Biochem. J. (1990) 270, 261–263 (Printed in Great Britain)

Alternative splicing of glucokinase mRNA in rat liver

David J. HAYZER*‡ and Patrick B. IYNEDJIAN†§

*Department of Medical Biochemistry and †Division of Clinical Biochemistry, University of Geneva School of Medicine, Geneva, Switzerland

The sequences of two near full-length cDNAs encoding rat liver glucokinase are reported. One of the cDNAs is essentially identical to the cDNA cloned by Andreone, Printz, Pilkis, Magnuson & Granner. [(1989) J. Biol. Chem. **264**, 363–369]. The other cDNA contains a 151 bp insertion and a downstream 52 bp deletion. The inserted block of bases has been shown to originate from an optional cassette exon, termed 2A, between the previously described exons 1 and 2. The conceptual translation product from the variant mRNA is identical to the original glucokinase protein for the first 15 amino acids. Next there is a novel polypeptide sequence of 87 residues, comprising 50 residues encoded by the cassette exon and 37 residues specified by an altered reading frame in exon 2. Due to the 52 bp deletion, 17 amino acids of the reference sequence are then missing, after which the sequence reverts to the original. Northern blot analysis with oligonucleotide probes has shown that alternatively spliced mRNA represents about 5% of total glucokinase mRNA. Alternative splicing of glucokinase mRNA in liver may explain earlier findings of minor isoforms of hepatic glucokinase.

INTRODUCTION

The mammalian enzyme glucokinase belongs to the family of the hexokinases (ATP: D-hexose 6-phosphotransferases; EC 2.7.1.1). It differs from the three other characterized hexokinases of mammalian tissues by its comparatively low affinity for glucose, the lack of product inhibition by glucose 6-phosphate and a molecular mass of about 50 kDa, as compared with 100 kDa for the other hexokinases [1]. Glucokinase was discovered in rat liver, where it plays an important regulatory role in glucose metabolism [2]. An enzyme with closely related biochemical and immunological properties was subsequently identified in the islets of Langerhans of the pancreas [3,4]. Recently the primary structure of hepatic glucokinase has been determined from the nucleotide sequence of a cDNA clone [5]. In addition, the mRNAs encoding glucokinase in liver and islet tissues have been shown to be identical over the 3′ untranslated region and most of the protein coding region, but to differ with respect to the 5′ leader sequence and the first 45 coding nucleotides [6–8]. At the gene level, the sequence divergence has been traced to transcription from two distinct tissue-specific promoters, with alternate splicing of different 5′-end exons in liver and islet mRNAs [7]. In this report, we present evidence for alternative splicing of glucokinase mRNA within the liver itself. A rat liver cDNA, which differs from the published sequence by insertion and deletion of large blocks of bases, is described. The genetic origin of the novel insert has been determined in relation to previously known exons and introns of the glucokinase transcriptional unit.

MATERIALS AND METHODS

The glucokinase cDNA clone pUC-GK1 and the cDNA library from rat liver from which it was originally isolated have been described previously [9]. Several additional clones were isolated by rescreening the library with the cDNA insert of pUC-GK1 as hybridization probe. Library screening, plaque purification of positive signals, isolation of λ DNA and subcloning into plasmid or M13 vectors were performed by conventional techniques [10]. For sequencing, entire cDNAs were inserted as single fragments into M13mp19. Progressively deleted subclones with appropriate overlaps were generated by the method of Dale *et al.* [11]. A rat genomic DNA clone containing part of the glucokinase gene was isolated from a library in the λ EMBL3 vector (Clontech, Palo Alto, CA, U.S.A.) by screening with a cDNA probe. A 2.5 kbp *Hin*dIII–*Hin*dIII fragment from this clone was introduced in both orientations in M13mp19. Sequencing of the region of interest was performed for the upper strand with the universal M13 primer and for the lower strand with sequence-specific oligonucleotide primers.

