Biol. Chem., Vol. 386, pp. 101–109, February 2005 • Copyright © by Walter de Gruyter • Berlin • New York. DOI 10.1515/BC.2005.013

Expression of the plasma prekallikrein gene: utilization of multiple transcription start sites and alternative promoter regions

Peter Neth[1], Marianne Arnhold[1], Viktoryia Sidarovich[1], Kanti D. Bhoola[2] and Edwin Fink[1,*]

[1] Abteilung für Klinische Chemie und Klinische Biochemie, Chirurgische Klinik Innenstadt, Ludwig-Maximilians-Universität München, Nussbaumstr. 20, D-80336 München, Germany

[2] Centre for Asthma, Allergy and Respiratory Research, School of Medicine and Pharmacology, The University of Western Australia, Sir Charles Gairdner Hospital, Hospital Avenue, Nedlands WA 6009, Perth, Australia

* Corresponding author
e-mail: fink@clinbio.med.uni-muenchen.de

Abstract

The plasma prekallikrein gene is expressed in many different human tissues at distinctly different levels and therefore tissue-specific control of the gene transcription is likely. In this study we demonstrate that transcription of the plasma prekallikrein gene can be initiated at multiple sites, for which at least four different promoters are utilized. A comparison of the genomic and mRNA sequences of mouse plasma prekallikrein revealed that the sequence segment that was formerly regarded as the first exon of the mouse plasma prekallikrein gene consists of three exons, with the first exon localized 14.2 kbp upstream of the translation start. For the rat and human plasma prekallikrein genes, *in silico* analysis suggested an analogous exon-intron organization. Determination of the transcription start sites showed that in both mouse and human, the proximal and distal regions could be utilized for transcription initiation; however, the proximal region is preferred. A deletion mutation analysis of the proximal promoter region using a 1.7-kbp segment revealed a strong activating region immediately upstream of the known mRNA, followed by both a modest repressor and an enhancer region.

Keywords: alternative promoter; alternative transcription start site; Inr element; plasma kallikrein; plasma prekallikrein; promoter activity; TATA box; transcription start site; 5′-untranslated region heterogeneity.

Introduction

Plasma prekallikrein (PPK), the zymogen of the serine protease plasma kallikrein, is synthesized as a single-chain glycoprotein in hepatocytes and secreted into the bloodstream (Bhoola et al., 1992) where it circulates mainly as a complex with high-molecular-weight kininogen (HMWK). *In vitro*, conversion of plasma prekallikrein to active plasma kallikrein is achieved by cleavage of the peptide bond Arg371-Ile372 either in a fluid phase by a factor XII fragment or on negatively charged surfaces by activated factor XII (factor XIIa) with HMWK as the cofactor (contact activation). Factor XIIa-independent PPK activation has been shown to occur on the surface of endothelial cells by a protease (Colman and Schmaier, 1997; Motta et al., 1998, 2001) which has been identified by Shariat-Madar et al. (2002) as prolylcarboxypeptidase. In addition, it has been observed that activation of prekallikrein complexed to HMWK on endothelial cell surfaces may also be catalyzed by heat shock protein 90 (Joseph et al., 2002).

Biological activities attributed to PPK circulating in the blood are: involvement in the activation of neutrophils (Schapira et al., 1982, 1983; Wachtfogel et al., 1983); activation of the C3 convertase of the alternative pathway (DiScipio, 1982); induction of the fibrinolytic cascade by converting the pro-urokinase plasminogen activator to an active molecule (Ichinose et al., 1986; Hauert et al., 1989); and the release from HMWK of bradykinin, which regulates cellular events through specific receptor-coupling of second messengers (Blais et al., 2000; Blaukat, 2003). The formerly accepted role of PPK, together with factor XII and HMWK, in the coagulation pathway is no longer sustainable, since deficiencies of these proteins in plasma do not result in bleeding disorders (Colman and Schmaier, 1997; Davie, 2003).

Recently we demonstrated that the *PPK* gene is transcribed not only in the liver, but also at distinctly different levels in non-hepatic human tissues (Neth et al., 2001). Furthermore, in immunocytochemical studies plasma kallikrein/PPK has been specifically visualized in several human tissues, such as pancreas, kidney, testis and stomach (unpublished data). Since PPK in the blood originates essentially or totally from the liver (Colman and Wong, 1979), we concluded that the extrahepatically synthesized PPK has special functions at or near the site of its synthesis.

