Rat Gene Coding for Heart 6-Phosphofructo-2-kinase/Fructose-2,6-bisphosphatase: Characterization of an Unusual Promoter Region and Identification of Four mRNAs[†,‡]

Mohamed Chikri[§] 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

Received February 27, 1995; Revised Manuscript Received May 1, 1995[⊗]

ABSTRACT: We have cloned previously a 22-kb rat gene which codes for heart 6-phosphofructo-2-kinase/fructose-2,6-bisphosphatase from an ATG located in exon 2. To characterize the promoter of the gene, we have now cloned and sequenced 1.9 kb of its 5′ region and show here that it has an unusual structural and functional organization. By S1 nuclease mapping and primer extension we found that this region contains the first, noncoding, exon of a mRNA that we call R3. The sequence upstream from this exon behaved as a promoter in transient transfection assays. These assays also suggested that the gene possesses more than one promoter. Indeed, by reverse transcription−polymerase chain reaction techniques we identified three additional mRNAs that differ by their 5′ noncoding exons upstream from the common, coding, exon 2. mRNA R1 contains two 5′ noncoding exons located upstream from the first exon of mRNA R3. mRNA R2 contains one 5′ noncoding exon located upstream from, and partially overlapping with, the first exon of mRNA R3. mRNA R4 contains one 5′ noncoding exon located downstream from the first exon of mRNA R3 but overlapping partially with it. The distribution of these mRNAs in rat tissues was evaluated by reverse transcription−polymerase chain reaction. We conclude that the gene contains four more exons than the 16 previously described and at least three promoters, two of which correspond to exonic sequences. The gene gives rise to at least four mRNAs which are expressed not only in heart but also in most tissues.

6-Phosphofructo-2-kinase[1] (PFK-2)[2] /fructose-2,6-bisphosphatase (FBPase-2) is a bifunctional enzyme that catalyzes the synthesis and degradation of fructose 2,6-bisphosphate, a ubiquitous allosteric stimulator of 6-phosphofructo-1-kinase. There are several PFK-2/FBPase-2 isozymes, which differ in tissue distribution and in sensitivity to hormonal regulation [reviewed by Rousseau and Hue (1993)]. Two PFK-2/FBPase-2 rat genes have been cloned.

Gene A (at least 55 kb) gives rise to a fetal- (F-) type mRNA, a muscle- (M-) type mRNA, and a liver- (L-) type mRNA. These three mRNAs originate from different primary transcripts through the use of distinct promoters and differ by their 5′ ends. The PFK-2 and FBPase-2 domains are respectively coded by exons 2−8 and 8−13, which are common to the three mRNAs (Darville *et al.*, 1989; Dupriez *et al.*, 1993).

Gene B (at least 22 kb) has a similar organization, exons 3−14 displaying 71% amino acid sequence identity with exons 2−13 of gene A (Darville *et al.*, 1991). Examination of gene B showed that it encodes, from an ATG located in exon 2, the heart (H) isozyme. Indeed, the amino acids encoded by exon 15 include a stretch of 17 residues obtained by direct sequencing of the PFK-2/FBPase-2 protein purified from bovine heart (Kitamura *et al.*, 1988). Second, gene B accounts faithfully for the nucleotide sequence of rat heart cDNA RH1-9. This is a partial cDNA (exons 1−9) obtained by reverse transcription−polymerase chain reaction (RT-PCR) from rat heart poly(A)-rich RNA (Darville *et al.*, 1991). Third, the 481 amino acid sequence encoded by exons 3−15 of gene B displays 95% identity with the sequence derived from a fully coding bovine heart cDNA (Sakata and Uyeda, 1990).

To study the transcriptional control of the H isozyme, we decided to characterize the promoter region of the rat PFK-2/FBPase-2 gene B. We found that this sequence has peculiar features and that the gene actually gives rise to at least four mRNAs, called here R1−R4, through the use of different promoters which are organized in an unusual fashion.

MATERIALS AND METHODS

Tissues, Cells, and RNA Preparation. Tissues taken from 200-g male albino Wistar rats sacrificed under Nembutal anesthesia were frozen in liquid N_2 and kept at −80 °C. Fetal liver was taken at day 21 of gestation. The mouse Sertoli cell line TM4 was obtained from the ATCC (CRL 1715) through Professor G. Verhoeven, KUL, Leuven, Belgium.

[†]This work was supported by an ARC of the D. G. Higher Education and Scientific Research−French Community of Belgium and by the Belgian Programme on Interuniversity Poles of Attraction initiated by the Belgian State S.S.T.C., the scientific responsibility being assumed by the authors.

[‡]The gene sequences mentioned in this paper have been submitted to GenBank/EMBL Data Bank under Accession Numbers X83725 and X83735.

* To whom correspondence should be addressed at ICP-UCL, Box 7529, 75 Avenue Hippocrate, B-1200 Brussels, Belgium.

[§]Fellow of the Fonds de Développement Scientifique de l'UCL.

[⊗] Abstract published in *Advance ACS Abstracts,* June 15, 1995.

[1] Enzymes: 6-phosphofructo-1-kinase (EC 2.7.1.11); 6-phosphofructo-2-kinase (EC 2.7.1.105); fructose-2,6-bisphosphatase (EC 3.1.3.46).

[2] Abbreviations: bp, base pair(s); dNTP, a mixture of dATP, dCTP, dGTP, and dTTP; FBPase-2, fructose-2,6-bisphosphatase; Inr, initiator sequence; kb, kilobase(s); PBS, phosphate-buffered saline; PFK-2, 6-phosphofructo-2-kinase; (RT-)PCR, (reverse transcription) polymerase chain reaction; Sp1, stimulatory protein 1.

0006-2960/95/0434-8876$09.00/0

The cells were grown as monolayers at 37 °C (5% CO_2) in Dulbecco's modified Eagle's medium (DMEM)/Ham's F-12 medium (1:1) supplemented with $NaHCO_3$ (1 g/L), 1.5 mM Hepes, 5% horse serum, and 2.5% fetal calf serum (Gibco-BRL). Total cellular RNA was extracted from rat tissues by the procedure of Chirgwin *et al.* (1979) and from TM4 cells according to Auffray and Rougeon (1980). Total RNA from rat Sertoli cells was a generous gift of Prof. G. Verhoeven. Poly(A)-rich RNA was prepared by oligo(dT)–cellulose chromatography (Aviv and Leder, 1972).

