BIOCHEMICAL AND BIOPHYSICAL RESEARCH COMMUNICATIONS **249,** 53–55 (1998)
ARTICLE NO. RC989086

Structural Organization of the Human Selenium-Dependent Phospholipid Hydroperoxide Glutathione Peroxidase Gene (GPX4): Chromosomal Localization to 19p13.3

Michael J. Kelner[1] and Mark A. Montoya
Department of Pathology, University of California, San Diego, 200 West Arbor Drive, San Diego, California 92103-8320

Received June 22, 1998

The primary structure of human selenium-dependent phospholipid hydroperoxide glutathione peroxidase (GPX4) was determined by genomic cloning. The gene structure of GPX4 spans only 2.8 kb and consists of 7 exons. The coding sequence resides on all 7 exons, and the mitochondrial leader sequence is contained entirely within the first exon. The selenocysteine coding nucleotide resides on the third exon. The introns all commenced with the consensus nucleotide sequence GTR and ended with the consensus nucleotide sequence YAG. Analysis of the GPX4 gene sequence identified a potential alternative tissue-specific first exon. Chromosomal FISH studies placed the GPX4 gene at 19p13.3 location, and downstream of the 23 kDa polypeptide DNA-directed RNA polymerase gene.

The phospholipid hydroperoxide glutathione peroxidase, GPX4, belongs to the family of selenium-dependent peroxidases, and was first described in 1982 (1,2). In contrast to the classical cellular glutathione peroxidase, GPX1, the GPX4 enzyme is a monomeric enzyme with two distinct forms (2). It consists of a mitochondrial form of approximately 22 kd and a non-mitochondrial form of approximately 19 kd (3). The tissue distribution of GPX4 is also distinct from GPX1. GPX4 is highly expressed in the testes and thyroid, whereas GPX1 is highly expressed in the liver, the lungs and certain human tumor cell lines (4-6). The GPX4 enzyme is capable of reducing peroxidized phospholipids, cholesterol hydroperoxides, and thiamine hydroperoxides (2,7,8). While the GPX1 enzyme can reduce free fatty acid hydroperoxides, it cannot reduce phospholipid hydroperoxides like GPX4 (9). Thus, GPX4 is considered to be the primary enzymatic defense against oxidation for biomembranes.

Subcellular distribution studies have identified the GPX4 as residing in the mitochondria, cytoplasma and the nuclear envelope (6,10,11). It is known that the import of GPX4 into the mitochondria does require a leader sequence (12). The human, rat, mouse, and pig cDNA sequences have been determined, and the pig gene was also isolated (3,13-18). Studies in the rat demonstrated that GPX4 uses multiple transcription and translation start sites (6). Transgenic overexpression of GPX4 protects host cells from oxidative stress and against various lipid hydroperoxide mediated injury (19,20). In addition to this antioxidant role, recent studies indicate GPX4 is also capable of modulating 15-lipoxygenase products (15,21).

Thus, GPX4 displays several unusual properties including tissue expression, subcellular localization, hormonal dependence and regulation of lipoxygenase products (4,6,11,21,22). Pursuing the mechanism(s) behind these unusual properties of GPX4 requires knowledge of the upstream nontranscribed region (also known as the promoter region), intron/exon structure, and down stream nontranscribed region which is often implicated in tissue-specific regulation. Therefore, to further investigate the unusual tissue distribution and hormonal regulation of GPX4, we determined the genomic structure of the human GPX4 gene which was previously determined to be a single copy gene residing on chromosome 19 (23).

MATERIALS AND METHODS

A probe corresponding to the last 653 through 875 nucleotides of human GPX4 cDNA (16) was generated from a pCD human cDNA library (Dr. Paul Berg, Stanford, CA) using PCR amplification. Oligonucleotides used for PCR screening were gga cct gcc cca cta ttt cta (upstream) and tgt ctg ttt att ccc acc cgg (downstream). The entire PCR probe was sequenced to confirm nucleotide identity to human GPX4 cDNA sequence (16). Amplification of human genomic DNA

[1] Corresponding author. Fax: 619-543-3730. E-mail: mkelner@ucsd.edu.

