CLONING AND NUCLEOTIDE SEQUENCE OF A FULL-LENGTH cDNA FOR HUMAN LIVER γ-GLUTAMYLCYSTEINE SYNTHETASE¹

Jerry J. Gipp, Chawnshang Chang, and R. Timothy Mulcahy²

Department of Human Oncology, University of Wisconsin-Madison Madison, Wisconsin 53792

Received April 13, 1992

SUMMARY: We have cloned and sequenced a full-length cDNA for human liver γ -glutamylcysteine synthetase (GCS), the rate-limiting enzyme in glutathione biosynthesis. The cDNA consists of 2634 bp containing an open reading frame encoding a protein of 367 amino acids and having a calculated $M_r = 72,773$. The nucleotide sequence of the cDNA for human liver GCS shares an 84% overall similarity with the composite rat GCS sequence deduced from three overlapping partial cDNAs (Yan and Meister, JBC 265: 1588-1593, 1990). The deduced amino acid sequences are 94% similar. Comparison of Northern blots of total RNA isolated from rat kidney or liver with that from human kidney revealed the GCS mRNA to be larger in the human tissue (~ 4.0 kb vs. ~ 3.7 kb). (The sequence for the human liver GCS cDNA has been assigned accession number M90656 in GenBank/EMBL databases.) • 1992 Academic

Glutathione (L- γ -glutamyl-cysteinyl-glycine; GSH) is a ubiquitous non-protein thiol involved in cellular metabolism, maintenance of cellular redox state, protection against oxidative injury and the detoxification of xenobiotics (1). Elevation of intracellular GSH has also been implicated in the resistance of human tumor cells to certain chemotherapeutic agents (2-5). Although GSH elevations are commonly detected in drug-resistant tumor cells, the mechanism(s) responsible for the increase has not been determined. This is at least partially attributable to a lack of appropriate molecular reagents for examining the expression and regulation of key enzymes involved in GSH homeostasis; thus prompting the current investigation.

Glutathione is synthesized from its constituent amino acids in two sequential ATP-requiring reactions catalyzed by γ -glutamylcysteine synthetase (GCS) and glutathione synthase, respectively (1). Under normal conditions, the upper limit of intracellular GSH concentration is regulated by GSH inhibition of GCS, the rate-limiting step in de novo GSH biosynthesis (1,6).

¹Supported from a General Research Support Grant to the University of Wisconsin Medical School from the NIH, Division of Research Facilities and Resources.

²To whom correspondence should be addressed at K4/338, 600 Highland Avenue, Madison, WI 53792.

 γ -Glutamylcysteine synthetase purified from rat kidney can be dissociated into two distinct subunits of M_r 73,000 and 27,700 (7). The larger subunit possesses all the catalytic activity of the isolated protein and can be inhibited by GSH (7). Recently, Yan and Meister (8) determined the nucleotide sequence of the M_r 73,000 subunit of rat kidney GCS following isolation and sequencing of three overlapping partial cDNA clones. The amino acid sequence derived from these cDNAs showed little overall (8%) similarity with the sequence for \underline{E} . \underline{coli} GCS deduced from sequencing of genomic DNA (9).

In this report, we describe the cloning and sequencing of a 2634 bp full-length cDNA for GCS isolated from an adult human liver cDNA library. The sequence of the human GCS cDNA and the deduced amino acid sequence of the corresponding enzyme are compared with those of the rat. Hybridization of total RNA isolated from rat kidney and liver and human liver with probes prepared from both the rat and human sequences was compared.

MATERIALS AND METHODS

Isolation of cDNA clones and DNA sequencing. A λ gt11 cDNA library was constructed using poly(A)⁺ RNA isolated from human kidney. A 1010 bp probe (rkGCS_{PCR}), corresponding to nucleotides 51-1061 of the rat GCS (8) sequence, was generated from rat kidney RNA by polymerase chain reaction using synthetic 20-mer oligonucleotide primers corresponding to nucleotides 51-70 and 1042-1061 of the rat sequence. The probe was radiolabeled with ³²P using the random primer technique and used to screen the human kidney library (10). cDNA was prepared from phage stock by a plate lysis method (10).