RESULTS AND DISCUSSION

Isolation and sequencing of two hepatic glucokinase cDNAs

A glucokinase cDNA termed GK1, 1.8 kbp in length and lacking the 5′ region of the mRNA, was previously isolated by immunological screening of an expression cDNA library constructed in this laboratory from size-selected polysomal mRNA from the livers of carbohydrate-fed rats [9]. Additional clones were isolated from the same library by hybridization screening with the GK1 cDNA. Two of the new clones, which appeared to be essentially full-length copies of the 2.4 kb mRNA encoding liver glucokinase, were selected for further investigation. By sequencing, the cDNA clone termed GK2 was found to be 2325 bp in length (exclusive of 13 A residues at the 3′ end) and virtually identical to the cDNA independently cloned by Andreone *et al.* [5]. Minor differences, e.g. a single base replacement (not resulting in an amino acid change) in the coding region, are listed in the legend to Fig. 1. The second cDNA, termed GK3, was shown to be 2441 bp in length [not including the poly(A) tail] and to differ in two major respects from GK2 or the published sequence (Fig. 1). The first difference is the insertion at position 154 of a stretch of 151 nucleotides not found in the previous cDNAs. The second is a deletion, at position 416, of a block of 52 nucleotides present in GK2 as well as in the sequence of Andreone *et al.* [5]. Downstream of this block, the GK2 and GK3 sequences were completely identical.

‡ Present address: Department of Medicine, Division of Cardiology, Emory University School of Medicine, Atlanta, GA 30322, U.S.A.

§ To whom correspondence should be addressed, at: Division of Clinical Biochemistry, Centre Medical Universitaire, 1, rue Michel-Servet, 1211 Geneva 4, Switzerland.

The nucleotide sequence data reported will appear in the EMBL, GenBank and DDBJ Nucleotide Sequence Databases under the accession numbers X53588, X53589 and X53590.

Fig. 1. Nucleotide sequences of two rat liver glucokinase cDNAs and predicted amino acid sequences

Both strands of GK3 cDNA and the upper strand of GK2 cDNA were sequenced entirely as described in the Materials and methods section. Numbering is from the 5′-end of the GK3 cDNA. Dashes denote sequence identity between the cDNAs. Underlined blocks of bases are present in only one cDNA. Arrows indicate boundaries between exons as deduced from this work and from [11]. The double underline shows the sequence complementary to the synthetic oligonucleotide used as probe for exon 2A. Beyond position 480, the GK2 and GK3 cDNA sequences are completely identical and differ from the cDNA sequence published in [5] as follows: (1) at position 1466 (according to numbering of [5]), G instead of A, no amino acid change; (2) at position 1937 (untranslated), A instead of G; (3) at position 2344, insertion of an extra CTA triplet; (4) at position 2347, insertion of an extra C immediately upstream of the poly(A) tract. Position 1 in the Figure corresponds to position 9 in [5].

The translation product specified by the GK3 sequence, when compared with the original glucokinase amino acid sequence, would be identical for the first 15 residues. Next there is a novel polypeptide sequence of 87 residues, comprising 50 residues encoded by the insert and 37 residues encoded by nucleotides shared with the original sequence, but read in a different frame. The 52 bp deletion then results in the loss of 17 residues, after which the amino acid sequence reverts to the original as defined by Andreone *et al.* [5] (Fig. 1). The variant glucokinase protein would have a molecular mass of 54773 Da, with a total 498 amino acid residues (compared with 51919 Da for 465 amino acids in the reference sequence) and an isoelectric point of 4.93 (compared with 4.85).

Identification of an optional cassette exon in the glucokinase gene

The extra block of bases unique to GK3 cDNA is inserted at the boundary between exons 1 and 2, as defined by Magnuson *et al.* [12]. It might represent a previously unrecognized cassette exon in the glucokinase transcriptional unit, or else reflect the use of alternate donor sites at the 3′-end of exon 1 or of alternate acceptor sites at the 5′-end of exon 2. To resolve this question, cloned rat genomic DNA was subjected to Southern blot analysis with a synthetic 20-mer oligonucleotide specific for the insert of GK3. This led to the localization of the region of interest to a 2.5 kbp *Hin*dIII–*Hin*dIII fragment extending from upstream of exon 2 to downstream of exon 4. The nucleotide sequence between the 5′ *Hin*dIII site and exon 2 is presented in Fig. 2. The block of 151 nucleotides first identified in GK3 cDNA lies within this region, starting 195 bases downstream of the *Hin*dIII site. The intervening sequence between this block and exon 2 is 296 bp in length. The sequences cag/G and G/gtgagt on the 5′ and 3′ sides of this block conform with the consensus sequences at acceptor and donor splice sites respectively [13]. Moreover, a polypyrimidine tract, as typically found at the 3′-end of introns, is noticeable upstream of the block (Fig. 2). From this analysis, we conclude that the extra sequence found in our GK3 cDNA clone reflects the alternative splicing of an optional cassette exon, exon