Primers and Probes. The sequence of all the primers and probes used (synthesized when necessary by Pharmacia-LKB) is given in the text with reference to Figure 1 except for the following that are specific for exon 2, Chik-3, 5′-AGCATTTCTTCCCTGCCCTT-3′ (nucleotides 180–199 of cDNA RH1-9); Chik-4, 5′-TTTGAGGCGTGGGGTTTATA-3′ (nucleotides 154–173 of cDNA RH1-9); Chik-7, GCT-GCTGTCTTCTGTGGAAA-3′ (nucleotides 125–144 of cDNA RH1-9), or specific for exon 9, AB1222, 5′-ATCAC-CTTTATGAACGAGAGGTCC-3′ (nucleotides 747–770 of cDNA RH1-9). Probes supplied with the 5′ AmpliFinder Race kit (Clontech) have the following sequence: anchor, 3′NH_2-GGAGACTTCCAAGGTCTTAGCTATCACT-TAAGCAC-P5′; anchor primer, 5′-CTGGTTCGGCCCAC-CTCTGAAGGTTCCAGAATCGATAG-3′. PCR sense primers for studying the tissue distribution of mRNAs R1, R2, R3, and R4 were as follows. TD-R1 corresponds to nucleotides 41–60 of the sequence shown in Figure 1, TD-R2 to nucleotides 481–500, TD-R3 to nucleotides 1082–1105, and TD-R4 to nucleotides 1253–1272.

Cloning and Sequencing of Genomic and cDNA Clones. The 1.9-kb *Sal*I–*Hin*dIII fragment of the genomic clone λ16 and the cDNAs described were subcloned in the pBluescript KS^- vector. DNA sequences were determined by the dideoxy termination method (Sanger *et al.*, 1977) using double-stranded plasmid DNA as a template. Both strands were sequenced with specific and universal primers using a T7 sequencing kit (Pharmacia LKB) according to the manufacturer's instructions.

S1 Nuclease Mapping. This was performed according to Ausubel *et al.* (1987). Briefly, a synthetic oligonucleotide probe (AB743) was end-labeled with [γ-^{32}P]ATP in the presence of T4 polynucleotide kinase (Boehringer) and hybridized (10^6 cpm) with single-stranded DNA purified from the pBluescript vector containing the 1.9-kb *Sal*I–*Hin*dIII fragment. Primer extension was completed in the presence of unlabeled dNTP and the Klenow fragment of DNA polymerase I (New England Biolabs). After cleavage at the *Bam*HI restriction site, the single-stranded probe (502 bases) was isolated on low melting point agarose and hybridized (5×10^5 cpm) to 10 μg of poly(A)-rich RNA or of yeast tRNA as a control. The samples were heated at 65 °C for 10 min. Hybridizations were carried out for 12 h at 30 °C in a solution containing 80% formamide, 0.4 M NaCl, 1 mM EDTA, and 40 mM Pipes (pH 6.4). The samples (20 μL) were then diluted to 0.3 μL with ice-cold buffer containing 0.28 M NaCl, 50 mM sodium acetate (pH 4.5), and 4.5 mM $ZnSO_4$, after which 600 units of S1 nuclease (Boehringer) were added. The tubes were incubated for 60 min at 30 °C. S1 nuclease-resistant products were analyzed by electrophoresis on denaturing gels (7 M urea and 6% polyacrylamide) followed by autoradiography.

Primer Extension Analysis. This was carried out as described (Sambrook *et al.*, 1989) with slight modifications. Primers were end-labeled with [γ-^{32}P]ATP in the presence of T4 polynucleotide kinase and hybridized (10^6 cpm) with 25 μg of total RNA or 5 μg of poly(A)-rich RNA in a solution containing 0.5 M Tris-HCl (pH 8.3), 80 mM $MgCl_2$, 0.3 M KCl, and 0.1 M dithiothreitol. The mixture (30 μL) was denatured at 80 °C for 5 min and incubated for 3 h at 50 °C. To initiate primer extension, 0.5 mM of dNTP and 500 units of ribonuclease H-free SuperScript reverse transcriptase (Gibco-BRL) were added and the incubation (60 μL) was transferred to 45 °C for 1 h. After extraction in phenol/chloroform, the products were analyzed by electrophoresis on 6% polyacrylamide–7 M urea gels followed by autoradiography.

Transfections and Determination of Reporter Gene Activity. To make construct *a*, the 1.9-kb *Sal*I–*Hin*dIII genomic fragment was subcloned upstream of the firefly luciferase coding sequence in the polylinker of the vector pXp1. This promoterless vector has been designed for the detection of eucaryotic promoters. It contains, upstream of the cloning sites, SV40 sequences that inhibit the generation of transcripts from cryptic initiation sites in the plasmid (Nordeen, 1988). The other constructs (*b*–*l*) were prepared by 5′ or 3′ deletions of construct *a* using the restriction enzymes mentioned in the text. All subcloning and transformation techniques were performed as described (Sambrook *et al.*, 1989) and the plasmids were purified using a kit from Qiagen. Transfection of the TM4 cells was performed by the method of Wigler *et al.* (1978). TM4 cells (500 000/6-cm dish) were incubated for 16 h in culture medium with a coprecipitate of calcium phosphate–DNA (10 μg of test plasmid plus 2 μg of pRSVβgal or pSV_2cat as an internal control to normalize for transfection efficiency). The cells were washed with phosphate-buffered saline (PBS) and incubated with fresh culture medium followed by addition of glycerol to a final concentration of 15%. After 2 min, the cells were washed and reincubated in culture medium for 24 h. To quantify the activity of reporter genes, each dish was washed with ice-cold PBS and the cells were collected in 200 μL of the cell culture lysis reagent supplied by Promega. After centrifugation, 50 μL of each cell extract was mixed with 100 μL of the luciferase assay reagent supplied by Promega and placed in a Lumac luminometer biocounter M2000. The light produced was measured (Lumac light units) within 10 s of mixing the extract and assay reagent. β-Galactosidase activity was determined on 50 μL of cell extract according to Sambrook *et al.* (1989). Chloramphenicol acetyltransferase was assayed with the Boehringer CAT-ELISA kit and a microplate reader 3550-UV (Bio-Rad).

