Proc. Natl. Acad. Sci. USA
Vol. 86, pp. 6543–6547, September 1989
Biochemistry

5′ flanking sequence and structure of a gene encoding rat 6-phosphofructo-2-kinase/fructose-2,6-bisphosphatase

(dual promoters/fructose 2,6-bisphosphate/glycolysis/isozymes/X chromosome)

MARTINE I. DARVILLE, KARINE M. CREPIN, LOUIS HUE, AND GUY G. ROUSSEAU*

Hormone and Metabolic Research Unit, International Institute of Cellular and Molecular Pathology, and University of Louvain Medical School, 75 avenue Hippocrate, B-1200 Brussels, Belgium

Communicated by Christian de Duve, June 2, 1989 (received for review April 25, 1989)

ABSTRACT **The synthesis and degradation of fructose 2,6-bisphosphate, a ubiquitous stimulator of glycolysis, are catalyzed by 6-phosphofructo-2-kinase (EC 2.7.1.105) and fructose-2,6-bisphosphatase (EC 3.1.3.46), respectively. In liver, these two activities belong to separate domains of the same 470-residue polypeptide. Various mRNAs have been described for this bifunctional enzyme, which is controlled by hormonal and metabolic signals. To understand the origin and regulation of these mRNAs, we have characterized rat genomic clones encoding the liver isozyme, which is regulated by cAMP-dependent protein kinase, and the muscle isozyme, which is not. We describe here a 55-kilobase gene that encodes these isozymes by alternative splicing from two promoters. Each of the putative promoters was sequenced over about 3 kilobases and found to include nucleotide motifs for binding regulatory factors. The two isozymes share the same 13 exons and differ only by the first exon that, in the liver but not in the muscle isozyme, contains the serine phosphorylated by cAMP-dependent protein kinase. The gene was assigned to the X chromosome. An analysis of the exon limits of 6-phosphofructo-2-kinase/fructose-2,6-bisphosphatase in relation to its functional domains and to its similarity with other proteins plus its G+C content at the third codon position suggests that this gene originates from several fusion events.**

Fructose 2,6-bisphosphate is the most potent allosteric stimulator of 6-phosphofructo-1-kinase (EC 2.7.1.11), a key enzyme of glycolysis. The synthesis of fructose 2,6-bisphosphate is catalyzed by 6-phosphofructo-2-kinase (PFK-2; EC 2.7.1.105) and its degradation is catalyzed by fructose-2,6-bisphosphatase (FBPase-2; EC 3.1.3.46). In liver, these two activities belong to separate domains of each subunit (470 amino acids) of the same homodimeric protein. This bifunctional enzyme integrates a number of hormonal and metabolic signals including cAMP. In fibroblasts, PFK-2 activity is stimulated by growth factors, tumor promoters, and tyrosine-specific oncogenic protein kinases (1, 2).

Biochemical and immunologic data suggest the existence of distinct PFK-2/FBPase-2 isozymes. Using PFK-2/FBPase-2 cDNA probes from rat liver (4), we have identified a 2.1-kilobase (kb) mRNA coding for the liver (L) isozyme, a 1.9-kb mRNA coding for the skeletal muscle (M) isozyme, and a 6.8-kb mRNA presumably coding for the heart isozyme (5). The L and M isozymes have a common sequence of 438 amino acids and differ only at the N terminus. In the L isozyme, the divergent sequence is 32 residues long and includes the serine (Ser-32) phosphorylated by cAMP-dependent protein kinase. In the M isozyme, the N terminus is 10 residues long and is unrelated to the unique part of the L isozyme. Expression of the L, but not the M, isozyme of PFK-2/FBPase-2 is controlled by diet and insulin (6, 7).

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

An analysis of the PFK-2 and FBPase-2 domains in relation to the polypeptide sequence suggested that they originate from a gene fusion event (2). This bifunctional enzyme also exhibits intriguing similarities with phosphoglycerate mutase [PGAM; EC 5.4.2.1] and bisphosphoglycerate mutase [BPGAM; EC 5.4.2.4] (8), serine proteases (8), and oncogenic virus proteins. To gain insight into the origin and regulation of the PFK-2/FBPase-2 gene(s) and isozymes, we have analyzed rat genomic clones identified with our cDNA probes.† We have characterized a gene that encodes the L and M isozymes by alternative use of different promoters. This gene is 55 kb long and contains 15 exons, 13 of which are common to the two isozymes. Part of this work has been presented in the form of an abstract (9).

