mouse, and surprisingly little was seen in human placenta. However, in mRNA from heart and skeletal muscle there were very high levels of FPGS-specific mRNA, in direct contrast to the levels in these tissues in the mouse. Although the basis of this difference is not clear at this time, this result has been reproduced on a second human multiple tissue blot. The major species of FPGS message in human normal tissues was 2.4 kb, but a minor band at 2.2 kb was also detected which represented 5–10% of the total hybridization, as estimated by PhosphorImager scanning. When RNA from eight human tumor cell lines was hybridized with the human FPGS probe, there was a high level of FPGS mRNA detected in all of the samples (Fig. 1c). The level of FPGS mRNA in these eight samples did not differ by more than a factor of four. Likewise, RNA extracted from the circulating lymphoblasts of a pediatric patient with acute lymphoblastic leukemia was found to have a similarly high level of FPGS mRNA. The lower molecular weight hybridizing band seen in the normal human tissues was not detected in any of these tumor specimens.

**5′-RACE Analysis of FPGS mRNA from CEM Cells**—We had previously used 5′-RACE and ribonuclease protection assays to define the transcriptional start sites for FPGS in a series of...
human leukemic and carcinoma cell lines (15). In that analysis, all of the 25 RACE products sequenced originated within what we defined as exon 1 (15). Recently, sequence heterogeneity has been reported at the 5' end of mouse and human FPGS transcripts resultant from alternative use of upstream exons spliced to exon 2 (21, 22). Hence, we extended our previous 5'-RACE analysis to more thoroughly define the initial exon used in CEM cells and, therefore, the position of the promoter used in this cell line. Of the 203 5'-RACE colonies picked, 140 hybridized with an exon 2 probe and 126 (90%) also hybridized with exon 1. The 14 clones which hybridized to exon 2 but not to the exon 1 probe were sequenced. None represented cDNAs with alternative 5' sequence; the clones sequenced either truncated within exon 2 or were false negatives, which contained exon 1 sequence, or corresponded to an unspliced precursor mRNA (1 clone). We concluded from this extensive 5'-RACE analysis that all of the transcripts from the FPGS gene in CEM cells initiated transcription using the previously defined (15) start sites in exon 1, and that DNA sequences close to exon 1 were uniquely promoting transcription of this gene in this human tumor cell. Hence, we initiated an analysis of the exon 1 proximal promoter with the concept that it was either the only or, at least, the major promoter used for the FPGS gene in this model of a human tumor.

**Design of FPGS-Luciferase Reporter Constructs and Definition of the Minimal Promoter Region**—In a previous study, we demonstrated how differential usage of two start methionines located within exon 1 of the human FPGS gene leads to targeting of FPGS activity to the mitochondrial and cytosolic compartments (15). Transcription was initiated at multiple sites spread over 80 bp, which divide the transcripts into two classes differing by the presence or absence of an additional upstream translation initiation site. In designing promoter reporter constructs to study transcription from this gene, segments of the sequence immediately upstream of the first exon were placed in front of a luciferase gene in the pGL3 plasmid, including the sequence at which transcription of the endogenous gene is initiated, but not the start methionine, which initiates translation of the coding region of the protein. Hence, in a series of such constructs, translation of luciferase would initiate from the same ATG downstream of any putative promoter sequences, allowing comparisons of luciferase activity as transcriptional efficiencies. Thus, problems associated with upstream translational initiation and early termination of
transcripts in the pG3 constructs would be avoided. However, in the FPGS gene, inclusion of all of the transcriptional initiation sites would also place the codon for the upstream start methionine (which initiates translation of the mitochondrial form of human FPGS) in some of the transcripts. Although any peptides initiated from this codon would be terminated by in-frame stop codons in the pG3 vector prior to the luciferase gene, the presence of this methionine had an effect on translation initiated downstream at the luciferase start methionine (Fig. 1). Thus, when the luciferase production was compared using two constructs differing only by the presence of the intact upstream start codon (pH3-BsHIII), or a mutated start codon (AGT) (pH3-BsHIII-Mut), the presence of the intact ATG was found to decrease luciferase activity by over 5-fold. Two other constructs also tested the effect of the presence of this ATG, pHinc-Eae and pH3-Eae; these constructs extended 3′ almost to the downstream start methionine, and stretched upstream for a length of 1.5 and 1.8 kb, respectively. The construct with the intact ATG (pHinc-Eae) produced 2.5-fold lower luciferase activity than did pH3-Eae, an effect which could be ascribed to the presence of the upstream start codon, because further deletion analysis demonstrated that the additional upstream sequence did not result in more efficient transcription (see below). Subsequently, all FPGS constructs that contained this region were designed to have an ATG codon in place of the upstream start codon to ensure that translation of the transcripts initiated only from the luciferase cassette start methionine. Overall, the sequence immediately upstream of exon 1 of the FPGS gene has very substantial promoter activity as can be seen from the results in Fig. 2, in which luciferase production in transient transfections driven by these constructs in CEM cells is expressed relative to the level of luciferase produced from the SV40 minimal promoter. The pH3-Eae construct produced 5.6-fold more luciferase activity than the SV40 promoter, a surprising level given that the latter is such an active promoter.