Identification of the 5′ Region of mRNAs. For mRNAs R1, R2, and R3, the Clontech 5′-AmpliFinder Race method was used as described in the text, according to the manufacturer's instructions. Total RNA (5 μg) was incubated with the Chik-3 primer at 65 °C for 5 min. After 20 μL of the reverse transcription master mix was added the solution was incubated at 52 °C for 30 min. The RNA template was then hydrolyzed at 65 °C for 30 min after 2 μL of 6 N NaOH was added. The incubation was neutralized by adding the same amount of 6 N acetic acid. Excess primer was removed after binding the cDNAs to a glass matrix support (Geno-Bind particles) as described by the manufacturer. The cDNAs were concentrated by ethanol precipitation. The

FIGURE 1: Nucleotide sequence of the 5′ region of the rat PFK-2/FBPase-2 gene B. Overlapping regions of the *Sal*I−*Hin*dIII fragment of the λ16 genomic clone were sequenced on both strands. The exons (1a′, 1b , 1c, and 1d) are indicated by capital letters and the regions that are common to exons 1b and 1c and to exons 1c and 1d are in bold-type. Restriction sites are indicated. The transcription initiation sites determined experimentally (one for exon 1b, two for exon 1c, and four for exon 1d) are indicated by an asterisk. An initiator sequence (Inr) is underlined twice. Sp1 consensus sequences (Faisst and Meyer, 1992) (one mismatch tolerated outside the GGC core) are underlined. The dots show the 5′ and 3′ ends of the sequence corresponding to rat heart cDNA RH1-9. The top of the figure shows the 5′ end (exon 1a) of a cDNA corresponding to mRNA R1, which is transcribed from a gene region located upstream from the *Sal*I−*Hin*dIII fragment.

single-stranded 5′-AmpliFinder anchor was ligated to the 3′-end of the cDNAs in the presence of 10 units of T4 RNA ligase in the single-stranded ligation buffer at room temperature for 18−20 h. The cDNAs were used as templates for PCR amplification between the AmpliFinder anchor primer and the Chik-4 primer. The incubation (1 μL) was diluted to 50 μL with DNase-free water after addition of 5 μL of concentrated (10×) PCR reaction buffer, 4 μL of 10 mM dNTP, 0.5 μL of AmpliTaq DNA polymerase (5 units/μL) supplied with the GeneAmp RNA PCR kit (Perkin Elmer Cetus). The samples were heated at 82 °C for 1 min prior to addition of the primers. The PCR parameters (35 cycles with a final extension time of 7 min) were denaturation at 94 °C for 45 s, annealing at 60 °C for 45 s, and extension at 72 °C for 2 min. The PCR products were separated on a 2% agarose gel, blotted on Hybond N^+ membranes (Amersham), and hybridized with the Chik-7 probe. The PCR products were subcloned in pBluescript and sequenced.

For mRNA R4, total RNA (10 μg) was incubated for 1 h at 30 °C in 50 μL of a buffer containing 100 mM sodium acetate (pH 5.0), 5 mM $MgSO_4$, and 3 units of RQ1 RNase-free DNase (Promega) to remove traces of cellular DNA. Following phenol/chloroform extraction, the RNA was precipitated with 2-propanol. Reverse transcription with primer AB1222 and PCR reactions with the Chik-9 and AB1222 primers were performed as described in the GeneAmp RNA PCR kit. To do this, 35 cycles were performed at 94 °C (60 s), 50 °C (1 min), and 72 °C (2 min). The PCR products were separated on a 2% agarose gel, blotted on Hybond membranes, and hybridized with the 644-bp *Ban*I−*Ava*I fragment of cDNA RH1-9 labeled in the presence of [α-^{32}P]dCTP using the Megaprime DNA labeling system of Amersham. The band detected by autoradiography was cut out, isolated with a Qiaex Gel Extraction kit (Qiagen), subcloned into the pGEM-T vector system I (Promega), and sequenced.

Detection of the mRNAs by RT-PCR. Total RNA (10 μg) was prepared as described for identification of the 5′ region of mRNA R4. Reverse transcription with random hexamers and PCR reactions between the Chik-4 primer and one of the TD-R1, TD-R2, TD-R3, and TD-R4 primers were performed with the GeneAmp RNA PCR kit as recommended by the manufacturer. Equal amounts of the PCR products were electrophoresed on a 2% agarose gel, blotted on Hybond membranes, and hybridized with the γ-^{32}P end-labeled Chik-7 probe.

RESULTS AND DISCUSSION

Cloning and Sequencing of the 5′ End of Gene B. The 5′ region of gene B is contained in a 16-kb λ clone called λ16 (Darville *et al.*, 1991). To characterize the promoter(s) of gene B, the 5′ end of λ16, namely, a 1.9-kb *Sal*I−*Hin*dIII fragment, was subcloned in the pBluescript phagemid and sequenced (Figure 1). Beforehand, we verified that this fragment included the 5′ end of rat heart cDNA RH1-9 (see introduction) by Southern blotting with a probe (AB247) complementary to nucleotides 6−25 of this cDNA. As indicated in Figure 1, the first 85 bp of cDNA RH1-9 were indeed found between the *Afl*II and *Ava*II sites of the gene fragment and thus defined an exonic sequence called here 1c. As shown below, this sequence corresponds to the first exon of a mRNA that we call R3. Exon 2, which contains the translation initiation site, is located downstream from the *Sal*I−*Hin*dIII fragment shown in Figure 1 (Darville *et al.*, 1991).

Identification of the Promoter of mRNA R3. We could not exclude that cDNA RH1-9 was incomplete at the 5′ end. We therefore sought to determine the transcription initiation site of mRNA R3, namely, the 5′ end of exon 1c, by S1 nuclease mapping. To do so, a 5′ labeled primer (AB743) complementary to nucleotides 29−49 of cDNA RH1-9 (coordinates 1132−1152 in Figure 1) was hybridized to the single-stranded *Sal*I−*Hin*dIII genomic fragment. After reverse transcription and restriction with *Bam*HI, a 502 base long single-stranded cDNA probe was obtained. This probe was hybridized to poly(A)-rich RNA purified from rat heart, brain, or lung, and incubated with S1 nuclease to digest the single-stranded material. While the free probe was correctly digested (Figure 2A, lane 3), a fragment of about the size of the probe was protected against this digestion by the three RNA preparations but not by yeast tRNA. The signal was more intense with lung RNA (Figure 2A, lane 2) than with brain RNA (lane 1) and was much fainter with heart RNA (not shown), indicating that this transcript is not restricted to heart. This suggested that the gene B transcript that had yielded cDNA RH1-9 actually extends up to the vicinity of the *Bam*HI site.

To determine more precisely the 5′ end of this transcript, we performed a primer extension experiment on total RNA from rat heart and other tissues. The 5′ end of the 24-mer antisense primer Chik-2 was chosen to hybridize 121 bp downstream from the *Bam*HI site (coordinates 748−771 in Figure 1). A 114-bp fragment was generated with RNA from heart (Figure 2B, lane 1). This confirmed the S1 nuclease mapping data and assigned a length of 531 bp to the first exon (exon 1c) of the mRNA (R3) that yielded cDNA RH1-9. This exon begins with the first adenine downstream from the *Bam*HI site (Figure 1). A fainter primer-extended band (86 bases) was seen, which corresponds to a transcript initiated at a C located in an octamer identical (Figure 1) to the initiator (Inr) sequence of the terminal deoxynucleotidyltransferase gene (Smale and Baltimore, 1989). The mRNA containing exon 1c was detected by primer extension not only in heart, brain, and lung, as expected, but also in the other tissues tested, namely, spleen and testis (Figure 2B), confirming that mRNA R3 is not tissue-specific.