MATERIALS AND METHODS

Screening of the Rat Genomic Library. The rat genomic library, prepared from liver DNA in λEMBL 3 (10), was a gift from W. Schmid and G. Schütz (Krebsforschungzentrum, Heidelberg). Approximately 3×10^6 phages were screened by plaque hybridization. Replicas were made on nylon membranes (Amersham) according to the supplier's instructions. The membranes were prehybridized at 65°C in 6× SSC (1× SSC = 0.15 M NaCl/0.015 M sodium citrate, pH 7)/10× Denhardt's solution (11). Hybridization with multiprime-labeled (Amersham) cDNA probes was carried out for 16 hr at 65°C in 3.5× SSC/1× Denhardt's solution/25 mM sodium phosphate, pH 7/2 mM EDTA/0.5% $NaDodSO_4$/heat-denatured salmon sperm DNA (100 μg/ml), using 5×10^5 cpm/ml (50 μl/cm² of membrane). Washing was performed at 50°C in 0.2× SSC/0.1% $NaDodSO_4$. Phage DNA was extracted by the plate lysate method (11).

Restriction Mapping and Sequencing of the Clones. Mapping of recombinant phages for the restriction endonucleases *Bam*HI, *Bgl* II, *Eco*RI, and *Hin*dIII was performed by the Southern Cross Restriction Mapping System following the instructions of NEN. The order of some fragments was deduced from partial digestions. DNA fragments containing exons were identified by Southern blot analysis (12) under the same hybridization conditions used for the screening of the library. These fragments were purified on agarose gels (13) and subcloned, some of them after digestion by another restriction endonuclease, in the phages M13mp8 and/or -mp9 (14). The recombinant M13 clones were sequenced by the dideoxynucleotide chain-termination method (15), using successive primers (Eurogentec, Liège, Belgium). Some clones,

Abbreviations: FBPase-2, fructose-2,6-bisphosphatase; PFK-2, 6-phosphofructo-2-kinase; PGAM, phosphoglycerate mutase; BPGAM, bisphosphoglycerate mutase; L, liver; M, muscle.

*To whom reprint requests should be addressed at: UCL-ICP, Box 7529, 75 avenue Hippocrate, B-1200 Brussels, Belgium.

†The sequence reported in this paper has been deposited in the GenBank data base (accession no. M26215, for the 2117-bp fragment of clone λ20, and M26216, for the 3720-bp fragment of clone λ20).

including those containing the 5′ end of the gene, were sequenced after progressive unidirectional deletions of the insert, generated by exonuclease III (16) according to the instructions of Promega.

S1 Nuclease Mapping. This procedure was performed as described (17). Synthetic oligonucleotides (Eurogentec) were 5′ labeled with T4 polynucleotide kinase. They were hybridized to the appropriate recombinant M13 template and extended with unlabeled dNTPs and the Klenow fragment of DNA polymerase I. The DNA was digested with a restriction enzyme and the single-stranded probes were isolated on alkaline low-melting-point agarose gels. The probes (5×10^4 cpm) were hybridized with 50 μg of total rat liver or muscle RNA (5) at 30°C for 16 hr in 20 μl of 80% (vol/vol) formamide/400 mM NaCl/1 mM EDTA/40 mM Pipes·NaOH, pH 6.4. The S1 nuclease digestion was performed at 30°C for 1 hr in 300 μl of S1 buffer (280 mM NaCl/4.5 mM $ZnSO_4$/50 mM sodium acetate, pH 4.5/20 μg of heat-denatured salmon sperm DNA per ml) containing 100–1000 units of enzyme. The products were analyzed on 5% or 6% acrylamide sequencing gels, with a M13 sequencing product as marker.

Chromosomal Assignment of the Gene. This procedure was performed by Southern blot analysis of DNA (generous gift of C. Szpirer, Free University of Brussels) from mouse–rat somatic hybrids that segregate rat chromosomes (18). The blots were washed at 50°C in 2× SSC/0.1% $NaDodSO_4$.