Following EcoR1 digestion of phage DNA, a partial (1679 bp) cDNA insert (hkGCS-1) was isolated from low melting agarose gels, purified by phenol/chloroform extraction and cloned into the EcoR1 site of the phagemid pSK-Bluescript. The nucleotide sequence was determined using Sequenase (U.S Biochemicals) according to the manufacturer's instructions, using T3, SK, T7 and KS primers as well as synthetic primers corresponding to internal GCS sequences. Nucleotide sequence was verified by bi-directional sequencing reactions. Sequence analysis was conducted using DNASTAR (Madison, WI) software.

An 0.8 kb Pst1 fragment isolated from hkGCS-1 (Figure 1) was labeled with ³²P and subsequently used to screen a human liver cDNA library (Stratgene).

Northern analysis. Total RNA was isolated from rat kidney and rat liver by the guanidinium method according to Chomczynski and Sacchi (11). Human kidney RNA was purchased from Clontech. Northern blots were prepared according to Sambrook, Fritsch and Maniatis (10).

RESULTS AND DISCUSSION

Initially, four clones were isolated from the human kidney cDNA library following screening with the radiolabeled rkGCS_{PCR} probe (solid box in rkGCS, Figure 1). However, of the four only one, the 1679 bp hkGCS-1, shared significant sequence homology with the published rat GCS sequence. Comparison with the rat sequence revealed that hkGCS-1 consisted of coding sequences for 387 amino acids on the carboxyl end of the protein and 398 nucleotides

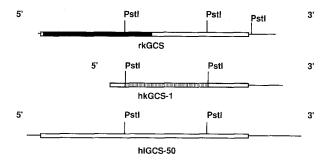


Figure 1. Restriction maps for rat (rkGCS) kidney cDNA, human kidney clone hkGCS-1 and the full-length GCS cDNA clone, hlGCS-50 isolated from human liver. Boxes represent open reading frame; the solid box in rGCS corresponds to the PCR fragment (rkGCS_{PCR}) used to screen the human kidney cDNA library. The striped box in hkGCS-1 corresponds to the 0.8 kb Pst1 fragment utilized to screen the human liver library.

of non-translated sequence containing a putative polyadenylation signal (AATAAC) (8). Repeated attempts to clone a full-length cDNA from the human kidney library were unsuccessful so the cloning strategy was modified to include screening of a human liver library. An 0.8 kb Pst1 fragment (stripped box in hkGCS-1, Figure 1) corresponding to sequences within the open reading frame of hkGCS-1 were used to identify positive clones. A total of 50 clones were isolated from the liver cDNA library and subjected to restriction analysis. Of these positive clones, only one, hlGCS-50, contained a full length insert. A schematic representation of the rat GCS cDNA, hkGCS-1, and hlGCS-50 are shown in Figure 1.

The nucleotide sequence of the sense strand of hlGCS-50 and the deduced amino acid sequence are compared with the corresponding rat sequences in Figure 2. The human cDNA consists of 92 bp of untranslated 5' sequence, a 1911 bp open reading frame starting at the ATG codon at position 93 and 628 bp of 3' sequence. The two cDNAs share an 89% nucleotide sequence similarity within the open reading frame, and 84% overall. The sequence AATAAC, suggested to be the putative polyadenylation signal in the rat GCS cDNA (8), was located at nucleotides 2221-2226 in hlGCS-50. However, the polyadenylation concensus sequence, AATAAA, was identified at nucleotides 2493-2498 in the distal 3' region of hlGCS-50; a sequence extending beyond that included in the rat cDNA as reported by Yan and Meister (8).