Functional Study of the Promoter Region. If transcription of mRNA R3 actually starts in the *Sal*I−*Hin*dIII fragment at the sites determined above, this fragment should drive transcription of a promoterless reporter gene in transfected cells. To test this prediction, the *Sal*I−*Hin*dIII gene fragment and deletants thereof were cloned upstream of the luciferase cDNA and these constructions were transiently transfected in the mouse Sertoli cell line TM4. The latter was chosen because, of all the tissues that scored positive for mRNA R3, testis was the one from which a readily transfectable cell line was available. Moreover, we found that not only purified rat Sertoli cells but also TM4 cells contain mRNA R3 (Figure 2B). Construct *a* (Figure 3) contains the R3 initiation site, a sequence of 657 bp upstream from this site, exon 1c (531 bp), and 732 bp of the first intron. This construct behaved as a strong promoter, since its transfection yielded up to 14 000 luciferase units as compared to no more than 10 units for the insertless plasmid or for the construct containing the insert in the opposite orientation. As shown

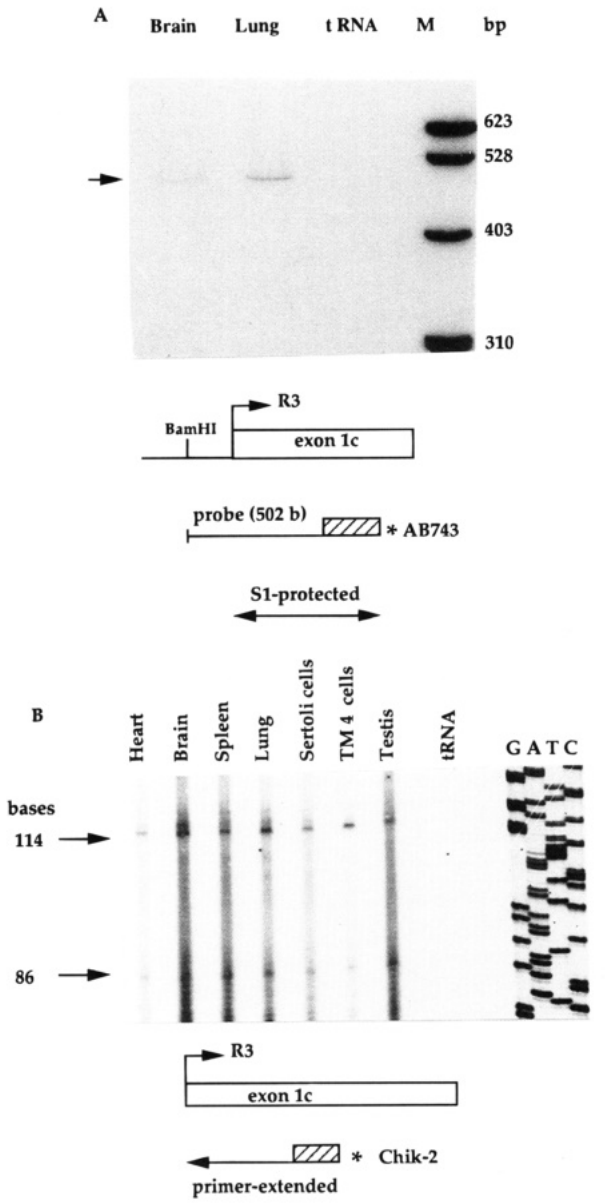

FIGURE 2: Determination of the 5′ end of mRNA R3 by S1 nuclease mapping (A) and by primer extension (B). (A) Poly(A)-rich RNA (10 μg) purified from rat brain or lung, or yeast tRNA (10 μg) as a negative control, was hybridized with the 502-base single-stranded probe labeled by the AB743 primer (coordinates 1132−1152 in Figure 1) and was digested with S1 nuclease. The figure shows an autoradiogram of the products separated by gel electrophoresis. M, *Hpa*III−digested pBR322 DNA. The arrow points to the protected fragment. (B) Total RNA (25 μg) extracted from the rat tissues indicated or from the mouse Sertoli cell line TM4, or yeast tRNA (25 μg) as a negative control, was hybridized with the labeled antisense Chik-2 primer (coordinates 748−771 in Figure 1) and reverse-transcribed. The figure shows an autoradiogram of the products of the reaction and of the sequencing of the *Sal*I−*Hin*dIII fragment (see Figure 1) with the Chik-2 primer. The elongation products are indicated by an arrow.

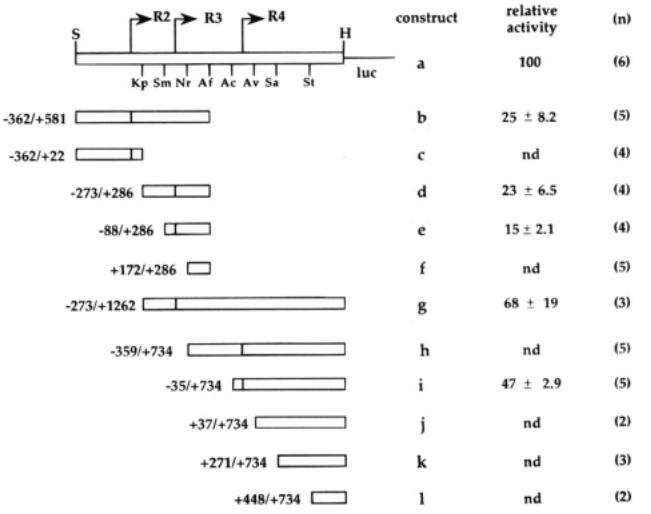

FIGURE 3: Promoter activity in the 5′ end of PFK-2/FBPase-2 gene B. The *Sal*I−*Hin*dIII genomic fragment and deletants thereof inserted in a promoterless luciferase cDNA reporter plasmid were transiently transfected in TM4 cells, together with a pRSVβgal or pSV2CAT control plasmid to correct for transfection efficiency. The coordinates of the constructs relate to the putative transcription initiation sites indicated by the arrows as follows: R2 (asterisk in exon 1b on Figure 1) for constructs *b* and *c*, R3 (most 5′ asterisk in exon 1c) for constructs *d*−*g*, R4 (most 3′ asterisk in exon 1d) for constructs *h*−*l*. The relative activities (mean ± SEM) are expressed as the ratio of luciferase over β-galactosidase or chloramphenicol acetyltransferase activities normalized for construct *a*. Absolute luciferase activities were 3000−14 000 Lumac units, depending on the experiment (*n* refers to the number of independent experiments). nd, relative activity not statistically above background. S, *Sal*I; H, *Hin*dIII; Kp, *Kpn*I; Sm, *Sma*I; Nr, *Nru*I; Af, *Afl*II; Ac, *Acc*I; Av, *Ava*II; Sa, *Sac*I; St, *Sty*I.

