Vol. 209, No. 1, 1995 BIOCHEMICAL AND BIOPHYSICAL RESEARCH COMMUNICATIONS
April 6, 1995 Pages 227-233

IDENTIFICATION OF A PUTATIVE ANTIOXIDANT RESPONSE ELEMENT IN THE 5'-FLANKING REGION OF THE HUMAN γ-GLUTAMYLCYSTEINE SYNTHETASE HEAVY SUBUNIT GENE[1]

R. Timothy Mulcahy[2] and Jerry J. Gipp

Department of Human Oncology, University of Wisconsin Medical School
600 Highland Avenue K4/334 CSC
Madison, WI 53792

Received February 13, 1995

SUMMARY: We have cloned the human γ-glutamylcysteine synthetase heavy subunit gene (GCS_h) from a P1 library and isolated a 5.5kb fragment (P1-GCS5') from the 5'- end of the P1 clone. P1-GCS5' has been sequenced from -1460 to +547. Multiple transcription start sites were identified by primer extension and S1 nuclease protection. Two start sites were identified by primer extension analysis within 23 bp (+1 and +10) of a consensus TATAAAA box; all sequences were numbered relative to the 5'-most of these two sites. Two additional major start sites were identified at -106 and +398. This latter site was the most prominent of all the initiation sites. In addition to a TATA box, the promoter contains a CCAAT box at -125 and GC boxes up- and down-stream of the TATAAAA. In addition, the first few hundred base pairs of the sequence are highly GC-rich (~75%). This sequence also contains several Sp-1 binding sites, a consensus AP-1 site and several AP-1-like binding sites, as well as putative AP-2 sites. A consensus metal responsive element (MRE) was identified at position +198. Sequence analysis also identified a putative core (5'-TGACnnnGCA-3') antioxidant response element (ARE) at -862 to -853. As is typical of other AREs, a second AP-1-like sequence is located adjacent to the core sequence. These results suggest that GCS_h gene expression in response to oxidative challenge may be regulated through an antioxidant response element similar to those recently detected in the promoter region of several Phase II enzymes. (The sequence of the 5'-flanking region of the human GCS heavy subunit gene has been assigned accession number L39773.)

Glutathione (GSH) is an abundant cellular non-protein sulfhydryl which serves as an important intracellular antioxidant providing protection against reactive oxygen species; functions to maintain cellular redox balance and is involved in the detoxification of xenobiotics, organic peroxides and heavy metals either through direct conjugation with reactive species or in reactions catalyzed by glutathione S-transferases and glutathione peroxidase (1,2). In addition to its protective role as the predominant cellular antioxidant GSH may also be involved in the regulation of stress-response gene expression either directly, or by virtue of its capacity to regulate the redox

[1]Supported by NIH grant CA57549, awarded to RTM.

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

0006-291X/95 $5.00

status of transcription factors (3-5). Elevations of intracellular GSH levels frequently result from xenobiotic exposure. Recent evidence suggests that this increase is related to increased activity of GCS (6-10), the rate-limiting enzyme in the *de novo* synthesis of GSH (2). Although altered GCS activity could result from post-translational modifications, accumulating evidence suggests a prominent role for an increase in the rate of gene transcription.

In cells the GCS holoenzyme exists as a dimer composed of a heavy (GCS_h; 73 kDa) and a light (GCS_l; 28 kDa) subunit which can be dissociated under non-denaturing conditions (11). The heavy subunit possesses all of the catalytic activity and is the site of GSH feedback inhibition (11). The light subunit has recently been shown to play an important regulatory function controlling the kinetic properties of the heavy subunit (12). We previously reported the cloning of the cDNAs for both subunits of human GCS (13,14), and now report cloning and sequencing of the promoter and 5'-flanking sequence of the GCS heavy subunit gene. These studies suggest the presence of a putative Antioxidant Responsive Element (ARE) within the promoter region of the gene.

MATERIAL AND METHODS

P1 library screening. A single clone (DMPC-HFF#1,545H11) containing the genomic sequence for the human heavy GCS subunit was isolated from the DuPont Merck Pharmaceutical Company human foreskin fibroblast P1 library #1 by Genome Systems, Inc. (St. Louis, MO) utilizing PCR primers corresponding to sequences in the GCS_h cDNA (13). The clone was designated as clone 762. The sequence of the sense and anti-sense primers used to identify clone 762 were: 5'-ACGAGGCTGAGTGTCCGTCT-3' and 5'-GGCGCTTGGTTTCCTCCCAG-3', respectively. PCR amplification utilizing these primers resulted in amplification of a single product 141bp in length. P1-GCS5', which contains a portion of the coding sequence for GCS_h and >5.4 kb of 5'-flanking sequence, was isolated from clone 762 DNA by BamH1 digestion and subcloned into pSK-Bluescript.