RESULTS AND DISCUSSION

Isolation and Analysis of Genomic Clones. The rat genomic library was screened with two rat liver PFK-2/FBPase-2 probes derived from the fully coding L-type cDNA RL2K-22c (4). Probe 22c2 extends from the 5′ end to the *Eco*RI site at nucleotide 446 (amino acid 90) and probe RL2K-8 spans nucleotides 674 (amino acid 167) to 1759 (within the 3′ untranslated sequence) of this cDNA. Eleven independent clones were isolated. Their orientation and the assignment of the exons was performed by Southern blot hybridization with probe 22c2 and with the three fragments of RL2K-8 generated by *Bgl* II and *Pst* I digestion. Eight clones were chosen to establish, by the Southern Cross method, the gene restriction map for *Bam*HI, *Bgl* II, *Eco*RI, and *Hin*dIII (Fig. 1). These clones overlapped, except for a gap of ≈0.2 kb between clones λ36 and λ41, as estimated by Southern blot hybridization of rat genomic DNA. Indeed, two genomic probes on both sides of the gap detected the same 10.2-kb *Bam*HI, 8.4-kb *Eco*RI, and 8.2-kb *Hin*dIII fragments.

Organization of the Gene and Origin of the L and M Isozymes. Southern blot hybridization of restricted genomic clones with the cDNA probes and sequencing of exon-containing fragments showed that the gene is at least 55 kb long and contains 15 exons (Fig. 1). The L isozyme is entirely accounted for by 14 exons (exons 1′ and 2–14) that encode amino acid sequences from 19 (exon 6) to 69 (exon 8) residues long (Table 1) and no known exon lies in the 0.2-kb gap between clones λ36 and λ41. Thus, the abnormally long liver cDNA clone RL2K-4c (4) is a L-type mRNA precursor containing an unspliced intron c.

The genomic clones were hybridized with probe 5c2, which corresponds to the first 303 nucleotides of the M-type cDNA RL2K-5c (5). This showed the occurrence, 4.3 kb upstream from exon 1′, of another exon that encodes the unique 5′ sequence of the M-type mRNA (Fig. 1). Therefore, the L- and M-type mRNAs arise by alternative splicing of either one of the first two exons with the 13 other exons. The four cysteines that are essential for the kinase activity of the enzyme (8) are contributed by exon 4 (Cys-107), exon 6 (Cys-160), and exon 7 (Cys-183 and -198). The histidine (His-258) that is phosphorylated by fructose 2,6-bisphosphate during the bisphosphatase reaction (8) is encoded by exon 8. Exon 1′ contains the amino acid consensus sequence for recognition of the L isozyme by cAMP-dependent protein kinase, and the codon for the phosphorylated serine (Ser-32) results from the splicing of this exon with exon 2 (Table 1). In agreement with the literature (19), all the introns have a GT at their 5′ end and an AG at their 3′ end. The origin of the mRNA for the heart isozyme, which appears to share homology with the 5′ but not the 3′ moiety of the L-type mRNA (5), remains unknown. Our identification (M.I.D., unpublished data) of another rat gene that resembles the one described here might help to answer this question.

5′ End of the Gene. We have shown by primer extension (5) that cDNAs 22c and 5c, respectively, lack about 75 base pairs (bp) and 22 bp at the 5′ end. To identify the transcription initiation (cap) site(s) by S1 nuclease mapping (Fig. 2), we first sequenced the 3.7-kb *Hin*dIII fragment of genomic clone λ20 that contains exon 1′ and the 2.9-kb *Bam*HI fragment of this clone that contains exon 1. For the L-type mRNA, we used a primer (Fig. 3) complementary to nucleotides 80–103 of cDNA 22c. The *Hpa* II site upstream from exon 1′ in the 3.7-kb *Hin*dIII fragment was chosen to generate a 622-base probe. This single-stranded probe was hybridized to rat liver RNA. Most of the fragments protected against S1 nuclease were 178–180 bases long. Less abundant populations of 200- and 161-base fragments were also seen (Fig. 2*A*). This pattern was identical to that obtained (5) by extension of L-type mRNA with the same primer. In this genomic fragment, there is no splice site consensus upstream from cDNA 22c that would correspond to the 5′ end of the protected fragments. Thus, the major cap site for the L-type mRNA is tentatively assigned, in this fragment, to the adenines in the CAGA sequence that contains the well-conserved (20) CA dinucleotide 78 bp upstream from the sequence corresponding to the 5′ end of cDNA 22c (Fig. 3). A sequence compatible with a TATA box (21) is found 30 bp upstream and a CCAAT homology on the complementary strand is found 107 bp upstream from this putative cap site. No TATA box is found at an appropriate position upstream from the other protected fragments. This putative promoter also contains consensus sequences (22) for binding the phorbol ester-activated transcription factor AP-1 (c-*jun*) and the glucocorticoid receptor. This is interesting because phorbol esters (23) and glucocorticoids (7) stimulate PFK-2 activity.

