THE JOURNAL OF BIOLOGICAL CHEMISTRY
Vol. 271, No. 22, Issue of May 31, pp. 13077–13087, 1996

Printed in U.S.A.

Purification and Properties of Human Cytosolic Folylpoly-γ-glutamate Synthetase and Organization, Localization, and Differential Splicing of Its Gene*

(Received for publication, February 28, 1996)

Linda Chen‡, Hong Qi‡, Julie Korenberg§, Timothy A. Garrow‡¶, Yun-Jung Choi‡, and Barry Shane‡‖

From the ‡Department of Nutritional Sciences, University of California, Berkeley, California 94720 and the §Ahmanson Department of Pediatrics and Medical Birth Defects Center, Cedars-Sinai Medical Center, UCLA, Los Angeles, California 90048

Human cytosolic folylpolyglutamate synthetase (FPGS) was expressed in *Escherichia coli* and purified to homogeneity. Tetrahydrofolate and dihydrofolate were the most effective substrates, while 5-substituted folates were poor substrates. Most pteroyldiglutamates were better substrates than monoglutamates.

The human FPGS gene spans 12 kilobases and contains 15 exons and 14 introns. A single FPGS gene was located to chromosome region 9q34.1. Four exon 1 variants were identified, each of which was spliced to exon 2. The exon 1 variant corresponding to the isolated cDNA contains two ATG codons and multiple transcription start sites in this region generates mitochondrial and cytosolic FPGS (Freemantle, S. J., Taylor, S. M., Krystal, G., and Moran, R. G. (1995) *J. Biol. Chem.* 270, 9579–9584). Exons 1B and 1C, generated by alternate splicing in intron 1, and exon 1A, which is 5′ to exon 1 and may encode an additional mitochondrial isoform, are preceded by a number of potential promoter sites.

Chinese hamster ovary cell transfectants expressing FPGS activity in the mitochondria contained normal mitochondrial and low cytosolic folylpolyglutamate pools. Mitochondrial FPGS activity is required for mitochondrial folate accumulation, while cytosolic FPGS activity is needed for establishment of normal cytosolic folate pools. The reconstructed FPGS gene restored normal cytosolic and mitochondrial folate metabolism in hamster cells.

Folylpolyglutamate synthetase (FPGS[1]; tetrahydrofolate:L-glutamate γ-ligase (ADP-forming), EC 6.3.2.17) activity is required for the synthesis and cellular retention of functional folate coenzymes and for the conversion of many antifolates to more active forms (1–6). FPGS activity is highest in proliferating tissues and activity, and mRNA levels increase after mitogen stimulation and decline during differentiation (7–10). It has been proposed that a coordinate down-regulation of folate enzymes including FPGS, and a consequent limitation in macromolecule synthesis, may be an early programmed event in cell maturation (10). The ability of human leukemia blast cells to metabolize methotrexate to polyglutamate derivatives has been correlated with efficacy of methotrexate treatment (11), and a wide variation in FPGS activity and in FPGS mRNA levels has been found in leukemia blast cells (8, 12). Increased expression of FPGS activity in model cells leads to increased sensitivity to antifolate drugs[2] while decreased FPGS activity is a mechanism for resistance to many antifolates, both in model cell systems and in clinical samples (5, 13–16).

Mammalian FPGS is a low abundance protein which has hindered its isolation in sufficient quantities for detailed physical analysis. Pig liver FPGS is the only mammalian enzyme to have been purified to homogeneity (17) although some characterizations of other crude or partially purified mammalian FPGS enzymes have been reported (18–22). In each case, only small amounts of protein have been obtained which has limited studies to kinetic analyses. The low abundance and instability of mammalian FPGS has complicated its purification in sufficient quantity to carry out mechanistic studies.

Mammalian cells possess both mitochondrial and cytosolic FPGS isozymes (23), and additional isoforms may be expressed in different tissues or in tumor cell (24). The Chinese hamster ovary (CHO) cell mutant AUXB1 lacks FPGS activity and as a result requires exogenous purines, thymidine, and glycine for growth (25–28). Studies using AUXB1 cells transfected with *Escherichia coli folC* gene (FPGS) constructs that target the expressed protein to the cytosol or mitochondria have demonstrated that cytosolic FPGS activity is required for cytosolic folate accumulation and purine and thymidylate biosynthesis, while mitochondrial FPGS activity is required for mitochondrial folate accumulation and glycine biosynthesis (23, 29). Expression of *E. coli* FPGS solely in the mitochondria of AUXB1 cells restored normal folate accumulation and metabolism in the cytosol as well as in the mitochondria, demonstrating that pteroyltriglutamates synthesized in the mitochondria can be released into the cytosol, although these polyglutamate species cannot enter the mitochondria (14, 29).

We previously isolated a human FPGS cDNA (30) which encoded cytosolic FPGS and suggested that it might lack 5′

* This study was supported in part by United States Public Health Service Grant CA-41991 from the National Cancer Institute, Department of Health and Human Services. Preliminary accounts of some of these studies were presented at the Tenth International Symposia on Pteridines and Folic Acid Derivatives, March 21–23, 1993, Orange Beach, Alabama. 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.

The nucleotide sequence(s) reported in this paper has been submitted to the GenBank™/EMBL Data Bank with accession number(s) U24252 (upstream region to intron 4) and U24253 (intron 4 to intron 11).

¶ Present address: Division of Foods and Nutrition, 445 Bevier Hall, 905 South Goodwin Ave., Urbana, IL 61801.

‖ To whom correspondence should be addressed. Tel.: 510-642-5202; Fax: 510-642-0535.

[1] The abbreviations used are: FPGS, folylpoly-γ-glutamate synthetase; PteGlu, pteroylglutamic acid, folic acid; $H_4PteGlu_n$, tetrahydropteroylpoly-γ-glutamate, *n* indicating the number of glutamate moieties; CHO, Chinese hamster ovary; 5′-RACE, rapid amplification of 5′-cDNA ends; RT-PCR, reverse transcription polymerase chain reaction; bp, base pair(s); kb, kilobase(s); PAGE, polyacrylamide gel electrophoresis; IPTG, isopropyl-1-thio-β-D-galactopyranoside.

[2] J.-C. Hsu, T. Garrow, and B. Shane, unpublished data.

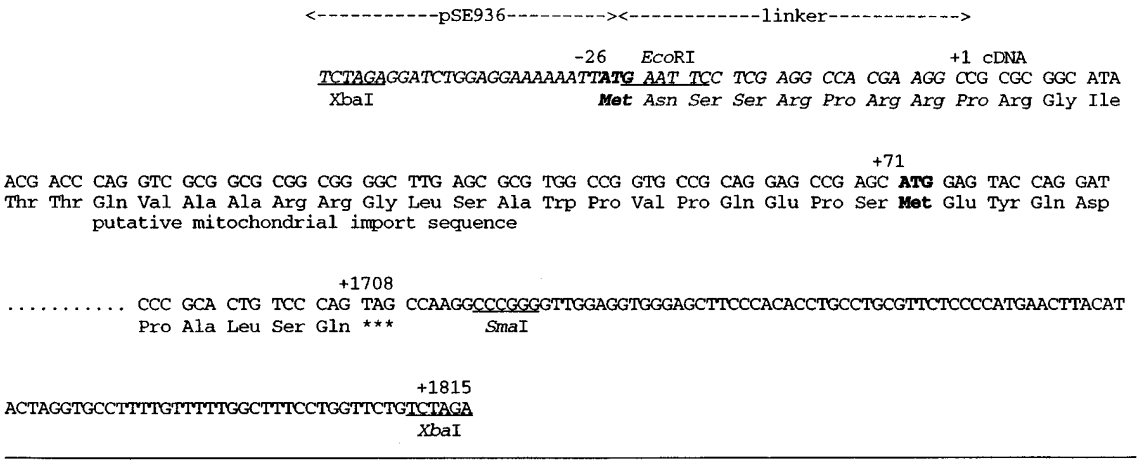

FIG. 1. **Plasmid constructs pSVK-*hFPGScyt* and pSVK-*hFPGSmit*.** The first 1815 bp of the human cDNA are represented. The 1833-bp *Eco*RI/*Xba*I fragment from pSE936-3 (21) was directionally subcloned into similarly treated pSVK-3 to create pSVK-*hFPGScyt*. In this construct, the first ATG (in *bold* at +71) originates from within the human cDNA and bypasses the partial mitochondrial leader sequence encoded by nucleotides +1 to +70 of the cDNA. The 1860-bp *Xba*I fragment from pSE936-3 was subcloned into similarly treated pSVK-3 to create pSVK-*hFPGSmit*. The orientation of the *inset* was checked by *Sma*I digestion (unique *Sma*I site in vector not shown). In this construct, the origin of the first ATG (at −26) is from the pSE936 vector and linker fusion. The subsequent 7⅔ codons are encoded by the linker. The human cDNA begins at +1 and codes for the partial mitochondrial leader sequence. Further experimental details are described under "Experimental Procedures."

sequences that could also encode the mitochondrial isozyme. Recently, Freemantle *et al.* (31) have described the organization of the 5′ region of the human FPGS gene and have shown that the isolated cDNA lacks the 5′ region of the mitochondrial leader sequence and that cytosolic and mitochondrial isozymes could be generated by alternate transcription start sites.

In this report, we describe the expression of human cytosolic FPGS in *E. coli* and the purification and properties of the enzyme. As a prelude to studying factors regulating the expression of FPGS in mammalian tissues and the role of the different isozymes, we have isolated and characterized the human gene. In this report, we describe the organization of the complete human FPGS gene, its localization, and splicing variants, and present a rationale for the role and need for mitochondrial and cytosolic FPGS isozymes.

