Proc. Natl. Acad. Sci. USA
Vol. 93, pp. 10134–10138, September 1996
Biochemistry

Human γ-glutamyl hydrolase: Cloning and characterization of the enzyme expressed *in vitro*

RONG YAO, ERASMUS SCHNEIDER, THOMAS J. RYAN, AND JOHN GALIVAN

Division of Molecular Medicine, Wadsworth Center, New York State Department of Health, Albany, NY 12201

Communicated by Gertrude B. Elion, Glaxo Wellcome, Inc., Research Triangle Park, NC, July 3, 1996 (received for review June 3, 1996)

ABSTRACT **A cDNA encoding human γ-glutamyl hydrolase has been identified by searching an expressed sequence tag data base and using rat γ-glutamyl hydrolase cDNA as the query sequence. The cDNA encodes a 318-amino acid protein of M_r 35,960. The deduced amino acid sequence of human γ-glutamyl hydrolase shows 67% identity to that of rat γ-glutamyl hydrolase. In both rat and human the 24 amino acids preceding the N terminus constitute a structural motif that is analogous to a leader or signal sequence. There are four consensus asparagine glycosylation sites in the human sequence, with three of them conserved in the rat enzyme. Expression of both the human and rat cDNA in *Escherichia coli* produced antigenically related proteins with enzyme activities characteristic of the native human and rat enzymes, respectively, when methotrexate di- or pentaglutamate were used as substrates. With the latter substrate the rat enzyme cleaved the innermost γ-glutamyl linkage resulting in the sole production of methotrexate as the pteroyl containing product. The human enzyme differed in that it produced methotrexate tetraglutamate initially, followed by the triglutamate, and then the diglutamate and methotrexate. Hence the rat enzyme is an endopeptidase with methotrexate pentaglutamate as substrate, whereas the human enzyme exhibits exopeptidase activity. Another difference is that the expressed rat enzyme is equally active on methotrexate di- and pentaglutamate whereas the human enzyme has severalfold greater activity on methotrexate pentaglutamate compared with the diglutamate. These properties are consistent with the enzymes derived from human and rat sources.**

γ-Glutamyl hydrolase (GH; E.C. 3.4.19.9) catalyzes the hydrolysis of folylpoly-γ-glutamates and antifolylpoly-γ-glutamates by the removal of γ-linked polyglutamates and glutamate (1). The specificity of the enzyme with regard to the pteroyl portion of the $PteGlu_n$ is rather low, and often the enzyme is active on poly-γ-glutamic acid (1, 2). Specificity is displayed in the site of cleavage of the poly-γ-glutamate tail; enzymes from different sources can have exo- or endopeptidase activity (1). In one study it was shown that GH is not active on protein bound poly-γ-glutamates (2). Recent studies from this laboratory have focused on GH derived from tumor cell lines in culture, and a number of novel properties of the enzyme have been discovered. As in other cell systems, the intracellular enzyme from the rat H35 hepatoma cell line is lysosomal (3). It has also been shown to be a glycoprotein with extensive asparagine-linked carbohydrates (4). Using the H35 cell line, it was demonstrated that GH activity in cultured cells is reduced by insulin (5) and its elevation is a mechanism of acquired drug resistance (6). In H35 cells and tumor cell lines from other mammals the vast majority of the enzyme activity was found to be secreted (7). Human breast cancer cell lines and the HEPG2 hepatoma cell line secrete notably high levels of GH (8). While large amounts of data are available about the catalytic activity of GH, the factors that control or regulate its function are not well understood. Furthermore, only small amounts of protein produced by purification procedures have been available for mechanistic and structural studies on GH. To address these deficiencies, the identification of a cDNA for human GH was undertaken along with the development of expression systems for the rodent and human enzymes.

We have previously cloned a rat GH cDNA and determined the deduced sequence of the rat enzyme (4). That cDNA was used to identify the analogous human sequence. Both were expressed in an *Escherichia coli* system and the enzymes produced were compared with regard to their catalytic activities.

MATERIALS AND METHODS

Data Base Search and Sequence Analysis. The full-length cDNA of rat GH was used as a query sequence for a sequence homology search by the data base searching program BLAST (Genetics Computer Group, Madison, WI). The FRAMESEARCH program was used to identify possible open reading frames (ORFs) for the expressed sequence tag (EST) (9) clones identified as homologous to rat GH by the BLAST program. Using the GAP program, the deduced human amino acid sequence was aligned and compared with rat GH.

