Short Communications

Mammalian Genome
Incorporating *Mouse Genome*

Mammalian Genome 11, 1034–1036 (2000).
DOI: 10.1007/s003350010189

The gene encoding rat 3-phosphoglycerate dehydrogenase

Mariette Robbi,[1] Younes Achouri,[1] Claude Szpirer,[2] Emile Van Schaftingen[1]

[1]Laboratoire de Chimie Physiologique, Christian de Duve Institute of Cellular Pathology and Université Catholique de Louvain, Avenue Hippocrate 75, B-1200 Brussels Belgium
[2]Laboratoire de Biologie du Développement, I.B.M.M., Université Libre de Bruxelles, B-6041 Gosselies, Belgium

Received: 17 March 2000 / Accepted: 23 June 2000

The first step of L-serine synthesis is catalyzed by 3-phosphoglycerate dehydrogenase (PHGDH), a ubiquitous enzyme in prokaryotes and eukaryotes. The sequence encoding this enzyme, now known for many species including rat (Achouri et al. 1997) and human (Cho et al. 2000), shows that it belongs to the family of D-hydroxyacid dehydrogenases (Vinals et al. 1993), which also comprises enzymes acting on D-glycerate and glycolate. In mammals, PHGDH is widely distributed in tissues (Willis and Sallach 1964). However, its activity in rat liver, but not in other tissues, strongly depends on the nutritional status, being low in animals fed a normal diet and increasing more than 10-fold after shifting to a protein-poor, carbohydrate-rich diet (Fallon et al. 1966; Mauron et al. 1973). Previous work has shown that this control is exerted at a pre-translational level by cysteine and glucagon, which appear to act on the stability of the mRNA and on the transcription of the gene, respectively (Achouri et al. 1999). PHGDH appears also to be quite active in proliferating cells (Snell 1984), most likely owing to transcriptional regulation (Cho et al. 2000); this high activity is probably related to the precursor role of serine in the biosynthesis of nucleotides. Although the sequence of the cDNAs encoding the rat and human enzymes has been published (Achouri et al. 1999; Cho et al. 2000), the structure of the corresponding gene is not known for any metazoan species.

The purpose of this work was to determine the structure of the gene encoding rat PHGDH and check whether the same promoter was used in the liver as in tissues where no dietary regulation of the expression is observed.

Screening of 10^6 plaques of a rat genomic library (from Stratagene, in Lambda Dash) with cDNA probes corresponding to the 5′ (nucleotides 1–514) and 3′ (nucleotides 772–1780) ends of the rat PHGDH cDNA (Achouri et al. 1997) yielded about 50 positive clones. When digested with *Hin*dIII and analyzed by Southern blotting, a majority of them showed one or two small (<4 kb) bands hybridizing with both 5′ and 3′ probes derived from the cDNA (not shown). Subcloning and sequencing (Achouri et al. 1997) showed that these clones contained intronless sequences with multiple mutations compared with the cDNA (not shown), allowing the conclusion that these sequences corresponded to pseudogenes.

Southern blotting showed that other clones contained several fragments that hybridized either with the 5′ probe or with the 3′ probe (not shown). Their inserts were subcloned in pBluescript and sequenced, allowing the identification of all exons except exons 6–8. The ≈4-kb region that was not covered by the clones was obtained by PCR with primers corresponding to the missing exons or intronic primers, and the genomic DNA library as template. Except for a ≈250-bp, highly repetitive region in the second half of intron 5, the whole gene and 5 kbp upstream of the 1st exon were completely sequenced on both strands (≈20 kbp) or on one strand with large overlaps between successive starting points (≈14 kbp).

Figure 1 shows the structure of the gene. It contains 13 exons, one of which (exon 1′) is expressed in testis but not in other tissues (see below). As shown in Table 1, all splicing sites conform to the known consensus (Padgett et al. 1986). The introns range in size from 0.1 to 10.5 kb, the largest one being between exons 5 and 6. The structure of the gene markedly differs from those of *Caenorhabditis elegans* (accession number Z83219) and *Arabidopsis thaliana* (AB010407.1), where only three exons are found. The rat *Phgdh* gene and its 5′ flanking sequence contain several dinucleotide or trinucleotide repeats (e.g., a 24-TG repeat 380 bp upstream of the initiator ATG), which could be helpful as polymorphic markers.