in Figure 3, a 5′ deletion of 400 bp (construct *g*) decreased promoter activity by only 30%. However, a 5′ deletion that eliminated the R3 initiation site (construct *h*) also eliminated promoter activity. To study the promoter region proper, the intron and part of exon 1c were removed by 3′ deletion. This (construct *b*) decreased transcriptional activity 4-fold. No further decrease was observed when 388 bp was deleted from the 5′ end (construct *d*). When only 88 bp was left upstream from the R3 initiation site (construct *e*), there was a further drop in activity, and all activity was lost with loss of the R3 site (construct *f*). These results indicate that a functional transcription initiation site is present between the *Sma*I and *Nru*I sites and confirm the S1 nuclease mapping and primer extension data.

The transfection data reported above showed that constructs *b* and *d* are 3−4-fold less active than constructs *a* and *g*. This suggested the presence of another promoter downstream of the *Afl*II site (Figure 3). Consistent with this hypothesis, we found that construct *i*, purported to correspond to an intronic sequence, drove transcription of the luciferase reporter gene twice as efficiently as constructs *b* and *d* (Figure 3). Further 5′ deletions (constructs *j*, *k*, and *l*) eliminated this activity. We concluded that gene B contains a second promoter, which is present between the *Afl*II and *Ava*II sites. In this case, additional mRNA(s) originating from this promoter should be transcribed from gene B. This was compatible with our finding (Vidal *et al.*, 1993) that the bovine PFK-2/ FBPase-2 gene B gives rise to several mRNAs, one of which differs from the others by the 5′ end.

Identification of mRNAs R1 and R2. We first attempted to clone fully coding rat mRNAs transcribed from gene B by screening with probe RH1-9 cDNA libraries (Clontech) from heart, brain, and testis. This approach failed, presumably because gene B transcripts are very rare. We therefore searched for the 5′ region of such mRNAs with the more direct technique called 5′ AmpliFinder Race (see Materials and Methods). We chose an antisense oligonucleotide (Chik-3) specific for a sequence (nucleotides 180−199 of cDNA RH1-9) located in exon 2, which contains the translation initiation site of the H isozyme. This primer was used to

prime cDNA synthesis by reverse transcription of total RNA from heart and fetal liver, a tissue reported to contain the H isozyme (Martin-Sanz *et al.*, 1992). The AmpliFinder anchor was ligated to the 3′ end of the cDNAs and the latter were amplified by PCR between a primer complementary to that anchor and a primer (Chik-4) complementary to exon 2 (nucleotides 154−173 of cDNA RH1-9). The PCR products amplified from heart and fetal liver yielded several bands by Southern blotting with a probe (Chik-7) located in exon 2 upstream (nucleotides 125−145 of cDNA RH1-9) from the Chik-4 primer (not shown). The PCR products were cloned in pBluescript and sequenced. Three distinct cDNAs were identified. One contained exon 1c and the expected part of exon 2. It therefore corresponds to the mRNA R3 discussed above. The two other cDNAs also originated from gene B but contained, upstream from exon 2, a sequence that differs from exon 1c.

The first (253 bp) of these 5′ cDNA fragments corresponds to a mRNA that we call R1. Upstream from exon 2, it contains 165 bp that belong to two novel exons called here 1a and 1a′. Exon 1a′ (100 bp) starts in the gene 18 bp downstream from the *Sal*I site (Figure 1). Exon 1a (65 bp, see Figure 1) is not found in the *Sal*I−*Hin*dIII genomic fragment. It therefore lies upstream from the 5′ end of the λ16 clone, in a region of the gene that has not been cloned. The splice site at the 3′ intron−exon 1a′ junction does not fit the classical consensus, as the G usually present at the 3′ end of introns is replaced by a C (Figure 1). The sequence of the 1a−1a′ exon junction is correct since Wanatabe *et al.* (1994) have recently cloned from rat brain a fully coding PFK-2/FBPase-2 cDNA whose sequence fits with that of gene B and is identical to exons 1a and 1a′ upstream from exon 2.

The other 5′ cDNA fragment (654 bp) corresponds to a mRNA that we call R2. Upstream from exon 2 one finds in this cDNA 566 bp that belong to a novel exon called 1b. The 5′ end of this cDNA lies 199 bp upstream from the R3 transcription initiation site (Figure 1). In the R2 transcript, the splice site at the 5′ end of the intron that separates exon 1b from exon 2 lies within exon 1c. Thus, 367 bp of exons 1b and 1c are overlapping (Figure 1). The actual length of exon 1b was determined by primer extension of poly(A)-rich RNA purified from rat heart, testis, brain, and fetal liver. The primer, SI−R2, was complementary to nucleotides 29−52 of the cDNA corresponding to mRNA R2 (coordinates 487−510 in Figure 1). Instead of the expected band of 52 bases, a major band of 149 bases was found with the RNA from the four tissues examined (Figure 4). The mRNA R2 therefore starts at least 97 bp upstream from the 5′ end of the cDNA. The missing portion of mRNA R2 most likely corresponds to the genomic sequence that prolongs exon 1b in the *Sal*I−*Hin*dIII fragment. Indeed, there is no splice site consensus in this sequence. Thus, the 5′ end of exon 1b would lie 26 bp upstream from the *Kpn*I site (Figure 1), and the genomic sequence upstream from exon 1b is the promoter of mRNA R2.