0006-291X/98 $25.00

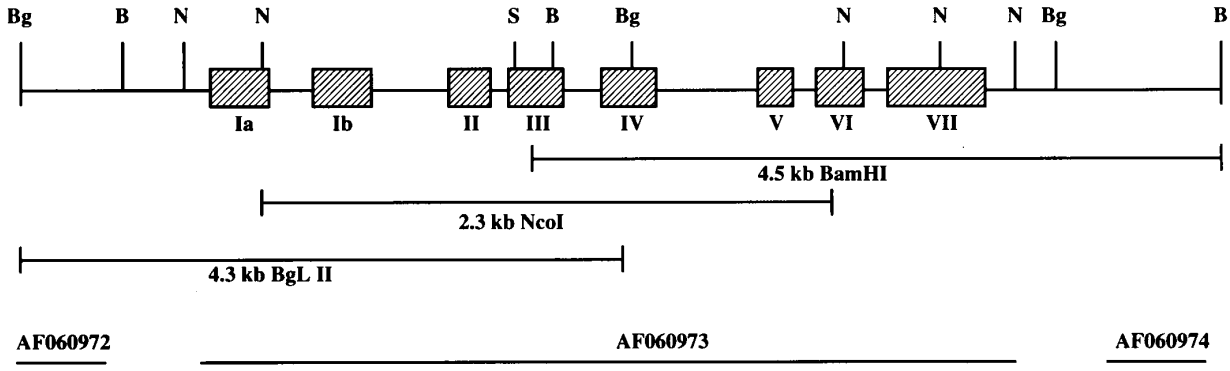

FIG. 1. Structural organization of the human selenium-dependent phospholipid hydroperoxide glutathione peroxidase (GPX4) gene. Striped boxes indicate exons of the gene. Restriction enzyme sites are identified as BglII (Bg), BamHI (B), NcoI (N), and SalI (S). The three overlapping clones and their relative location are provided below the GPX4 map. The regions sequenced and their corresponding GenBank accession numbers are noted at the bottom of the diagram.

with these PCR primers yielded a single product of correct size, which was also sequenced to confirm nucleotide identity. A human genomic foreskin fibroblast bacteriophage P1 library (Genome Systems, St. Louis, MO) was screened by PCR using these oligonucleotides. Three clones were identified as positive by PCR amplification. The PCR amplification product from one of the clones, 13648, was sequenced to confirm that the product corresponded to nucleotides 633 to 875 of the GPX4 cDNA sequence (16). The nucleotide sequence analysis was performed by the dideoxynucleotide chain-termination method with sequencing grade *Taq* DNA polymerase and the fmol thermocycling DNA sequencing kit by Promega Corp. (Madison, WI) using the Direct Incorporation SEQ protocol. The reagent 7deazaGTP was added when necessary to minimize band compression often associated with the high G + C content in genomic DNA. All DNA inserts were sequenced in both directions to ensure accuracy of the sequencing data, the gene was sequenced in each direction in overlapping fragments, and autoradiograms containing sequence letters were proofread after data entry.

The chromosomal localization of the P1 Clone 13648 was determined by the fluorescence in situ hybridization (FISH) technique as previously described (24). The entire P1 clone #13648 was labeled with digoxigenin dUTP by nick translation. The initial experiment resulted in specific labeling of the short arm of a group F chromosome believed to be chromosome 19 on the basis of size, morphology, and banding pattern. A second experiment was conducted in which a genomic probe which had been previously mapped to chromosome 19q, and confirmed by cohybridization with a probe from the E2A locus, was cohybridized with the P1 clone. This experiment resulted in the specific labeling of the long arm and the short arm of chromosome 19. Measurements of 10 specifically labeled chromosomes demonstrated that GPX4 is located at the terminus of the short arm of chromosome 19, an area corresponding to the 19p13.3 band A total of 80 metaphase cells were analyzed with 75 exhibiting specific labeling. Thus, the FISH chromosomal localization studies localized the human GPX4 gene to the 19p13.3 region, which is in agreement with a previous report using somatic cell hybrids that localized GPX4 to chromosome 19 (23).