To localize the cap site of the M-type mRNA, we used a primer (Fig. 3) complementary to nucleotides 30–53 of cDNA 5c. The *Hae* III site upstream from exon 1 in the 2.9-kb

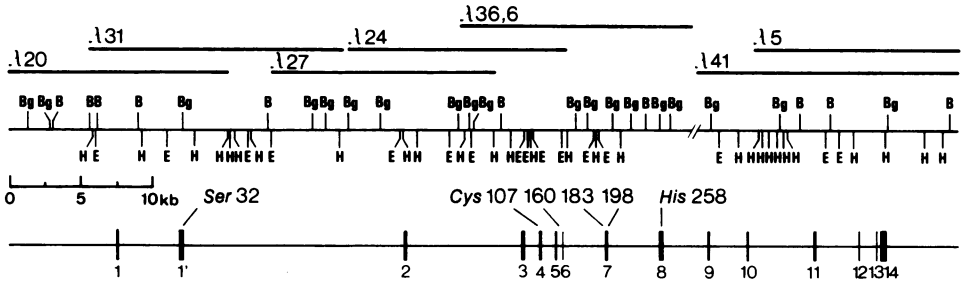

FIG. 1. Gene for PFK-2/FBPase-2. Upper horizontal lines refer to the rat genomic clones indicated. Restriction map shows sites for *Bam*HI (B), *Bgl* II (Bg), *Eco*RI (E), and *Hin*dIII (H). Location and size of the 15 exons, with position of the residues critical for activity, are indicated by vertical bars on the lower line.

Table 1. Localization and sequence of exon–intron junctions

Exon	Position in cDNA*		5′ splice site	Intron	3′ splice site		
1	1–126	(Ala-10)	cag GTAGGT	--[a]--	TGTCTTCCAG	cct	(Ala-10)
1′	1–269	(Ser-32)	gct GTAAGT	--[a]--	TGTCTTCCAG	cct	(Ser-32)
2	270–395	(Val-74)	aag GTACGT	--[b]--	CTGTTTTCAG	tgt	(Val-74)
3	396–489	(Lys-105)	gaa GTAAGT	--[c]--	TGTTTTTCAG	gca	(Lys-105)
4	490–556	(Ala-127)	gcg GTAAAT	--[d]--	ATGTCTCCAG	gtt	(Val-128)
5	557–631	(Lys-152)	aag GTAGAG	--[e]--	TCTCTTATAG	gtc	(Val-153)
6	632–688	(Lys-171)	aag GTAAGG	--[f]--	CTTCTCACAG	caa	(Gln-172)
7	689–810	(Ser-212)	cag GTAAGA	--[g]--	TCTGGCCCAG	cca	(Ser-212)
8	811–1018	(Gln-281)	cag GTAAGG	--[h]--	GTACCCCTAG	tat	(Tyr-282)
9	1019–1165	(Ala-330)	gcg GTAAGA	--[i]--	CTTTTTATAG	ggt	(Gly-331)
10	1166–1270	(Glu-365)	gag GTAACG	--[j]--	TCTTTTTTAG	tcc	(Ser-366)
11	1271–1400	(Asp-409)	cag GTAACA	--[k]--	TGCCTTTCAG	atg	(Asp-409)
12	1401–1463	(Gly-430)	atg GTAAGT	--[l]--	TTTTTTTTAG	gct	(Gly-430)
13	1464–1528	(Glu-451)	gag GTAGGT	--[m]--	CTCTTCTTAG	aat	(Asn-452)
14	1529–1830	(Tyr-470)					
Rodent consensus			nag GTRAGT		YYYYYYNCAG	g	