A series of 3′-deletion constructs of the FPGS promoter region were made to assess the importance of sequences in this region on transcription of the FPGS gene. As shown by the data of Fig. 3b, removal of 82 bp at the 3′ end of the very active pH3-Eae construct to form pH3-BsHIII-Mut reduced the luciferase activity ∼3-fold. This construct does not include the 3′-most transcriptional start site (Fig. 3a). However, from ribonuclease protection analysis (15), this start site represents only about 5–10% of the total FPGS message in CEM cells, so that the 3-fold drop in activity appears to be due to the loss of sequences that enhance transcription rather than to the loss of a minor transcriptional start site. The pH3-SmaI construct has another 60 bp removed from the 3′ end and includes only the two upstream transcriptional start sites. Transfection of this construct produces approximately 18- and 5-fold less luciferase activity than the pH3-Eae and pH3-BsHIII-Mut constructs, respectively. Again, the drop in activity does not correspond to the loss of transcriptional start sites; although only two major start sites remained in this construct, they represented over 50% of the transcripts as judged by ribonuclease protection analysis (15). We concluded from this analysis that the majority of exon 1 was essential for maximal transcriptional activity from the FPGS promoter.

5′-Deletions of the FPGS promoter region were constructed and analyzed to define the upstream boundary of the minimal promoter region (Fig. 3c). The 1.6 kb of sequence upstream of exon 1 contained in the pH3-Eae construct could be reduced to 43 bp (pF11-Eae and pRC-Mut3) without a reduction in luciferase activity in transient assays in CEM cells. In fact, as the construct size was decreased, there was a gradual increase in luciferase activity. The construct pRC1-Mut3, in which both start methionines have been mutated, was designed to determine whether extension of what appeared to be the minimal promoter (pF11-Eae) further 3′ to the end of exon 1 enhanced transcription. Transcription levels driven by the pRC-Mut3 and pF11-Eae constructs did not differ significantly, leading to the conclusion that sequences downstream of the Eae site did not contribute substantively to the function of this promoter.

This conclusion was supported by a comparison of transcription from the pStuI-Eae and pRc2-Mut3 constructs (Fig. 3c). The minimal portion of the FPGS promoter required to drive transcription, then, consists of bp 1 to 150 of exon 1 and the 43 bp of the 5′ region flanking it. This region contains two forward and one reverse GC-boxes, the core binding sequence for the transcription factor Sp1, and a single inverted CCAAT motif, i.e., a potential Y-box.