RACE experiments were performed to determine the position of the start site(s) in different tissues (liver, kidney, brain, lung, and testis). The primers used for these experiments were derived from exon 5 and exon 2, to allow the possible detection of an alternative 1st exon. Sequencing of 24 clones derived from different tissues (8 from liver and 4 from each of the four other tissues) indicated that 10 of the RACE products started at position −21, compared with the published cDNA sequence, 7 of them at position −15, and 6 of them at different positions between −42 and +8 (Fig. 1), without any apparent tissue preference. These results indicated that the *Phgdh* gene, like other housekeeping genes, has no precise transcription start, in agreement with the lack of TATA box. Furthermore, the promoter used in the liver did not appear to be different from the one used in other tissues where the expression of the gene is not under dietary control.

One of the four RACE clones from testis started at position −247, but in this clone, nucleotides −206 to +38 had been excised, in agreement with the presence of splice site consensus sequences. This alternative transcript has the same initiator ATG codon as the known cDNA and encodes, therefore, the same protein. To check the use of the more upstream start site, we analyzed total RNA from different tissues by RT-PCR by using a sense primer in exon 1′ (Fig. 1B) and an antisense primer in exon 2. PCR amplification resulted in the formation of a product with the expected size (≈ 240 bp) only when it was performed with cDNA from testis. In contrast, amplification of a ≈270-bp fragment was observed with cDNA from all tissues when a primer corresponding to exon 1 (shown in Fig. 1) was used together with the antisense primer in exon 2. These results indicated that the more upstream start site was used in testis but not in liver, kidney, brain, or lung.

The rat *Phgdh* gene was localized by using mouse × rat cell hybrids, which segregate rat chromosomes (Szpirer et al. 1984). The probe used, a *Pst*I intronic fragment corresponding to nucleotides 176–837 of the first intron and labelled by random priming,

Database accession number: RNO271975

Correspondence to: E. Van Schaftingen; E-mail: vanschaftingen@bchm.ucl.ac.be

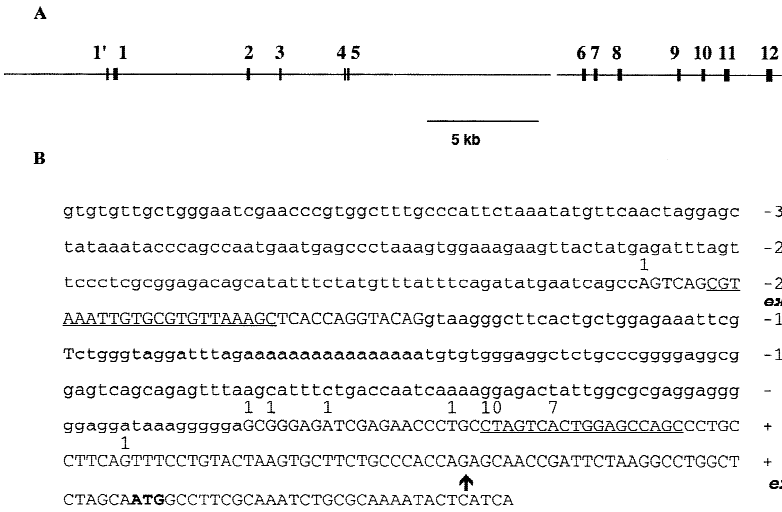

Fig. 1. Structure of rat *Phgdh* gene. **(A)** shows the general structure and **(B)** the sequence of the region of exon 1′ and 1. In B, exonic sequences are capitalized. Nucleotide 1 is the first nucleotide of the published cDNA sequence (Achouri et al. 1997). The positions of the first nucleotide of the RACE products are indicated above the sequence with the number of occurrences. The sequences used to design 5′ primers for the specific amplification of transcripts starting with exon 1′ or 1 are underlined. The arrow indicates the position of the acceptor site for exon 1′.

Table 1. Sequence of the intron/exon borders and localisation of exons in the protein sequence.