Nucleotides in exons are lowercase letters and in introns are capital letters; introns names are in boldface type. N, any nucleotide; R, a purine; Y, a pyrimidine. The codons underlined are those for the amino acids indicated, which refer to those in the M isozyme for exon 1 and in the L isozyme for the other exons. Sizes (kb) of introns are as follows: a, 15.3; b, 8; c, 1; d, 1; e, 0.4; f, 2.7; g, 3.6; h, 2.8; i, 2.4; j, 4.5; k, 3; l, 1.2; m, 0.3.

*Nucleotides in cDNA 5c for exon 1 and in cDNA 22c for the other exons.

*Bam*HI fragment of clone λ20 was chosen to generate a 686-base probe that was hybridized with rat muscle RNA. Most of the fragments protected against S1 nuclease had 73 or 74 bases; fragments of 61 and 97 bases were also seen. This pattern (data not shown) was identical to that obtained (5) by extension of muscle mRNA with the same primer. This would locate the cap site for the M-type mRNA at the adenine 21 bp upstream from the 5′ end of cDNA 5c and 32 bp downstream from a TATA box in this genomic fragment. However, this genomic sequence contains a 3′ splice site consensus at the junction with the 5′ end of cDNA 5c. To confirm this cap site, we repeated the experiment with a primer (Fig. 3) complementary to nucleotides 72–88 of cDNA 5c. The *Sca* I site upstream from exon 1 was used to generate a 230-base probe. The majority of the fragments protected were 109 bases long and a minority were 130 and 95–97 bases long, consistent with the previous experiment (Fig. 2*B*). However, some fragments had 86 bases, as if the 5′ end of a less abundant muscle PFK-2/FBPase-2 mRNA population was coded for by genomic sequences upstream from the proposed cap site. This putative promoter for the M-type mRNA also contains an AP-1 consensus sequence as well as three CCAAT homologies, two of which are on the other strand (Fig. 3). Finally, the region between exons 1 and 1′ contains the sequence $[d(GT)]_{28}$, typical of Z-DNA, and might therefore play a regulatory role (24).

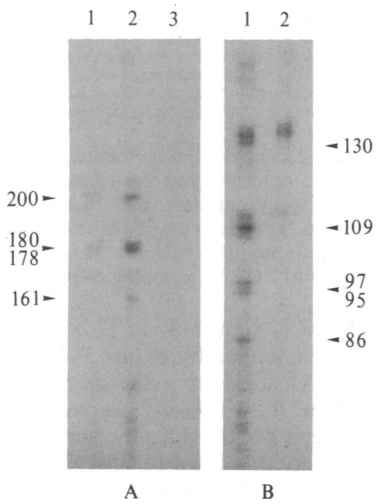

FIG. 2. S1 nuclease mapping analysis of exons 1′ (*A*) and 1 (*B*). (*A*) Radioactive probe (see text and Fig. 3) was incubated with rat liver RNA (lanes 1 and 2) or yeast tRNA (lane 3) and digested with 300 (lanes 1 and 3) or 100 (lane 2) units of S1 nuclease prior to loading on the gel. (*B*) Radioactive probe (see text and Fig. 3) was incubated with rat muscle RNA (lane 1) or yeast tRNA (lane 2) and digested with 200 units of S1 nuclease prior to loading on the gel. Arrowheads point to the length, in nucleotides, of protected fragments.

We conclude that the L and M isozymes arise through alternative use of contiguous promoters and consequent differential splicing. Our detection of fragments protected against S1 nuclease that are longer than those corresponding to each major cap site is probably due to promoter heterogeneity. This is a frequent feature of housekeeping genes that lack a TATA box, the sequence required for univocal positioning of RNA polymerase II (25). The putative TATA boxes shown in Fig. 3 differ considerably from the typical TATAAAA sequence (21). Consistent with the housekeeping character of the gene described here, its two promoters are active in many tissues (5). Still, one of these promoters was called L type because it controls expression of a mRNA encoding the sequence of the protein actually purified from liver. The other promoter was called M-type because it controls expression of a mRNA that is more abundant in muscle than in other tissues and encodes a protein sequence whose expected properties are typical of the M isozyme.