**SP1 Transactivation of the FPGS Promoter**—Because several potential binding sites for Sp1 mapped to the region immediately upstream of transcriptional initiation, we sought to determine the functional activity of binding of Sp1 to the FPGS promoter in the control of transcription from this gene. To address this question, a subset of the FPGS-luciferase reporter constructs were transfected into Drosophila SL2 cells, which lack endogenous Sp1, and the effect of cotransfection with the Sp1 expression vector, pPacSP1, was tested. Transfection of the luciferase gene driven by the SV40 control promoter showed a
260-fold activation of expression by Sp1, compared with that seen in cells cotransfected with the control plasmid, pPac + NdeI (Fig. 4), in agreement with previous experience with this system (23). The longest FPGS promoter construct studied (pH3-Eae) was as effective as the SV40 control promoter in stimulating luciferase activity in this system (approximately 300-fold). As the size of the FPGS promoter region was decreased from the 5’ end, the degree of activation by Sp1 also decreased. However, even the construct found to define the minimal promoter in human cells, pF11-Eae, which contained only two Sp1 binding sites, showed a 22-fold induction compared with that observed with the vector only control. This difference between the transactivation of luciferase by upstream sequences of pF11-Eae in Drosophila and human leukemic cells was very clear, but our results do not allow us to distinguish whether there is a silencer element operative upstream in the human cells (as perhaps suggested by the data of Fig. 3c) or an upstream enhancer that became functional in Drosophila cells. Of course, it cannot be excluded that the effect is due to the high levels of Sp1 expression expected from the Drosophila actin promoter used in pPacSp1. Nevertheless, we could conclude that Sp1 binding to the GC boxes immediately upstream from exon 1 in this gene would greatly stimulate transcription from the FPGS promoter and that they constitute functional Sp1 binding sites.

DNase I Footprint Analysis of the Minimal FPGS Promoter—DNase I footprint analysis of the FPGS promoter region showed four regions protected by purified Sp1 transcription factor that were also protected by HeLa cell total nuclear extract (Fig. 5a). Each of these sites corresponded to GC-boxes which were expected to bind Sp1 (Fig. 5b). The HeLa extract also protected another site not bound by Sp1, which is labeled H1 in Fig. 5b. This site lies just downstream of the first start methionine codon and is therefore in a transcribed and translated region of the FPGS gene. The H1 site spanned 29 bp and contains a GC-rich region at the 5’ end and a consensus E-box (CCACCTGC) at the 3’ end. The sequences of this region of the mouse (19) and hamster (15) FPGS genes were compared with the human sequence (Fig. 5c). The 5’ end of H1 was conserved across these species, and the core E-box (CACCTG) was present in the human and hamster sequences but was not found in the mouse gene.

DISCUSSION

FPGS serves the function of trapping folate cofactors in the cell and, as a result, is essential for normal cellular proliferation. Earlier studies (5, 6) had shown that FPGS activity was high in cells undergoing rapid proliferation and dropped when cells either committed to differentiation or entered an extended G0 phase. In addition, enzyme levels have been shown to range from undetectable to rather high among differentiated tissues (4, 6), leading to the concept that control of FPGS expression was obeying a tissue-specific regulatory mechanism. The Northern analysis presented here shows a similar pattern. FPGS mRNA levels are high in all of the in vitro cultured tumor cell lines, which are dividing rapidly, whereas, in both the murine and human normal tissues, FPGS message levels showed distinct tissue-specific patterns. It would appear that at least some of these controls operate at the transcriptional level.
level. Having determined the sites of transcriptional initiation of the FPGS gene, we set out to learn the first level of information about the transcriptional control of this locus. Our previous analysis (15) of the sequence directly upstream of the first exon in CEM leukemic cells revealed a GC-rich region with no canonical TATA sequence, a structure usually characteristic of housekeeping genes and proto-oncogenes. In this study, we demonstrate that this region has substantial activity as a transcriptional promoter.

The proximal 120 bp of the promoter studied contained seven forward and one reverse GC-boxes, the core binding site for Sp1 as well as for several other transcription factors. DNase I footprinting analysis with Sp1 recombinant protein and with HeLa nuclear extract indicated that each of the forward GC-boxes bound Sp1 and were bound in Hela cell nuclear extracts. When the reporter constructs were transfected into Drosophila SL2 cells, which lack endogenous Sp1, expression was transactivated by the cotransfection of an Sp1-expressing plasmid. Hence, it would appear that Sp1 plays a major role in transcription from this gene, as with other TATA-less promoters that have multiple GC-boxes immediately upstream of the major site of transcriptional initiation.