Both the rat and human cDNAs contain an open reading frame 1911 bp in length, encoding proteins of 637 amino acid residues. The calculated molecular weight of the protein encoded by hIGCS-50 is 72,773 Daltons, in good agreement with the estimated molecular weight of the large subunit of rat GCS (M_r =73,000). The deduced amino acid sequences share 94% similarity with the deduced sequence for the rat enzyme. These results indicate that, unlike the situation with E coli GCS (8), a high degree of conservation exists between the human and rat proteins.

hlgcs rkgcs	1: 1:	GGCACGAGGCTGAGTGTCCGTCTCGCGCCCGGAAGCGGGCGACCGCCGTCAGCCCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	
hGCS hlGCS rkGCS rGCS	1: 93: 36: 1:	Met Gly Leu Leu Ser Gln Gly Ser Pro Leu Ser Trp Glu Glu Thr Lys Arg His Ala Asp His Val Arg Arg His ATG GGG CTG CTG TCC CAG GGC TCG CCG CTG AGC TGG GAA ACC AAG CGC CAT GCC GAC CAC GTG CGG CAC	
hGCS hlGCS rkGCS rGCS	26: 168: 111: 26:	Gly Ile Leu Gln Phe Leu His Ile Tyr His Ala Val Lys Asp Arg His Lys Asp Val Leu Lys Trp Gly Asp Glu GGG ATC CTC CAG TTC CTG CAC ATC TAC CAC GCC GTC AAG GAC CGG CAC AAG GAC GTT CTC AAG TGG GGC GAT GAGC	
hGCS hlGCS rkGCS rGCS	51: 243: 186: 51:	Val Glu Tyr Met Leu Val Ser Phe Asp His Glu Asn Lys Lys Val Arg Leu Val Leu Ser Gly Glu Lys Val Leu GTG GAA TAC ATG TTG GTA TCT TTT GAT CAT GAA AAT AAA AAA GTC CGG TTG GTC CTG TCT GGG GAG AAA GTT CTT G GC	
hGCS hlGCS rkGCS rGCS	76: 318: 261: 76:	Glu Thr Leu Gln Glu Lys Gly Glu Arg Thr Asn Pro Asn His Pro Thr Leu Trp Arg Pro Glu Tyr Gly Ser Tyr GAA ACT CTG CAA GAG GAG GAA AGG GAA AGG ACA AAC CCA AAC CCT ACC CTT TGG AGA CCA GAG TAT GGG AGT TAC	
hGCS hlGCS rkGCS rGCS	101: 393: 336: 101:	Met Ile Glu Gly Thr Pro Gly Gln Pro Tyr Gly Gly Thr Met Ser Glu Phe Asn Thr Val Glu Ala Asn Met Arg ATG ATT GAA GGG ACA CCA GGA CAG CCC TAC GGA GGA ACA ATG TCC GAG TTC AAT ACA GTT GAG GCC AAC ATG CGA	
hGCS hlGCS rkGCS rGCS	126: 468: 411: 126:	Lys Arg Arg Lys Glu Ala Thr Ser Ile Leu Glu Glu Asn Gln Ala Leu Cys Thr Ile Thr Ser Phe Pro Arg Leu AAA CGC CGG AAG GAG GCT ACT TCT ATA TTA GAA GAA AAT CAG GCT CTT TGC ACA ATA ACT TCA TTT CCC AGA TTA	