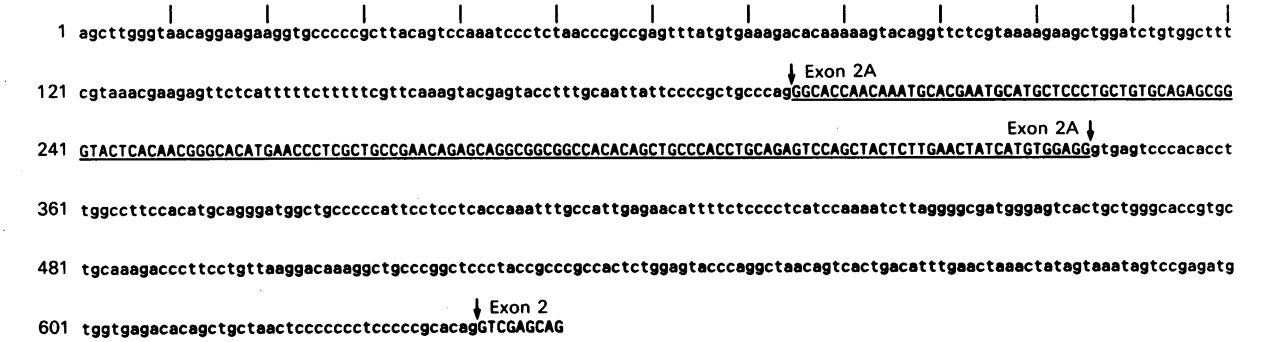

Fig. 2. Alternatively spliced exon and flanking intron sequences in glucokinase transcription unit

Rat genomic DNA was cloned and sequenced as described in the Materials and methods section. Numbering is from the *Hin*dIII site about 3.9 kbp downstream of exon 1 (see [12]). Intron sequences are in lower case, exon sequences in upper case. The optional cassette exon 2A is underlined.

2A, between the previously described exons 1 and 2. It should be noted, however, that no typical branch-point consensus motif (TRAY, where R = purine and Y = pyrimidine [13]) is to be found within 50 nucleotides of flanking sequence upstream of the new exon. On the other hand, the sequence CTAAC 25 nucleotides upstream of exon 2 (Fig. 2) would be expected to serve as strong branch point. Perhaps this accounts for frequent skipping of exon 2A during processing of glucokinase mRNA (see below).

Alternatively spliced glucokinase mRNA is minor species

Poly(A)-containing RNA from rat liver was subjected to Northern blot analysis with two distinct oligonucleotide probes. One probe mapped to a region of exon 2 found in both the GK2 and GK3 sequences and would therefore not discriminate between the different forms of mRNA. As expected [9], this probe hybridized with a 2.4 kb mRNA which is strongly induced in rat liver upon glucose feeding, but which is undetectable in the fasted state (Fig. 3*a*). The other probe, specific for exon 2A, hybridized with a minor RNA of similar size, as well as with several larger RNA species, all induced by glucose refeeding (Fig. 3*b*). Note that the autoradiographic exposure time was 3 times longer for blot (*b*) than for blot (*a*). From the relative exposure times and intensities of the 2.4 kb bands, we infer that only about 5% of mature glucokinase mRNA contains the exon 2A region.