Sequencing. Sequencing reactions utilized Sequenase (U.S. Biochemical) and synthetic oligonucleotides (20-mers) primers corresponding to internal sequences. The nucleotide sequence was verified by multiple bi-directional seqeuncing reactions.

Primer extension and S1 nuclease protection. Nuclease protection and primer extension studies employed standard techniques. The relative position of the four different primers used for extension assays and two probes used for S1 nuclease protection are illustrated in Figure 2. Both assays made use of total RNA isolated from DU-145/M4.5 human prostate carcinoma cells (15), which express high levels of GCS_h mRNA (16). In addition, total RNA isolated from HepG2 human hepatoblastoma cells, HT-29 colon carcinoma cells and Mel 28 human melanoma cells was used in some assays.

RESULTS

Sequence of the 5'-flanking region of the hGCS$_h$ subunit gene: A 5.5 kb BamH1 fragment was isolated from a human P1 library clone (clone 762), which by Southern analysis was determined to contain the entire genomic sequence of the human GCS_h gene. Sequence analysis indicates that this fragment, designated P1-GCS5', contains 82 bp of coding sequence, including the ATG start codon, and approximately 5.4 kb of 5'-flanking sequence. The sequence of the 3'-most 2007 bases of P1-GCS5' is shown in Figure 1. The first several hundred bases upstream of the start codon are extremely GC-rich (>75%). Canonical TATA and CAAT boxes were identified, as were putative binding sites for AP-1, Sp-1 and Ap-2 transcription factors.

Figure 1. Sequence of P1-GCS5' from -1437 to +548. Relative locations of the TATAAAA (elipse) and CCAAT sequences are shown, as are the *Inr* sequence, metal responsive element (MRE) and the putative Antioxidant Responsive Element (ARE). AP-1 binding sites are indicated by solid arrows; AP-1-like binding sites by dashed arrows. Sp-1 sites are designated by solid underlining while AP-2 sites are indicated as dashed underlines. Location of transcription start sites are indicated by bent arrowheads. The position of the start site designated as +1 for numbering convention is indicated. The start codon and coding sequence included in P1-GCS5' are shown as the dark box at the 3' end of the sequence. The sequence was confirmed by multiple overlapping, bi-directional sequencing reactions.

Consensus GC-box motifs were located 15 bases upstream and 230 bases downstream of the TATA box. A sequence homologous to the functional Initiator (*Inr*) element (5'-CTCAnTCT-3'), which can function alone or in concert with a TATA box and other upstream elements to direct precise transcription initiation by RNA polymerase II from a single base within the element itself (17,18), was also identified.

Transcription start sites: Primer extension and S1 nuclease protection assays were performed to determine the sites of GCS_h gene transcription initiation using total RNA isolated from DU-145/M4.5 prostate carcinoma cells. The results of these experiments are summarized in Figure 2. Primer extension analyses revealed that multiple start sites were utilized in transcription of the gene. As is typical of TATA-directed initiation, a start site was localized 20 bases downstream of the TATA box. In addition, a second start site was also localized in close proximity to the TATA sequence, 9 bases downstream from the first. S1 nuclease protection identified a protected species originating two bases downstream from this second TATA-associated start site. The site nearest the TATA box is located just upstream of the 5' end of the S1 probe utilized in protection assays so its utilization was not confirmed by nuclease protection using this probe. Nevertheless, this site was evident in all primer extension determinations, including those utilizing RNA isolated from several other cell lines (data not shown).