RESULTS AND DISCUSSION

Analysis of three overlapping clones (Figure 1) revealed that the human GPX4 gene (AF 060973) consists of 7 exons over a 2.8 kb region, which is in contrast to the other classical selenium-dependent GPX enzymes (GPX1 and GPX2) which contain only two exons. The exon lengths are identical to that previously reported in the pig gene (15) with the exception of the 5' and 3' untranslated regions. The 27 amino acid mitochondrial leader sequence (12) resides entirely in the first exon and the selenocysteine codon (15) resides in the third exon. Analysis of the exon/intron junctions revealed that GPX4 splicing followed the general consensus splice sequences for exon-intron junctions with one major exception (Table 1).

Analysis of the GenBank nonredundant database revealed the upstream region (AF060972) of the GPX4 gene (Figure 1) was homologous to the downstream region of a cosmid clone (AC004151), also previously localized to chromosome 19p13.3 region, which contained the 23 k-Da polypeptide DNA-directed RNA polymerase II gene. There was no homology of the region (AF060974) downstream of GPX4 to any GenBank submission.

Analysis of dBEST and TIGR Human-Gene Index (THC201938) databases revealed the presence of a potential alternative first exon that was tissue-specific. The clones obtained from human testis-derived tissue (such as AA400687), where GPX4 is highly expressed, all contain the same first exon (exon 1a, figure 1) pre-

TABLE 1

Sequences at the Intron/Exon Splice Junctions

Exon # and size (bp)	Donor	Intron size (bp)	Acceptor
1, 161	ACCATG gtgagc	1068	ttgcag TGCGCG
2, 95	GTACCG gtgggc	86	tcgcag GGGCTT
3, 145	AAGCAG gtgggc	147	ccacag GAGCCA
4, 152	GGGAAA gtgcgt	433	ttccag TGCCAT
5, 25	ACCAAG gtaagg	133	ccacag AGTTCC
6, 60	CCCCTG gtgggt	80	cgacag GTGATA
7, 240			
Consensus	NAG gtra		yag NNN

viously identified when the mRNA sequence of GPX4 was first determined using a human testis cDNA library (16). Several dBEST clones from non-testis libraries, such as heart or colon tissue where GPX4 is minimally expressed (such as AA733049 or AA587357) all contain a different first exon (exon 1b, figure 1), with an out-of-frame start codon, located slightly downstream from the testis-associated first exon. The use of a tissue-specific alternative first exon to control expression is compatible with a previous report describing different transcription and translational start sites for testis versus somatic tissue (6). Computer analysis of the GPX4 sequence did not identify any repetitive elements. Interestingly, while the GRAIL program identified 7 exons, only the first 3 were correctly identified. In contrast, while the GENIE Program only recognized 5 exons it did identify the last 4 exons correctly (exons 1,4,5,6,7).

In summary, the GPX4 gene resides at chromosomal location 19p13.3, consists of 7 exons spanning 2.8 kb (in contrast to GPX1 and GPX2), the exon/intron junctions follow the general consensus splice sequences (with one exception), a tissue-specific alternative exon splicing appears to be present, and the upstream neighboring gene is the 23 k-Da polypeptide DNA-directed RNA polymerase II gene. There is no homology of the region downstream of GPX4 to any GenBank of other databank submission, so the downstream neighboring gene cannot be identified at this time.

ACKNOWLEDGMENT

This work was supported by NIH Grant CA52310.