Exon No.	Exon size (bp)	5′ Splice donor	Intron No.	Intron size (kb)	3′ Acceptor site	Splicing site in the amino acid sequence[a]
1′	41	GTACAGgtaagggctt	1′	0.2	tgcttctgcccaccagAGCAAC	
1	>206 168[b]	CTCCAGgtgaggcgag	1	5.9	gtcttcctctgcccagGACTGT	44 LQ/DC
2	152	CATGAAgtaagttgtg	2	1.4	tcttctgcatctttagCACCCC	95 MN/TP
3	66	GGCCAGgtaagtccct	3	2.7	tgccctgttcttgcagGCAGAT	117 AR/QI
4	55	AAGAAGgtgagctgtt	4	0.1	acctttgcggttgcagTTCATG	135 KK/FM
5	99	ATGAAGgtaagaggtt	5	10.5	cttctcttctgaccagACTGTA	168 MK/TV
6	133	CTACAGgtaggctctc	6	0.4	ctcctcttgcctccagGCTTGC	212 TT/GL
7	149	ACAGAAgtaagtgctt	7	0.9	atatctctctgggcagGAGCCA	262 TE/EP
8	153	GGGGTTgtaagtatgt	8	2.5	tcccctctttctgcagGTAAAC	313 GV/VN
9	133	CACAAGgtgagcctag	9	0.9	ttctctgtgcctccagGAACAT	357 TQ/GT
10	131	CTCAATgtgcgtgtgc	10	0.9	tgtcgtttctctgcagGTCACC	401 LN/VT
11	238	TGATCGgtgagaggct	11	0.7	tttcctttcctcccagGCCTAC	480 MI/GL
12	247					

[a] Codons split by the splicing sites are underlined; numbers indicate the position of the first indicated residue. Nucleotides in capitals are from exons.
[b] If exon 1′ present.

detected a single ≈25-kbp *Sca*I restriction fragment in rat DNA. Mouse DNA showed a fragment at 7.5 kb (data not shown). The rat fragment cosegregated with rat Chromosome (Chr) 2 (no discordant clone); at least three discordant clones were counted for each of the other chromosomes. The rat *Phgdh* gene thus resides on rat Chr 2. Regional chromosome localization of the rat gene was obtained by fluorescent in situ hybridization (FISH) as described previously (Pinkel et al. 1988; Szpirer et al. 1998). The probe used was a full-length cDNA fragment (Achouri et al. 1997). As illustrated in Fig. 2, the cDNA probe generated double chromatid signals on Chr 2 only and unambiguously mapped at 2q34. Several rat ESTs corresponding to PHGDH have been identified (UIowa Cluster: RN.UI.3760, available at: http://ratest.eng.uiowa.edu); their radiation hybrid mapping is not yet available and may be difficult to obtain owing to the presence of several pseudogenes in the rat genome.

Radiation hybrid mapping indicates that the human PHGDH gene is localized close to marker D1S514 in the centromeric region of Chr 1 (Unigene Hs. 3343). Another localization mentioned in Unigene, on the long arm of Chr 2, is due to the fact that the human clone encoding PHGDH (NM_006623) reported by Cho et al. (2000) is chimeric, containing at its 5′ end 590 nucleotides of an unrelated cDNA. The localization of the genes encoding PHGDH in the human and the rat genomes is in agreement with the finding that several genes that are found in the rat Chr region 2q34 map to the human Chr interval 1p21–1q21 (Szpirer et al. 1998).

To check the activity of the promoter, various portions of the

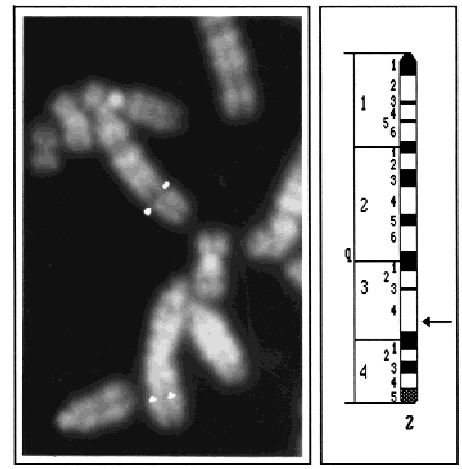

Fig. 2. Regional localization of the rat *Phgdh* gene by FISH. A portion of a metaphase is shown. The signals generated by the cDNA probe map to 2q34.