hGCS hlGCS rkGCS rGCS	151: 543: 486: 151:	Gly Cys Pro Gly Phe Thr Leu Pro Glu Val Lys Pro Asn Pro Val Glu Gly Gly Ala Ser Lys Ser Leu Phe Phe GGC TGT CCT GGG TTC ACA CTG CCC GAG GTC AAA CCC AAC CCA GTG GAA GGA GGA GCT TCC AAG TCC CTC TTC TTT CC A CAG TAT	
hGCS hlGCS rkGCS rGCS	176: 618: 561: 176:	Pro Asp Glu Ala Ile Asn Lys His Pro Arg Phe Ser Thr Leu Thr Arg Asn Ile Arg His Arg Gly Glu Lys CCA GAT GAA GCA ATA AAC AAG CAC CCT CGC TTC AGT ACC TTA ACA AGA AAT ATC CGA CAT AGG AGA GGA GAA AAG C C Gly	
hGCS hlGCS rkGCS rGCS	201: 693: 636: 201:	Val Val Ile Asn Val Pro Ile Phe Lys Asp Lys Asn Thr Pro Ser Pro Phe Ile Glu Thr Phe Thr Glu Asp Asp GTT GTC ATC AAT GTA CCA ATA TIT AAG GAC AAG AAT ACA CCA TCT CCA TTT ATA GAA ACA TTT ACT GAG GAT GAT G	
hGCS hlGCS rkGCS rGCS	226: 768: 711: 226:	Glu Ala Ser Arg Ala Ser Lys Pro Asp His Ile Tyr Met Asp Ala Met Gly Phe Gly Met Gly Asn Cys Cys Leu GAA GCT TCA AGG GCT TCT AAG CCG GAT CAT ATT TAC ATG GAT GCC ATG GGA TTT GGA ATG GGC AAT TGC TGT CTCGAACACCC	
hGCS hlGCS rkGCS rGCS	251: 843: 786: 251:	Gln Val Thr Phe Gln Ala Cys Ser Ile Ser Glu Ala Arg Tyr Leu Tyr Asp Gln Leu Ala Thr Ile Cse Pro Ile CAG GTG ACA TTC CAA GCC TGC AGT ATA TCT GAG GCC AGA TAC CTT TAT GAT CAG TTG GCT ACT ATC TGT CCA ATT	
hGCS hlGCS rkGCS rGCS	276: 918: 861: 276:	Val Met Ala Leu Ser Ala Ala Ser Pro Phe Tyr Arg Gly Tyr Val Ser Asp Ile Asp Cys Arg Trp Gly Val Ile GTT ATG GCT TTG AGT GCT GCA TCT CCC TTT TAC CGA GGC TAT GTG TCA GAC ATT GAT TGT CGC TGG GGA GTG ATT	
hGCS hlGCS rkGCS rGCS	301: 993: 936: 301:	Ser Ala Ser Val Asp Asp Arg Thr Arg Glu Glu Arg Gly Leu Glu Pro Leu Lys Asn Asn Asn Tyr Arg Ile Ser TCT GCA TCT GTA GAT GAT AGA ACT CGG GAG GGA GGA CTG GAG CCA TTG AAG AAC AAT AAC TAT AGG ATC AGT	
hGCS hlGCS rkGCS rGCS	326: 1068: 1011: 326:	Lys Ser Arg Tyr Asp Ser Ile Asp Ser Tyr Leu Ser Lys Cys Gly Glu Lys Tyr Asn Asp Ile Asp Leu Thr Ile AAA TCC CGA TAT GAC TCA ATA GAC AGC TAT TTA TCT AAG TGT GGT GAG AAA TAT AAT GAC ATC GAC TTG ACG ATA GTG	
hGCS hlGCS rkGCS rGCS	351: 1143: 1086: 351:	Asp Lys Glu Ile Tyr Glu Gln Leu Leu Glu Glu Gly Ile Asp His Leu Leu Ala Gln His Val Ala His Leu Phe GAT AAA GAG ATC TAC GAA CAG CTG TTG CAG GAA GGC ATT GAT CAT CTC CTG GCC CAG CAT GTT GCT CAT CTC TTTC -CG	

