Cloning, sequencing and characterization of the human Alpha glutathione S-transferase gene corresponding to the cDNA clone pGTH2

Antje KLÖNE, Rolf HUSSNÄTTER and Helmut SIES*

Institut für Physiologische Chemie I, Heinrich-Heine-Universität Düsseldorf, Moorenstrasse 5, W-4000 Düsseldorf, Germany

The human Alpha glutathione S-transferase gene corresponding to the human liver cDNA clone pGTH2 was isolated from a cosmid genome library. The gene, represented by the clone cosGTH2, spans nearly 12 kb and contains seven exons. The intron/exon borders conform to the standard rules, and an open reading frame is present, starting at position 67 in exon 2, the double-stop codon being at position 733 in exon 7. Exons 1, 2 and 7 differ in length from the known rat gene coding for the Ya enzyme. A 209 bp 5'-upstream region contains TATA and CAT boxes and, in addition, motifs for Sp1-, NF1- and HNF1-binding factors. Clone cosGTH2 represents the less basic subunit, α_x , of two Alpha glutathione S-transferase subunits (α_x and α_y) expressed in liver, which is identical with the kidney subunit α_2 .

INTRODUCTION

The soluble glutathione S-transferases (GSTs) form a supergene family of dimeric enzymes which in man is comprised of different multigene families referred to as Alpha, Mu and Pi [1]. These enzymes catalyse the conjugation of GSH to a variety of electrophilic compounds and exhibit tissue- and development-specific expression patterns [2–4]. Differences in Alpha GST expression were observed between normal and malignant tissue, with very low levels present in the latter [5,6].

H.p.l.c. analysis demonstrates two GST Alpha subunits in liver and kidney, referred to as α_x and α_y or α_1 and α_2 respectively; the former is more basic than the latter [5,7], the pI being 8.9 for α_1 and 8.4 for α_2 [1].

Corresponding to the h.p.l.c. data, two full-length cDNA clones for the human Alpha GST family have been derived from liver cDNA libraries, namely pGTH1 [8], corresponding to clone λ GST2-3 [9], and pGTH2 [10]. This indicates that the human Alpha GSTs are products of at least two different genes [10]. A cDNA clone from human kidney, GST α 12k, synthesized by the PCR method [5] is equivalent to the liver cDNA clone pGTH1 [8] or λ GST2-3 [9], except for a base change (G→A) at position 442 in pGTH1 (position 430 in λ GST2-3).

The human Alpha GST locus (GST-2) was localized on chromosome 6p12 [9]. So far, one Alpha GST gene of rat [11] and of mouse [12] have been described, and their gene structure has been shown to be highly conserved [8,10]. Regions of human GST loci have been isolated for the Mu [13] and Pi genes [14] but so far not for the Alpha GST genes.

We have analysed a chromosomal region of the human Alpha GST gene locus (GST-2). The structural analysis revealed a gene as the genomic equivalent of the cDNA clone pGTH2 [10].

MATERIALS AND METHODS

Human genomic cosmid library and screening

The human cosmid library was constructed by Poustka *et al.*

[15] using high-molecular mass DNA of human blood; 5×10^5 independent recombinants of this library were screened by using 32 P-labelled fragments of the cDNA clone GST α 12k, which contained the entire coding region of the human Alpha GST gene. The hybridization conditions were those described by Maniatis *et al.* [16].

Labelling *in vitro*

All probes used in this study were 32 P-labelled by oligonucleotide random priming [17], and had a specific radioactivity greater than 10^9 c.p.m./ μ g of DNA.

DNA sequencing and restriction mapping

Clones showing a positive signal in hybridization experiments with the cDNA clone GST α 12k were digested with different restriction enzymes to generate smaller fragments. Fragments that continued to hybridize to the cDNA clone GST α 12k were isolated and digested further with several enzymes creating overlapping fragments to generate the restriction map. These fragments were cloned into the sequencing vector Bluescript pKS and sequenced according to the method of Sanger *et al.* [18]. The published sequences of cosGTH2 were determined in both strands.

RESULTS

Isolation of a human Alpha GST clone

By using the human kidney Alpha GST cDNA clone GST α 12k as a probe [5], a human cosmid library was screened, revealing a positive clone called cosGTH2.

Southern blots of overlapping fragments of cosGTH2 revealed hybridization to the 3' region (exons 6 and 7) and 5' region (exon 1–5) of GST α 12K. These fragments were subcloned into the Bluescript vector and hybridized with either the 3' or the 5' region of GST α 12K and afterwards sequenced. Restriction analysis of the overlapping fragments compared with the hybridization pattern of the 3' and 5' region of GST α 12K,

Abbreviation used: GST, glutathione S-transferase.