DNA Sequencing of Clone R07771 and H09442. The two human EST clones, R07771 and H09442, were obtained from Genome Systems (St. Louis). The initial sequencing for clone R07771 was done using T3 and T7 primers on the pT7T3D-Pac vector and for clone H09442 using M13R and M13F primers on the Lafmid BA vector by the method of Sanger *et al.* (10) using Sequenase (United States Biochemical). Sequencing of the entire insert was done with GH-specific primers synthesized by the Molecular Genetics Core Facility at the Wadsworth Center (Albany, NY).

Cloning and Expression of Human and Rat GH in *E. coli* Cells. The rat and human cDNAs were cloned into the pET24a vector followed by transformation into *E. coli* Bl21DE3pLysS competent cells (Novagen). The *E. coli* cells carrying the GH gene were grown in terrific broth medium in the presence of kanamycin (final concentration, 30 μg/ml) and chloramphenicol (final concentration, 34 μg/ml) at 37°C to OD_{600} of ≈0.6 and expression was induced with 1 mM isopropyl β-D-thiogaloctoside for 16 h. Cells were sonicated in 0.1 M NaAc (pH 5.0) and stored on ice at 4°C.

Western Blot of Expressed GH. The *E. coli* cells containing expressed human and rat GH were collected by centrifugation and the cell pellet was suspended in SDS/PAGE sample buffer. A typical yield was about 1 mg total protein per ml of culture under the experimental conditions used. Aliquots of cell lysates (2 μl, ≈20 μg protein) of each sample were electrophoresed and electroblotted onto nitrocellulose filters.

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Abbreviations: GH, γ-glutamyl hydrolase; $4\text{-}NH_2\text{-}10\text{-}CH_3PteGlu_n$, methotrexate poly-γ-glutamate, with *n* the total number of glutamates; EST, expressed sequence tag.

Data deposition: The sequence reported in this paper has been deposited in the GenBank data base (accession no. U55206).

The antibody raised against the purified rat GH was purified on protein A-agarose and used at 1:10,000 dilution (4).

Catalytic Activity Assay of GH. GH activity of cell extracts was assayed by incubating with 100 μM methotrexate poly-γ-glutamate (4-NH_2-10-CH_3PteGlu$_2$ or -Glu$_5$) as substrate in the presence of 50 mM 2-mercaptoethanol and 0.1% Triton X-100 at 37°C (7). The cell extract of expressed human GH was used as a 1:2500 dilution of the sonicated cell suspension; for expressed rat GH a 1:100 dilution was used. The catalytic products were detected by HPLC analysis (2, 3, 11).

RESULTS AND DISCUSSION

Studies on GH isolated from rat H35 cells have established the GH cDNA and deduced amino acid sequence (4). Using the BLAST program, two clones with GenBank accession numbers R07771 and H09442 were identified from the data base of human ESTs with 80% and 75% homology, respectively, to the rat GH cDNA sequence. The cDNA for clone R07771 was isolated from female placenta and for H09442 was from infant brain. The sequence homology between rat GH and these human EST clones was further confirmed with deduced amino acid sequences. The FRAMESEARCH showed that clone R07771 had an ORF from nucleotide 2 to 302 in a 0.8-kb sequence and an ORF for H09442 is from nucleotide 59 to 348 in a 1.3-kb sequence. The amino acid sequence of clone R07771 showed a 75% identity to rat GH amino acids 153–253, while clone H09442 was 77% homologous to the 5′ end of rat GH. These results strongly suggested that the two ESTs identified were parts of the human homolog to rat GH. Restriction enzyme digestion by *Not*I and *Hin*dIII confirmed the presence of a 1.3-kb insert in clone H09442, suggesting that it contained the full-length cDNA for human GH, whereas clone R07771 has a 0.8-kb insert by *Not*I and *Eco*RI digestion. Sequencing of both inserts showed that the 5′ sequence of clone R07771 overlaps with H09442 at nucleotide 518 (Fig. 1) and confirmed that H09442 contained the full-length cDNA for human GH. The nucleotide sequences of human and rat cDNA showed 77% identity based upon a cDNA described for rat hepatoma GH (4).