5′ flanking sequence of the *Phgdh* gene were inserted in front of the sequence encoding firefly luciferase (in pGL_2 basic from Promega). FTO2B hepatoma cells (200,000 cells/3-cm dish; grown in DMEM/Ham's F12 medium containing 10% fetal calf serum)

were transfected with 10 μg of these constructs (with the calcium phosphate method; Sambrook et al. 1989), together with a second plasmid (pRL-CMV, from Promega, 2 μg) containing a control promoter driving the expression of *Renilla* luciferase. After an overnight incubation, the cells were washed with phosphate buffer saline and further incubated with DMEM medium for 24 h. At this time, they were washed and lysed, and the activities of both firefly and *Renilla* luciferases were assayed in cytosolic extracts using the Dual-Luciferase™ Reporter Assay system (Promega). The following values were obtained (triplicates ± SEM): no promoter: 0.08 ± 0.06; nucleotides −580 to +22 (same numbering as in Fig. 1): 0.71 ± 0.04; nucleotides −765 to +22: 0.75 ± 0.07; nucleotides −1560 to +22: 5.5 (2 values). These experiments indicated that the 5′ region of the PHGDH gene has promoter activity and that it contains a stimulatory element between positions −1560 and −765. No difference in activity was observed if the transfected cells were incubated in the absence of cysteine. This is in agreement with previous results showing that the presence of cysteine does not affect transcription of the PHGDH gene in isolated hepatocytes, as assessed by nuclear run-on assays, but that it decreases the stability of the mRNA (Achouri et al. 1999).

Acknowledgments. We thank M. Rivière and P. Van Vooren for excellent technical assistance. This work was supported by the Actions de Recherche Concertées, by the Belgian Federal Service for Scientific, Technical and Cultural Affairs, by the Fonds National de la Recherche Scientifique, and the Fonds de la Recherche Scientifique Médicale. C. Szpirer is a Research Director of the FNRS (Belgium).

References

Achouri Y, Rider MH, Van Schaftingen E, Robbi M (1997) Cloning, sequencing and expression of rat liver 3-phosphoglycerate dehydrogenase. Biochem J 323, 365–370

Achouri Y, Robbi M, Van Schaftingen E (1999) Role of cysteine in the dietary control of the expression of 3-phosphoglycerate dehydrogenase in rat liver. Biochem J 344, 15–21

Cho HM, Jun DY, Bae MA, Ahn JD, Kim YH (2000) Nucleotide sequence and differential expression of the human 3-phosphoglycerate dehydrogenase gene. Gene 245, 193–201

Fallon HJ, Hackney EJ, Byrne WL (1966) Serine biosynthesis in rat liver: regulation of enzyme concentration by dietary factors. J Biol Chem 241, 4157–4167

Mauron J, Mottu F, Spohr G (1973) Reciprocal induction and repression of serine dehydratase and phosphoglycerate dehydrogenase by proteins and dietary-essential amino acids in rat liver. Eur J Biochem 32, 331–342

Padgett RA, Grabowski PJ, Konarska MM, Seiler S, Sharp PA (1986) Splicing of messenger RNA precursors. Annu Rev Biochem 55, 1119–1150

Pinkel D, Landegent J, Collins C, Fuscoe J, Segraves R et al. (1988) Fluorescent in situ hybridization with human chromosome-specific libraries: detection of trisomy 21 and translocations of chromosome 4. Proc Natl Acad Sci USA 85, 9138–9141

Sambrook J, Fritsch EF, Maniatis T (1989) *Molecular Cloning: A Laboratory Manual*, 2nd edn. (Cold Spring Harbor, N.Y.: Cold Spring Harbor Laboratory)

Snell K (1984) Enzymes of serine metabolism in normal, developing and neoplastic rat tissues. Adv Enzyme Regul 22, 325–400

Szpirer C, Szpirer J, Van Vooren P, Tissir F, Simon JS et al. (1998) Gene-based anchoring of the rat genetic linkage and cytogenetic maps: new regional localizations, orientation of the linkage groups and insights into mammalian chromosome evolution. Mamm Genome 9, 721–734

Szpirer J, Levan G, Thörn M, Szpirer C (1984) Gene mapping in the rat by mouse-rat cell hybridization: synteny of the albumin and alpha-foetoprotein genes and assignment to chromosome 14. Cytogenet Cell Genet 38, 142–149

Vinals C, Depiereux E, Feytmans E (1993) Prediction of structurally conserved regions of D-specific hydroxy acid dehydrogenases by multiple alignment with formate dehydrogenase. Biochem Biophys Res Commun 192, 182–188

Willis JE, Sallach HJ (1964) The occurrence of D-3-phosphoglycerate dehydrogenase in animal tissues. Biochim Biophys Acta 81, 39–54