EXPERIMENTAL PROCEDURES

Materials—L-[U-^{14}C]Glutamate (specific activity 270 mCi/mmol) was obtained from Amersham, (6*S*)-5-formyltetrahydro[^{3}H]folate (folinate, 5-formyl-H_4PteGlu, specific activity 20 Ci/mmol) was obtained from Moravek, and [α-^{35}S]dATP (1000 Ci/mmol), [α-^{32}P]dCTP (3000 Ci/mmol), and [γ-^{32}P]ATP (6000 Ci/mmol) were obtained from DuPont NEN. Folic acid, methotrexate, and aminopterin were obtained from Sigma and (6*ambo*)-H_4PteGlu from Fluka. Reduced and oxidized folylpolyglutamates were synthesized as described previously (17, 32) or were obtained from Schirck Laboratories. Concentrations of folate derivatives were calculated from their absorption spectra (33). Deficient α-minimal essential medium lacking purines, glycine, thymidine, and folate was obtained from JRH BioSciences. DNA restriction and modifying enzymes and RNase A were obtained from Boehringer Mannheim, Promega, or New England Biolabs. AmpliTaq DNA polymerase was from Perkin-Elmer. Nitrocellulose was obtained from Schleicher and Schuell. Oligonucleotides were synthesized at the UC Berkeley Microchemical Facility. *E. coli* strain JM109(λDE3) was obtained from Promega. All other materials were obtained from commercial vendors.

Plasmid Construction for Expression of Human FPGS in E. coli—An *Nde*I restriction site was introduced at the first ATG in the human FPGS cDNA (pTZ18U-25 (30), accession number M98045) by mutagenesis using the procedure of Nakamaye and Ecksteine (34). The antisense primer 5′-CTGGTACTCCATATGCTCGGCTCCTG-3′ introduced an AT dinucleotide between bases 70 and 71 of the cDNA (Fig. 1). A 2117-bp *Nde*I-*Bam*HI fragment from pTZ18U-25-*Nde*I, containing the entire open reading frame of human FPGS, was ligated into similarly treated pET3A (Novagen). Plasmid pET3A has the bacteriophage T7 gene 10 promoter and transcriptional termination sequences. This new construct, pET3A-25 was transformed into *E. coli* JM109(λDE3) for the production of unfused human FPGS.

To improve expression in *E. coli*, 12 of the first 27 codons in the cDNA ORF were modified for optimal codon usage (35). The complementary oligonucleotides 5′-TATGGAATACCAGGACGCTGTTCGCATGCTGA-ACACCCTGCAGACCAACGCTGGCTACCTGGAACAGGTTAAACGCC-AGCGCG-3′ and 3′-ACCTTATGGTCCTGCGACAAGCGTACGACTTT-GTGGGACGTCTGGTTGCGACCGATGGACCTTGTCCAATTTGCGGC-GCGCCACTG-5′ were lyophilized, resuspended in 90% formamide, and purified by PAGE. Each primer was phosphorylated using T4 polynucleotide kinase, and the primers were annealed to each other by heating at 100 °C for 3 min followed by slow cooling to room temperature in 60 mM Tris buffer, pH 7.5, containing 6 mM $MgCl_2$, 10 mM dithiothreitol, and 400 μM ATP. Five micrograms of pET3A-25 was digested with *Nde*I and *Bst*EII and purified by agarose gel electrophoresis. Approximately 0.5 μg of this DNA was ligated with 10 μl of the annealed oligonucleotides in a total volume of 20 μl. Half of this reaction mixture was transformed into JM109(DE3) and 4 transformants, whose inserts lacked an *Nae*I restriction site, were subjected to double-stranded sequencing (Sequenase, version 2) to verify the new primary structure at the 5′ end of the cDNA (pET3A-25*).

Plasmid Construction for Expression of Human FPGS in CHO Cells—Two different constructs of the human FPGS cDNA in plasmid pSVK-3 (Pharmacia Biotech Inc.) were prepared as shown in Fig. 1. Plasmid pSVK-*hFPGScyt* was constructed by ligating the 2130-bp *Eco*RI fragment from pSE936–25 (30) into similarly treated pSVK-3. In some constructs, most of the 3′-untranslated region was deleted by digesting the plasmid with *Xba*I and religating (Fig. 1). Plasmid pSVK-*hFPGSmit* was made by ligating an 1860-bp *Xba*I fragment from pSE936-3 (30), which contained the human FPGS cDNA with the longest 5′ region, into similarly treated pSVK-3. This construct adds an in-frame ATG upstream of the cDNA sequence (Fig. 1). Plasmids were purified by cesium chloride density centrifugation. These plasmid constructs insert the human FPGS cDNA downstream from an SV40 early promoter and transcriptional start region and upstream from a small tumor antigen splice site and a poly(A) signal. The inserts provide the first ATG for translational initiation.

Transfection of CHO Cells and Subcellular Fractionation—CHO AUXB1 cells were cultured in deficient α-minimal essential medium containing 10% dialyzed fetal calf serum, glycine, thymidine, and hypoxanthine as described previously (29, 36). Cells were transfected with pSVK-*hFPGScyt* or pSVK-*hFPGSmit* (10 μg), and transfectants were selected by their ability to grow in medium containing folate and glycine but lacking thymidine and purines, and individual transfectants were cloned as described previously (36). The ability of the cloned transfectants to grow in the absence of glycine was assessed as described previously (28). Procedures for the subcellular fractionation of CHO

TABLE I
Primers used in 5′-RACE and primer extension studies

Primer	Sequence (5′ to 3′)	Direction	Gene location[a]	Exon
2J	GCGGATCCGCTCCCGAA	Antisense		5
L5	CATAGCTTCGGAGGATACATT	Antisense	1734 to 1755	4
L4	CCAGTGACGTGGATGATGTT	Antisense	1596 to 1615	3
L2	ATGGCTTCCAACTGTGTCTGA	Antisense	1438 to 1458	2
L5A1	GTACTCCATGCTCGGCTCCT	Antisense	116 to 135	1/1B
L5C1	CCTCTTTCCCCACTCTCACT	Antisense	461 to 480	1C
L5B1	AGCCCCCTTGCGCCACTAAA	Antisense	−180 to −161	1A
L5B2	CCACTAAAAAGCCGCCCCTTC	Antisense	−192 to −172	1A
L51	AGACGCCGCTGCCAGGAATAGA	Antisense	33 to 54	1
L52	GGCTCCGCGCCCGCGACATAGT	Antisense	−3 to 19	1
PR1	GCGGCACCGGCCACGCGCTCAAGCC	Antisense	91 to 115	1
PR2	ACCTGCTCCAGGTAGCCGGCATTGG	Antisense	1392 to 1416	2
T-1	CAGACTCCGGGATCACTGGAACTC	Antisense	430 to 453	1C
U2	CAGGAGCCGAGCATGGAGTA	Sense	116 to 134	1

[a] Numbered from the first base (+1) of the ATG codon of the mitochondrial leader sequence.

cells, the measurement of subcellular distributions of FPGS and marker enzyme activities, and the measurement of intracellular folate accumulation and folylpolyglutamate chain length distributions have been previously described (23, 29, 36).

Folylpolyglutamate Synthetase Assay—FPGS activity was monitored during enzyme purification procedures and in CHO cell extracts by the incorporation of [^{14}C]glutamate (250 μM) into folylpolyglutamate products using (6*ambo*)-H_4PteGlu as the substrate, as described previously (23). Assays were normally conducted for 1 h at 37 °C. In kinetic studies with pure enzyme, the assay mixtures contained various concentrations of the substrate under investigation and fixed concentrations of (6*ambo*)-H_4PteGlu (40 μM), ATP (1 mM), and L-glutamate (2 mM), as appropriate. The amount of FPGS was adjusted to ensure that less than 10% of the limiting substrate was converted to product at the lowest substrate concentration used.

Enzyme Purification—All buffer solutions were adjusted to the indicated pH at room temperature. Extracts were maintained at 0 °C, and all other procedures were performed at 0–4 °C.

JM109(λDE3) harboring plasmid pET3A-25 was grown to midlog phase in Luria media containing 50 μg/ml ampicillin and 0.5% glycerol (15 liters). The cells were collected by centrifugation at 650 × *g* for 10 min. The cell paste (114 g) was resuspended in 100 mM Tris-HCl buffer, pH 7.5 (300 ml), containing 2.5 mM EDTA, 50 mM mercaptoethanol, and 1 mM phenylmethylsulfonyl fluoride, and the cells were disrupted using a Branson sonicator (power setting 8 using a 50% duty cycle for 3 × 8 min). The extract was then centrifuged at 12,000 × *g* for 45 min, and the supernatant was decanted through a double layer of cheesecloth to give the crude extract (fraction 1).

The crude extract (325 ml) was applied to a hydroxylapatite (Bio-Rad) column (7.5 × 7.5 cm) that had been equilibrated with 100 mM Tris-HCl buffer, pH 7.5, containing 2.5 mM EDTA and 50 mM mercaptoethanol (Buffer A). The column was washed with the equilibration buffer (1000 ml) and eluted with a linear gradient (1500 ml) of potassium phosphate buffer, pH 7.5 (0–150 mM) in the same buffer. Fractions containing FPGS activity were pooled (fraction 2).

Fraction 2 enzyme (450 ml) was applied to an Affi-Gel Blue (Bio-Rad) column (20 × 1.5 cm) that had been equilibrated with Buffer A. The column was washed with Buffer A containing 150 mM KCl (300 ml) and a linear gradient (500 ml) of KCl (0.15–1 M) in the same buffer was used to elute the enzyme. Fractions containing FPGS activity were pooled (fraction 3).

Fraction 3 enzyme (224 ml) was applied to a phenyl-agarose (BRL) column (20 × 1 cm) that had been equilibrated with 100 mM Tris-HCl buffer, pH 7.5, containing 1 mM EDTA, 600 mM KCl, and 50 mM mercaptoethanol. The column was washed with the equilibration buffer (180 ml), and enzyme activity was eluted with 100 mM Tris-HCl buffer, pH 8.2, containing 1 mM EDTA, 50 mM mercaptoethanol, and 20% (v/v) ethylene glycol. Active fractions were pooled and dialyzed against 100 mM Tris-HCl buffer, pH 7.5, containing 0.5 mM EDTA and 50 mM mercaptoethanol (2 × 12 liter).