The deduced amino acid sequence of human and rat GH showed 67% identity (Fig. 2). N-terminal amino acid sequence analysis of the purified rat GH showed that mature rat GH starts at Gly-25. This suggested the possibility of a leader sequence encompassing the first 24 amino acids (4). In the deduced human sequence the first 24 amino acids of the N terminus showed 67% identity to the putative rat leader

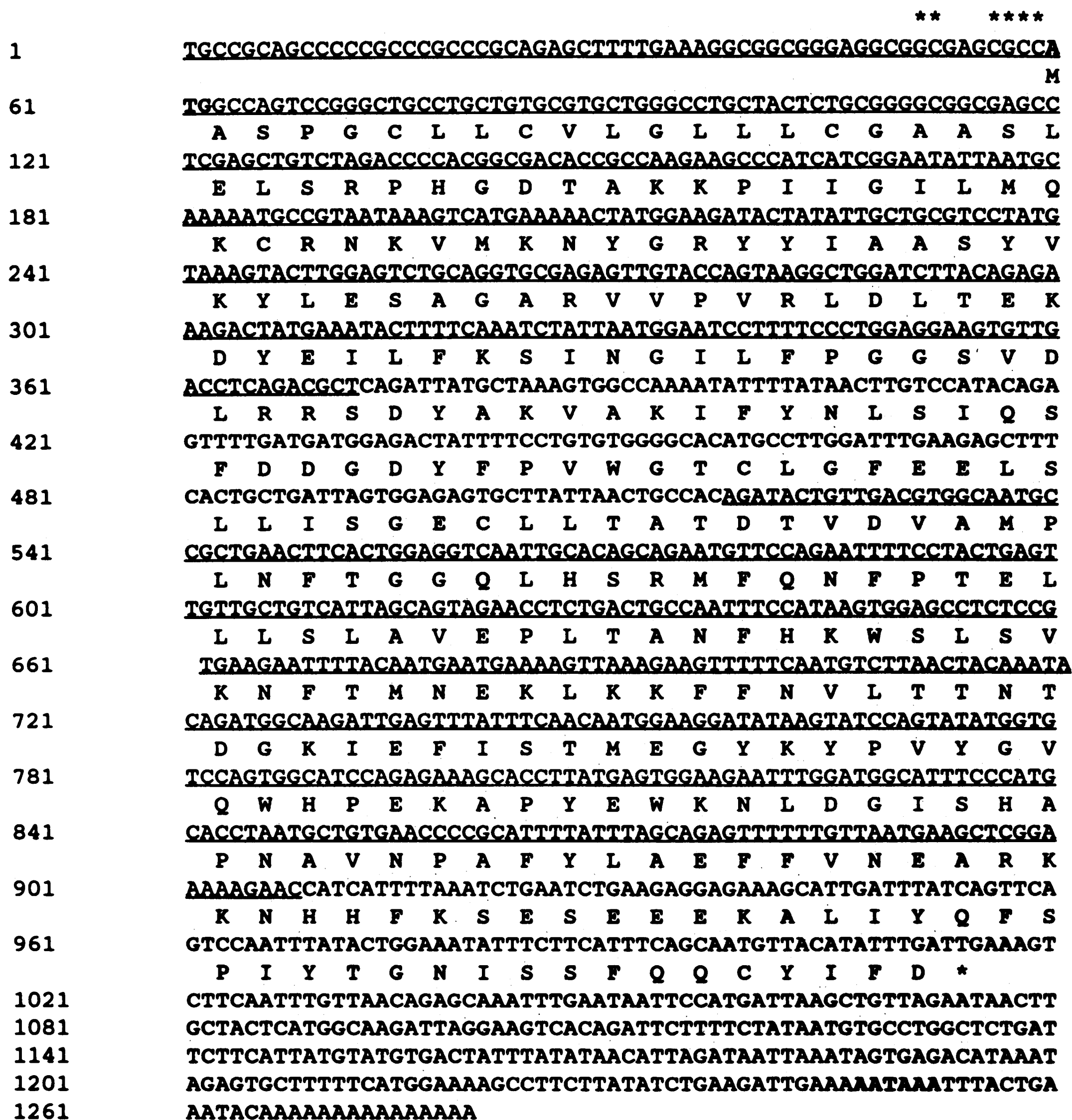

FIG. 1. Nucleotide and deduced amino acid sequence of human GH. The initiation codon and the polyadenylylation signal are in boldface type. Asterisks indicate the bases that match the consensus sequence for initiation of translation in vertebrates (12). The sequences in the EST data base for clones H09442 and R07771 are underlined from base 1 to 372 and base 518 to 908, respectively.

```
Human     1 MASPGCLLCVLGLLLCGAASLELSRPHGDTAKKPIIGILMQKCRNKVMKN 50
            ||| | |||.:.||||| || :|| ..: ..|:|||||:||.| .. |..
Rat       1 MASLGRLLCAWVLLLCGLASPGLSGSYERGSKRPIIGIIMQECYGN.MTK 49

Human    51 YGRYYIAASYVKYLESAGARVVPVRLDLTEKDYEILFKSINGILFPGGSV 100
            .||:||||||||::|||||||||:||||.: :||.||:||||:|:|||:.
Rat      50 LGRFYIAASYVKFIESAGARVVPIRLDLNDAQYETLFRSINGVLLPGGGA 99

Human   101 DLRRSDYAKVAKIFYNLSIQSFDDGDYFPVWGTCLGFEELSLLISGECLL 150
            :| :|:|.:|||||:. .::|||:||||||||||||:||||:|:|.: ||
Rat     100 NLTHSGYSRVAKIFFTKALESFDNGDYFPVWGTCLGLEELSVLVSNDNLL 149

Human   151 TATDTVDVAMPLNFTGGQLHSRMFQNFPTELLLSLAVEPLTANFHKWSLS 200
            | |:| .| :||||| :  :||||.|:|.||| ||| | |||||||||||
Rat     150 TLTNTSSVKLPLNFTRDSKQSRMFRNLPEELLNSLASENLTANFHKWSLS 199

Human   201 VKNFTMNEKLKKFFNVLTTNTDGKIEFISTMEGYKYPVYGVQWHPEKAPY 250
            ||||| |||||||||:||.|||||.||||.|||||||:|:|||||||||:
Rat     200 VKNFTENEKLKKFFNILTVNTDGKTEFISSMEGYKYPIYAVQWHPEKAPF 249

Human   251 EWKNLDGISHAPNAVNPAFYLAEFFVNEARKKNHHFKSESEEEKALIYQF 300
            |||.| |||||||||...||||.||:.|| |.:|||..| ||...|||||
Rat     250 EWKKLRGISHAPNAVKTSFYLAKFFISEALKNDHHFENELEETESLIYQF 299

Human   301 SPIYTGNISSFQQCYIFD* 319
            :|:||||||||||.|:|:|
Rat     300 CPVYTGNISSFQQAYMFN* 318
```