Chromosomal Assignment of the Gene. The DNA of mouse–rat somatic hybrids was digested with *Bam*HI since the hybridization patterns of the DNA of the two species with probe RL2K-8 were different. The rat-specific restriction fragments (10.2 kb and 8.2 kb) were not detected in the DNA of the cell hybrids LB 330.TG3, LB 330.TG6, and LB 1040.TG5, which lack rat chromosomes X and 13. However, these fragments were detected in the DNA of the other cell hybrids including LB 330, LB 780.8, and LB 1040, which retain chromosome X but not chromosome 13 (Table 2). No discordant hybrid was observed for chromosome X. The gene was, therefore, assigned to the X chromosome. We have shown that fructose 2,6-bisphosphate is present in sperm cells (26). Unless only the X chromosome-bearing male gametes contain this molecule, one has to postulate that all

FIG. 3. Sequence of the 5′ flanking region and two putative promoters of the PFK-2/FBPase-2 gene. (*A*) Exon 1-containing *Bam*HI fragment of genomic clone λ20. (*B*) Exon 1′-containing *Hin*dIII fragment of clone λ20. Numbers in the right margin refer to nucleotides. Exons are underlined, with the proposed translation initiation codon indicated by an overbar. *, Proposed major cap sites; ●, 5′ end of the corresponding cDNA. Above the corresponding sequences are shown the primers used for S1 mapping (see Fig. 2) and the putative binding sites for regulatory factors. GRE?, downstream half of a glucocorticoid-response element. A sequence containing a putative TATA box is shown 32 bp upstream from exon 1, and one such box is shown 30 bp upstream from exon 1′. Other boxes correspond to CCAAT homologies on either DNA strand.

spermatozoa possess PFK-2 activity. This activity might result from translation of an mRNA produced prior to meiosis, acquired through post-meiotic intercellular cytoplasmic bridges, or transcribed from a homologous gene present on an autosome and activated during spermatogenesis, as is the case for phosphoglycerate kinase (27).

Evolutionary Considerations. Our data support the exon shuffling hypothesis (28) and shed light on the evolutionary origin of the PFK-2/FBPase-2 gene. Pilkis *et al.* (2) postulated that liver PFK-2/FBPase-2 arose from the fusion of two genes. Our work not only supports this hypothesis but also delineates the exons probably recruited in this process. The four cysteines that are critical for PFK-2 activity lie between residues 106 and 199. A region of papillomavirus protein E1 that contains cysteines at positions corresponding to Cys-107, -183, and -198 is similar to PFK-2/FBPase-2 residues 95–200, which are encoded by exons 3–7. Thus, these five exons might have clustered to make up the PFK-2 domain. By the same token, the FBPase-2 domain might have been contributed by exons 8 and 9 since the similarity with PGAM and BPGAM (8) extends from residue 250 to residue 330. Recruitment of exon 1′ has conferred sensitivity to cAMP-dependent protein kinase. Exon 10 contains Tyr-359, which is preceded by a motif for phosphorylation by tyrosine-specific protein kinase (4). Thus, the gene described here might result from assembly of two catalytic domains—i.e., PFK-2 (exons 3–7) and FBPase-2 (exons 8 and 9)—and of several regulatory domains encoded by the other exons. This interpretation is supported by the difference in G+C content at the third position of codons (29) in the exons. Indeed, this

Table 2. Assignment of the rat PFK-2 gene to chromosome X by molecular hybridization and synteny analysis