Interestingly, novel biological roles of plasma kallikrein that fit into the concept of local PPK functions have been reported recently. Hecquet et al. (2002) demonstrated that plasma kallikrein, as well as some other proteases, activates bradykinin B2 receptor directly, independent of kinin release. Selvarajan et al. (2001) recently reported that plasma kallikrein is required for adipogenesis. The enzyme mediates activation of a plasminogen cascade, which promotes adipocyte differentiation by degrading the fibronectin-rich extracellular matrix of preadipocytes. Peek et al. (2002) found that plasma kallikrein, as well as coagulation factor XI, can activate pro-hepatocyte growth factor. Hepatocyte growth factor (HGF) is the

Bereitgestellt von | Universitaetsbibliothek der LMU Muenchen
Angemeldet | 129.187.254.47
Heruntergeladen am | 18.11.13 11:20

Table 1 Alignment of mouse mRNA to mouse, rat and human genomic sequences.

Exon and intron numbering	E1a	I1a	E1b	I1b	E1c
Length of newly identified mouse exons and introns	87	1271	226	11852	291
Segments of mouse exons that map to rat genomic sequence	3–87		1–226		2–290
Length of mapping rat segments	85		228		290
Identity, %	94		92		85
Gap, %	0		1		1
Length of rat intervening sequences		1271		16400	
Segments of mouse exons that map to human genomic sequence	6–85		1–175		80–258
Length of mapping human segments	80		175		188
Identity, %	81		85		74
Gap, %	2		0		9
Length of human intervening sequences		1198		16734	

Mouse exons mE1a, mE1b and mE1c were identified by mapping mouse PPK mRNA to the mouse genomic sequence. The analogous rat and human exons E1a, E1b and E1c result from BLAST alignment of mouse PPK mRNA to rat and human genomic sequences (see the materials and methods section).

ligand for the tyrosine kinase receptor c-Met, and thus plasma kallikrein may also regulate processes that involve the HGF/c-Met signaling pathway. Akita et al. (2002) investigated the molecular mechanism underlying impaired liver regeneration by circulating endotoxin after partial hepatectomy and provided convincing evidence that this impairment is caused by plasma kallikrein-mediated activation of latent transforming growth factor β (TGF-β). Interestingly, we recently found co-localization of PPK and TGF-β1 in the tubules and mesangial cells of normal renal parenchyma and the tissue adjacent to clear cell carcinoma of the human kidney (unpublished data).

Synthesis in multiple tissues and involvement in various physiological processes suggest that a distinct regulation of PPK synthesis in particular tissues and cells is required, and we set out to characterize the transcriptional control of human *PPK gene* expression. In the current study, we show that the transcription of human *PPK* gene can be initiated at multiple sites, which are up to 23 kbp apart, indicating that different promoters can be utilized, presumably in a tissue- and/or cell type-specific manner. Alternatively spliced mRNA variants were detected, some of them indicating the existence of N-terminally truncated PPK molecules. The major transcription start sites are close to the 5′-end of the known PPK mRNA (Chung et al., 1986). Examination by deletion analysis of the 1.7-kbp region upstream of the PPK mRNA for promoter activity showed that it contains both transcription-enhancing and -repressing segments.

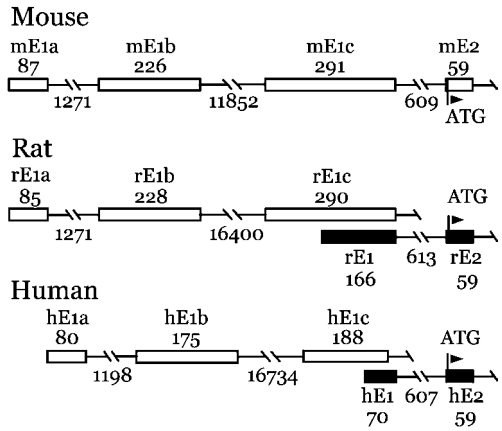

Figure 1 Organization of the 5′-regions of mouse, rat and human *PPK* genes.
The mouse exon-intron structure was obtained by mapping mouse PPK mRNA to the mouse genomic sequence. The putative exons E1a, E1b, and E1c (open boxes) for rat and human were assigned by aligning mouse mRNA to rat and human genomic sequences. The known human and rat exons 1 and 2 are shown as filled boxes; the start codons located at position 2 of exons mE2, rE2 and hE2 are indicated.