Figure 2. cDNA sequences and deduced amino acid sequences for rkGCS and hlGCS-50. The nucleotide and amino acid sequences are numbered on the left. The rkGCS residues that are identical to corresponding residues in hlGCS-50 are indicated by hyphens (-). Missing residues and gaps are shown as dots (.). * indicates the stop codon. The sequence for the human GCS cDNA has been assigned accession number M90656.

```
Ile Arg Asp Pro Leu Thr Leu Phe Glu Glu Lys Ile His Leu Asp Asp Ala Asn Glu Ser Asp His Phe Glu Asn ATT AGA GAC CCA CTG ACA CTG TTT GAA GAG AAA ATA CAC CTG GAT GAT GCT AAT GAG TCT GAC CAT TTT GAG AAT
hGCS
      376:
hlgcs
     1218:
rkGCS
rGCS
hGCS
      401:
             Ile Gln Ser Thr Asn Trp Gln Thr Met Arg Phe Lys Pro Pro Pro Asn Ser Asp Ile Gly Trp Arg Val Glu
             ATT CAG TCC ACA AAT TGG CAG ACA ATG AGA TTT AAG CCC CCT CCA AAC TCA GAC ATT GGA TGG AGA GTA GAA
rkGCS
     1236:
rGCS
             Phe Arg Pro Met Glu Val Gln Leu Thr Asp Phe Glu Asn Ser Ala Tyr Val Val Phe Val Val Leu Leu Thr Arg TTT CGA CCC ATG GAG GTG CAA TYA ACA GAC TTT GAG AAC TCT GCC TAT GTG GTG TTT GTG GTA CTG CTC ACC AGA
hGCS
      426-
hlgcs
     1368:
rGCS
      426:
             hGCS
      451:
HIGGS
     1443:
rkGCS
rGCS
hGCS
             Arg Asp Ala Val Leu Gln Gly Met Phe Tyr Phe Arg Lys Asp Ile Cys Lys Gly Gly Asn Ala Val Val Asp Gly AGA GAT GCT GTC TTG CAG GGA ATG TTT TAT TTC AGG AAA GAT ATT TGC AAA GGT GGC AAT GCA GTG GTG GAT GGT
higcs
     1518:
rkGCS
             rGCS
             501:
hGCS
hlgcs
     1593:
                Ser - - - Thr - Ser - Pro Ser -
rGCS
             Gly Lys Glu Gly Val Phe Pro Gly Leu Ile Pro Ile Leu Asn Ser Tyr Leu Glu Asn Met Glu Val Asp Val Asp GGG AAG GAA GGT GTG TTT CCT GGA CTG ATC CCA ATT CTG AAC TCT TAC CTT GAA AAC ATG GAA GTG GAT GTG GAC
hlgcs
     1668:
rkGCS
     1611:
rGCS
hGCS
      551:
             Thr Arg Cys Ser Ile Leu Asn Tyr Leu Lys Leu Ile Lys Lys Arg Ala Ser Gly Glu Leu Met Thr Val Ala Arg ACC AGA TGT AGT ATT CTG AAC TAC CTA AAG CTA ATT AAG AAG AGA GCA TCT GGA GAA CTA ATG ACA GTT GCC AGA
hlgcs
rkGCS
     1686:
              rGCS
      551:
      576:
             hGCS
hlgcs
     1818:
rkGCS
hGCS
      601:
             Leu Ile Leu Lys Cys Asn Gln Ile Ala Asn Glu Leu Cys Glu Cys Pro Glu Leu Leu Gly Ser Ala Phe Arg Lys
             CTT ATT TTG AAG TGT AAC CAA ATT GCA AAT GAA TTA TGT GAA TGC CCA GAG TTA CTT GGA TCA GCA TTT AGG AAA
hlGCS
     1893:
rkGCS
     1836:
rGCS
      601:
             hGCS
      626:
higgs
     1968:
rkGCS
     1911:
rGCS
      626:
             hlgcs
rkGCS
     1980:
               GCTTTCTTTGGTAGGTAAATCTAGAGTTTATACAGTGTACATGTACATAGTAAAGTATTTTT..G...ATTAACAATGTATTTTAATAACA.T..ATC
rkGCS
     2056:
                          TAAAGTCATCATGAACTGGCTIGTACATTTTTAAATTCTTACTCTGGAGCAACCTACTGTCTAAGCAGTTTTGTAAATGTACTGGAAATTGTACAATAC
rkGCS
     2147:
             TIGCATTCCAGAGTTAAAATGTTTACTGTAAATTTTTGTTCTTTTAAAGACTACCTGGGACCTGATTTATTGAAAATTTTTCTCTTTAAAAAACATTTTCT
higcs 2331:
             hlgcs 2430:
hlgcs 2529:
             hlGCS 2628:
             AAAAAA
```