FIG. 2. Amino acid sequence comparison of GH deduced from human and rat cDNA. Sequence identity (67.3%) is indicated by a vertical line, and the conserved amino acids are shown in double dot (score, >0.5) and dot (score, >0.1). The leader sequence of rat GH is underlined. The four potential glycosylation sites in the human enzyme are shaded. The GAP program from the Genetics Computer Group was used for sequence comparison.

sequence. This suggested that the first 24 amino acids of the human sequence also may be a leader sequence. As such, the leader sequence would mediate translocation of GH into the endoplasmic reticulum in agreement with its secretion. Analogous leader sequences are responsible for targeting proteins for export via this pathway (13). Analyses of a large number of leader sequences have revealed a common structural motif that occurs in the absence of significant amino acid sequence identity (14). The consensus sequence consists of 20–30 amino acids in which there is usually a positive charge at the N terminus, a central hydrophobic core followed by a more polar region at the carboxyl terminus. The first 24 amino acids in both human and rat GH contain a hydrophobic core region starting at Leu-7 and ending at Leu-15 followed by a more polar carboxyl terminus with serine at position 24 in human and rat sequences and a Glu at position 22 in the human leader sequence. At the N terminus the rat sequence contains a positively charged Arg at position 6 but a positive charge is not seen in the human sequence.

Other common domains occur in the human and rat sequences. In rat GH, there are seven consensus Asn-linked glycosylation sites (4). There are four potential Asn-containing glycosylation sites in the deduced human GH sequence (15), three of which are the same as those in the rat GH sequence (Fig. 2, shaded boxes). The catalytic dependence of these enzymes on thiol reagents suggested the possibility that they are cysteine proteases (2, 4). In the mature human enzyme there are four cysteines, two of which (Cys-43 and Cys-134) are also conserved in the mature rat enzyme. Cys-134 is in a 17-amino acid sequence that is identical in both enzymes except at position 138 where a Phe is replaced by a Leu in the rat enzyme. This high sequence conservation suggests that Cys-134 may be catalytically essential.

