Structural Organization of the Human Folyypoly-γ-glutamate Synthetase Gene: Evidence for a Single Genomic Locus

Shirley M. Taylor, Sarah J. Freemantle, and Richard G. Moran

Departments of Microbiology/Immunology and Pharmacology/Toxicology and the Massey Cancer Center, Medical College of Virginia, Richmond, Virginia 23298

Abstract

The cytotoxicity, and probably the selectivity, of folate antimetabolites depend upon the expression of the enzyme folylpoly-γ-glutamate synthetase in tumor cells. Evidence for the existence of multiple forms of this enzyme and the need to define the control mechanisms of expression levels in normal and neoplastic cells has focused attention on the gene(s) encoding these forms. The organization of the genomic locus for the human folylpoly-γ-glutamate synthetase (FPGS) gene has been determined. The complete 2256 nucleotides of cDNA for the 5'-untranslated region, mitochondrial leader sequence, coding region, and 3'-untranslated region were distributed on 15 exons stretching over 11.2 kb of genomic DNA. All of the restriction fragments found in diploid human genomic DNA could be accounted for by fragments contained on the isolated genomic clones. Likewise, Southern analysis of the transfected human genomic DNA that complemented the FPGS' phenotype of a hamster cell line indicated that the same gene had been integrated in all of three independently derived transfecteds. We conclude that the genomic locus that we now report appears to be the only gene encoding FPGS-related sequences in the human complement.

Introduction

The classical folate antimetabolite methotrexate was the first agent to be rationally applied to the treatment of childhood cancers and the first to cause complete remissions of acute lymphocytic leukemia of childhood. A key feature of the biochemistry of MTX3 has proved to be essential for proliferation of mammalian cells. These MTX polyglutamates are at least as active as the parent drug as inhibitors of dihydrofolate reductase; however, in contrast to MTX itself, they are glutamates are at least as active as the parent drug as inhibitors of the enzyme and the need to define the control mechanisms determinant of these MTX polyglutamates for the human folylpoly-γ-glutamate synthetase (FPGS) gene has been determined. The expression of FPGS is high in pre-B-cell acute lymphocytic leukemia samples and significantly lower in T-cell leukemias, a distribution that correlates with the responsiveness of these leukemias to MTX (7). Likewise, the ability to accumulate polyglutamates of MTX has emerged as an indicator of the responsiveness of individual leukemias to chemotherapy with multidrug regimens containing MTX (9), presumably an indication of the centrality of MTX in the therapeutic activity of these combinations.

Two types of polymorphism have been suggested for FPGS expressed in mammalian cells. FPGS has been found to be distributed in both cytosolic and mitochondrial compartments of mammalian cells (10). Two forms of mRNA encoded by the FPGS gene in human CEM leukemia cells have been shown to differ by the use of an upstream translational start site encoding an NH2-terminal 42-amino acid peptide that functions as a mitochondrial leader sequence (11). Alternative transcriptional start sites (11), rather than alternative exons, are used for the generation of mitochondrial and cytosolic protein from this single locus. More indirect evidence has also suggested a polymorphism in the FPGS expressed in some mouse tumors and normal mouse tissues (12, 13) and that differences exist in the catalytic properties of the putative isoforms.

These observations have raised the questions of whether there are multiple genes for FPGS and whether alternative exon usage would allow the generation of species of enzyme with therapeutically exploitable differences in substrate preference. We now report results supporting the concept that there is only one genomic locus corresponding to, or closely homologous to, the reported (11, 14) cDNA sequences. We also furnish the detailed genomic mapping information that is a necessary prerequisite for investigation of the existence and position of alternative exons used to generate catalytically active isoforms of FPGS.

Materials and Methods

Library Screening. Clones representing the human genomic locus of the FPGS gene were isolated from a λ FIX II human male placental genomic library (Stratagene) as described (11). Hybridization of the nitrocellulose filters was at 63°C in 5X Denhardt's solution, 5X SSC [0.15 M sodium chloride, 0.015 M sodium citrate (pH 7.0)], 50 mM sodium phosphate (pH 6.5), 0.1% SDS, and 0.1 mg/ml denatured salmon sperm DNA. The filters were washed at 55°C in 0.5X SSC-0.1% SDS.