Results

Comparison of the mRNA and genomic sequences of human, mouse and rat PPK

A comparison of the known PPK mRNAs from mouse (Seidah et al., 1990; NM_008455), rat (Seidah et al., 1989; Beaubien et al., 1991; NM_012725) and human (Chung et al., 1986; NM_000892) showed that the three 5′-UTRs vary significantly in length, encompassing 605, 167 and 71 bp, respectively. Therefore, the question arose as to whether the known human mRNA represents a full-length transcript of the *PPK* gene, and hence we examined this aspect by *in silico* analysis.

Mapping of the mRNAs to their respective genomic sequences confirmed for human (Yu et al., 2000) and rat (Beaubien et al., 1991) that the *PPK* genes consist of 15 exons with the coding sequence starting at position 2 of exon 2. In contrast, for mouse our alignment revealed that what was previously regarded as mouse exon 1 consists of three exons of 87, 226 and 291 bp (tentatively designated mE1a, mE1b and mE1c), which are separated by introns I1a and I1b of 1271 and 11 852 bp, respectively (Table 1, Figure 1).

The finding that mouse PPK mRNA contains two additional exons led to the view that analogous transcripts may also be formed in human and rat. In order to examine this possibility, we aligned the mouse exons mE1a, mE1b and mE1c to human chromosome 4 and rat chromosome 16, where the *PPK* genes are localized. As

Bereitgestellt von | Universitaetsbibliothek der LMU Muenchen
Angemeldet | 129.187.254.47
Heruntergeladen am | 18.11.13 11:20

shown in Table 1 and Figure 1, nearly the complete exons mE1a, mE1b and mE1c mapped to the rat genomic sequence with identities between 85% and 94%. In analogy to the new mouse exons, these segments were provisionally designated rE1a, rE1b and rE1c. The intervening sequences rI1a and rI1b were 1271 and 16 400 bp in length, respectively.

By alignment of the three mouse exons to the human genomic sequence, three segments of high similarity were also found (designated hE1a, hE1b, hE1c; Table 1, Figure 1), whereby hE1c did not overlap the complete conventional human exon 1, but only 48 bp of its 5′-end. The lengths of the intervening sequences corresponding to mouse introns 1a and 1b (1271 and 11 852 bp) were 1198 and 16 734 bp, respectively.

Taken together, comparison of the mRNA sequences of human, mouse and rat, and the alignment of the newly identified first three mouse exons to the rat and human genomic sequences suggested that the known rat and human mRNA sequences do not represent full-length transcripts of the *PPK* gene. Transcripts of the *PPK* gene with 5′-extensions (as compared to the known mRNAs) containing two additional exons might also exist in these two species. Conversely, it seemed possible that transcription of the mouse *PPK* gene can also be initiated within exon mE1c at start sites analogous to those of rat or human.

Table 2 Transcription start sites in human liver, pancreas, kidney and HepG2 cells.

Position of TSS		Liver	Pancreas	Kidney	HepG2
UTR	-127			1	
UTR	-36			1	
UTR	-11				4
UTR	-5				1
E1	+1	5	3	2	1
E1	+3	1			
E1	+4	1			
E1	+6	3	3	1	
I1	136			1	
I1	243			1	
I1	248			1	
I1	265			1	
I1	334			1	
I1	436			1	
E2	+75			1	
E2	+76			2	
E2	+77			1	

TSS, transcription start site; UTR, 5′ untranslated region; E1, E2, conventional exons 1 and 2; I1, intron 1. The numbering of positions in the 5′-upstream region and in exons 1 and 2 is relative to the 5′-end of the known mRNA sequence; in intron 1 the numbering refers to the 5′-end of intron 1. Numbers in columns 3–6 indicate the number of clones with the respective transcription start site.

The known human PPK mRNA is a full-length transcript and the transcriptional start sites in human liver, pancreas and kidney are diverse

With the aim of characterizing the promoter of the human plasma prekallikrein gene, we determined first whether the known mRNA sequence (Chung et al., 1986) represented a full-length transcript and whether the same transcription starts were utilized in different tissues. To answer this question, transcription initiation sites were determined by RNA ligase-mediated rapid amplification of 5′-cDNA ends (RLM-RACE) with the PPK-specific primers positioned in exon 2 using poly(A)$^+$ mRNA from human liver, pancreas and kidney, the three tissues with the highest PPK mRNA expression. The relative expression levels in liver, pancreas and kidney are 100%, 68% and 25%, respectively (Neth et al., 2001).