The human FPGS promoter gives rise to multiple transcription initiation sites spread over 80 bp (15), a typical characteristic of a TATA-less GC-rich promoter lacking an initiator element (24). In the absence of an initiator element, precedent would predict that start site selection would be directed by a proximal Sp1 binding site (25) that generates multiple start sites between 30 and 55 bp downstream. The positioning of start sites is of significance for the FPGS gene as it is this which determines the subcellular localization of enzymatic activity. The first exon of the FPGS gene in CEM cells contains 2 in-frame start methionine residues; translation initiated from the upstream methionine directs enzyme to the mitochondria and from the downstream methionine directs it into the cytosol. Additional studies will be required to define the role of the factor-binding elements within the FPGS promoter in the regulation of start site distribution in this gene.

Within the minimal promoter region, only one footprint, designated herein H1, could not be attributed to GC-boxes. The 3′ region of this 29-bp sequence was homologous with an E-box (5′-CACCTGTC-3′), an element first recognized in a tissue-specific enhancer of the murine heavy chain immunoglobulin gene (26). The CANNTG portion of this motif has since been shown to be the core binding site for the basic helix-loop-helix transcription factors (27). The E-box-binding proteins represent several families of proteins, one group of which is the muscle regulatory factors involved in transcriptional activation of muscle-specific genes (27, 28). We note that the position corresponding to the H1 element in the human gene did not contain an E-box in the mouse gene (Fig. 5c), an observation which suggests that this element is involved in the expression of FPGS mRNA in human, but not mouse, heart and skeletal muscle.

There has been circumstantial evidence (29, 30) for the existence of isoforms of FPGS, which differ in kinetic characteristics, yet there appears to be only one FPGS genomic locus (13). We had previously mapped the multiple transcriptional start sites for this locus in a number of human cancer cell lines and reported that transcripts for both a mitochondrial and cytosolic forms of FPGS arose using the same promoter (15). It is not yet clear whether these forms of enzyme explain the different catalytic features seen in some studies (29, 30), but it would seem unlikely, given that the mitochondrial leader is expected to be cleaved after passage into this organelle. Our analysis of the 5′-termini of FPGS transcripts in human cancer cell lines (Ref. 15 and this study) has identified only a single form of exon 1. Others have subsequently reported that this same exon 1 initiated either the only or the most prevalent form of FPGS transcripts in HepG2, MCF-7, and human fetal liver (21). In more recent reports, alternate forms of exon 1 have been detected in cDNAs for human (21) and mouse FPGS (22), each of which was spliced to exon 2. To date, the alternative transcripts identified in human cells do not appear capable of translation to active FPGS, and the use of downstream translational start sites also did not lead to active enzyme (21).

On the other hand, the divergence in sequence reported recently at the 5′-terminus of FPGS transcripts in mouse normal and malignant tissues could indeed lead to isoforms of enzyme (22), and the structure of the mouse gene suggested would indirectly support the existence of a second upstream promoter used in normal mouse tissues. It is now a central issue whether the use of alternative promoters or a level of regulation not yet defined on the promoter we herein describe is involved in the tissue-specific control of FPGS activity and/or the generation of functional isoforms of FPGS.

Acknowledgments—We thank Dr. Shirley Taylor for considerable advice and encouragement during this study and Drs. Taylor, Peter Melera, and Kathleen Scotto for critiques of the manuscript. We also thank Dr. Robert Tjian for providing the Drosophila Sp1 constructs.