Restriction Mapping, Subcloning, and Sequencing. The λ genomic clones were mapped using Southern blots of insert DNA digested with SacI, HindIII, and BamHI and analyzed with radiolabeled cDNA and oligonucleotide probes. Nine subclones, some of them overlapping and two corresponding to the same fragment isolated from two different genomic clones, were subcloned into pBluescript II SK+ for mapping and sequencing of intron/exon
Measurement of Intron Sizes by PCR. Placement of exons relative to restriction sites was achieved by PCR using genomic subclones as target DNA. Gene-specific primers within an exon and M13 forward or reverse vector primers were used to amplify the sequence extending to each restriction site. PCR was carried out using 10 pg/µg of denatured subclone DNA, 3 ng/µl of each primer, 0.2 mM deoxynucleotide triphosphates, and 1 unit of Taq polymerase in a total volume of 25 µl. Cycling conditions were 30 cycles of 94°C for 60 s, 58–62°C for 60 s, and 72°C for 2 min in a MJ Research thermal cycler. Sizes of introns were measured by PCR using pairs of gene-specific primers located in adjacent exons. Target DNA was 10 pg/µl genomic subclone DNA (introns 8, 11, and 13) or 1 ng/µl genomic DNA (introns 4, 9, and 14). The lengths of these PCR products were measured on 1% agarose gels.

Southern Analysis of the Human FPGS Genomic Locus. Genomic DNA from peripheral blood lymphocytes from two normal individuals, CEM leukemia cells, AUX B1 cells, and four AUX B1 derivatives transfected with total genomic DNA from CEM cells (11) were digested with SacI restriction endonuclease, resolved on a 1% agarose gel, transferred to nylon membrane (Biotrans; ICN) in 1× SSC and probed with a cDNA fragment containing the entire coding sequence and 3' UTR (11, 14). Hybridization conditions were 63°C using 0.5 M sodium phosphate (pH 7.0), 7% SDS, 1% BSA, and 1 mM EDTA, and filters were washed to high stringency with 0.5× SSC and 0.1% SDS at 55°C.

Results

Isolation of A Clones for the Human FPGS Genomic Locus. A cDNA fragment corresponding to the first 690 bp of the published (14) human FPGS open reading frame was used as a probe to screen a human male placental genomic library constructed in A-FIXII vector (Stratagene). Forty positive λ plaques were isolated from a screen of 1.5 × 10⁶ recombinants. Three overlapping clones were found sufficient to encompass the previously identified cDNA sequences (11, 14) for the 3'-UTR, the mitochondrial leader sequence, the open reading frame for the structural gene, and the 5'-UTR of the human FPGS gene and were used for all of our subsequent analyses. These clones, ABL, λ21, and λ22, were determined to be contiguous by the construction of a restriction map with three endonucleases, BamHI, HindIII, and SacI; this map indicated overlaps of ABL and λ22 and of λ22 and λ21 by 5–6 kb, with distinctive placement of restriction sites on both overlapping segments (Fig. 1). Sequence analysis of intronic regions from the overlapping segments of ABL and segments λ22 (on subclones pFBLNH6.0 and pF22B3.5) and of λ22 and λ21 (on subclones pF21S3.0 and pF22S3.0 and on pF21S2.0 and pF22S2.0) confirmed the contiguity of these clones. Subsequently, another recombinant phage isolated from this same screen, λ28, was found to contain all of the exons shown in Fig. 1 and to have an identical map to the regions of overlap with the other clones. The use of a 6.0-kb NotI-HindIII fragment of ABL, subcloned as pFBL.0H6 (Fig. 1), to define the first three exons of this gene and the transcriptional start sites used for cytotoxic and mitochondrial FPGS messages has previously been described (11). The several restriction fragments that were found to hybridize with cDNA were subcloned into pBluescript for sequence analysis, as detailed in Fig. 1.

Partial sequencing of the subclones shown in Fig. 1 using primers corresponding to the FPGS cDNA and subsequent sequencing in the opposite direction with intronic primers identified all of the intron/exon borders of the published cDNAs (Fig. 1). The size of the genomic locus, as defined from the major transcriptional start site (11) to the previously identified polyadenylation signal sequence (14), was 11.2 kb. Of the total of 15 exons into which the gene was distributed, the 5'-UTR, both previously described translational start codons, the sequence encoding the mitochondrial leader peptide, and the first 122 nucleotides of coding region (11) were located on exon 1, exons 2–14 were exclusively coding region, and exon 15 contained 410 nucleotides of coding region and 450 nucleotides of 3'-UTR. The coding region exons internal to the gene ranged in size from 54 to 151 nucleotides (Table 1), well within the expected distribution of exon boundaries. Sequence analysis was performed using double-stranded DNA and Sequenase 2.0 (U.S. Biochemical Corp.).