By analysis of 31 clones from the three tissues, 15 different sites for transcription initiation were detected (Table 2); the distance on the genomic sequence between the most upstream and most downstream transcription start sites was 811 bp. Transcription start sites at or near position +1 of the known mRNA were detected in liver, pancreas and kidney, proving that the known human PPK mRNA represents a full-length transcript. Additional transcription start sites, further upstream or downstream of position +1, were found exclusively in the kidney (Table 2): two transcripts had 5′-extensions of 36 and 127 bp, and six had initiation sites within intron 1. Accordingly, the segment of intron 1 upstream of the respective transcription start sites, as well as the conventional exon 1, are included in the promoter region. In this type of transcript the conventional first exon is missing, even though the complete PPK protein is encoded. Most interestingly, three additional transcription start sites detected in the kidney were located in exon 2 in positions +75, +76 and +77 (the numbering refers to the known mRNA sequence); hence, the promoter region includes the complete intron 1 (607 bp, NT_022792; or 608 bp, AF232734 and our own sequencing results) and the conventional exon 1 (70 bp). Furthermore, since the start sites are downstream of the start codon at +72, translation has to start at an ATG codon further downstream, causing a truncation by at least 38 amino acids. As a result, a PPK molecule would be produced that would be lacking not only the signal peptide, but also at least 19 amino acids of the mature protein.

The human *PPK* gene has transcription initiation sites analogous to mouse and *vice versa*

To address the question as to whether transcription start sites analogous to mouse *PPK* can be used in human and *vice versa*, we carried out a series of 5′-RLM-RACE analyses with mRNA from human and mouse liver. Using human liver mRNA and the gene-specific primer located in hE1a (cf. Figure 2), we detected two types of transcripts (types I and II, cf. Figure 2) with the transcription start sites 18.5 and 22.5 kbp upstream of the previously reported human PPK mRNA (Yu et al., 2000). Type I transcripts (n=2) started at positions resulting in lengths of hE1a of 83 and 110 bp (the 3′-end of hE1a identified by the alignment with mouse exon 1, cf. Table 1, was corrected following the GT-AG rule). The type II transcript (n=1) contained two additional exons upstream of hE1a. The first upstream exon, hE1a-1 (135 bp), was separated from hE1a by 93 bp, and the second upstream exon, hE1a-2 (285 bp), from hE1a-1 by a 3562-bp intron.

Using mouse liver mRNA and the gene-specific primer positioned in mE1c, we identified five different transcrip-

Bereitgestellt von | Universitaetsbibliothek der LMU Muenchen
Angemeldet | 129.187.254.47
Heruntergeladen am | 18.11.13 11:20

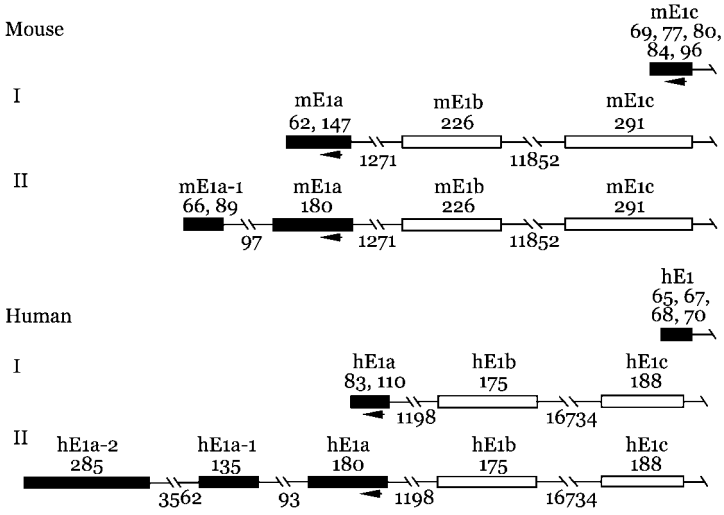

Figure 2 Determination of transcription initiation sites of the mouse and human *PPK* genes by RLM-RACE.
The positions of gene-specific primers (arrowheads) were located in mouse exons mE1c and mE1a, and in human exons hE2 (not shown) and hE1a. Numbers indicate the lengths of exons and introns in bp. Filled boxes represent the exons detected by RLM-RACE and open boxes the exons identified by alignment of mouse mRNA to the genomic sequences. Roman numerals indicate the different types of transcripts (cf. results section).