The dialyzed enzyme (fraction 4; 213 ml) was applied to a heparin-agarose (BRL) column (20 × 1 cm) equilibrated with 100 mM Tris-HCl buffer, pH 7.5, containing 1 mM EDTA and 50 mM mercaptoethanol. The column was washed with 150 ml of equilibration buffer and eluted with a linear gradient (300 ml) of KCl (0–500 mM) in the same buffer. Fractions containing FPGS activity were pooled and dialyzed against 2 × 4 liters of 50 mM Tris-HCl buffer, pH 8.4, containing 0.5 mM EDTA, 50 mM mercaptoethanol, and 10% ethylene glycol (fraction 5).

Fraction 5 enzyme (20 ml) was applied to a DE52 (Whatman) column (20 × 1 cm) equilibrated with 50 mM Tris-HCl buffer, pH 8.4, containing 0.5 mM EDTA, 10% (v/v) ethylene glycol, and 50 mM mercaptoethanol. The column was washed with equilibration buffer (150 ml) and eluted with a linear gradient (300 ml) of KCl (0–500 mM) in the same buffer. Fractions containing FPGS activity were pooled giving fraction 6 (11 ml).

Protein purity was analyzed by polyacrylamide electrophoresis in sodium dodecyl sulfate (SDS-PAGE) using a 4% stacking gel and a 12.5% separating gel (37). Protein bands were visualized by a silver staining procedure. Total protein was determined on pooled fractions by a modified Lowry procedure (38) using bovine serum albumin as the standard.

Isolation of Genomic Clones—A human lung fibroblast cell line W138 genomic library in the Lambda FIX II vector (obtained from Stratagene), was screened (10^6 plaques) with [^{32}P]dCTP-labeled primers generated using the Random Primed DNA Labeling Kit (Boehringer Mannheim) and a human FPGS cDNA (30) as the template. Following plaque purification, 3 positive clones were obtained. The phage DNA from these clones were purified and characterized by restriction mapping and Southern hybridization using an *Eco*RI-*Bam*HI 375-bp fragment of the 5′ region and a 355-bp *Xba*I-*Eco*RI 3′ region fragment of the human cDNA. *Xba*I fragments of the genomic clones, and the entire clones bordered by *Not*I sites, were subcloned into pBKS$^+$ (Stratagene) for further analysis.

DNA Sequencing and Intron Size Determination—DNA was sequenced by the method of Sanger *et al.* (39) using Sequenase version 2.0 (United States Biochemical Corp.). Exon-intron junctions were determined by direct sequencing across the junctions using oligonucleotide primers based on the cDNA sequence. Intron sizes were determined by sequencing through the region or by PCR using flanking primers. The FPGS gene sequence, with the exception of parts of 4 introns, was determined and verified by sequencing both DNA strands.

5′-RACE Analysis of FPGS cDNA Ends—Total RNA was isolated from human HepG2 cells with guanidinium thiocyanate followed by cesium chloride centrifugation (40) or using Tri Reagent (Molecular Research Center) according to the manufacturer's instructions. Primers used in this study are shown in Table I. cDNA corresponding to the 5′ end of HepG2 mRNA was synthesized and amplified (41) using a Marathon cDNA Amplification Kit (Clontech) and following kit instructions with the following modifications. Antisense primer L5 was used, instead of oligo(dT), to synthesize the first strand cDNA. After the anchor (Clontech) ligation, the cDNA was first PCR-amplified with the anchor primer (Clontech) and an inner gene specific antisense primer (L4). One μl of the resulting PCR product was reamplified with the anchor primer and another nested gene-specific antisense primer (L2). PCR was carried out at 94 °C for 5 min, and for 35 cycles at 94 °C for 1 min, 57 °C for 0.5 min, and 72 °C for 0.5 min, with a final extension at 72 °C for 5 min. PCR products were analyzed on a 2.0% agarose gel with ϕX174/*Hae*III-digested DNA (Promega) as size standards. Bands of interest were excised, and the DNA was purified using a QIAEX gel extraction kit (Qiagen) and cloned into a pGEM-T vector (Promega).

In other studies, a 5′AmpliFINDER RACE kit (Clontech) was used. Primer L5 was used to initiate first strand cDNA synthesis. Anchor primer from the kit and primer L4 were used to amplify the cDNA with the same cycling profile as described above. The resulting PCR products were ethanol-precipitated and directly cloned into the pGEM-T vector. Ampicillin-resistant colonies were screened with a probe generated by

amplification of the human FPGS cDNA using primers U2 and L2 labeled with [α-^{32}P]dCTP. Positive colonies were further analyzed by hybridization with ^{32}P-labeled antisense oligonucleotides L5A1, L5B1, or L5C1 which are specific for human FPGS exons 1 plus 1B, 1A, and 1C, respectively (Fig. 3). Positive colonies recognized by each probe were further characterized by restriction enzyme digestion, and clones with larger size inserts were sequenced.

The 5′ transcriptional ends of fetal liver FPGS mRNA was also assessed by RACE using human fetal liver 5′-RACE-Ready cDNA (Clontech). Antisense primers 2J, L5, L4, and L2 were used successively to PCR-amplify the 5′ transcriptional end.

Primer Extension Analysis—Oligonucleotide primers complementary to the 5′ end of the human FPGS cDNA sense strand sequence (PR1, PR2, L5A1, L51, L52, L5B1, L5B2, T-1, and L5C1) were labeled at the 5′ end with [γ-^{32}P]ATP (6000 Ci/mmol) using T4 polynucleotide kinase and purified on Sephadex G-25 Quick Spin columns (Boehringer Mannheim). Labeled probe (2×10^6 cpm) was hybridized to 50 μg of HepG2 total RNA in 1 × reverse transcriptase buffer (BRL) at 75 °C for 10 min and 30 °C for 60 min. Primer and RNA were coprecipitated with ethanol, and the pellet was resuspended in 50 μl of reaction solution containing 1 × reverse transcriptase buffer, avian myeloblastosis virus or Moloney murine leukemia virus reverse transcriptase (40 units), 10 mM dithiothreitol, and 0.2 mM concentration of each dNTP with or without 5% dimethyl sulfoxide (26, 27). The primer was first extended at 42 °C for 30 min. Additional reverse transcriptase (40 units) was then added, and the mixture was incubated at 52 °C for 30 min. The RNA was digested with 20 μg of DNase-free RNase A at 37 °C for 30 min. Ammonium acetate (final concentration 2 M) was added, and the samples were extracted with 50 μl of phenol/chloroform/isoamyl alcohol (25:24:1). Following ethanol precipitation, the pellet was washed with 70% ethanol, air-dried, and resuspended in 5 μl of DNA sequencing stop solution (U. S. Biochemical Corp.). The DNA products were analyzed on a 6% polyacrylamide-urea sequencing gel and compared to DNA sequence reaction products obtained with human FPGS genomic DNA (clone 9-4) using the same ^{32}P-labeled oligonucleotide as the sequencing primers (42).

Construction of the FPGS Gene—The 4.5-kb *Xba*I fragment of genomic clone 9 (Fig. 3) corresponding to the 5′ end of the cDNA, was subcloned into pBKS$^+$ (Stratagene). The 2.7-kb fragment of the 5′-region of this insert, from the *Not*I site of the vector to the *Sac*II site of the insert, was ligated to the *Sac*II-*Not*I 3′-fragment (17 kb) of genomic clone 5 to construct the entire FPGS gene. The FPGS gene was subcloned into the *Not*I site of pBKS$^+$.

Chromosomal Localization of FPGS Gene—A FPGS cDNA probe were labeled with biotin-11-dUTP by nick translation and hybridized to metaphase chromosomes prepared from normal male peripheral blood by the bromodeoxyuridine synchronization method (43). *In situ* hybridization, amplification, and detection with avidin-conjugated fluorescein isothiocyanate was as described previously (44). Two amplifications were carried out using biotinylated anti-avidin. Metaphase chromosomes were counterstained with Chromomycin A_3 followed by Distamycin A, by a modification of the procedure of Magenis *et al.* (45), to generate clear reverse bands.

RESULTS

Expression and Purification of Human FPGS—A variety of constructs of the human FPGS cDNA were cloned into various *E. coli* expression vectors in an attempt to overexpress the human protein. In the studies described in this report, an *Nde*I site was introduced at the first ATG in the cDNA, and an *Nde*I-*Bam*HI fragment encoding the entire open reading frame was subcloned into pET3A and expressed in JM109(λDE3). Plasmid pET3A has a bacteriophage T7 gene 10 promoter and ribosome binding site and transcriptional termination sequences. JM109(λDE3) is a λ lysogen that has the bacteriophage T7 RNA polymerase linked to the *lacUV5* promoter. FPGS was expressed at about 0.5% of the total soluble protein in *E. coli* crude extracts. Use of a modified cDNA in which the first 27 codons were optimized for preferred codon usage in *E. coli* (pET3A-25*) increased expression up to a maximum of about 1.2% of soluble protein. Although a T7 promoter vector was used in these studies, IPTG induction of T7 polymerase reduced or eliminated expression of active human FPGS. High level expression of human FPGS in *E. coli* appeared to be toxic and loss of plasmid, plasmid rearrangement, and deletions were observed after IPTG induction.[3] However, expression of active enzyme above 1% soluble protein could not be achieved and consistent expression at this level required the use of freshly transformed cells. Expression of the cDNA in *trp-lac* vectors, such as pTRC99 (Pharmacia), resulted in similar or lower levels of expression. We have now modified the culture conditions to allow IPTG induction of the T7 vectors and obtain over 50 mg of human FPGS per liter of culture medium, most of which is in inclusion bodies,[3] and have also overexpressed the enzyme using modified *Baculovirus* expression vectors,[3] but report here the purification of enzyme expressed in *E. coli* as the kinetic data described in this report were obtained with FPGS purified from *E. coli*.

TABLE II
Purification of human folylpolyglutamate synthetase
Experimental conditions are described under "Experimental Procedures."