* To whom correspondence should be addressed.

The nucleotide sequence data reported have been submitted to the EMBL, GenBank and DDBJ Nucleotide Sequence Databases under the accession numbers X65726–X65734.

EXON 2

pGTH2 MAEKPKLHYS NIRGRMESIR WLLAAAGVE
cosGTH2 *****

EXON 3

pGTH2 FEEKFIKSAE DLDKLRN
cosGTH2 *****
cosGTH2⁺ *DQ**LE*R* E*****

EXON 4

pGTH2 GYLMFQQVP MVEIDGMKLV QTRAILNYIA SKYNLYGKDI KEKAL
cosGTH2 *****
cosGTH2⁺ *S***L*S *****R*****T **R**

EXON 5

pGTH2 IDMYIEGIAD LGEMILLLPF TQPEEQDAKL ALIQEKTNR YFPFAFEK
cosGTH2 *****

EXON 6

pGTH2 VLKSHGQDYL VGNKLSRADI HLVELLYYVE ELDSSLISF PLLK
cosGTH2 *****

EXON 7

pGTH2 ALKTRISNLP TVKKFLQPGS PRKPPMDEKS LEESRKIFRF
cosGTH2 *****

Fig. 3. Comparison of the amino acid sequences of different exons with the liver cDNA clone pGTH2

The deduced amino acid sequence resulting from Fig. 2 for clone cosGTH2 corresponds to pGTH2. The open reading frame starts at nucleotide +31 of exon 2 (see Fig. 2). The superscript plus sign (+) indicates deduced amino acid sequences of exons that show base changes compared with the cDNAs (Fig. 2).

including the sequence data, generated the organization of the isolated cosmid as shown in Fig. 1.

Nucleotide sequence analysis

Sequence data of cosGTH2 revealed absolute homology to the human liver cDNA clone, pGTH2 [10] (Figs. 1 and 2). Like its homologue in rat [11] and mouse [12] the human GST Alpha gene contains seven exons (E1–E7) interrupted by six introns. Furthermore, clone cosGTH2 contains additional regions corresponding to exons 3 and 4 (E3⁺ and E4⁺) (Figs. 1 and 2). These exons show a homology of 90% to pGTH2 and to the other known cDNAs. The exons (E⁺) with the decreased homology are downstream from the exons (E) that correspond to pGTH2 (Fig. 2).

The rules for the intron/exon junctions are fulfilled for all exons, E1–E7 and E3⁺ and E4⁺, and an open reading frame is present which starts at position 67 in exon 2 (E2) with the codon for initiation methionine. The sequence of the exons, the intron/exon junctions and the sequenced portions of the introns are shown in Fig. 2, and the resulting amino acid sequence is shown in Fig. 3. Table 1 shows the length of the exons (E) and introns of the isolated human Alpha gene compared with the Ya gene of rat [11].

Upstream region of cosGTH2

Fig. 4 shows the 5' upstream nucleotide sequence from –1 to –209 of cosGTH2, taking position 1 of the cDNA [10] as the transcription initiation site. A TATA box with the sequence TATAAA was located at position –21 to –26 relative to the first nucleotide of exon 1. Two CAT-box-like elements were located, one with the sequence 5'-CTATT-3' at –44 to –48, and a second with the sequence 5'-CCTAACT-3' at position –75 to –80. The TATA box conforms to the consensus sequence [19], whereas both potential CAT box sequences have alterations

Table 1. Length of the exons (a) and introns (b) in cosGTH2

For comparison, data on the Ya gene of the rat [11] are also presented (data on the Yc gene are not available in the literature). The sequences for exons 1–7 of the human gene cosGTH2 are given in Fig. 2.

(a)

		Length (bp)						
Species	Exon...	1	2	3	4	5	6	7
Human		36	117	52	133	142	132	317
Rat		43	109	52	133	142	132	234

(b)

		Length (kbp)						Reference
Species	Intron...	1	2	3	4	5	6	
Human		4.1	1.7	1.1	2.0	1.3	1.0	The present work
Rat		2.3	3.5	0.65	2.1	1.5	0.8	[11]

tggttaacttgataccaaccagtaagga –181

acata^{tt}aaacagtgctctctgataagcagatcacttgccctcatgtcttagaatccagt –121

aggtgccccttgccatgaaatgtgtggagtgcccttttcgctaacttgacccttcttctc –61

agtgaggagggaactattgagaggaacaaagagattataatacattaggacctggaattc –1

Fig. 4. Nucleotide sequence of the 5' upstream region of cosGTH2

The numbers represent the last nucleotide in each lane. The sequences corresponding to the TATA box, the CAT box, the Sp1-binding site, the NF1- and HNF1-recognition sites are underlined (see the text).