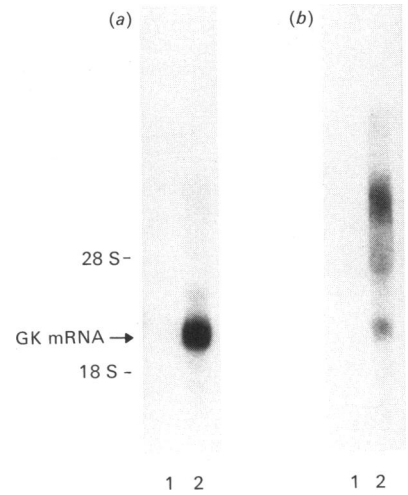

Fig. 3. Northern blot analysis of liver glucokinase (GK) mRNA

Poly(A)-containing RNA was isolated from total liver RNA by chromatography on oligo(dT)–cellulose. After denaturation with glyoxal and dimethyl sulphoxide, samples of poly(A)-containing RNA (12 μg) were electrophoresed in a 1% agarose gel. The resolved RNA was transferred electrophoretically to nylon membranes. Blot (*a*) was hybridized with an oligonucleotide probe mapping in exon 2, between positions 317 and 335 in Fig. 1. Blot (*b*) was hybridized with a probe mapping in exon 2A (double underline in Fig. 1). Both probes were labelled to similar specific activity with polynucleotide kinase and [γ-^{32}P]ATP. Conditions for hybridization and washing of the blots were as described in [6]. Lane 1, RNA from a fasted rat; lane 2, RNA from a rat fasted for 48 h and re-fed on glucose for 6 h.

Received 25 April 1990; accepted 12 June 1990

The larger hybridizable RNA species probably represent splicing intermediates.

Conclusion

Glucokinase transcripts are subject to alternative splicing in rat liver. A novel exon of 151 bp, exon 2A, has been identified in the glucokinase transcription unit between the previously described exons 1 and 2. Inclusion of this exon in the coding region of the message results in downstream frameshift. Interestingly, the cDNA clone in which this insert was discovered also exhibits a 52 bp deletion with respect to the reference sequence, corresponding to the activation of an internal donor site in exon 2. This combined insertion–deletion mechanism results in a frameshift over only a restricted region of the mRNA (part of exon 2), followed by reversal to the original frame for the largest part of the message (exons 3–10). Northern blot analysis has shown that the novel form of mRNA is quantitatively minor. However, the fact that the cDNA clone was isolated from a library generated with polysomal mRNA supports the contention that the minor message may be engaged in translation. The altered polypeptide domain resides outside the region to which putative ATP- and glucose-binding sites have been assigned [5], consistent with the maintenance of catalytic activity. Alternative splicing of glucokinase mRNA may be related to earlier biochemical evidence for multiple isoforms of hepatic glucokinase [14].

This research was supported by grant 32-25664.88 of the Swiss National Science Foundation (to P.B.I.).

REFERENCES

1. Colowick, S. P. (1973) in The Enzymes, (Boyer, P. D., ed.), pp. 1–48, Academic Press, New York
2. Weinhouse, S. (1976) Curr. Top. Cell Regul. **11**, 1–50
3. Meglasson, M. D., Burch, P. T., Berner, D. K., Najafi, H., Vogin, A. P. & Matschinsky, F. M. (1983) Proc. Natl. Acad. Sci. U.S.A. **80**, 85–89
4. Iynedjian, P. B., Mobius, G., Seitz, H. J., Wollheim, C. B. & Renold, A. E. (1986) Proc. Natl. Acad. Sci. U.S.A. **83**, 1998–2001
5. Andreone, T. L., Printz, R. L., Pilkis, S. J., Magnuson, M. A. & Granner, D. K. (1989) J. Biol. Chem. **264**, 363–369
6. Iynedjian, P. B., Pilot, P.-R., Nouspikel, T., Milburn, J. L., Quaade, C., Hughes, S., Ucla, C. & Newgard, C. B. (1989) Proc. Natl. Acad. Sci. U.S.A. **86**, 7838–7842
7. Magnuson, M. A. & Shelton, K. D. (1989) J. Biol. Chem. **264**, 15936–15942
8. Newgard, C. B., Quaade, C., Hughes, S. D. & Milburn, J. L. (1990) Biochem. Soc. Trans. **18**, 851–853
9. Iynedjian, P. B., Ucla, C. & Mach, B. (1987) J. Biol. Chem. **262**, 6032–6038
10. Sambrook, J., Fritsch, E. F. & Maniatis, T. (1989) Molecular Cloning: A Laboratory Manual, 2nd edn., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY
11. Dale, R. M. K., McClure, B. A. & Houchins, J. P. (1985) Plasmid **13**, 31–40
12. Magnuson, M. A., Andreone, T. L., Printz, R. L., Koch, S. & Granner, D. K. (1989) Proc. Natl. Acad. Sci. U.S.A. **86**, 4838–4842
13. Smith, C. W. J., Patton, J. G. & Nadal-Ginard, B. (1989) Annu. Rev. Genet. **23**, 527–577
14. Allen, M. B. & Walker, D. G. (1976) Biochem. Soc. Trans. **4**, 1057