Fraction	Volume	Activity	Protein	Specific activity	Purification	Yield
	ml	*units/ml*[a]	*mg/ml*	*units/mg*	*-fold*	*%*
Crude	325	2,560	30.6	82.5	1	100
Hydroxylapatite	450	1,540	8.90	173	2	83
Affi-Gel blue	224	2,660	1.39	1,910	23	71
Phenyl-agarose	213	1,180	0.156	7,580	92	30
Heparin-agarose	20	7,920	0.671	11,800	143	19
DEAE-cellulose	11	12,900	0.701	18,400	223	17

[a] Nanomoles of glutamate incorporated into folate product per h. Assay mixtures contained 5 mM ATP, 40 μM H_4PteGlu, and 250 μM glutamate.

Table II shows a typical purification obtained with the pET3A-25 construct, which expressed human FPGS at about 0.5% of soluble protein in *E. coli*. Chromatography on hydroxylapatite resulted in a 2-fold purification of the enzyme and also removed nucleic acids from the crude preparation. FPGS was purified another 10-fold by chromatography on Affi-Gel Blue. Most of the applied protein did not bind to this column, and FPGS was eluted with a KCl gradient. The final purification was achieved by chromatography on phenyl-agarose, heparin-agarose, and DEAE-cellulose, respectively. Purified enzyme was judged to be homogeneous by SDS-PAGE (Fig. 2). A single band was observed with an apparent M_r of about 61,000, which is consistent with the deduced amino acid sequence (545 amino acids, M_r = 60,128). This purification procedure has been repeated multiple times, starting with FPGS enrichments in crude extracts ranging from 0.1 to 1.2% and has resulted in homogeneous enzyme each time.

General Properties—Activity required a monovalent cation. K^+ (20 mM) was most effective, followed by NH_4^+ and Rb^+, while Na^+, Li^+, and Cs^+ were ineffective. There was an absolute requirement for a reducing agent and the enzyme displayed a high pH optimum (pH 9.6). The K_m values for L-glutamate and MgATP were 201 μM and 200 μM. Activity was stimulated by sodium bicarbonate (10 mM), which caused a 4-fold decrease in the K_m for MgATP (54 μM) and a 20% increase in V_{max}.

Folate and Analog Substrate Specificity—The kinetic constants for a variety of folates and folate analogs are shown in Table III. Enzyme concentrations and incubation times were adjusted to minimize the addition of more than one glutamate moiety to the folyl or antifolyl substrate and to ensure initial rate conditions. Marked substrate inhibition was observed with many of the substrates that displayed a high affinity for the enzyme. V_{max}/K_m values shown in Table III are the on rates for

[3] A. Brenner, I. Atkinson, T. Garrow, and B. Shane, unpublished data.

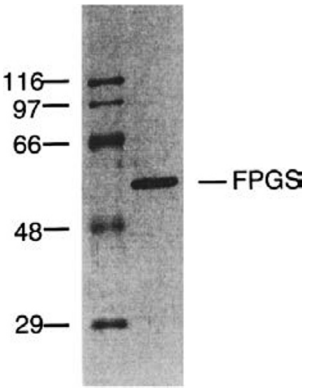

FIG. 2. **Sodium dodecyl sulfate gel electrophoresis of DE52 purified human folylpolyglutamate synthetase.** Experimental details are described under "Experimental Procedures." M_r ($\times 10^{-3}$) of protein standards are shown.

TABLE III
Kinetic constants of folates and folate analogs for human folylpolyglutamate synthetase

Enzyme activity was assayed as described under "Experimental Procedures." Assay mixtures contained various concentrations of the indicated substrate, ATP (1 mM) and L-glutamate (2 mM).

Substrate	K_m	V_{max} (rel)[a]	V_{max}/K_m (rel)[a]
	μM		
PteGlu	59	66	4.9
$PteGlu_2$	16	83	23
$PteGlu_3$	20	56	12
$PteGlu_4$	12	34	13
$PteGlu_5$	64	1	0.07
$H_2PteGlu$	0.81	85	460
$H_2PteGlu_2$	47	96	8.3
(6*ambo*)-H_4PteGlu	1.6	83	230
(6*S*)-H_4PteGlu	4.4	100	100
(6*S*)-$H_4PteGlu_2$	3.3	82	109
(6*S*)-$H_4PteGlu_3$	1.4	19	60
(6*S*)-$H_4PteGlu_4$	1.6	10	29
(6*S*)-$H_4PteGlu_5$	1.4	0.6	1.9
(6*R*)-10-formyl-H_4PteGlu	3.7	23	27
(6*R*)-10-formyl-$H_4PteGlu_2$	2.7	30	49
(6*S*)-5-formyl-H_4PteGlu	105	85	3.6
(6*S*)-5-formyl-$H_4PteGlu_2$	13	68	23
(6*S*)-5-methyl-H_4PteGlu	48	83	7.6
Aminopterin	4.4	118	118
Methotrexate (Glu-1)	71	94	5.8
Methotrexate (Glu-2)	50	16	1.4
Methotrexate (Glu-3)	148	3.2	0.086
5-Deazaacyclotetrahydrofolate	5.3	107	89
2-Methyl-5,8-dideazaisofolate	2.8	101	158

[a] Values are relative to that obtained with (6*S*)-H_4PteGlu (equals 100). V_{max} 59 μmol/h/mg, k_{cat} 0.99 s^{-1}; k_{cat}/K_m 0.225 $\times 10^{-6}$ $M^{-1} \cdot s^{-1}$.

the folate substrates (32) and are a comparison of the reaction rates with the different substrates at concentrations considerably below the K_m for the substrate.

Unsubstituted reduced folates are the preferred substrates for FPGS. H_2PteGlu was the most effective substrate with the lowest K_m value. H_4PteGlu was almost as effective and the 6*ambo* mixture displayed a lower K_m than the 6*S* isomer, suggesting a preference for the unnatural 6*R* diastereoisomer of H_4PteGlu. PteGlu was a fairly poor substrate due to its high K_m value. 10-Formyl-H_4PteGlu displayed a low K_m value but 10-substitution caused a drop in V_{max}. Substitution at the 5-position of folates resulted in a large increase in K_m values. (6*S*)-5-Formyl-H_4PteGlu was a poor substrate although previous studies (21, 46), using partially purified enzyme, had suggested that it was an effective substrate for human FPGS.

$H_4PteGlu_n$ derivatives were the most effective polyglutamate substrates. Extension of the glutamate chain decreased K_m values but also caused a decrease in V_{max} with chain lengths beyond the diglutamate. With most folates, the diglutamate derivative was a more effective substrate than the monoglutamate due primarily to a decrease in K_m value, a decrease that was most pronounced for 5-formyl-$H_4PteGlu_2$. However, $H_2PteGlu_2$ was a poor substrate due to a large increase in K_m value.

The folate analogs 5-deazaacyclotetrahydrofolate, an inhibitor of transformylases involved in purine synthesis (47), and 2-methyl-5,8-dideazaisofolate, an inhibitor of thymidylate synthase (48), were very effective substrates of human FPGS (Table III). Polyglutamates of these compounds are much more potent inhibitors of their target enzymes than the monoglutamate derivatives. Aminopterin was a very good substrate for the enzyme with the 4-amino substitution decreasing K_m values by about 15-fold compared to PteGlu, while methotrexate was less effective due to an increased K_m. Extension of the glutamate chain of methotrexate decreased substrate effectiveness primarily due to a drop in V_{max} values, and the triglutamate derivative was an extremely poor substrate.

Organization of the Human FPGS Gene—A Lambda Fix II library was screened as described under "Experimental Procedures," and three clones were obtained, two of which were found to be identical by restriction enzyme mapping and Southern analysis. One clone (C5, Fig. 3) lacked the 5′ region of the FPGS cDNA. The second clone (C9) overlapped clone C5 and contained an additional 10 kb of 5′ region, but lacked the region corresponding to the 3′ end of the cDNA (Fig. 3). *Xba*I fragments were subcloned into $pBKS^+$ and intron/exon junctions were sequenced using primers to various regions of the cDNA (Table IV). All intron/exon splice junctions follow the GT-AG rule (49). The *Xba*I fragments were sequenced in both directions, and the sequence of the gene, with the exception of parts of four intronic regions, was obtained. The human FPGS gene spans about 12 kb and consists of 15 exons and 14 introns (Fig. 3, Table IV). Two sequences, corresponding to the 5′ region of the gene to the 5′ end of intron 4 (accession number U24252) and the 3′ region of intron 4 to the 5′ region of intron 11 (accession number U24253), have been deposited in the GenBank™ data base. The coding sequence of the gene was in agreement with our previously published cDNA sequence (30).

Freemantle *et al.* (31, U14939) recently reported the sequence of the 5′ genomic region of human FPGS and demonstrated that the originally isolated cDNA lacked 56 bp of coding region, including an upstream ATG which was preceded by Sp1 sites, and that multiple transcription start sites in this region can result in cytosolic and mitochondrial forms of FPGS. Our sequence is in agreement except for some differences in intron 1. These include additional C residues at +160, +167, and +337 (numbered from the ATG codon of the mitochondrial leader sequence) and GC transpositions at +215/216 and +350/351, and G instead of A at +982. An additional 170-bp *Bam*HI fragment (residues +396 to +566) is not present in U14939. This additional region is not a cloning artifact as it contains part of an alternate exon 1 identified by 5′-RACE of human mRNA (see below). The C5 clone starts at a *Sau*3aI (*Bam*HI) site at residue +566.