In keeping with the convention commonly applied to numbering sequences relative to a TATA-associated initiation start site, we designated the start site nearest the TATA sequence as the +1 position, despite the fact that an additional start site 86 bases upstream of the TATAAAA sequence is apparently also utilized in GCS_h gene transcription. Using this numbering convention, GCS_h transcripts are initiated from positions -106, +1, +10, +398 and +414 (Fig. 2). Position +414 is a relatively minor start site which was not always evident by primer extension and

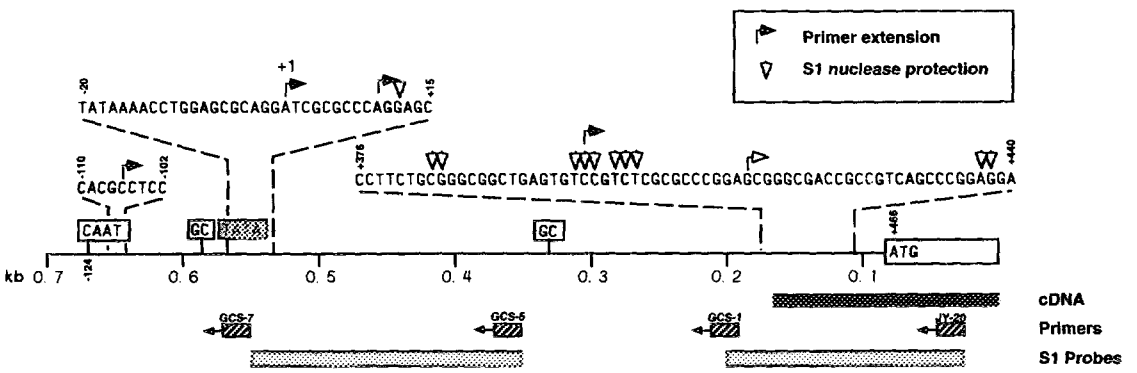

Figure 2. Detailed mapping of transcription start sites in the 5'-flanking region of the GCS_h subunit gene. Relative position of primers and probes used for pimer extension (bent arrowheads) and S1 nuclease protection (open downward arrowheads), respectively, are shown. The solid horizontal box indicates the limits of the 5'-end of the human GCS_h subunit cDNA previously isolated from a human liver cDNA library (13); the open box represents coding sequence. The solid bent arrowheads indicate the position of major transcription start sites, while a minor site at position +404 is indicated by an open bent arrowhead. The first start site downstream from the TATA box was designated +1.

which was not confirmed in S1 nuclease protection assays. In contrast, +398 represents the most prominent band detected by primer extension in each of the cell lines examined. Several protected species corresponding to +398 and bases immediately adjacent to this position were evident in all S1 protection assays. This site is located within a highly GC-rich region of P1-GCS5'. Initiation from position -106 was likewise consistently evident in each extension reaction regardless of the cellular source of RNA. No evidence of *Inr*-associated transcription initiaion was detected. However, the primers selected were not optimal for detecting initiation from this site, so its utilization cannot be ruled out on the basis of these experiments.

Putative Antioxidant Response Element (ARE): In addition to the multiple transcription binding sites located in the 5'-flanking region of the GCS_h gene, a consensus metal responsive element (MRE) was located at +198. Of potential significance with respect to regulation of the GCS_h gene in response to conditions which induce oxidative stress, a putative ARE complex was identified at -853 to -879 (Fig. 1). This complex consists of the consensus ARE core motif (5'-TGACnnnGC-3') described by Rushmore et al (19), an adjacent AP-1-like sequence and an ETS binding site, similar to the ARE arrangements in several other oxidative-stress responsive genes (20 and references therein). The sequence of the putative GCS_h ARE and arrangement of its constituent components is shown in Figure 3 aligned with the sequence of AREs from other genes. The orientation of the ARE in the GCS_h gene is opposite of that of AREs described in other genes. Consequently, the sequence shown for the GCS_h ARE in Figure 3 is that of the non-coding strand in order to facilitate sequence comparisons. The function of the ARE in other genes has been shown to be independent of orientation (29).

DISCUSSION

Transcription of the GCS_h subunit gene is apparently complex and involves transcription from multiple distinct initiation sites, utilizing TATA-dependent and -independent mechanisms.