On the basis of homology to the N-terminal amino acid sequence of rat GH, mature human GH cDNA, which is predicted to start at Arg-25 (base 132, Fig. 1), was generated by PCR and cloned into the pET24a vector. This expressed human GH was immunoreactive with the antibody raised against the purified rat GH from H35 cells (Fig. 3, lane b). The cDNA encoding mature rat GH without the leader sequence was also cloned into the pET24a vector and expressed in *E. coli* cells. The expressed rat enzyme was immunoreactive with the same antibody against rat GH (Fig. 3, lane c). The molecular weight of the immunoreactive bands is ≈35 kDa for both the expressed rat and human enzyme. The calculated molecular weights for the rat and human enzymes are 33 and 36 kDa, respectively. Thus, the size of the expressed protein agrees with the theoretical values for the human and rat GH.

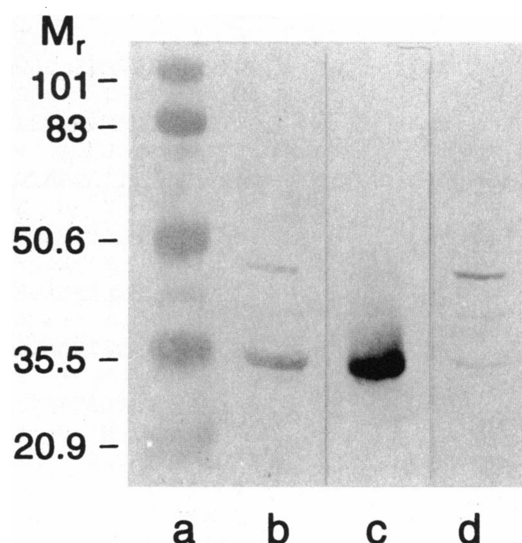

FIG. 3. Western blot analysis of expressed rat and human GH demonstrating the cross reactivity of the antibody with both rat and human enzyme. The antibody raised against rat GH was used at 1:10,000 dilution. Lanes: a, prestained marker proteins; b, expressed human GH (1:10 dilution of cell lysate); c, expressed rat GH (1:100 dilution of cell lysate), d, *E. coli* control (1:10 dilution of cell lysate). Cell lysates were prepared as described.

The catalytic conversion of the substrates methotrexate di- and pentaglutamate was examined with the human and rat GH expressed in *E. coli* cells. No activity was detected in nontransformed or pET24a-only transformed *E. coli* controls, whereas both the expressed human and rat enzyme cleaved the polyglutamates indicating catalytic activity (Table 1). When the substrate preference with regard to polyglutamate chain length was evaluated, it was found that the rat enzyme was equally active on the di- and pentaglutamate of methotrexate, consistent with the properties of GH purified from H35 cells (2, 7). In contrast, the expressed human enzyme exhibited greater activity on 4-NH_2-10-CH_3PteGlu$_5$ relative to 4-NH_2-10-CH_3PteGlu$_2$, consistent with the properties of the enzyme secreted by human breast cancer, HL60, and HEPG2 hepatoma cell lines (8).

A second feature of catalysis that is different between human and rat enzymes is the pattern of cleavage of 4-NH_2-10-CH_3PteGlu$_5$ (Fig. 4). With the expressed rat enzyme, the only pteroyl-containing product was the monoglutamate (4-NH_2-10-CH_3PteGlu$_1$). The same result was observed with the purified H35 enzyme and has been shown to be due to the selective cleavage of the innermost γ-linkage of the polyglutamate chain via an endoproteolytic mechanism (2). In contrast, the human enzyme produced 4-NH_2-10-CH_3PteGlu$_4$ through -Glu$_1$, consistent with the properties of the enzyme from human tumor cells (8, 16). This cleavage pattern (Fig. 4) can be accomplished if human GH has exopeptidase activity, but some internal cleavage cannot be ruled out. The fact that the first pteridine-containing product to appear is 4-NH_2-10-CH_3PteGlu$_4$, followed by 4-NH_2-10-CH_3PteGlu$_3$, suggests that expressed human GH is primarily an exopeptidase. Precedent exists for this in human systems since GH from human sarcoma cells (16) and jejunal brush border (17) act as exopeptidases. More detailed studies on the expressed human GH are required to determine whether any internal cleavage occurs during the reaction. The results of the experiments shown in Table 1 and Fig. 4 led to the conclusion that the cloned and expressed cDNAs were derived from the respective genes for human and rat GH.