Cell hybrid	Rat chromosome																					PFK-2 gene*
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	X	
LB 150.1	−	−	+	+	−	−	+	−	+	−	+	+	+	−	−	+	+	+	+	−	+	+
LB 161	−	+	+	+	+	+	+	−	+	+	−	−	+	+	+	+	+	+	+	+	+	+
LB 210-I	−	−	−	−	−	−	−	−	−	−	−	−	+	+	−	+	−	+	−	−	+	+
LB 251	+	+	−	+	−	−	+	−	−	+	−	+	+	−	−	−	+	−	+	−	+	+
LB 330	−	+	−	+	−	+	−	−	−	+	−	+	−	−	−	+	+	−	−	−	+	+
LB 330.TG3	−	−	−	+	−	−	−	−	−	+	−	+	−	−	−	−	−	−	−	−	−	−
LB 330.TG6	−	+	−	+	−	+	−	−	−	+	−	+	−	−	−	−	−	−	−	−	−	−
LB 510.6	−	+	+	+	−	−	+	−	−	−	−	+	+	+	+	+	+	+	−	−	+	+
LB 600	+	+	+	+	+	+	+	−	−	+	+	+	+	+	+	+	−	+	+	−	+	+
LB 630	−	−	+	+	+	+	+	−	+	−	+	+	+	+	+	+	−	+	+	−	+	+
LB 780.6	−	+	−	−	−	−	−	−	−	+	+	−	+	−	−	−	−	+	−	−	+	+
LB 780.8	−	+	−	−	−	−	+	−	−	+	+	−	−	−	−	−	−	+	−	−	+	+
LB 810	−	+	+	+	−	+	+	+	−	+	+	+	+	+	+	+	+	−	+	+	+	+
LB 860	−	+	+	+	−	−	+	−	+	−	+	+	+	−	+	+	+	+	−	+	+	+
LB 1040	−	−	+	+	−	+	+	−	−	+	+	+	−	−	+	+	−	+	−	+	+	+
LB 1040.TG5	−	−	−	+	−	+	−	−	−	−	+	−	−	−	+	+	−	+	−	−	−	−
BS 511	−	+	+	+	−	+	+	+	−	−	+	−	+	−	+	+	+	−	+	−	+	+
Concordant hybrids, no.	5	12	12	11	6	8	14	5	7	9	11	10	14	9	10	13	11	12	10	7	17	
Concordant frequency, %	29	71	71	65	35	47	82	29	41	53	65	59	82	53	59	76	65	71	59	41	100	

*Presence (+) or absence (−) of the hybridizing rat bands in the Southern blots.

value is about the same (36–52%; mean = 46%) in exons 3–7, it is high (64% and 61%, respectively) in exons 8 and 9, and it is lower in exons 2 (55%), 10 (46%), 11 (49%), 12 (48%), and 14 (53%). High values are found in exons 1′ (61%) and 13 (77%). Thus, the two ancestral genes contributing the PFK-2 and FBPase-2 domains came from quite different genomic environments, and unrelated exons also seem to have been involved in the fusion events. Tauler *et al.* (30) have shown that, besides a sequence similarity with PGAM and BPGAM, the FBPase-2 domain displays some of the catalytic properties of these enzymes, strengthening the idea of an evolutionary kinship. Sakoda *et al.* (3) have proposed that BPGAM originated as a separate entity very soon after the appearance of the ancestral yeast PGAM gene. Indeed, only 50% of the residues of mammalian BPGAM and PGAM are identical. A comparison of amino acids 250–330 in PFK-2/FBPase-2 with the corresponding sequence in the PGAM family shows 38% identity with yeast PGAM but only 30% identity with the mammalian enzyme whether it is PGAM or BPGAM. Since the identity between yeast PGAM and mammalian PGAM or BPGAM over the same sequence is 49%, the ancestral gene for the FBPase-2 domain may have arisen from the PGAM tree before the split between yeast PGAM and the ancestor of the present-day PGAM and BPGAM.

We thank M. Place and I. Antoine for assistance, Dr. M. H. Rider for helpful suggestions, and V. Henry and T. Lambert for typing the manuscript. K.M.C. held a fellowship from the Institut pour l'Encouragement de la Recherche Scientifique dans l'Industrie et l'Agriculture. This work was supported in part by the Belgian State–Prime Minister's Office–Science Policy Programming (Incentive Program in Life Sciences 88/93-122 Grant 20) and by the Caisse Générale d'Epargne et de Retraite (Belgium).