to the consensus sequences [20,21]. A potential binding site for the transcription factor SP1 has been identified with the sequence 5'-TGGGAGGGAAC-3' at –48 to –57, which resembles the consensus sequence 5'-(G/T)GGGCGG(G/A) (G/A) (C/T)-3' [22]. Upstream of this region at position –86 and –107 the nucleotide sequence contains two half-recognition sites for the nuclear factor NF1: consensus sequence, 5'-TGG(A/C)N5G-CCAA-3' [23]. An absolute homology was found to the ATTAAC motif reported to mediate liver-specific expression by hepatocyte nuclear factor HNF1 [24] at position –171 to –176.

DISCUSSION

H.p.l.c. data revealed that two GST Alpha isoenzymes are present in human liver and kidney tissue [5,7]. Recently, we showed that the kidney isoenzymes, α_1 and α_2 , have retention times similar to the two liver forms, referred to as α_x and α_y [5]. Hayes and co-workers [25] showed that the more basic subunit α_x (or α_1) is coded by the liver cDNA pGTH1 [8], whereas α_y (or α_2) is coded by the second liver cDNA clone, pGTH2 [10].

The human Alpha GST locus located on chromosome 6p12 contains at least two separate genes [9]. These two genes are represented by two different cDNAs, namely GST α 12k [5], pGTH1 [8] or λ GST2-3 [9] on the one hand and pGTH2 [10,12] on the other. The existence of two genes correlates with the fact that two signals of 0.8 kb and 1.0 kb are detected on Northern-blot analysis [5].

The sequence data of one of these two different cDNAs, pGTH2, conform with the sequences of the clone cosGTH2 (Fig. 2). Therefore cosGTH2 represents the second Alpha GST isoenzyme α_2 (or α_1) which is the less basic subunit. It is of interest to note that in an enzymological study [26] of the human kidney enzymes, there was no difference in substrate specificity, but the $V_{\max}$ with 1-chloro-2, 4-dinitrobenzene was about double in α_1 compared with α_2 . With regard to the Ya rat gene [11], there are differences in the length of exons 1, 2 and 7, which are 43, 109 and 234 bp in the rat respectively (see Table 1).

The absolute homology to the cDNA, pGTH2, an open reading frame as well as integrity of the intron/exon junction would be unlikely if this region represented by E1–E7 were part of a pseudogene.

Furthermore, a TATA box and other 5'-regulatory elements are located at the upstream region (Fig. 4). It is interesting that the 5' region of the human GST gene presented here shows no striking similarity in the analysed 5' region to the corresponding regions of the rat [11] and mouse [12,27] Alpha GST genes, although the NF1 and the half ATTAAC motifs are also located at the 5' region of the mouse gene [27]. The SP1 site was also found in the human GST Pi gene [14]. An altered control of expression is discussed for the human and rat GST Pi gene [28], and this also might be the reason for the low sequence similarity between the 5' regions of the Alpha GST gene of rodents and man.

Rushmore *et al.* [29] recently described further a regulatory element in the rat Alpha gene upstream of the region presented here, called antioxidant-responsive element (ARE). It will be of interest to see whether such elements are located in the human gene described here.

Clone cosGTH2 also contains stretches of diminished similarity to the known cDNAs, indicated by the superscript plus sign in Fig. 2; these exons could be members of an Alpha GST-like pseudogene family. The decreased similarity to the known cDNAs could support this hypothesis; however, the open reading frame and the integrity of the intron/exon borders contradict this idea. So far, analysed GST pseudogenes do not contain introns or open reading frames [30]. These criteria make it probable that the genetic regions are kept under selection pressure, maybe because these regions are genes or are brought under pressure by gene conversion. Gene conversion has been described for the GST Mu family in man [13,31,32].

Note added in proof (received 12 May 1992)

The human gene equivalent of pGTH1 has been described recently [33].

Expert technical assistance by Peter Graf and help with DNA screening by Kerstin Bellmann is gratefully acknowledged. We thank A. Poustka, DKFZ, Heidelberg, Germany, for the human cosmid library. This work was supported by the Doktor Robert Pflieger-Stiftung, Bamberg, Germany, and by the National Foundation for Cancer Research, Bethesda, MD, U.S.A.