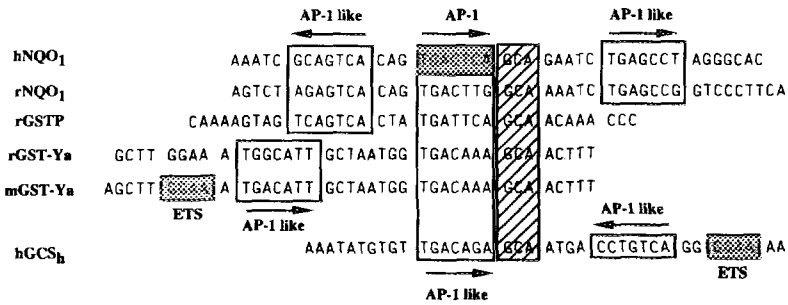

Figure 3. Sequence comparisons of AREs. The sequence of the putative ARE identified in the 5'-flanking region of the GCS_h subunit gene is aligned with ARE sequences from several other genes to illustrate sequence similarities. hNQO1, human NADPH quinone oxidoreductase gene; rNQO1, rat NADP(H) quinone oxidoreductase gene; rGSTP, rat GST-P gene; r GST-Ya, rat GST-Ya gene; mGST-Ya, murine GST-Ya gene; hGCSh, human GCS_h subunit gene. In the case of $hGCS_h$, the sequence shown is of the non-coding strand. (Modified from Jaiswal, (20)).

Despite the presence of a functional TATA box, the promoter region also has features common to many "houskeeping genes", including a high GC content, multiple Sp-1 binding sites and multiple start sites spread over a region of several hundred bases. The studies reported here examined constitutive expression of the gene in cells maintained in exponential growth. It is conceivable that differential utilization of the various start sites we have identified or the use of additional sites might result when transcription of the gene is induced, as for example by chemotherapeutic drugs, heavy metals or antioxidants.

Increases in cellular GSH levels frequently accompany exposure to agents or conditions which result in up-regulation of other GSH-associated enzymes, including glutathione S-transferases (GST), which conjugate specific xenobiotics with GSH as part of detoxification pathways. As many of the same agents which result in elevated GSH levels up-regulate GST expression and the expression of other Phase II enzymes by activation of AREs (20,21), we hypothesized that increased GSH levels might result from transcriptional up-regulation of GCS expression mediated by an ARE-like element in the promoter region of the GCS_h and/or GCS_l genes. The ARE and the EpRE, a functional ARE-equivalent found in the regulatory region of the murine GST-Ya gene (22,23), activate gene expression in response to diverse chemical inducers (19,21,24), including Michael reaction acceptors, quinones, diphenols, peroxides, isothiocyanates, heavy metals and arsenicals, apparently via a redox sensitive signal transduction pathway. The cloning data presented in this report do indeed suggest that an ARE is present in the regulatory region of the GCS_h subunit gene. It is possible therefore that this element might exert control over transcription following exposure of cells to various representatives of the aforementioned classes of compounds, several of which have been documented to increase intracellular GSH levels (6,7,25-27). Although the functional properties of this putative responsive element have not as yet been established in the case of the GCS_h gene, the close homology with functional AREs cloned from other Phase II enzymes of human and rodent origin (Figure 3) strongly supports the hypothesis that the complex is functional.

In addition to our cloning data, the hypothesis that a functional ARE exists in the regulatory region of either or both of the GCS subunit genes is also substantiated by recent experimental evidence. Shi et al (26) hypothesized that GCS would be inducible by oxidative stress and demonstrated that exposure of rat lung epithelial cells to the pro-oxidant, 2,3-dimethoxy-1,4-napthoquinone, results in elevated GSH levels and transcriptional up-regulation of GCS_h expression. These observations were interpreted as suggesting the posible existence of an ARE-like element in the 5'-flanking region of the gene. Similarly, Liu et al (6) predicted the existence of an EpRE in the murine GCS gene based on results demonstrating that GCS activity in mouse hepatoma cells was increased following exposure to the synthetic indolic antioxidant, 5,10-dihydroindeno[1,2-*b*]indole. Borroz et al also speculated that the increase in GCS_h mRNA detected in mouse liver following dietary treatment with the antioxidant 2(3)-tert-butyl-4-hydroxyanisole (BHA) might result via activation of an ARE (27). Although not conclusive, the collective evidence, particularly the cloning data, are consistent with the existence of a functional ARE among the *cis*-acting elements directing expression of the human GCS_h gene.

Like AREs in other promoters, the GCS_h ARE consists of multiple components including two adjacent AP-1-like sequences, one containing the core sequence shown to be required for basal

and induced expression (19), and an ETS site which has recently been shown to strongly influence the activity of the adjacent AP-1-like sites in the mouse GST-Ya subunit gene (24,28). As described by Jaiswal (20) in a recent review, the functional properties of individual AREs may, in fact, be dependent on the type, organization and orientation of the various elements within an ARE complex. In this respect, the putative ARE in the GCS_h subunit gene is unique in that it contains AP-1-like sequences in a "face-to-face" orientation with the second AP-1-like sequence located on the 3'-side of the conserved ARE core sequence. Determination of the functional properties of the putative GCS_h ARE and determination of the significance of the unique features of its organization are the objectives of on-going experiments.