We therefore determined whether the sequence upstream from exon 1b can drive transcription of the luciferase reporter gene in transfected cells. This was not the case (construct *c* in Figure 3) under conditions where the cells were responsive to other constructs. The lack of stimulatory or inhibitory transcriptional activity of construct *c* was confirmed by the fact that it had no effect when linked to

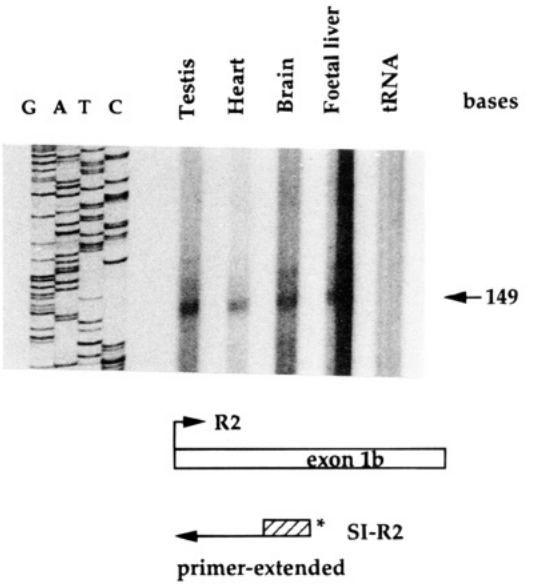

FIGURE 4: Determination of the 5′ end of mRNA R2 by primer extension. Poly(A)-rich RNA (5 μg) purified from the rat tissues indicated, or yeast tRNA (5 μg) as a negative control, was hybridized with the labeled SI−R2 antisense primer (coordinates 487−510 in Figure 1) and reverse-transcribed as indicated at the bottom. The figure shows an autoradiogram of the products of the reaction and of the sequencing of the genomic fragment with the same primer. The arrow points to the elongation product.

construct *d* to yield construct *b* (Figure 3). Therefore, one possibility is that the promoter of the R2 transcript lies in construct *c* but is not active in our assay system. The other possibility is that this promoter is located further upstream, beyond an additional 5′ exon (or exons) that belongs to mRNA R2. Indeed, the genomic sequence (Py)11 AC (Figure 1) that flanks the 5′ end of exon 1b is compatible with a splice site. Although it ends with AC rather than AG, it would be similar to the splice site at the 3′ intron−exon 1a′ junction. As a consequence, exon 1b could be spliced with exon 1a or 1a′, in which case transcription of mRNA R2 could be driven by the same promoter as mRNA R1. This would require that the primer extension products seen in Figure 4 correspond to transcripts undergoing splicing and on which the reverse transcriptase stops at the spliceosome, the extension products of the mature mRNA 2 being too long to be detected under our electrophoretic conditions. Thus, these results confirmed the 5′ structure of mRNA R3 and showed the existence of two additional mRNAs, R1 and R2. They also provided evidence for the presence in gene B of at least two distinct promoters, one for mRNA R1 and one for mRNA R3.

Identification of mRNA R4. At that point of our investigations, we had found that the rat gene B produces mRNAs R1, R2, and R3. However, none of these mRNAs contained the exon we had postulated around the *Ava*II site on the basis of transfection data (Figure 3). Perhaps the mRNA, called here R4, containing this putative exon was absent from the fetal liver RNA we had used to identify mRNAs R1, R2, and R3, or it had escaped detection by the AmpliFinder Race technique. We therefore searched for mRNA R4 by reverse transcription and PCR of total RNA extracted from rat testis, a tissue which is relatively rich in B-gene transcripts. Gene B-specific cDNAs were prepared by reverse transcription with an antisense primer (AB1222) complementary to nucleotides 747−770 of cDNA RH1-9, i.e., at the beginning

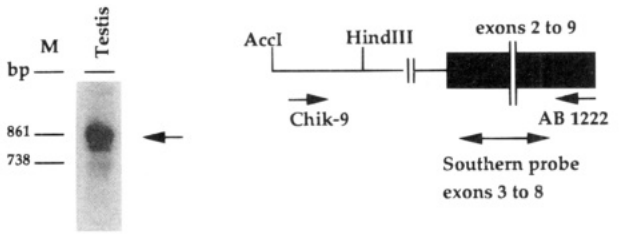

FIGURE 5: Identification of mRNA R4 by RT-PCR and Southern blotting. Autoradiogram of a Southern blot of the cDNA amplified by RT-PCR of total RNA (30 μg) from rat testis using the strategy depicted at the right. See text for details.

of exon 9. To avoid potential problems with ligation or tailing reactions at the 5′ end prior to PCR, we used a sense primer Chik-9 located at the beginning of the putative exon specific for mRNA R4, namely, 27 bp downstream from the *Ava*II site (coordinates 1253−1272 in Figure 1). The cDNAs were amplified between Chik-9 and AB1222 and the specific PCR products were identified by Southern blotting with a 664-bp probe that corresponds to exons 3−8 of cDNA RH1-9. This yielded a strong signal with testis cDNA (Figure 5). No signal was seen when the same experiment was conducted with RNA prepared from rat brain or from TM4 cells.

The positive material from testis was extracted from the gel, cloned, and sequenced. A 767-bp fragment was obtained whose sequence showed that it corresponds to mRNA R4. Starting from the 3′ end, one finds exons 9−2. Upstream from exon 2 the sequence, called here exon 1d, corresponds to none of the exons of mRNA R1, R2, or R3. In the genomic sequence (Figure 1), exon 1d extends from the Chik-9 primer to a classical splice site (for exon 2) located 82 bp downstream. Because the 5′ end of the cDNA corresponding to mRNA R4 was imposed by the Chik-9 primer, we determined the actual 5′ end of mRNA R4 by extension with a labeled antisense primer, SI−R4, located just downstream from the Chik-9 primer (coordinates 1279−1300 in Figure 1). Primer extension of total RNA (30 μg) showed three elongation products (129, 141, and 144 bases) with testis and one product (115 bases) with heart. There was no signal with brain, consistent with our failure to detect the corresponding cDNA in this tissue. This suggests that four transcription initiation sites (Figure 1) for mRNA R4 occur within 50 bp of the *Acc*I−*Ava*II fragment, in keeping with our transfection data. Thus, 30 bp of exon 1d and 1*c* are overlapping. It is noteworthy that construct *h* was devoid of promoter activity in our assay system, despite the fact it contains the promoter for mRNA R4 (Figure 3). A comparison with construct *i* suggests the presence in the *Nru*I−*Acc*I fragment of sequence(s) that inhibit the R4 promoter. This inhibition might result from binding of a factor, called GCF, known to repress transcription of several genes. Indeed, the *Nru*I−*Acc*I fragment contains five consensus-binding sites ($NN^G/_CCG^G/_C^G/_C^G/_CCN$) (Kageyama and Pastan, 1989; Philippe *et al.*, 1994) for this factor.