Figure 2-Continued

The 1.0 kb rkGCS_{PCR} fragment and hkGCS-1 cDNA were used to probe Northern blots of total RNA isolated from rat kidney, rat liver and human kidney. Both probes identified similar RNA species (Figure 3). GCS mRNA was more abundant in rat kidney than in rat liver, consistent with previous observations (8). Both probes hybridized with a single 4.0 kb band in

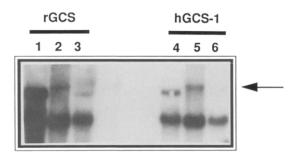


Figure 3. Northern blot analysis of total RNA isolated from rat kidney (lanes 1 and 4), rat liver (lanes 2 and 5) and human kidney (lanes 3 and 6). Ten μ g of total RNA was loaded per lane. Lanes 1-3 were probed with the 1.0 kb rkGCS_{PCR} fragment while lanes 4-6 were probed with hkGCS-1. The arrow indicates the position of the human kidney GCS message. The lower band in each of the lanes corresponds to β -actin message.

RNA isolated from human kidney. This RNA was approximately 300 base pairs larger than the species identified in the RNA isolated from the rat tissues. This species-specific size differential is similar to that reported for γ -glutamyl transpeptidase, another GSH metabolizing enzyme (12).

Glutathione plays a prominent role in cellular homoestasis and is particularly important in the response of cells to various types of injurious agents. The ability of cells to survive exposure to xenobiotics capable of inducing oxidative damage, ionizing radiation, heavy metals or alkylating agents has been correlated with intracellular levels of GSH. While the regulation of GSH levels in response to these insults has not been defined, there is ample evidence in the toxicological literature to support the role of GCS in many GSH-related adaptive responses (13-16). Defects in GSH synthesis and/or metabolism, particularly in red blood cells, have also been associated with various disease states (17). The availability of a full-length cDNA clone for human GCS will permit the investigation of GSH regulation at the molecular level; an approach which should provide insights to alternative means of manipulating GSH levels to experimental and therapeutic advantage.

REFERENCES

- 1. Meister, A. (1989) In: Coenzymes and Cofactors. Glutathione: Chemical, Biochemical and Medical Aspects (D. Dolphin, R. Poulson, O. Avramovic, Eds.), Vol. III, pp. 367-474. John Wiley and Sons, New York.
- 2. Suzukake, K., Vistica, B. P. and Vistica, D. T. (1983) Biochem. Pharm. 32, 165-167.
- 3. Arrick, B. A. and Nathan, C. F. (1984) Cancer Res. 44, 4224-4232.
- 4. Russo, A. and Mitchell, J. B. (1985) Cancer Treat. Reports 69, 1293-1296.
- 5. Russo, A., Carmichael, J., Friedman, N., DeGraff, W., Tochner, Z., Glatstein, E. and Mitchell, J.B. (1986) Int. J. Radiat. Oncol. Biol. Phys. 12, 1347-1354.

- 6. Deneke, S. M. and Fanburg, B. L. Am. J. Phys. 257, L163-L173.
- 7. Seelig, G. F. and Meister, A. (1985) Methods Enzymol. 113, 379-390.
- 8. Yan, N. and Meister, A. (1990) J. Biol. Chem. 265, 1588-1593.
- 9. Watanabe, K., Yamano, Y., Murata, K. and Kimura, A. (1986) Nucleic Acids Res. 14, 4393-4400.
- 10. Sambrook, J., Fritsch, E. F. and Maniatis, T. (1989) Molecular Cloning: A Laboratory Manual. Cold Spring Harbor Laboratory Press.
- 11. Chomzynski, P. and Sacchi, N. (1987) Anal. Biochem. 162, 156-159.
- 12. Goodspeed, D. C., Dunn, T. J., Miller, C. D. and Pitot, H. C. (1989) Gene 76, 1-9.
- 13. Chung, A. and Maines, M. D. (1981) Biochem. Pharmacol. 23, 3217-3223.
- 14. Ogino, T., Kawabata, T. and Awai, M. (1989) Biochim. Biophys. Acta 1006, 131-135.
- 15. Ohno, K. and Hirata, M. (1990) Biochem. Biophys. Res. Commun. 168, 551-557.
- 16. Sriram, R. and Ali-Osman, F. (1991) Proc. Amer. Assoc. Cancer Res. 32:428.
- Mentzer, W. C. and Glader, B. E. (1989) In The Hereditary Hemolytic Anemias (W. C. Mentzer and G. M. Wagner, Eds.), pp. 267-318. Churchill Livingstone, New York, NY.