REFERENCES

1. Deneke, S.M. and Fanburg, B.L. *Am.J.Physiol.* **257**:L163-L173, 1989.
2. Meister, A. Metabolism and Function of Glutathione. In: *Coenzymes and Cofactors. Glutathione: Chemical, Biochemical and Medical Aspects. Vol. III, Part A*, edited by Dolphin, D., Poulson, R. and Avramovic, O. New York: John Wiley and Sons, 1989, p. 367-474.
3. Abate, C., Patel, L., Rauscher III, F.J. and Curran, T. *Science* **246**:870-872, 1990.
4. Bergelson, S., Pinkus, R. and Daniel, V. *Cancer Res.* **54**:36-40, 1994.
5. Lautier, D., Luscher, P. and Tyrrell, R.M. *Carcinogenesis* **13**:227-232, 1992.
6. Liu, R.-M., Vasiliou, V., Zhu, H., et al. *Carcinogenesis* **15**:2347-2352, 1994.
7. Woods, J.S., Davis, H.A. and Baer, R.P. *Arch.Biochem.Biophys.* **296**:350-353, 1992.
8. Chung, A.S. and Maines, M.D. *Biochem.Pharmacol.* **30**:3217-3223, 1981.
9. Shi, M.M., Iwamoto, T. and Forman, H.J. *Am.J.Physiol.* **267**:L414-L421, 1994.
10. Eaton, D.L. and Hamel, D.M. *Toxicol.Appl.Pharmacol.* **126**:145-149, 1994.
11. Seelig, G.F., Simondsen, R.P. and Meister, A. *J.Biol.Chem.* **259**:9345-9347, 1984.
12. Huang, C.-S., Anderson, M.E. and Meister, A. *J.Biol.Chem.* **268**:20578-20583, 1993.
13. Gipp, J.J., Chang, C. and Mulcahy, R.T. *Biochem.Biophys.Res.Commun.* **185**:29-35, 1992.
14. Gipp, J.J., Bailey, H.H. and Mulcahy, R.T. *Biochem.Biophys.Res.Commun.* **206**:584-589, 1995.
15. Ripple, M., Mulcahy, R.T. and Wilding, G. *J.Urol.* **150**:209-214, 1993.
16. Bailey, H.H., Gipp, J.J., Ripple, M., Wilding, G. and Mulcahy, R.T. *Cancer Res.* **52**:5115-5118, 1992.
17. Smale, S.T. and Baltimore, D. *Cell* **57**:103-113, 1989.
18. Smale, S.T., Schmidt, M.C., Berk, A.J. and Baltimore, D. *PNAS* **87**:4509-4513, 1990.
19. Rushmore, T.H., Morton, M.R. and Pickett, C.B. *J.Biol.Chem.* **266**:11632-11639, 1991.
20. Jaiswal, A.K. *Biochem. Pharm.* **48**:439-444, 1994.
21. Daniel, V. *Crit.Revs.Biochem.Mol.Biol.* **28**:173-207, 1993.
22. Friling, R.S., Bensimon, A., Tichauer, Y. and Daniel, V. *PNAS* **87**:6258-6262, 1990.
23. Friling, R.S., Bergelson, S. and Daniel, V. *PNAS* **89**:668-672, 1992.
24. Prestera, T., Holtzclaw, W.D., Zhang, Y. and Talalay, P. *Proc.Natl.Acad.Sci.USA* **90**:2965-2969, 1993.
25. Huang, H., Huang, C.F., Wu, D.R., Jinn, C.M. and Jan, K.Y. *Toxicol.* **79**:195-204, 1993.
26. Shi, M.M., Kugelman, A., Iwamoto, T., Tian, L. and Forman, H.J. *J.Biol.Chem.* **269**:26512-26517, 1994.
27. Borroz, K.I., Buetler, T.M. and Eaton, D.L. *Toxicol.Appl.Pharmacol.* **126**:150-155, 1994.
28. Bergelson, S. and Daniel, V. *Biochem.Biophys.Res.Commun.* **200**:290-297, 1994.
29. Li, Y. and Jaiswal, A.K. *Eur.J.Biochem.* **226**:31-39, 1994.