tional start sites within mE1c (Figure 2). The resulting mE1c lengths of 69, 77, 80, 84 and 96 bp are similar in size to the conventional human exon 1. No transcripts corresponding to the known mouse mRNA (i.e., starting with mE1a) were detectable in this experiment. Therefore, we carried out a second 5′-RACE with the gene-specific primer in mE1a (Figure 2). We found two types of transcripts: type I (n=2) started with mE1a, whereby the lengths of mE1a were 62 and 147 bp. Transcripts of type II (n=2) had an additional upstream exon, mE1a-1, which was separated from mE1a (180 bp) by an intron of 97bp. For mE1a-1, two transcription start sites were obtained, causing mE1a-1 lengths of 66 and 89 bp. No transcription start site representing exactly the 5′-end of the published mouse PPK mRNA (Seidah et al., 1990) was detected.

Additional evidence for the use of multiple transcription initiation sites was gained from a search of EST databases. For human, two sequences were found corresponding to our type II transcript (two additional exons upstream of hE1a; Figure 2), and two sequences starting at positions +3 and +9 of the known mRNA. For mouse, three sequences were detected with the 5′-ends in mE1a, whereby none of the sequence start sites coincided exactly with that of the known mouse PPK mRNA. Furthermore, six sequences were found with one (our type II, cf. above) to three additional upstream exons. A total of 13 EST sequences had start sites within mE1c; the resulting exon lengths would be 26, 28, 35, 50, 54, 55, 70, 71, 76, 98 and 119 bp, and thus would be similar to the size of the conventional human exon 1. A search of the Database of Transcriptional Start Sites (Suzuki et al., 2002) produced no results for human, whereas for mouse 10 transcription start sites within mE1c are available, resulting in lengths of mE1c of 26, 50, 54, 55, 70, 76 and 96 bp, but no distal transcription start sites are given.

Taken together, the RLM-RACE and the data mining results demonstrate that both in mouse and human, transcription of the *PPK* gene can be initiated at sites in regions that are distantly apart. Form the experimental results it can be concluded that transcription of the human *PPK* gene in the liver can be driven by three different promoters, namely the proximal promoter preceding position 1 of the conventional human exon 1, hE1, and two distal promoters located upstream of hE1a and hE1a-2, approximately 18.5 and 22.5 kbp upstream of hE1 (Figure 2).

Transcriptional start sites in HepG2 cells

The human hepatoma cell line HepG2, which is widely used as a model system for hepatocytes, expresses PPK mRNA, as we had previously ascertained by RT-PCR (data not shown). Since we intended to use this cell line for promoter activity studies, we examined which transcription start sites are utilized in HepG2 cells. DNA sequence analysis of six PCR products obtained by RLM-RACE revealed that transcription of the *PPK* gene in HepG2 cells starts in the same region as in liver and pancreas (Table 2). The initiation sites identified represented positions -11, -5 and +1. The similarity or identity of the start sites to those found in liver and pancreas indicates that in HepG2 cells the same promoter region is effective, which justified employing HepG2 cells as a model for studying the activity of the proximal *PPK* promoter.

Deletion-mutation analysis of the proximal *PPK* promoter region

A segment comprising 1675 bp of the 5′-flanking region and 22 bp of the conventional exon 1 of human *PPK* gene was cloned by employing the 'Genome Walker' technique. The segment was linked to the 5′-end of the promoterless secretory alkaline phosphatase reporter gene of the pSEAP2-Basic plasmid. This construct was transiently transfected into HepG2 cells and the promoter activity of the construct was determined by measuring

Bereitgestellt von | Universitaetsbibliothek der LMU Muenchen
Angemeldet | 129.187.254.47
Heruntergeladen am | 18.11.13 11:20