1. Hue, L. & Rider, M. H. (1987) *Biochem. J.* **245**, 313–324.
2. Pilkis, S. J., El-Maghrabi, M. R. & Claus, T. H. (1988) *Annu. Rev. Biochem.* **57**, 755–783.
3. Sakoda, S., Shanske, S., DiMauro, S. & Schon, E. A. (1988) *J. Biol. Chem.* **263**, 16899–16905.
4. Darville, M. I., Crepin, K. M, Vandekerckhove, J., Vandamme, J., Octave, J. N., Rider, M. H., Marchand, M. J., Hue, L. & Rousseau, G. G. (1987) *FEBS Lett.* **224**, 317–321.
5. Crepin, K. M., Darville, M. I., Hue, L. & Rousseau, G. G. (1989) *Eur. J. Biochem.*, in press.
6. Crepin, K. M., Darville, M. I., Hue, L. & Rousseau, G. G. (1988) *FEBS Lett.* **227**, 136-140.
7. Colosia, A. D., Marker, A. J., Lange, A. J., El-Maghrabi, M. R., Granner, D. K., Tauler, A., Pilkis, J. & Pilkis, S. J. (1988) *J. Biol. Chem.* **263**, 18669–18677.
8. Lively, M. O., El-Maghrabi, M. R., Pilkis, J., D'Angelo, G., Colosia, A. D., Ciavola, J. A., Fraser, B. A. & Pilkis, S. J. (1988) *J. Biol. Chem.* **263**, 839–849.
9. Darville, M. I., Crepin, K. M., Szpirer, J., Levan, G., Hue, L. & Rousseau, G. G. (1989) *Arch. Int. Physiol. Biochim.* **97**, B19 (abstr.).
10. Shinomiya, T., Scherer, G., Schmid, W., Zentgraf, H. & Schütz, G. (1984) *Proc. Natl. Acad. Sci. USA* **81**, 1346–1350.
11. Maniatis, T., Fritsch, E. F. & Sambrook, J. (1982) *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Lab., Cold Spring Harbor, NY).
12. Southern, E. M. (1975) *J. Mol. Biol.* **98**, 503–517.
13. Lizardi, P. M., Binder, R. & Short, S. A. (1984) *Gene Anal. Tech.* **1**, 33–39.
14. Messing, J., Crea, R. & Seeburg, P. H. (1981) *Nucleic Acids Res.* **9**, 309–321.
15. Sanger, F., Nicklen, S. & Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. USA* **74**, 5463–5467.
16. Henikoff, S. (1984) *Gene* **28**, 351–359.
17. Ausubel, F. M., Brent, R., Kingston, R. E., Moore, D. D., Seidman, J. G., Smith, J. A. & Struhl, K. (1987) *Current Protocols in Molecular Biology* (Wiley–Interscience, New York), pp. 4.6.1–4.6.13.
18. Szpirer, J., Islam, M. Q., Cooke, N. E., Szpirer, C. & Levan, G. (1988) *Cytogenet. Cell Genet.* **47**, 42–45.
19. Shapiro, M. B. & Senapathy, P. (1987) *Nucleic Acids Res.* **15**, 7155–7174.
20. Bucher, P. & Trifonov, E. N. (1986) *Nucleic Acids Res.* **14**, 10009–10026.
21. Breathnach, R. & Chambon, P. (1981) *Annu. Rev. Biochem.* **50**, 349–383.
22. Wingender, E. (1988) *Nucleic Acids Res.* **16**, 1879–1902.
23. Bosca, L., Rousseau, G. G. & Hue, L. (1985) *Proc. Natl. Acad. Sci. USA* **82**, 6440–6444.
24. Rich, A., Nordheim, A. & Wang, A. H. J. (1984) *Annu. Rev. Biochem.* **53**, 791–846.
25. La Thangue, N. B. & Rigby, P. W. J. (1988) in *Transcription and Splicing*, eds. Hames, B. D. & Glover, D. M. (IRL, Oxford), pp. 1–42.
26. Philippe, B., Rousseau, G. G. & Hue, L. (1986) *FEBS Lett.* **200**, 169–172.
27. McCarrey, J. R. & Thomas, K. (1987) *Nature (London)* **326**, 501–505.
28. Gilbert, W. (1985) *Science* **228**, 823–824.
29. Aota, S. & Ikemura, T. (1986) *Nucleic Acids Res.* **14**, 6345–6355.
30. Tauler, A., El-Maghrabi, M. R. & Pilkis, S. J. (1987) *J. Biol. Chem.* **262**, 16808–16815.