Differential Splicing and Alternate Exon Usage—5′-RACE analysis of RNA from HepG2 cells using nested antisense primers to exon 3 and exon 2 regions resulted in multiple species of different length. The products were cloned and sequenced, and the longest form obtained for the exon 1 region started at −43 (Fig. 3), which agrees with the transcription start site suggested by Freemantle *et al.* (22). The sequences of the RACE products were identical with the gene and cDNA sequences except HepG2 mRNA had a G at position +64 instead of an A. Additional alternate exon 1 regions (Fig. 3, Table IV), that were

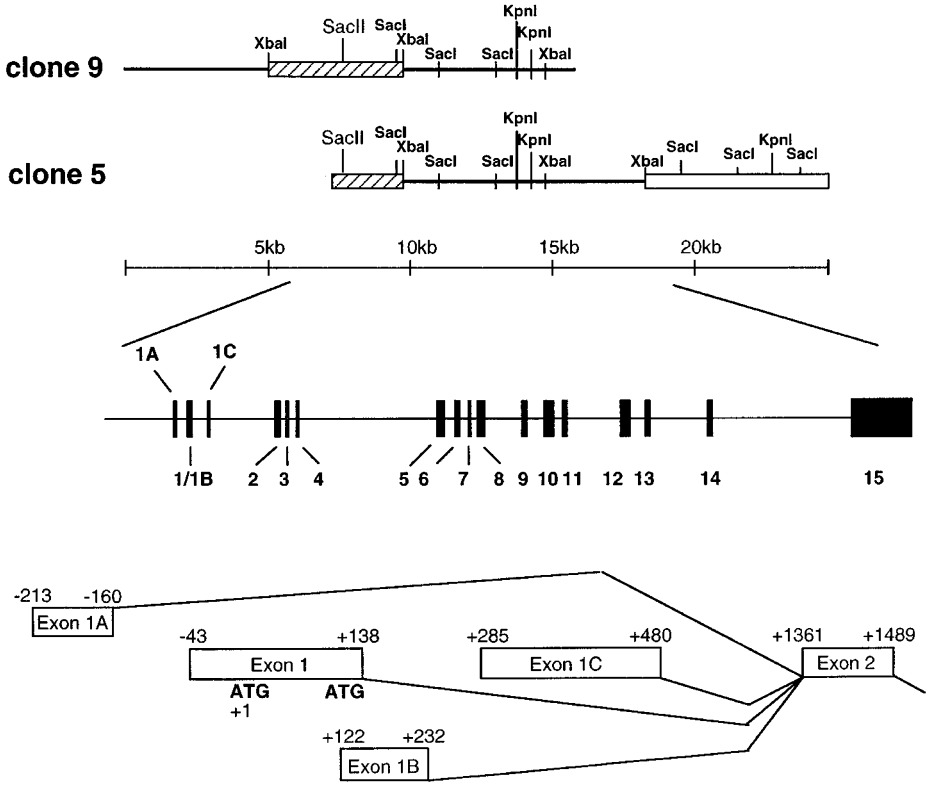

FIG. 3. **Organization of the human FPGS gene.** The positions of the overlapping genomic clones (C5 and C9), exon organization, and alternate exon 1 regions are shown. The *hatched* and *open blocks* on the clones indicate fragments that hybridized to the 5′ and 3′ regions of the human FPGS cDNA. In the expanded diagram of the alternate exon 1 variants, nucleotides are numbered from the ATG encoding the start methionine of the mitochondrial leader sequence in the exon 1 variant.

distinct from exon 1 and from any previously isolated cDNA sequence, were also detected and, in each case, the 5′ sequence of the product obtained (Fig. 3) was identical with a region of the gene and the junction of the 5′ region with exon 2 matched consensus splice junctions (Table IV). Several of these species were due to alternate splicing in intron 1 at positions +233 (exon 1B) and +481 (exon 1C) (Fig. 3).

The 5′ end of the longest 5′-RACE products obtained encoding exon 1B (+122, Fig. 3) extended into exon 1 suggesting that some of the exon 1B forms arose by splicing of an exon 1 transcript at +233 rather than at +139. However, some of the exon 1B RACE products did not extend into exon 1 and may represent separate transcripts. It may be that transcripts that start late in the exon 1 region do not get spliced at position +139. If exon 1B arose by extension of exon 1, and the translation product started at the exon 1 cytosolic ATG (+127), the resulting mRNA would be out of frame with codons in the exon 2 region. However, translation of this variant could start at an alternate ATG at the beginning of exon 2 at position +1373 (Table IV) equivalent to amino acid residue 9 in the cytosolic form of the protein, although this unlikely (see below).

None of the 5′-RACE products obtained for exon 1C, which varied in size, extended as far upstream as exon 1B and the longest form obtained started at +285 (Fig. 3). Primer extension analysis, using primers specific to the exon 1C region (L5C1 and T-1), indicated multiple species with major starts at +330 and +373 and failed to indicate longer species that extended into exon 1B. This region is preceded by potential Sp1 binding sites (+264 to +269 and +356 to +361) and an E2A site (+331 to +338). Exon 1C lacks an in-frame start ATG although translation of this variant could also start at the alternate ATG at the beginning of exon 2 (Table IV).

An additional splice variant encoded an upstream exon (exon 1A, Table IV, Fig. 3) spliced to exon 2. Again, 5′-RACE products of various lengths, with different 5′ ends, were obtained, the longest starting at position −213 (Fig. 3). All 5′ sequences terminated at residue −160, a consensus splice junction (Table IV). Primer extension analysis using exon 1A specific probes (L5B1, L5B2, Table I) indicated several starts equivalent to positions −288 and −226, assuming no additional splicing in this region (see below).

Exon 1A precedes all the Sp1 binding sites but is preceded by a computer identified Adh1 promoter site (−256 to −251), imperfect TATA sites (−240 to −235, −434 to −429), an APRT promoter site (−469 to −462), and an E-box insulin-responsive element (−456 to −449) (50). Exon 1A does not contain a start ATG, although translation of the longest 5′-RACE product for exon 1A would give a peptide (SPGWCTRKGRLFSGARGL) with characteristics of a mitochondrial leader sequence (51). Thr-6 in this sequence is encoded by an ACG codon (−198), which is a possible alternate translation start codon. An examination of the sequence upstream of exon 1A also indicates a potential splice site 9 bases upstream of the longest 5′-RACE product (−223, Table IV). The upstream region also contains various ATG codons, one of which at position −565 is followed by a potential splice junction (−556). Translation from this ATG, which removal of a potential intron (−556 to −223) would generate the peptide MLAACPSPGWCTRKGRLFSGARGL. The longest primer extension products obtained with exon 1A specific primers were of sufficient size to start prior to this potential ATG. The region upstream of this potential start ATG contains a consensus CTF/NF-1 site (−803 to −795) (52).

Abundance and Role of Splice Variants—In initial 5′-RACE studies, the major transcripts obtained from HepG2 mRNA encoded the exon 1A variant while exon 1 and exon 1C variants predominated in the human fetal liver library. Further 5′-RACE studies indicated all four variants in HepG2 mRNA. Differences in distributions between experiments probably reflected PCR variations. To further investigate the distribution of these variants and to check for other variants, 5′-RACE was performed on HepG2 mRNA using an exon 4 primer (L5) for first strand cDNA synthesis and PCR was performed with an exon 3 antisense primer (L4), but no nested primer was used. The products obtained were cloned into a pGEM-T vector and

TABLE IV
Exon-intron splice junctions of the human FPGS gene

Exon	Position[a]	Intron	5′-Splice donor	Intron size	3′-Splice acceptor
				bp	
1	1–138	1	G GAG TAC CAG gtatcaggcg Glu Tyr Gln	1222	tggttgccag GAT GCC GTG CGC ATG Asp Ala Val Arg Met
1A		1A	CGCAAGGGGGCTG gtcggtggta	1519	tggttgccag GAT GCC GTG CGC ATG Asp Ala Val Arg MET
1B		1B	GCTAGCCGAGAGG gtatcgggag	1128	tggttgccag GAT GCC GTG CGC ATG Asp Ala Val Arg MET
1C		1C	GTGGGGAAAGAGG gtagggaaga	880	tggttgccag GAT GCC GTG CGC ATG Asp Ala Val Arg MET
2	139–267	2	T GGG CTG CAG gtaaggtaga Gly Leu Gln	85	tctgtcccag GTG GAG GAC T Val Glu Asp
3	268–321	3	G AAG GGG AAG gtgaggggca Lys Gly Lys	83	ttctttcaag GGC TCC ACC T Gly Ser Thr
4	322–386	4	GA TTC TTT AG gtactggctt Phe Phe Ser	~2000[b]	gtcccggcag C TCT CCC CAC Ser Pro His
5	387–501	5	G GAG ACC AAG gtgccgcatg Glu Thr Lys	120	tgtgccacag GAT GGC AGC T Asp Gly Ser
6	502–579	6	C CAA GAG AAG gtgtgtgccc Gln Glu Lys	134	tctttgacag GTG GAC CTG G Val Asp Leu
7	580–641	7	AC ATC ATC AG gtgagcgcag Ile Ile Arg	101	tcccctgcag G AAG CCT GTG Lys Pro Val
8	642–744	8	C ATC TTT AAG gtgaccaggc Ile Phe Lys	545	ccctccacag CAA GGT GTC C Gln Gly Val
9	745–822	9	G CAG ATC TCA gtaagtctga Gln Ile Ser	246	tctcccctag TGT CCT CTA T Cys Pro Leu
10	823–970	10	GAC CGC CAT G gtgagtgggc Asp Arg His	94	caccccgcag GT GCT GGG GA Gly Ala Gly
11	971–1060	11	ATG CGG CTC G gtgagttaga Met Arg Leu	~800[b]	cctcccccag GG CTT CGG AA Gly Leu Arg
12	1061–1211	12	GG CCG AGC GG gtgaggggca Pro Ser Gly	196	tctgtcgcag T GGC CCC GAG Gly Pro Glu
13	1212–1287	13	CTG CTG CAG gtgaggggcc Leu Leu Gln	~850[b]	ggccttctag CCC TGC CAG T Pro Cys Gln
14	1288–1354	14	GGC AAC GCA G gtgagaggtg Gly Asn Ala	~2000[b]	tgccccacag AC CAA CAG AA Asp Gln Gln
15	1355–2214		AATTAAAGCCTTTGTTTTT		
pre-1A[c]			*ATTTCATGCTAGCA gtggggcca*	334	*tggtcccgag GCCTGCCCCTCAC*

[a] Position in cDNA starting from first base of ATG of mitochondrial leader sequence.
[b] Intron sizes were assessed by PCR. All others were determined by PCR and direct sequencing.
[c] Potential upstream exon to exon 1A. Intronic region bases −556 to −223, numbered from ATG of mitochondrial leader sequence (+1).