Fig. 1. Restriction map and exon structure of the human FPGS genomic locus. Three overlapping genomic clones, ABL, λ22, and λ21, were used to determine the intron/exon structure of the human FPGS gene. These clones had insert sizes of 19, 17, and 15 kb, respectively; ABL extends upstream of the map shown and contained 14 kb of sequence upstream of the major transcriptional start site (11). Genomic restriction mapping and sequencing of putative overlapping subclone pairs (pFBLNH6.0,pF22B3.5, pF21S2.0,pF22S2.0, and pF21S3.0,pF22S3.0) verified the contiguity of this restriction map. A pair of closely spaced BamHI and SacI sites were displaced slightly for clarity; they map to cDNA sequence corresponding to exon 5. Subclones used in the analysis of intron/exon boundaries and for the determination of intron sizes are depicted below the organizational map of the locus. Exons 5–11 map to subclone pF21S2.0, but they have been displaced slightly for clarity.
sizes (15). Splice junctions were identified by departure from the sequence of cDNA coincident with the position of consensus sequences for splice donors and acceptors. All of the splice sites identified agreed well with published (16) consensus sequences (Table 1). Seven of the 14 introns were sufficiently small (83–166 nucleotides) that they were sequenced in toto during intron/exon identification. In addition, intron 1 was sequenced in its entirety to allow identification of regulatory sequences sometimes found in the initial intron of TATA-less genes (17). The sequences of these introns are not listed herein but have been submitted to GenBank.4 The sizes of introns 4, 8, 9, 11, 13, and 14 (Table 1) were determined by PCR using both λ clones and plasmid subclones as target DNAs; these PCR measurements were unambiguous (Fig. 2).

Attempts to Detect Other FPGS-related Genes. Two lines of experimentation were used to investigate whether there were multiple genes encoding functional FPGS or closely related sequences in the human genome. The first approach used cell lines in which the FPGS phenotype of the hamster AUXB1 cell, a CHO derivative, was complemented by the transfection of high molecular weight DNA from human CEM cells. The different substrate preferences of hamster and human FPGS (ratio = 1.41 ± 0.11) allow distinction of reversion events from transfomants in such experiments. Thus, FPGS from CEM cells had a higher enzyme activity using dATP as a substrate than using ATP (the ratio of the activity with ATP to that with dATP = 0.73 ± 0.08), but FPGS from CEM cells had the opposite preference (ratio = 1.41 ± 0.11). An evaluation of the kinetic characteristics of FPGS expressed in these independently derived transfectants indicated that the gene corresponding to the locus diagrammed in Fig. 1 had been stably integrated into the genome of each cell line (Fig. 3, Lanes a and b) and in CEM DNA (Fig. 3, Lane c) was compared with that predicted from the map shown in Fig. 1. The patterns of restriction fragments from the human genomic DNA could be accounted for by the bands arising from this single locus, as depicted by the hybridization pattern in the cloned phage (Fig. 3, Lanes BL, 21, and 22). The one band in FB1

---

![Intron #](image)

---

4 The nucleotide sequences reported in this paper have been submitted to the GenBank/EMBL Data Bank with accession numbers U40863 through U40868.