GH has been studied in a number of systems and found to have a variety of hydrolytic sites with regard to pteroylpolyglutamates (1, 2, 14–23). Depending upon the source, it can cleave at the external glutamate (exopeptidase) or at any of the internal amide bonds (endopeptidase). The pteroyl portion of the molecule may affect this process since rat hepatoma GH

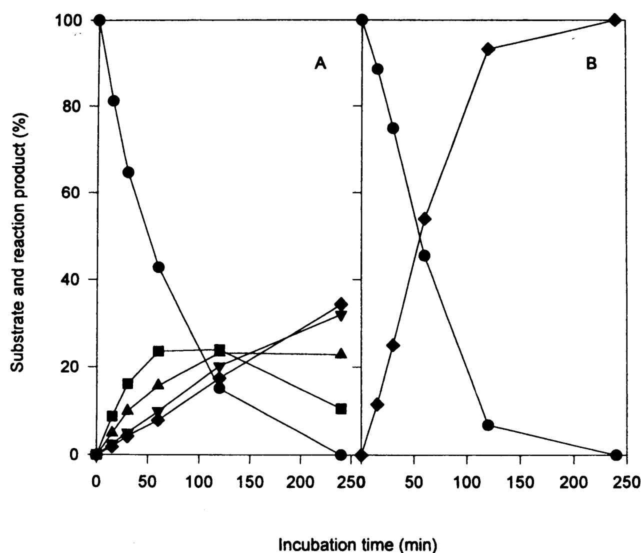

FIG. 4. Time course of the hydrolysis of 4-NH_2-10-CH_3PteGlu$_5$ (●) and appearance of 4-NH_2-10-CH_3PteGlu$_4$ (■), 4-NH_2-10-CH_3PteGlu$_3$ (▲), 4-NH_2-10-CH_3PteGlu$_2$ (▼), and 4-NH_2-10-CH_3PteGlu$_1$ (◆) catalyzed by expressed human (*A*) and rat (*B*) GH.

Table 1. Catalytic characteristics of GH expressed in *E. coli* cells

GH	Substrate	% conversion/h
Human	4-NH_2-10CH_3PteGlu$_2$	15.3
	4-NH_2-10CH_3PteGlu$_5$	74.8
Rat	4-NH_2-10CH_3PteGlu$_2$	28.0
	4-NH_2-10CH_3PteGlu$_5$	27.2
E. coli	4-NH_2-10CH_3PteGlu$_2$	<0.1
	4-NH_2-10CH_3PteGlu$_5$	<0.1

E. coli cells were cultured (10 ml cultures) and induced with 1 mM isopropyl β-D-thiogalactoside overnight in terrific broth medium containing kanamycin (30 μg/ml final concentration) and chloramphenicol (34 μg/ml final concentration).

cleaves 4-NH_2-10-CH_3PteGlu$_5$ selectively at the innermost linkage but acts more randomly on γ-Glu$_5$ (2). These results show that highly homologous cDNA from rat and human sources encode GHs that have different substrate preference and specificity. Presumably these differences are due to the nonidentical regions of the proteins. Thus it appears that this elemental gene sequence may have potential to encode many if not all of the catalytic types of GHs that have been reported in mammalian systems (1, 2, 8, 16–23).

A wide variety of molecular weights have been reported for GH. From mammalian sources the enzyme has been reported to be between 25,000 and 700,000 (1, 2, 17, 18, 20–23). Many are 80,000–120,000 (1, 2, 18, 20, 21). However, these were often measured by gel filtration, and this procedure is known to produce artifacts with glycoproteins such as GH (2, 4). The human and rat enzyme that are described here are both about 35,000 Da for the peptide, which may be the real molecular weight of the protein component of the other GHs that have been characterized. However, in one case, human jejunal brush border, the weight of the native enzyme appears to be 700,000 and, when denatured with sodium dodecyl sulfate/urea, two bands of 145,000 and 115,000 Da are observed (18). The relationship of those structures to the one described here is not yet understood.