1000 individual colonies of *E. coli* transformants expressing this vector were screened with a labeled probe composed of the most 3′ 23 bp of exon 1 and the most 5′ 98 bp of exon 2. 548 clones gave a positive signal. The positive clones were then screened with labeled primers L5A1, L5B1, or L5C1 (Table I) which are specific for exons 1 and 1B, 1A, and 1C, respectively. The exon 1/1B probe hybridized with 287 colonies (52%). Nine of these clones were sequenced and seven encoded the exon 1 variant while 2 encoded the exon 1B variant. The exon 1A probe hybridized with 87 colonies (16%). However, only 3 of 11 clones sequenced encoded the exon 1A variant, the remainder encoded the exon 1C variant. The exon 1C probe hybridized to 82 colonies (15%), and all 6 clones sequenced encoded the exon 1C variant. Seventeen percent (92) of the clones recognized by the exon 1/exon 2 primer were negative when probed with the exon 1 variant specific primers. Five of these were sequenced. Three extended only 5 bp into exon 1, while 2 were unspliced at the 5′ end of exon 2 and contained a short region of the 3′ end of intron 1. Whether the latter represents a splicing intermediate or an additional variant was not ascertained.

Preliminary Northern analyses using exon 1 variant specific probes have been conducted to overcome the bias that might be introduced by PCR. The predominant variants in MCF7 cells were the exon 1 (2.5 kb) and exon 1C (2.4 kb) forms, which were approximately equally distributed, while exon 1 variants predominated over exon 1C variants in HepG2 cells. Transfection of CHO AUXB1 cells with a genomic construct that could only express the exon 1C variant failed to restore growth in medium lacking thymidine and purines but the transcribed message was not spliced in these cells. Expression of a cDNA construct containing exons 2 to 15, but lacking all exon 1 regions, under the control of an SV40 promoter, also failed to restore growth under these conditions. As synthesis of a functional FPGS from the exon 1B and 1C variants would require translation from an ATG at the start of exon 2 (Table IV), it is unlikely that exon 1B and 1C mRNA variants are translated to produce a functional FPGS.

Role of Mitochondrial and Cytosolic FPGS Isozymes—FPGS is located in the mitochondria and cytosol of eukaryotic cells and mitochondrial FPGS activity is required for mitochondrial 1-carbon metabolism (23, 29) and for a normal 1-carbon flux in the cytosol.[4] In our initial report on the human FPGS cDNA (30), expression of the cDNA in CHO AUXB1 cells restored cytosolic thymidylate and purine synthesis, but the cells remained auxotrophic for glycine, suggesting the absence of FPGS activity and a folate pool in the mitochondria. Freemantle *et al.* (31) showed that expression of the complete cDNA containing the entire mitochondrial leader sequence restored glycine prototrophy to these cells, suggestive of at least mitochondrial expression of FPGS.

cDNA constructs were prepared containing (*hFPGSmit*) or lacking (*hFPGScyt*) an upstream ATG codon (Fig. 1), and the constructs were expressed in AUXB1 cells. *hFPGSmit* was constructed prior to knowledge of the exact nature of the FPGS leader sequence to investigate whether the 5′-untranslated

[4] R.-F. Huang and B. Shane, unpublished data.

TABLE V

Subcellular distribution of FPGS activity and folylpolyglutamates in CHO cells

Folate-depleted cells were cultured in deficient α-minimal essential medium containing glycine, thymidine, hypoxanthine, and (6*S*)-5-formyl-H_4[^{3}H]PteGlu (10 nM) for 24 h, and subcellular fractions were prepared as described under "Experimental Procedures." Folylpolyglutamates in mitochondrial (mito) and cytoplasmic (cyto) fractions were identified as described under "Experimental Procedures." FPGS activity was measured in parallel experiments in cells cultured in identical medium lacking labeled folate. Glutamate dehydrogenase (GDH), a mitochondrial marker enzyme, was assayed as described under "Experimental Procedures."

Cell	Cell fraction	FPGS		GDH distribution	Folate distribution	Polyglutamate chain length distribution									
		Activity	Distribution			1	2	3	4	5	6	7	8	9	10
		pmol/h/mg	%	%	%	%									
CHO WTT2[a]	Cyto	155	54	26	62	3	1	1	3	24	42	20	6	1	0
	Mito	529	46	74	38	2	1	1	1	7	22	48	13	5	0
AUX-*hFPGScyt*1	Cyto	273	96	25	98	0	0	0	1	9	33	43	14	1	0
	Mito	35	4	75	2	0	0	0	0	9	32	43	16	0	0
AUX-*hFPGScyt*2	Cyto	3582	98	26	99	0	0	0	2	9	19	35	27	7	1
	Mito	79	2	74	1										
AUX-*hFPGSmit*	Cyto	44	25	27	31	0	0	3	2	31	41	18	5	0	0
	Mito	388	75	73	69	0	0	0	1	5	19	49	25	1	0

[a] Data for wild type CHO cell from Ref. 29.

region of the cDNA could encode part of a leader sequence and to develop model cells expressing human FPGS in the mitochondria. The subcellular distribution of FPGS activity and folylpolyglutamates and their chain length distribution in representative AUXB1 transfectants are shown in Table V. About 50% of the FPGS activity and 40% of cellular folate are associated with the mitochondrial fraction in wild type CHO cells. AUX-*hFPGScyt* transformants grew in the absence of thymidine and purines but not glycine and expressed a wide range of FPGS activities (*e.g. cyt1* and *cyt2*, Table V). No obvious differences were noted in the range of expression levels when the majority of the 3′-untranslated region of the cDNA was deleted (Fig. 1). Essentially all the FPGS activity and cellular folate was associated with the cytosol in these transformants (Table V). The trace amount of folate in the mitochondrial fraction was of identical glutamate chain length distribution to cytosolic folate, suggesting cytosolic contamination of this fraction. Cells expressing higher levels of FPGS contained longer folylpolyglutamate species.

Cells transfected with *hFPGSmit* constructs (AUX-*hFPGSmit*) were prototrophic for glycine, thymidine, and purines, and contained mitochondrial FPGS activity and folate (Table V). The latter half of the human FPGS mitochondrial leader sequence is sufficient to direct the protein to the mitochondria. The proportion of cellular FPGS activity recovered in the cytosolic fraction was similar to that of glutamate dehydrogenase, a mitochondrial matrix marker, suggesting that most, if not all, of the FPGS activity was mitochondrial in this transfectant. Similarly, most of the intracellular folate was mitochondrial although the proportion recovered in the cytosolic fraction was consistently slightly higher than that of glutamate dehydrogenase (Table V). AUX-*hFPGSmit* transfectants contain a small cytosolic folate pool of glutamate chain length distinct from, and shorter than, the mitochondrial folate pool, which was sufficient to allow growth in the absence of thymidine and purines.

AUXB1 cells co-transfected with the reconstructed complete FPGS gene and pCDNA3, a vector containing the neomycin gene, demonstrated similar transfection frequencies when selected for growth in the absence of glycine, purines, and thymidine or for G418 resistance. Most cells initially selected for G418 resistance were also prototrophic for glycine, purines, and thymidine.

Chromosomal Localization of FPGS—Human FPGS was previously localized to chromosome 9 by somatic cell hybridization (53). The 2.2-kb FPGS cDNA was used to localize the gene to human chromosome band 9q34.1 (Fig. 4). Three independent experiments were performed and over 100 metaphase cells were evaluated. Signals were clearly seen on two chromatids of at least one chromosome band 9q34.1 in 40% of cells and at no other sites in greater than one cell.

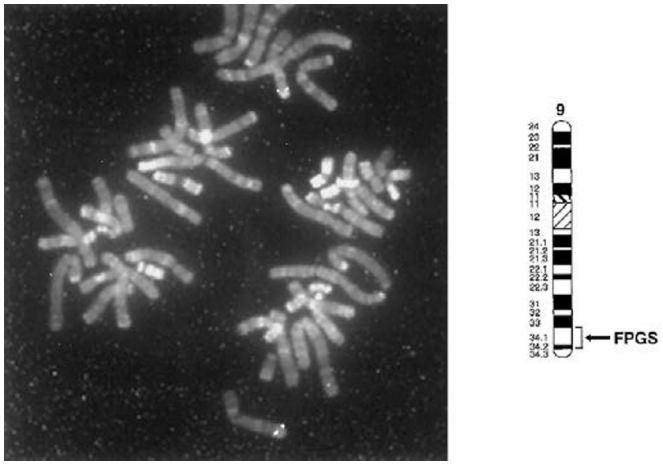

FIG. 4. **Fluorescence *in situ* localization of FPGS gene to human chromosomal region 9q34.1–34.2.** A human chromosomal preparation was hybridized with a 2.2-kb FPGS cDNA probe labeled with biotin-11-dUTP. The fluorescein isothiocyanate signals are clearly shown at the chromomycin and distamycin reverse banded chromosomal region of 9q34.1–34.2. The human chromosome 9 ideogram shows the location of the FPGS gene.

DISCUSSION

Mammalian FPGS enzymes are low abundance and unstable proteins which has hampered their purification or isolation in significant quantities. Overexpression of a FPGS cDNA has allowed the purification of the human enzyme to homogeneity and characterization of its properties. The only other mammalian FPGS enzyme to have been purified, the pig liver protein, required about a 200,000-fold purification (30). Human FPGS has similar general properties to other mammalian FPGS enzymes. The K_m for MgATP reported here is similar to values previously reported for the human enzyme (21, 54) and is about 20-fold higher than the value for the pig liver enzyme (32). Bicarbonate is a nonessential activator of human FPGS (21, 54) and increases the affinity of the nucleotide substrate.

The enzyme has a pH optimum of approximately 9.4, which is similar to the pH optima of bacterial (55–57), yeast,[5] and pig liver (17) enzymes. Other studies have reported a pH optima of about 8.3 (18, 21). This discrepancy is most likely due to the use of higher concentrations of glutamate in previous assays and

[5] A. Brenner and B. Shane, unpublished data.

possibly decreased stability of the enzyme in crude extracts at the higher pH. Studies with the pig liver enzyme showed that V_{max}/K_m values for glutamate demonstrated a steep pH profile with an optimum at pH 9.4 (17), similar to the p*K* of the amino group of glutamate, and that the enzyme functioned well at physiological pH provided higher levels of glutamate were provided. The K_m for glutamate reported here is lower than that reported for extracts from human leukemia cells (54) or liver (21) (0.4 to 1.2 mM), which presumably reflects the lower pH values of the assays used in the other studies.