---

THE HUMAN FPGS GENE

Table 1 FPGS intron-exon structure

<table>
<thead>
<tr>
<th>Exon no.</th>
<th>Exon size (nt)</th>
<th>5' intron sequence</th>
<th>Exon sequence</th>
<th>3' intron sequence</th>
<th>Intron length (nt)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>180</td>
<td>acgtggtgaccag</td>
<td>CATCAG</td>
<td>gtagcggc</td>
<td>1039b</td>
</tr>
<tr>
<td>2</td>
<td>129</td>
<td>tcctcggtccgag</td>
<td>GGCAG</td>
<td>gtaagcgg</td>
<td>85b</td>
</tr>
<tr>
<td>3</td>
<td>54</td>
<td>ccctcctccgag</td>
<td>GATCG</td>
<td>gtgaaggcg</td>
<td>83b</td>
</tr>
<tr>
<td>4</td>
<td>65</td>
<td>ccctcctccgag</td>
<td>GGCTCC</td>
<td>gtaggcgg</td>
<td>2309</td>
</tr>
<tr>
<td>5</td>
<td>115</td>
<td>acgtggtgaccag</td>
<td>CTTTAG</td>
<td>gccgggct</td>
<td>111b</td>
</tr>
<tr>
<td>6</td>
<td>78</td>
<td>tcctcctccgag</td>
<td>ACCAAG</td>
<td>gtagcgcgg</td>
<td>150b</td>
</tr>
<tr>
<td>7</td>
<td>62</td>
<td>tcctcctccgag</td>
<td>CATCAG</td>
<td>gtagcgcgg</td>
<td>99b</td>
</tr>
<tr>
<td>8</td>
<td>103</td>
<td>tcctcctccgag</td>
<td>TTTAAC</td>
<td>gtagcgcgg</td>
<td>580</td>
</tr>
<tr>
<td>9</td>
<td>78</td>
<td>tcctcctccgag</td>
<td>ATCTCA</td>
<td>gtagcgcgg</td>
<td>279</td>
</tr>
<tr>
<td>10</td>
<td>148</td>
<td>cgctcctccgag</td>
<td>GCCAGT</td>
<td>gtagcgcgg</td>
<td>125b</td>
</tr>
<tr>
<td>11</td>
<td>1078</td>
<td>aactgcggccag</td>
<td>GCTCG</td>
<td>gtagcgcgg</td>
<td>819</td>
</tr>
<tr>
<td>12</td>
<td>151</td>
<td>gctctcctccgag</td>
<td>GAGCCG</td>
<td>gtagcgcgg</td>
<td>172b</td>
</tr>
<tr>
<td>13</td>
<td>76</td>
<td>gctctcctccgag</td>
<td>TCACG</td>
<td>gtagcgcgg</td>
<td>867</td>
</tr>
<tr>
<td>14</td>
<td>67</td>
<td>cgctcctccgag</td>
<td>ACCAAC</td>
<td>gtagcgcgg</td>
<td>2195</td>
</tr>
<tr>
<td>15</td>
<td>860</td>
<td>tcctcctccgag</td>
<td>ACGCAC</td>
<td>gtagcgcgg</td>
<td></td>
</tr>
</tbody>
</table>

---

*kb* = kilobase.

The nucleotide sequences reported in this paper have been submitted to the GenBank/EMBL Data Bank with accession numbers U40863 through U40868.

---

they contained the hygromycin B resistance gene used as the coselectable marker in these transfections (data not shown).

The second approach used to evaluate the potential for redundant genes for FPGS involved determination of FPGS-related sequences by direct Southern blotting. When the pattern of restriction fragments in normal diploid peripheral blood lymphocytes (Fig. 3, Lanes a and b) and in CEM DNA (Fig. 3, Lane c) was compared with that predicted from the map shown in Fig. 1, all of the hybridizing bands in human genomic DNA could be accounted for by the bands arising from this single locus, as depicted by the hybridization pattern in the cloned phage (Fig. 3, Lanes BL, 21, and 22). The one band in FB1
cells that had a different mobility (7.0 kb) instead of that seen in human DNA and transfectants FA1 and FC2 (5.0 kb) represented the most 3' fragment, indicating that insertion into the genome had occurred immediately upstream of the most 3' SacI restriction site.

Discussion

The generation of mitochondrial and cytosolic forms of FPDS in human tissues appears explained by our previously published finding (11) of a mitochondrial leader sequence encoded by the same gene as that which Garrow et al. (14) have demonstrated to complement the human FPGS locus. λ clones and human genomic DNAs were digested with SacI, resolved on a 1% agarose gel, and hybridized with a radiolabeled cDNA corresponding to the full-length cDNA for FPGS, including the mitochondrial leader sequence, the coding region, and the 3'-UTR up to nucleotide 2173. Hybridization was carried out at 63°C for 18 h, and the final wash was at 55°C and 0.5× SSC with 0.1% SDS. Left lanes: Lane BL, 10 pg ABL; Lane 21, 100 pg of A21, Lane 22, 100 pg of A22. Center lanes: Lane 21, 10 pg of A21; Lane 22, 10 pg of A22. Lanes a-h, 10 μg genomic DNA isolated from the following sources: peripheral blood lymphocytes from two normal individuals (Lanes a and b), CEM leukemia cells (Lane c), AUX B1 cells transfected with total CEM DNA, cell lines FA1 (Lane d), FB1 (Lane e), FC2 (Lane f), and FD1 (Lane g), and AUX B1 cells (Lane h). Size markers are λ DNA restricted with HindIII and ΦX174 DNA restricted with HaeIII. Sizes of SacI fragments found in FPDS genomic clones are indicated on the left of the figure in kilobases (kb).