Functional Organization of Promoter Region of Gene B and Structure of the 5′ end of mRNAs R1−R4. A comparison of the sequence of the cDNA clones described above with that of the gene, as well as the primer extension, S1 nuclease mapping, and transfection data, lead us to propose for the 5′ end of gene B the structure shown in Figure 6. The gene gives rise to at least four mRNAs that differ upstream from a common exon 2 which contains the translation initiation

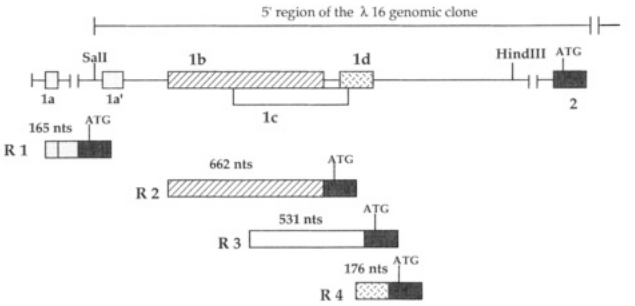

FIGURE 6: Functional organization of the promoter region of gene B and structure of the 5′ end of mRNAs R1, R2, R3, and R4. At the top is shown the part of rat gene B studied here. All the exons shown (boxes) are noncoding except exon 2, which contains the translation initiation codon at position 21 in this exon and is common to the four mRNAs. Exon 1a is located in the gene at an unknown distance upstream from the λ16 clone. The size in nucleotides (nts) refers to the length of the mRNA upstream from exon 2.

site. The mRNAs R1, R3, and R4 each originate from a distinct promoter and therefore from a different primary transcript.

The sequence of the R1 promoter is unknown since it lies upstream from the cloned region of gene B. It drives transcription of a mRNA (R1) that corresponds to the cDNA cloned by Wanatabe *et al.* (1994) from rat brain while this work was in progress. This cDNA, which has a 2-kb trailer sequence, extends to the poly(A) tail and includes all the exons from 1a to 16 except exon 15. Therefore, mRNA R1 corresponds to the exon 15-less mRNA (called B3) that we cloned earlier from bovine heart (Vidal *et al.*, 1993). This mRNA is interesting because it codes for an isozyme that lacks several sites of phosphorylation by protein kinases (Rider *et al.*, 1992). As the cDNA (3.4 kb) of mRNA R1 gave by Northern blotting of brain RNA a single 7.4-kb signal (Wanatabe *et al.*, 1994), this mRNA must contain an additional sequence of 4 kb upstream from exon 1a. Alternatively, this signal could be due to an immature transcript, which seems unlikely because it was the only band seen by Northern blotting with RNA from brain, heart, testis, liver, skeletal muscle (Wanatabe *et al.*, 1994), and lung (M. I. Darville and G. G. Rousseau, unpublished data).

The R3 and R4 promoters contain exonic sequences. This peculiar organization occurs in the c-*myc* gene, which contains at least two promoters, one being part of the exon transcribed from the most 5′ promoter (Battey *et al.*, 1983). A similar phenomenon has been described in the human prolactin gene (DiMattia *et al.*, 1990). This gene gives rise, from two promoters, to three mRNAs that differ by the 5′ end. The most 3′ promoter of the prolactin gene coincides with a region that corresponds to two overlapping exons of different lengths. The R3 promoter contains no obvious TATA box but an Inr sequence (Figure 1). Inr sequences are reported to play an important role in transcription initiation from TATA-less promoters (Weiss and Reinberg, 1992). The R4 promoter region is GC-rich and contains no TATA box but several consensus sequences for binding the Sp1 transcription factor (Figure 1). The presence of such binding sequences, which can stimulate transcription from TATA-less promoters, often correlates with multiple transcription initiation sites, as seen here (Blake *et al.*, 1990). They may occur downstream of these sites, as was the case in exon 1d.

As to mRNA R2, its promoter could lie immediately upstream from exon 1b. The lack of transcriptional activity of this region in our transfection assays could reflect the inadequacy of these assays for detecting the R2 promoter or the presence in the fragment tested of a sequence that inhibits specifically the R2, but not the R3, promoter. Another possibility is that mRNA R2 originates from the same primary transcript as mRNA R1, exon 1b being spliced with exons 1a or 1a′. This would involve a nonclassical splice site, as also postulated here for mRNA R1 (see above) and described in other genes (Shapiro and Senapathy, 1987; Citri *et al.*, 1987; Kozasa *et al.*, 1988; Kiss *et al.*, 1989; Larson *et al.*, 1990; Quan and Forte, 1990).

Tissue Distribution of mRNAs from Gene B. The tissue distribution of mRNAs R1−R4 was evaluated by reverse transcription−PCR. Total RNA was extracted from rat heart, lung, brain, spleen, thymus, testis, small intestine, skeletal muscle, kidney, liver, and fetal liver and from mouse TM4 Sertoli cells. This RNA was reverse-transcribed by random priming to avoid stops at secondary structures. The resulting cDNAs were amplified by PCR between a sense primer specific for exon 1a′ (TD-R1), 1b (TD-R2), 1c (TD-R3), or 1d (TD-R4) and an antisense primer (Chik-4, already described) specific for exon 2, which is common to the four mRNAs. Southern blotting of the PCR products with an internal probe (Chik-7, already described) located in exon 2 showed for each mRNA a single band of the expected size. mRNAs R1 and R3 were present in heart, lung, brain, spleen, thymus, testis, small intestine, skeletal muscle, kidney, and adult and fetal liver. mRNA R2 was present in all these tissues except thymus, muscle, and adult liver. mRNA R4 was present in all these tissues except brain, spleen, and fetal liver.

These data show that none of the gene B transcripts is strictly tissue-specific. However, these mRNAs might not be translated in all the tissues where they are present. Ruan *et al.* (1994) have shown that translation of the mRNA coding for *S*-adenosylmethionine decarboxylase is inhibited in T cells by the translation of a minicistron that codes for a hexapeptide starting 14 bp downstream from the cap site. An inhibition of translation by such minicistrons has been described for the genes coding for fibroblast growth factor 5 (Bates *et al.*, 1991), for proenkephalin (Rao and Howells, 1993), and for the β2-adrenergic receptor (Parola and Kobilka, 1994). A similar posttranscriptional regulation could apply to mRNAs R2 and R3. Indeed, the 5′ noncoding region of mRNA R2 contains five minicistrons encoding putative peptides of 9, 15, 23, 24, and 45 amino acids and that of mRNA R3 contains two minicistrons encoding putative peptides of 9 and 45 amino acids. Interestingly, genes that contain, like gene B, initiation codons within their 5′ transcript leader (less than 10% of the vertebrate genes) show a strong bias toward protooncogenes (Kozak, 1987). This feature is also found in the M-type mRNA coded by PFK-2/FBPase-2 gene A (Crepin *et al.*, 1992).