REFERENCES

- Mannervik, B. & Danielson, U. H. (1988) *Crit. Rev. Biochem.* **23**, 283–337
- Boyer, T. D. (1989) *Hepatology* **9**, 486–496
- Ketterer, B., Meyer, D. J. & Clark, A. G. (1988) in *Glutathione Conjugation: Mechanisms and Biological Significance* (Sies, H. & Ketterer, B., eds.) pp. 73–135, Academic Press, London
- Mannervik, B. (1985) *Adv. Enzymol. Relat. Areas Mol. Biol.* **57**, 357–417
- Klöne, A., Weidner, U., Hussnätter, R., Harris, J., Meyer, D., Peter, S., Ketterer, B. & Sies, H. (1990) *Carcinogenesis* **11**, 2179–2183
- Strohmeier, T., Klöne, A., Wagner, G., Hartmann, M. & Sies, H. (1992) *J. Urol.* **147**, 1424–1428
- Ostlund-Farrants, A.-K., Meyer, D. J., Coles, B., Southan, C., Aitken, A., Johnson, P. J. & Ketterer, B. (1987) *Biochem. J.* **245**, 423–428
- Tu, C.-P. D. & Quian, B. (1986) *Biochem. Biophys. Res. Commun.* **141**, 229–237
- Board, P. G. & Webb, G. C. (1987) *Proc. Natl. Acad. Sci. U.S.A.* **84**, 2377–2381
- Rhoads, D. M., Zarlengo, R. P. & Tu, C.-P. D. (1987) *Biochem. Biophys. Res. Commun.* **29**, 474–481
- Telakowski-Hopkins, C. A., Rothkopf, G. S. & Pickett, C. B. (1986) *Proc. Natl. Acad. Sci. U.S.A.* **83**, 9393–9397
- Daniel V., Sharon, R., Tichauer, Y. & Sarid, S. (1987) *DNA* **6**, 317–324
- Taylor, J. B., Oliver, J., Sherrington, R. & Pemble, S. E. (1991) *Biochem. J.* **274**, 587–593
- Cowell, I. G., Dixon, K. H., Pemble, S. E., Ketterer, B. & Taylor, J. B. (1988) *Biochem. J.* **255**, 79–81
- Poustka, A., Rackwitz, H.-R., Frischauf, A.-M., Hohn, B. & Lehrach, H. (1984) *Proc. Natl. Acad. Sci. U.S.A.* **81**, 4129–4133
- Maniatis, T., Fritsch, E. F. & Sambrook, J. (1989) *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY
- Feinberg, A. P. & Vogelstein, B. (1983) *Anal. Biochem.* **132**, 6–13
- Sanger, F., Nicklen, S. & Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. U.S.A.* **74**, 5463–5467
- Breathnach, R. & Chambon, P. (1981) *Annu. Rev. Biochem.* **50**, 349–383
- Chodosh, L. A., Baldwin, A. S., Carthew, R. W. & Sharp, P. A. (1988) *Cell* **53**, 11–24
- Santoro, C., Mermod, N., Andrews, P. C. & Tjian, R. (1988) *Nature (London)* **334**, 218–224
- Briggs, M. R., Kadonaga, J. T., Bell, S. P. & Tjian, R. (1986) *Science* **234**, 47–52
- Borgmeyer, U., Nowock, J. & Sippel, A. E. (1984) *Nucleic Acids Res.* **12**, 4295–4311
- Courtois, G., Morgan, J. G., Campbell, L. A., Fourel, G. & Crabtree, G. R. (1987) *Science* **238**, 688–692
- Hayes, J. D., Kerr, L. A. & Cronshaw, A. D. (1989) *Biochem. J.* **264**, 437–445
- Singh, S. V., Leal, T., Ansari, G. A. S. & Awasthi, Y. C. (1987) *Biochem. J.* **246**, 179–186
- Daniel, V., Sharon, R. & Bensimon, A. (1989) *DNA* **8**, 399–408
- Dixon, K. H., Cowell, I. G., Xia, C. L., Pemble, S. E., Ketterer, B. & Taylor, J. B. (1989) *Biochem. Biophys. Res. Commun.* **163**, 815–822
- Rushmore, T. H., Morton, M. R. & Pickett, C. B. (1991) *J. Biol. Chem.* **266**, 11632–11639
- Okuda, A., Sakai, M. & Muramatsu, M. (1987) *J. Biol. Chem.* **262**, 3858–3863
- Taylor, J. B., Oliver, J., Pemble, S. E. & Ketterer, B. (1990) in *Glutathione S-Transferase and Drug Resistance* (Hayes, J. D., Pickett, C. B. & Mantle, T. J., eds.), pp. 242–249, Taylor and Francis, London, New York and Philadelphia
- Vorachek, W. R., Pearson, W. R. & Rule, G. S. (1991) *Proc. Natl. Acad. Sci. U.S.A.* **88**, 4443–4447
- Rozen, F., Nguyen, T. & Pickett, C. B. (1992) *Arch. Biochem. Biophys.* **292**, 589–593