Fig. 3. Southern blot analysis of λ genomic clones and genomic DNA defining the human FPGS locus. λ clones and human genomic DNAs were digested with SacI, resolved on a 1% agarose gel, and hybridized with a radiolabeled cDNA corresponding to the full-length cDNA for FPGS, including the mitochondrial leader sequence, the coding region, and the 3'-UTR up to nucleotide 2173. Hybridization was carried out at 63°C for 18 h, and the final wash was at 55°C and 0.5× SSC with 0.1% SDS. Left lanes: Lane BL, 10 pg ABL; Lane 21, 100 pg of A21, Lane 22, 100 pg of A22. Center lanes: Lane 21, 10 pg of A21; Lane 22, 10 pg of A22. Lanes a-h, 10 μg genomic DNA isolated from: peripheral blood lymphocytes from two normal individuals (Lanes a and b), CEM leukemia cells (Lane c), AUX B1 cells transfected with total CEM DNA, cell lines FA1 (Lane d), FB1 (Lane e), FC2 (Lane f), and FD1 (Lane g), and AUX B1 cells (Lane h). Size markers are λ DNA restricted with HindIII and ΦX174 DNA restricted with HaeIII. Sizes of SacI fragments found in FPDS genomic clones are indicated on the left of the figure in kilobases (kb).

stringent hybridization conditions (which were needed to minimize the in-lane background from the very GC-rich 5'-end of the full-length cDNA probe) did not allow detection of additional restriction fragments. However, the conditions used allowed detection of the endogenous hamster FPDS gene (Fig. 3, Lanes d-h, open arrow), albeit as a weak signal; equivalent strength minor bands can be seen in the human DNA (Fig. 3, Lanes a-c, small arrows). Hence, we take our evidence that a single FPDS gene is present in the human genome (Fig. 3) as highly suggestive but not definitive. In theory, it would seem possible that a second gene might encode a protein capable of converting only monoglutamate folates to diglutamate derivatives or of foyl diglutamates to longer chain polyglutamates. However, in vitro studies with recombinant FPDS expressed in baculovirus-infected insect cells (19) have demonstrated that the single FPDS species corresponding to the published (14) cDNA for cytosolic FPDS easily catalyzes both reactions. Hence, it would appear most likely at this point that there is only one genomic locus for any and all species of FPDS required by mammalian cells. Isoforms other than those involving mitochondrial and cytosolic species, if formed, would be likely to be generated from the locus we herein report by alternative splicing of exons not represented in Fig. 1.

Studies on several other systems have led to the concept that the evolution of enzymes was a process of "exon shuffling," from which one would generalize that individual exons or small groups of exons often correspond to individual functional domains (20). For instance, the 12 transmembrane domains of the mdr1 gene are situated on 12 individual exons, and the two Walker A and B motifs, which together constitute the ATP-binding sites of this protein, are also present as four distinct exons (21). Similarly, the two components of the nucleotide binding fold of a number of dehydrogenases are contained on three and two adjacent exons (22). One would presume, then, that the functional domains of the FPDS protein would be distributed onto separate exons or small groups of exons. Two peptides have been suggested previously to constitute Walker A and B sites of the FPDS protein on the basis of consensus considerations (14). However, mutation of the analogous site in the Escherichia coli FPDS did not specifically eliminate ATP binding (23). Further molecular and structural study of the fragments represented by each exon may well define the functional domains involved in the FPDS protein and reaction.

The human FPDS gene was found on a rather small genomic locus distributed onto 15 exons; the majority of the exons was separated by small introns, with six more sizable intervening sequences. If alternative exons were to be found in this gene, they would be likely to be located within the larger introns. The sequences of introns 1–3, 5–7, 10, and 12 were determined and have been deposited in Genbank. With the availability of the genetic map shown in Fig. 1, assessment of the presence or absence of alternative exons and the sequence and position of any such putative exons should readily proceed.

Acknowledgments

We thank Jennifer Kilgore for her skillful assistance in the location of exon borders.

References


Structural Organization of the Human Folylpoly-γ-glutamate Synthetase Gene: Evidence for a Single Genomic Locus

Shirley M. Taylor, Sarah J. Freemantle and Richard G. Moran


Updated version
Access the most recent version of this article at:
http://cancerres.aacrjournals.org/content/55/24/6030

E-mail alerts
Sign up to receive free email-alerts related to this article or journal.

Reprints and Subscriptions
To order reprints of this article or to subscribe to the journal, contact the AACR Publications Department at pubs@aacr.org.

Permissions
To request permission to re-use all or part of this article, use this link http://cancerres.aacrjournals.org/content/55/24/6030. Click on "Request Permissions" which will take you to the Copyright Clearance Center's (CCC) Rightslink site.