REFERENCES

1. Ursini, F., Maiorino, M., Valente, M., and Gregolin, C. (1982) *Biochim. Biophys. Acta* **710,** 197–211.
2. Ursini, F., Maiorino, M., and Gregolin, C. (1985) *Biochim. Biophys. Acta* **839,** 62–72.
3. Pushpa-Rekha, T. R., Burdsall, A. L., Oleksa, L. M., Chisolm, G. M., and Driscoll, D. M. (1995) *J. Biol. Chem.* **270,** 26993–26999.
4. Maiorino, M., Chu, F. F., Ursini, F., Davies, K. J. A., Doroshow, J. H., and Esworthy, R. S. (1991) *J. Biol. Chem.* **266,** 7728–7732.
5. Roveri, A., Casasco, A., Maiorino, M., Dalan, P., Calligaro, A., and Ursini, F. (1992) *J. Biol. Chem.* **267,** 6142–6146.
6. Dreher, I., Schmutzler, C., Jakob, F., and Köhrle, J. (1997) *J. Trace Element Med. Biol.* **11,** 83–91.
7. Thomas, J. P., Maiorino, M., Ursini, F., and Gitrotti, A. W. (1990) *J. Biol. Chem.* **265,** 454–461.
8. Bao, Y., Jemth, P., Mannervik, B., and Williamson, G. (1997) *FEBS Letters* **410,** 210–212.
9. Weitzel, F., Ursini, F., and Wendel, A. (1990) *Biochim. Biophys. Acta* **1036,** 88–94.
10. Godeas, C., Sandri, G., and Panfili, E. (1994) *Biochim. Biophys. Acta* **1191,** 147–150.
11. Roveri, A., Maiorino, M., and Ursini, F. (1994) *Methods Enzymol* **233,** 202–212.
12. Arai, M., Imai, H., Sumi, D., Imanaka, T., Takano, T., Chiba, N., and Nakagawa, Y. (1996) *Biochem. Biophys. Res. Commun.* **227,** 433–439.
13. Schuckelt, R., Brigelius-Flohé, R., Maiorino, M., Roveri, A., Reumkens, J., Straßburger, W., Ursini, F., Wolf, B., and Flohé, L. (1991) *Free Rad. Res. Commun.* **14,** 343–361.
14. Sunde, R. A., Dyer, J. A., Moran, T. V., Evenson, J. K., and Sugimoto, M. (1993) *Biochem. Biophys. Res. Commun.* **193,** 905–911.
15. Brigelius-Flohé, R., Aumann, K-D., Blocker, H., Gross, G., Kiess, M., Kloppel, K-D., Maiorino, M., Roveri, A., Schuckelt, R., Ursini, F., Wingender, E., and Flohé, L. (1994) *J. Biol. Chem.* **269,** 7342–7348.
16. Esworthy, R. S., Doan, K., Doroshow, J. H., and Chu, F-F. (1994) *Gene* **144,** 317–318.
17. Imai, H., Sumi, D., Hanamoto, A., Arai, M., and Sugiyama, A. (1995) *J. Biochem.* **118,** 1061–1067.
18. Nam, S-Y., Nakamuta, N., Kurohmaru, M., and Hayashi, Y. (1997) *Gene* **198,** 245–249.
19. Imai, H., Sumi, D., Sakamoto, H., Hanamoto, A., Arai, M., Chiba, N., and Nakagawa, Y. (1996) *Biochem. Biophys. Res. Commun.* **222,** 432–438.
20. Yagi, K., Komura, S., Kojima, H., Sun, Q., Nagata, N., Ohishi, N., and Nishikimi, M. (1996) *Biochem. Biophys. Res. Commun.* **219,** 486–491.
21. Imai, H., Narashima, K., Arai, M., Sakamoto, H., Chiba, N., and Nakagawa, Y. (1998) *J. Biol. Chem.* **273,** 1990–1997.
22. Brigelius-Flohé, R., Friedrichs, B., Maurer, S., Schultz, M., and Streicher, R. (1997) *Biochem. J.* **328,** 199–203.
23. Chu, F.-F. (1994) *Cytogenet. Cell Genet.* **66,** 96–98.
24. Stokke, T., Collins, C., Kuo, W., Kowbel, D., Shadraven, F., Tanner, M., Kallioniemi, A., Kallioniemi, O., Pinkel, D., Deaven, L., and Gray, J. (1995) *Genomics* **26,** 134–137.