In conclusion, we have characterized here the structural and functional organization of the 5′ region of rat gene B. We have shown that it contains four more exons than the 16 previously described and we have defined four mRNAs that differ by their 5′ end. Our data also point to the existence in gene B of at least three promoters. This region of gene B exhibits several unusual features such as exon overlap, intronic promoters, and abnormal splicing consensus. Since fructose 2,6-bisphosphate is a ubiquitous molecule, it was somewhat surprising to find that the bifunctional enzyme that catalyzes its synthesis and degradation, PFK-2/FBPase-2, is coded in the rat by at least four genes, two of which (genes A and B) have been cloned (Rousseau and Hue, 1993). Moreover, gene A gives rise to at least three mRNAs through use of distinct promoters. We have now found that gene B also contains several promoters that give rise to different mRNAs. An additional level of complexity stems from the fact that, in bovine heart at least, distinct mRNAs originating from gene B are also produced by differential splicing (Vidal *et al.*, 1993). The relationships between these numerous mRNAs and the PFK-2/FBPase-2 isozymes described in different tissues remain to be clarified. Still, this emphasizes the importance of fructose 2,6-bisphosphate for the fine spatial and temporal control of carbon flux through the glycolytic pathway.

ACKNOWLEDGMENT

We thank L. Hue for continued interest, the members of the HORM unit for advice, and T. Lambert for secretarial assistance. The TM4 cells and rat Sertoli cell RNA were generously provided by G. Verhoeven.

REFERENCES

Auffray, C., & Rougeon, F. (1980) *Eur. J. Biochem. 107*, 303−314.

Ausubel, F. M., Brent, R., Kingston, R. E., Moore, D. D., Seideman, J. G., Smith, J. A., & Struhl, K. (1992) *Short Protocols in Molecular Biology*, 2nd ed., John Wiley & Sons, New York.

Aviv, H., & Leder, P. (1972) *Proc. Natl. Acad. Sci. U.S.A. 69*, 1408−1412.

Bates, B., Hardin, J., Zhan, X., Drickamer, K., & Goldfarb, M. (1991) *Mol. Cell. Biol. 11*, 1840−1845.

Battey, J., Moulding, C., Taub, R., Murphy, W., Stewart, T., Potter, H., Lenoir, G., & Leder, P. (1983) *Cell 34*, 779−787.

Blake, M. C., Jambou, R. C., Swick, A. G., Kahn, J. W., & Clifford Azikhan, J. (1990) *Mol. Cell. Biol. 10*, 6632−6641.

Chirgwin, J. M., Przybyla, A. E., MacDonald, R. J., & Rutter, W. J. (1979) *Biochemistry 18*, 5294−5299.

Citri, Y., Colot, H. V., Jacquier, A. C., Yu, Q., Hall, J. C., Baltimore, D., & Rosbach, M. (1987) *Nature 326*, 42−47.

Crepin, K. M., De Cloedt, M., Vertommen, D., Foret, D., Michel, A., Rider, M. H., Rousseau, G. G., & Hue, L. (1992) *J. Biol. Chem. 267*, 21698−21704.

Darville, M. I., Crepin, K. M., Hue, L., & Rousseau, G. G. (1989) *Proc. Natl Acad. Sci. U.S.A. 86*, 6543−6547.

Darville, M. I., Chikri, M., Lebeau, E., Hue, L., & Rousseau, G. G. (1991) *FEBS Lett. 288*, 91−94.

DiMattia, G. E., Gellersen, B., Duckworth, M. L., & Friesen, H. G. (1990) *J. Biol. Chem. 265*, 16412−16421.

Dupriez, V. J., Darville, M. I., Antoine, I. V., Gégonne, A., Ghysdael, J., & Rousseau, G. G. (1993) *Proc. Natl. Acad. Sci. USA 90*, 8224−8228.

Faisst, S., & Meyer, S. (1992) *Nucleic Acids Res. 20*, 3−26.

Kageyama, R., & Pastan, I. (1989) *Cell 59*, 815−825.

Kiss, I., Deak, F., Holloway, R. G., Jr., Delius, H., Mebust, K. A., Frimberger, E., Argraves, W. S., Tsonis, P. A., Winterbottom, N., & Goetinck, P. F. (1989) *J. Biol. Chem. 264*, 8126−8134.

Kitamura, K., Kangawa, K., Matsuo, H., & Uyeda, K. (1988) *J. Biol. Chem. 263*, 16796−16801.

Kozak, M. (1987) *Nucleic Acids Res. 15*, 8125−8148.

Kozasa, T., Itoh, H., Tsukamoto, T., & Kaziro, Y. (1988) *Proc. Natl. Acad. Sci. U.S.A. 85*, 2081−2085.

Larson, R. G., Henning, D., Haidar, M. A., Jhiang, S., Lin, W. L., Zhang, W. W., & Bush, H. (1990) *Cancer Commun. 2*, 63−71.

Martin-Sanz, P., Cascales, M., & Bosca, L. (1992) *Biochem. J. 281*, 457−463.

Nordeen, S. K. (1988) *BioTechniques 6*, 454−457.

Parola, A. L., & Kobilka, B. K. (1994) *J. Biol. Chem. 269*, 4497−4505.
Philippe, M., Larondelle, Y., Lemaigre, F., Mariamé, B., Delhez, H., Mason, P., Luzatto, L., & Rousseau, G. G. (1994) *Eur. J. Biochem. 226*, 377−384.
Quan, F., & Forte, M. A. (1990) *Mol. Cell. Biol. 10*, 910−917.
Rao, S. M., & Howells, R. D. (1993) *J. Biol. Chem. 268*, 22164−22169.
Rider, M. H., Vandamme, J., Lebeau, E., Vertommen, D., Vidal, H., Rousseau, G., Vandekerckhove, J., & Hue, L. (1992) *Biochem. J. 285*, 405−411.
Rousseau, G. G., & Hue, L. (1993) *Prog. Nucleic Acids Res. Mol. Biol. 45*, 99−127.
Ruan, H., Hill, J. R., Fatemie-Nainie, S., & Morris, D. R. (1994) *J. Biol. Chem. 269*, 17905−17910.
Sakata, J., & Uyeda, K. (1990) *Proc. Natl. Acad. Sci. U.S.A. 87*, 4951−4955.
Sambrook, J., Fritsch, E. F., & Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual*, 2nd ed., Vols. 1−3, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
Sanger, F., Nicklen, S., & Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. U.S.A. 74*, 5463−5467.
Shapiro, M. B., & Senapathy, P. (1987) *Nucleic Acids Res. 15*, 7155−7174.
Smale, S. T., & Baltimore, D. (1989) *Cell 57*, 103−113.
Vidal, H., Crepin, K. M., Rider, M. H., Hue, L., & Rousseau, G. G. (1993) *FEBS Lett. 330*, 329−333.
Watanabe, F., Sakai, A., Furuya, E., & Uyeda, K. (1994) *Biochem. Biophys. Res. Commun. 198*, 335−340.
Weiss, L., & Reinberg, D. (1992) *FASEB J. 6*, 3300−3309.
Wigler, M., Pollicer, A., Silverstein, S., & Axel, R. A. (1978) *Cell 14*, 725−731.

BI950444L