REFERENCES
1. McGuire, J. J., and Bertino, J. R. (1981) Mol. Cell. Biochem. 38, 19–48
2. McBurney, M. W., and Whitmore, G. F. (1974) Cell 2, 173–182
3. Taylor, S. M., and Hanna, M. L. (1977) Arch. Biochem. Biophys. 181, 331–334
4. Moran, R. G., and Colman, P. D. (1984) Anal. Biochem. 140, 326–342
5. Sirotnak, F. M., Johnson, T. B., Samuels, L. L., and Galivan, J. (1988) Biochem. Pharmacol. 37, 4239–4241
6. Barredo, J., and Moran, R. G. (1992) Mol. Pharmacol. 42, 687–694
7. Whitehead, V. M., Rosenblatt, D. S., and Vuchich, M. J., Jr., Shuster, J. J., Witte, A., and Beaulieu, D. (1990) Blood 76, 44–49
8. Whitehead, V. M., Vuchich, M. J., Lauer, S. J., Mahoney, D., Carroll, J. A., Shuster, J. J., Essentiel, D. W., Payment, C., Look, A. T., Akabutu, J., Bowen, T., Taylor, L. D., Camitta, B., and Pullen, D. J. (1992) Blood 80, 1316–1323
9. Barredo, J., Synold, T. W., Laver, J., Belling, M. V., Pui, C.-H., Priest, D. G., and Evans, W. E. (1994) Blood 84, 564–569
10. Pizzorzo, G., Mini, E., Corenoble, M., McGuire, J. J., Moroson, B. A., Cashmore, A. R., Dreyer, R. N., Lin, J. T., Mazzei, T., Periti, P., and Bertino, J. R. (1988) Cancer Res. 48, 2149–2155
11. McClisken, D. E., McGuire, J. J., Russell, C. A., and Bertino, J. R., Pizzorzo, G., and Mini, E. (1991) J. Biol. Chem. 266, 6184–6187
12. Li, W.-W., Lin, J. T., Schweitzer, B. I., Tong, W. P., Niedzwiecki, D., and Bertino, J. R. (1992) Cancer Res. 52, 3908–3913
13. Taylor, S. M., Freeman, S. J., and Moran, R. G. (1995) Cancer Res. 55, 6030–6034
14. Garrow, T. A., Adamo, A., and Shane, B. (1992) Proc. Natl. Acad. Sci. U. S. A. 89, 9151–9155
15. Freeman, S. J., Taylor, S. M., and Moran, R. G. (1995) J. Biol. Chem. 270, 9579–9584
16. Egan, M. G., Sirlin, S., Rumberger, B. G., Garrow, T. A., Shane, B., and Sirotnak, F. M. (1995) J. Biol. Chem. 270, 5462–5468
17. Fort, D. W., Lark, R. H., Smith, A. K., Marling-Carson, M., Weitman, S. D., Shane, B., and Sirotnak, F. M. (1995) J. Biol. Chem. 270, 4239–4241
18. Roy, K., Mitsugi, K., Sirlin, S., Shane, B., and Sirotnak, F. M. (1995) J. Biol. Chem. 270, 26918–26922
19. Spinella, M. J., Brlige, K., and Goldman, J. D. (1996) Biochim. Biophys. Acta 1306, 11–14
20. Frohman, M. A., Dush, M. K., and Martin, G. R. (1988) Proc Natl. Acad. Sci. U. S. A. 85, 8996–9002
21. Chen, L., Qh, H., Kerenberg, J., Garrow, T. A., Choi, Y.-J., and Shane, B. (1996) J. Biol. Chem. 271, 13077–13087
22. Roy, K., Mitsugi, K., and Sirotnak, F. M. (1997) J. Biol. Chem. 272, 5587–5593
23. Courey, A. J., and Tjian, R. (1988) Cell 55, 887–898
24. Smale, S. T., and Baltimore, D. (1989) Cell 57, 103–113
25. Lu, J., Lee, W., Jiang, C., and Keller, E. B. (1994) J. Biol. Chem. 269, 5391–5402
26. Ephrussi, E., Church, G. M., Tonegawa, S., and Gilbert, W. (1985) Science 227, 434–440
27. Edmondson, D. R., and Olson, E. N. (1983) J. Biol. Chem. 268, 755–758
28. Ludolph, D. C., and Kaniezy, S. F. (1995) FASEB J. 9, 1595–1604
29. Rumberger, B. G., Burreco, J. R., and Sirotnak, F. M. (1994) Cancer Res. 53, 4639–4645
30. Tse, A., and Moran, R. G. (1994) Proc. Am. Assoc. Cancer Res. 35, 304