The folate substrate specificity of the human enzyme was qualitatively similar to that reported for the purified pig liver enzyme (32) although some differences were noted from that reported for cruder preparations from leukemia cells (54) and human liver (21). This may reflect that the specific activity of the enzyme used in the current study was about 50,000-fold higher than that used in the previous studies (21). The maximal catalytic rate of human FPGS (1 s^{-1}) is about 40% that of the pig liver enzyme (32). The unsubstituted reduced folates, H_2PteGlu and H_4PteGlu, and aminopterin were the preferred substrates for the enzyme while PteGlu bound less effectively. 10-Formyl-H_4PteGlu was also a good substrate although with a reduced V_{max}, while 5-substitution of reduced folate caused a large elevation in K_m. Previous studies have suggested that (6*S*)- and (6*ambo*)-5-formyl-H_4PteGlu are very good substrates (K_m about 5 μM) for human FPGS while the 6*R* isomer is a poor substrate (K_m about 100 μM) (21, 46). However, the high K_m obtained for (6*S*)-5-formyl-H_4PteGlu in the current study (105 μM) is consistent with values obtained for other folates substituted at the 5-position, and also with the value obtained with homogeneous pig liver enzyme (32). The lower values obtained by other investigators may have been due to metabolism of (6*S*)-5-formyl-H_4PteGlu to 10-formyl-H_4PteGlu. This conversion is catalyzed by 5,10-methenyltetrahydrofolate synthetase and 5,10-methylene tetrahydrofolate cyclohydrolase, enzymes that would be present in crude tissue extracts. The FPGS reaction mixture contains all the substrates necessary for this interconversion, and this would also explain the poor substrate activity of (6*R*)-5-formyl-H_4PteGlu.

Studies with model systems have demonstrated that cellular accumulation of folate, which requires its conversion to polyglutamate derivatives of chain length at least three, is dependent on FPGS activity levels with physiological levels of poor substrates for the enzyme and becomes highly dependent on FPGS activity levels with pharmacological levels of all folates (28, 36). Pharmacological doses of 5-formyl-H_4PteGlu are used in combination with thymidylate synthase inhibitors in some chemotherapeutic regimens to elevate tissue folate levels (58). The poor substrate activity of 5-formyl-H_4PteGlu for human FPGS suggests that the ability of tissues to accumulate 5-formyl-H_4PteGlu will be highly dependent on the level of FPGS activity. In addition, 5-formyl-H_4PteGlu accumulation may also be dependent on the level of 5,10-methenyltetrahydrofolate synthetase activity, which metabolizes this folate to 10-formyl-H_4PteGlu, a more effective substrate for FPGS. This enzyme has a K_m of about 1 μM for 5-formyl-H_4PteGlu (59) and would be operating under V_{max} conditions when pharmacological levels of 5-formyl-H_4PteGlu are provided.

The specificity for folylpolyglutamate substrates is qualitatively similar to that reported for the pig liver enzyme (32) except that diglutamate derivatives, with the exception of H_4PteGlu$_2$, compared to their respective pteroylmonoglutamates, tended to be somewhat better substrates. This was most pronounced for 5-formyl-H_4PteGlu$_2$ which exhibited a 7-fold decrease in K_m compared to the monoglutamate. V_{max} values fell beyond the diglutamate and H_4PteGlu$_n$ derivatives were the most effective polyglutamate substrates.

The 4-amino substitution of pteroylmonoglutamate increases its substrate effectiveness for mammalian FPGS but greatly impairs substrate activity with di- and longer polyglutamate derivatives (60–62). Methotrexate is a fairly poor substrate for mammalian FPGS enzymes and activity drops significantly with extension of the glutamate chain. The major drop in activity for the pig liver and CHO enzymes occurs at the diglutamate. However, activity falls off less sharply with the human enzyme and the large fall off in activity occurs at the triglutamate. The improved substrate activity of the diglutamate makes human cells particularly sensitive to methotrexate. Methotrexate accumulation by mammalian cells, which involves metabolism to at least the triglutamate, and its cytotoxic efficacy are very sensitive to the level of FPGS activity (13). CHO cell transfectants expressing human FPGS activity accumulate methotrexate more effectively than wild type CHO cells expressing the same levels of CHO FPGS activity and are more sensitive to the antifolate (13).

AUX-*hFPGScyt* transfectants, which expressed human FPGS activity solely in the cytoplasm, were unable to accumulate folate in the mitochondria and remained glycine auxotrophs. These data support our earlier findings that mitochondrial folate accumulation is dependent on mitochondrial FPGS activity and that the defect in glycine biosynthesis was due to a lack of mitochondrial folate despite the presence of cytoplasmic folates (23, 29). AUX-*hFPGSmit* transfectants, which appeared to express FPGS activity solely in the mitochondria, also contained long chain cytosolic folylpolyglutamates and were prototrophic for thymidine and purines, which are synthesized in the cytosol, as well as glycine. These results mirrored our observations with transfectants expressing *E. coli* FPGS in the mitochondria of AUXB1 cells (AUX-*mcoli*) (14, 29). Pulse-chase studies with AUX-*mcoli* and wild type CHO cells indicated a slow release of folylpolyglutamates from the mitochondria to the cytosol of these cells, and this was more pronounced in the AUX-*mcoli* transfectants, which synthesize shorter polyglutamate species (predominantly triglutamate) than wild type cells. Although the possibility that AUX-*hFPGSmit* transfectants express trace levels of human FPGS in the cytosol can not be totally eliminated, previous studies suggest that cytosolic metabolism of folate to polyglutamates cannot account for the cytosolic folate pool in AUX-*hFPGSmit* cells. AUXB1 cells transfected with human genomic DNA and expressing human FPGS at the trace levels that could be present in AUX-*hFPGSmit* cells (36) contain folylpolyglutamates that are of shorter chain length than was found in this study.

The proportion of cellular folate in the cytosol of AUX-*hFPGSmit* transfectants was much lower than in AUX-*mcoli* transfectants, ranging from about 4 to 8% of total cell folate after adjustment for mitochondrial folate contamination (as judged by glutamate dehydrogenase distributions), and was of shorter glutamate length than mitochondrial folate, suggesting that efflux of polyglutamates from the mitochondria shows a preference for the shorter chain length species. About 60% of folate in wild type CHO cells is cytosolic. Although expression of mammalian FPGS in the mitochondria is sufficient to generate a cytosolic folylpolyglutamate pool and to allow cytosolic 1-carbon metabolism, expression of a cytosolic FPGS is required for the generation of a normal folate pool in the cytosol.

The FPGS gene was mapped to chromosome region 9q34.1 and no additional chromosomal signals, suggestive of a second closely related gene, were observed, and the reconstructed FPGS gene restored AUXB1 cells to the wild type phenotype. Reversion frequencies of the AUXB1 mutant are also consistent with a single genetic mutation causing the multiple auxotrophy

(25, 26, 63). The organization of the 5′ region of the human gene is in general agreement with that described recently by Freemantle *et al.* (31). The exon 1 variant contains two ATG codons and alternate transcription start sites in this region generate mRNAs encoding mitochondrial and cytosolic FPGS isozymes (31). However, we have identified two additional splice variants (exons 1B and 1C) that arise by alternate splicing in the intron 1 region and an additional variant (exon 1A) that is transcribed by a promoter upstream of the exon 1 promoter region. All variants are spliced to exon 2.

The exon 1C variant is not an extended form of exon 1, and primer extension and RACE analyses suggested multiple transcription start sites in the intron 1 region. This region contains a number of potential promoter sites including Sp1 binding sites. Preliminary studies suggest that exon 1B and 1C variants are unlikely to generate functional FPGS. Exon 1A is preceded by a number of potential promoter sites including a binding site for CTF/NF-1 (CCAAT-binding transcription factors, nuclear factor 1) with the consensus sequence GCCAAT (52) and an E-box, generally represented by the sequence CANNTG. We have not yet identified the transcription start site(s) of exon 1A, and the possibility of further upstream introns cannot be excluded. The deduced sequence of the peptide encoded by exon 1A is characteristic of a mitochondrial leader sequence and exon 1A may encode an additional mitochondrial isoform of FPGS.

In our original cloning of a human FPGS cDNA from an Epstein Barr virus-transformed lymphocyte library, we obtained four clones, all of which encoded the exon 1 variant (30). As the cDNAs were cloned by functional complementation of an *E. coli* FPGS mutant, this may suggest that the other splice variants may not encode functional protein or may reflect an absence of these variants in the library used. However, it should be noted that we did not obtain any exon 1 forms with a complete mitochondrial leader sequence and, if any of the variants encode additional mitochondrial isoforms, it is probable that they would not be isolated by the complementation cloning procedure used. A single 2.5-kb band was observed in Northern analyses of HepG2 and MCF-7 cell mRNA using the human FPGS cDNA as a probe. However, the other exon 1 variants would be of similar expected size (within 0.1 kb) and would not be distinguished by Northern analyses. We have shown, using exon 1 variant specific probes, that the exon 1C variant is a major species (2.4 kb) in MCF-7 cells.

Decreased FPGS activity has been identified as a mechanism for cellular resistance to antifolates (5, 13–16, 64). However, the decreased FPGS activity is not always accompanied by decreases in FPGS mRNA levels (65). Although this may reflect modulation of translation rates or the presence of a mutation in the mRNA, this could also be due to down-regulation of a functional mRNA species being masked by the presence of variant mRNAs which do not encode active FPGS. Future studies on the regulation of FPGS activity should quantitate individual mRNA variants before conclusions on the absence of transcriptional control are reached. If some of the different mRNA species encode functional FPGS, this may explain the report of differences in substrate specificity for enzyme from different tissues of the same animal (24). We are currently investigating whether the different splice variants observed in this study are translated and whether they encode functional FPGS.