The availability of a cDNA and reactive antibodies for human GH will allow several aspects about the role of these enzymes in human systems to be investigated. The mechanism of secretion of GH and the function of secreted GH can be investigated. Using molecular probes it should be possible to inhibit GH and evaluate its role in cellular folylpolyglutamate homeostasis and in the pharmacological activity of classical antifolates. These investigations as well as the large-scale preparation of the human enzyme are underway.

The molecular genetics core of the Wadsworth Center are thanked for the synthesis of oligonucleotides. We acknowledge the use of the Program Manual for the Wisconsin package, Version September 1994, Genetics Computer Group, Madison, WI. This work was supported by National Cancer Institute Grant CA25933 to J.G.

1. McGuire, J. J. & Coward, J. K. (1984) in *Folates and Pterins*, eds. Blakley, R. L. & Benkovic, S. J. (Wiley, New York), Vol. 1, pp. 135–190.
2. Wang, Y., Nimec, Z., Ryan, T. J., Dias, J. A. & Galivan, J. (1993) *Biochim. Biophys. Acta* **1164**, 227–235.
3. Yao, R., Rhee, M. & Galivan, J. (1995) *Mol. Pharmacol.* **48**, 505–511.
4. Yao, R., Nimec, Z., Ryan, T. & Galivan, J. (1996) *J. Biol. Chem.* **271**, 8525–8528.
5. Galivan, J. & Rhee, M. (1995) *Biochem. Pharmacol.* **50**, 1659–1663.
6. Rhee, M. S., Wang, Y., Nair, M. G. & Galivan, J. (1993) *Cancer Res.* **53**, 2227–2230.
7. O'Connor, B. M., Rotundo, R. F., Nimec, Z., McGuire, J. J. & Galivan, J. (1991) *Cancer Res.* **51**, 3874–3881.
8. Rhee, M., Ryan, T. & Galivan, J. (1995) *Cell. Pharmacol.* **2**, 289–292.

9. Lennon, G. G., Auffray, C., Polymeropoulos, M. & Soares, M. B. (1996) *Genomics* **33,** 151–152.
10. Sanger, F., Nicklen, S. & Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. USA* **74,** 5463–5467.
11. Wang, Y., Rotundo, R. F., Nimec, Z., Ryan, T. J. & Galivan, J. (1993) in *Chemistry and Biology of Pteridines and Folates*, eds. Ayling, J. E., Nair, M. G. & Baugh, C. M. (Plenum, New York), pp. 655–658.
12. Kozak, M. (1987) *Nucleic Acids Res.* **15,** 8125–8148.
13. Lyko, F., Martoglio, B., Jungnickel, B., Rapoport, T. A. & Dobberstein, B. (1995) *J. Biol. Chem.* **270,** 19873–19878.
14. Izard, J., Doughty, M. & Kendall, D. (1995) *Biochemistry* **34,** 9904–9912.
15. Winzler, R. J. (1973) in *Hormonal Proteins and Peptides*, ed. Li, C. H. (Academic, New York), pp. 1–15.
16. Li, W. W., Waltham, M., Tong, W., Schweitzer, B. I. & Bertino, J. R. (1993) in *Chemistry and Biology of Pteridines and Folates*, eds. Ayling, J. E., Nair, M. G. & Baugh, C. M. (Plenum, New York), pp. 635–638.
17. Chandler, C. J., Wang, T. T. & Halsted, C. H. (1986) *J. Biol. Chem.* **261,** 928–933.
18. Silink, M., Reddel, R., Bethel, M. & Rowe, P. B. (1975) *J. Biol. Chem.* **250,** 5982–5994.
19. Priest, D. G., Veronee, C. D., Mangum, M., Bednarek, J. M. & Doig, M. T. (1982) *Mol. Cell. Biochem.* **43,** 81–87.
20. Elsenhans, B., Ahmad, O. & Rosenberg, I. H. (1984) *J. Biol. Chem.* **259,** 6364–6368.
21. Wang, T. T., Chandler, C. J. & Halsted, C. H. (1986) *J. Biol. Chem.* **261,** 13551–13555.
22. Samuels, L. L., Goutas, L. J., Priest, D. G., Piper, J. R. & Sirotnak, F. M. (1986) *Cancer Res.* **46,** 2230–2235.
23. Bhandari, S. D., Gregory, J. F., Renuart, D. R. & Merritt, A. M. (1990) *J. Nutr.* **120,** 467–475.