Acknowledgments—We thank Drs. Alfred Brenner and Ian Atkinson (UCB) and Patrick Stover (Cornell University) for advice and helpful discussions.

REFERENCES

1. McGuire, J. J., and Coward, J. K. (1984) in *Folates and Pterins* (Blakley, R. L., and Benkovik, S. J., eds) Vol. 1, pp. 135–90, Wiley Interscience, New York
2. Shane, B. (1989) *Vitam. Horm.* **45,** 263–335
3. Chabner, B. A., Allegra, C. J., Curt, G. A., Clendeninn, N. J., Baram, J., Koizumi, S., Drake, J. C., and Jolivet, J. (1985) *J. Clin. Invest.* **76,** 907–912
4. Hanlon, M. H., Ferone, R., Mullin, R. J., and Keith, B. R. (1990) *Cancer Res.* **50,** 3207–3211
5. McCloskey, D. E., McGuire, J. J., Russell, C. A., Rowan, B. G., Bertino, J. R., Pizzorno, G., and Mini, E. (1991) *J. Biol. Chem.* **266,** 6181–6187
6. Beardsley, G. P., Moroson, B. A., Taylor, E. C., and Moran, R. G. (1989) *J. Biol. Chem.* **264,** 328–333
7. Sirotnak, F. M., Johnson, T. B., Samuels, L. L., and Galivan, J. (1988) *Biochem. Pharmacol.* **37,** 4239–4241
8. Barredo, J., and Moran, R. G. (1992) *Mol. Pharmacol.* **42,** 687–694
9. Fort, D. W., Lark, R. H., Smith, A. K., Marling-Cason, M., Weitman, S. D., Shane, B., and Kamen, B. A. (1993) *Br. J. Haematol.* **84,** 595–601
10. Egan, M. G., Sirlin, S., Rumberger, B. G., Garrow, T. A., Shane, B., and Sirotnak, F. M. (1995) *J. Biol. Chem.* **270,** 5462–5468
11. Whitehead, V. M., Rosenblatt, D. S., Vuchich, M.-J., Shuster, J. J., Witte, A., and Beaulieu, D. (1990) *Blood* **76,** 44–49
12. Lenz, H.-J., Danenberg, K., Schnieders, B., Goeker, E., Peters, G. J., Garrow, T., Shane, B., Bertino, J. R., and Danenberg, P. V. (1994) *Oncol. Res.* **6,** 329–335
13. Kim, J.-S., Lowe, K. E., and Shane B. (1993) *J. Biol. Chem.* **268,** 21680–21685
14. Kim, J.-S., and Shane B. (1994) *J. Biol. Chem.* **269,** 9714–9720
15. Pizzorno, G., Chang, Y.-M., McGuire, J. J., and Bertino, J. R. (1989) *Cancer Res.* **49,** 5275–5280
16. Li, W.-W., Lin, J. T., Schweitzer, B. I., Tong, W. P., Niedzwiecki, D., and Bertino, J. R. (1992) *Cancer Res.* **52,** 3908–3913
17. Cichowicz, D. J., and Shane, B. (1987) *Biochemistry* **26,** 504–512
18. McGuire, J. J., Hsieh, P., Coward, J. K., and Bertino, J. R. (1980) *J. Biol. Chem.* **255,** 5776–5788
19. Moran, R. G., and Colman, P. D. (1984) *Biochemistry* **23,** 4580–4589
20. Pristupa, Z. B., Vickers, P. J., Sephton, G. B., and Scrimgeour, K. G. (1983) *Can. J. Biochem. Cell Biol.* **62,** 495–506
21. Clarke, L., and Waxman, D. J. (1987) *Arch. Biochem. Biophys.* **256,** 585–596
22. Moran, R. G., Colman, P. D., Rosowski, A., Forsch, R. A., and Chan, K. K. (1985) *Mol. Pharmacol.* **27,** 156–166
23. Lin, B.-F., Huang, R.-F. S., and Shane, B. (1993) *J. Biol. Chem.* **268,** 21674–21679
24. Rumberger, B. G., Barrueco, J. R., and Sirotnak, F. M. (1990) *Cancer Res.* **50,** 4639–4643
25. McBurney, M. W., and Whitmore, G. F. (1974) *Cell* **2,** 173–182
26. Taylor, R. T., and Hanna, M. L. (1977) *Arch. Biochem. Biophys.* **181,** 331–344
27. Foo, S. K., and Shane, B. (1982) *J. Biol. Chem.* **257,** 13587–13592
28. Lowe, K. E., Osborne, C. B., Lin, B.-F., Kim, J.-S., Hsu, J.-C., and Shane, B. (1993) *J. Biol. Chem.* **268,** 21665–21673
29. Lin, B.-F., and Shane, B. (1994) *J. Biol. Chem.* **269,** 9705–9713
30. Garrow, T. A., Admon, A., and Shane, B. (1992) *Proc. Natl. Acad. Sci. U. S. A.* **89,** 9151–9155
31. Freemantle, S. J., Taylor, S. M., Krystal, G., and Moran, R. G. (1995) *J. Biol. Chem.* **270,** 9579–9584
32. Cichowicz, D. J., and Shane, B. (1987) *Biochemistry* **26,** 513–521
33. Blakley, R. L. (1969) in *The Biochemistry of Folic Acid and Related Pteridines* (Neuberger, A., and Tatum, E. L., eds) pp. 92–94, North Holland Publishing Co., Amsterdam
34. Nakamaye, K., and Ecksteine, F. (1986) *Nucleic Acids Res.* **14,** 9679–9698
35. de Boer, H. A., and Kastelein, R. A. (1986) in *Maximizing Gene Expression* (Reznikoff, W., and Gold, L., eds) pp. 225–235, Butterworths, Boston
36. Osborne, C. B., Lowe, K. E., and Shane, B. (1993) *J. Biol. Chem.* **268,** 21657–21664
37. Laemmli, U. K. (1970) *Nature* **227,** 680–685
38. Peterson, G. L. (1977) *Anal. Biochem.* **83,** 346–356
39. Sanger, F., Nicklen, S., and Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. U. S. A.* **74,** 5463–5467
40. Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989) in *Molecular Cloning: A Laboratory Manual*, 2nd Ed., Cold Spring Harbor Laboratory, Cold Spring Harbor, NY
41. Apte, A. N., and Siebert, P. D. (1993) *BioTechniques* **15,** 890–893
42. Ausubel, F. M., Brent, R., Kingston, R. E., Moore, D. D., Seidman, J. G., Smith, J. A., and Struhl, K. (1987) *Current Protocols in Molecular Biology*, John Wiley & Sons, New York
43. Zabel, B. U., Naylor, S. L., Sakaguchi, A. Y., Bell, G. I., and Shows, T. B. (1983) *Proc. Natl. Acad. Sci. U. S. A.* **80,** 6932–6936
44. Garrow, T. A., Brenner, A., Whitehead, V. M., Chen, X.-N., Duncan, R. G., Korenberg, J. R., and Shane, B. (1993) *J. Biol. Chem.* **268,** 11910–11916
45. Magenis, R. E., Donlon, T. A., and Tomar, D. R. (1985) *Hum. Genet.* **69,** 300–303
46. McGuire, J. J., and Russell, C. A. (1991) *J. Cell. Pharmacol.* **2,** 317–323
47. Kelly, J. L., McLean, E. W., Cohn, N. K., Edelstein, M. P., Duch, D. S., Smith, G. K., Hanlon, M. H., and Ferone, R. (1990) *J. Med. Chem.* **33,** 561–567
48. Hagan, R. L., Duch, D. S., Smith, G. K., Hanlon, M. H., Shane, B., Freisheim, J. H., and Hynes, J. B. (1991) *Biochem. Pharmacol.* **41,** 781–787
49. Padgett, R. A., Grabowski, P. J., Konarska, M. M., Seiler, S., and Sharp P. A. (1986) *Annu. Rev. Biochem.* **55,** 1119–1150
50. Kadesch, T. (1993) *Cell Growth & Differ.* **4,** 49–55
51. von Heijne, G. (1986) *EMBO J.* **5,** 1335–1342
52. Santoro, C., Mermod, N., Andrews, P. C., and Tjian, R. (1988) *Nature* **334,** 218–224
53. Jones, C., Kao, F.-T., and Taylor, R. T. (1980) *Cytogenet. Cell Genet.* **28,** 181–194
54. Bolanowska, W. E., Russell, C. A., and McGuire, J. J. (1990) *Arch. Biochem.*

Biophys. **281,** 198–203
55. Shane, B. (1980) *J. Biol. Chem.* **255,** 5655–5662
56. Bognar, A. L., Osborne, C., and Shane, B. (1987) *J. Biol. Chem.* **262,** 12337–12343
57. Bognar, A. L., and Shane, B. (1983) *J. Biol. Chem.* **258,** 12574–12581
58. Zhang, Z.-G., and Rustum, Y. M. (1991) *Cancer Res.* **51,** 3476–3481
59. Stover, P., Huang, T., Schirch, V., Maras, B., Valiante, S., and Barra, D. (1993) *Adv. Exp. Med. Biol.* **338,** 723–726
60. George, S., Cichowicz, D. J., and Shane, B. (1987) *Biochemistry* **27,** 522–529
61. Cook, J. D., Cichowicz, D. J., George, S., Lawler, A., and Shane, B. (1987) *Biochemistry* **26,** 530–539
62. Coll, R. J., Cesar, D., Hynes, J. B., and Shane, B. (1991) *Biochem. Pharmacol.* **42,** 833–838
63. Sussman, D. J., Milman, G., and Shane, B. (1986) *Somat. Cell Mol. Genet.* **12,** 531–540
64. Wang, F.-S., Aschele, C., Sobrero, A., Chang, Y.-M., and Bertino, J. R. (1993) *Cancer Res.* **53,** 3677–3680
65. Pizzorno, G., Moroson, B. A., Cashmore, A. R., Russello, O., Mayer, J. R., Galivan, J., Bunni, M. A., Priest, D. G., and Beardsley, G. P. (1995) *Cancer Res.* **55,** 566–573