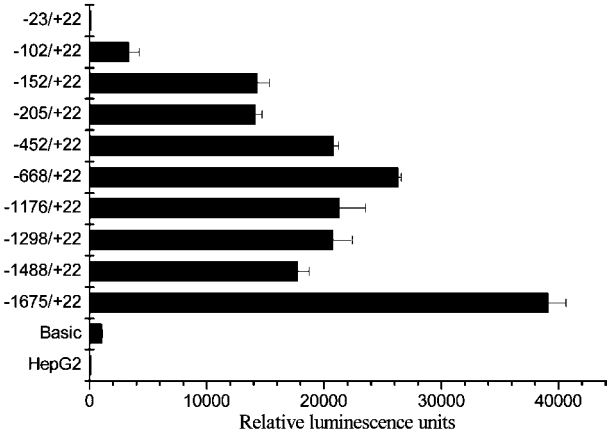

Figure 3 Deletion mutation analysis of the human *PPK* promoter region.
The 5′- and 3′-ends of each deletion mutant are given on the left-hand side. Each construct was cloned into the pSEAP2 reporter gene plasmid and transiently transfected into HepG2 cells (in triplicate) and the supernatant was analyzed for alkaline phosphatase activity. Controls, indicated as Basic and HepG2, were HepG2 cells transfected with the reporter gene plasmid without insert and non-transfected HepG2 cells. The activity is given in relative luminescence units; the bars indicate the mean±SD (n=3) of the SEAP activity.

the alkaline phosphatase activity released into the supernatant. The promoter activity of the -1675/+22 construct was approximately 43-fold higher than of pSEAP2-Basic (Figure 3), indicating that the sequence encompasses a promoter region. Subsequent *in silico* analysis using MatInspector (Quandt et al., 1995) revealed that numerous consensus sequences for activating and repressing transcription factors are present within the -1675/+22 segment. Taking into consideration the location of potential transcription factor binding sites, nine deletion variants of the full-length -1675/+22 construct were prepared and examined for transcriptional activity. Figure 3 shows that the 10 reporter gene constructs displayed significantly different promoter activity. The highest activity was observed for the full-length construct -1675/+22 (arbitrarily set to 100%). Deletion of the sequence -1675 and -1488 resulted in decrease in activity to approximately 45%. Subsequent deletion of -1488 to -1298, -1298 to -1176, and -1176 to -688 led to a gradual activity recovery to approximately 67%. Further deletion of sequences -668 to -452 and -452 to -205 caused a decrease to 36%; this activity remained unchanged by deletion to -152. Removal of the next 50 bp generated a sharp drop to 9% and deletion of the sequence -102/-23 completely abolished this residual promoter activity.

Discussion

Experimental results and data mining for transcription start sites indicated that transcription of the *PPK* gene in human and mouse liver could be initiated in three different regions (Figure 2). However, neither in mouse nor in human were the distal initiation sites detected when 5′-RLM-RACE experiments were carried out with the gene-specific primers located downstream of the proximal initiation sites, even though the distal start sites must in principle be detectable by such an experiment. However, the distal transcription start sites were detectable when the gene-specific primers in 5′-RLM-RACE experiments were located in the distal exons mE1a and hE1a, respectively (Figure 2), namely, when experimental conditions selective for detecting the distal transcription start sites were chosen. This finding indicates that the proximal start sites represent the major region for transcription initiation. If transcription is mainly initiated near the proximal start site and the gene-specific primer located downstream to this site, then the number of formed short transcripts will be larger than the number of the long ones, which start at the distal transcription start sites. Thus, it would have been necessary in our experiments to screen a much larger number of clones than the 16 we examined to identify the long transcripts in addition to the short ones. Interestingly, the known mouse PPK mRNA described by Seidah et al. (1990) represents one of the long transcripts. Obviously, the authors were quite fortunate in finding one of these rare long transcripts by chance and against statistical expectancy. Taken together, our experiments elucidated that in both mouse and human liver transcription of the *PPK* gene can be initiated in at least three different regions, but the major transcription start region is the proximal one.

A comparison of the proximal transcription start sites in the human liver, pancreas and kidney, the three tissues with the highest PPK mRNA expression levels (Neth et al., 2001), unveiled tissue-dependent utilization of transcription starts (Table 2) and promoters. In the liver and pancreas, all start sites coincided with or were close to position +1 of the known PPK mRNA, whereas in the kidney additional start sites upstream and downstream of +1 were detected. For all the start sites downstream of position +1 the promoter differs substantially from the one upstream of +1, as it includes part of intron 1 and exon 1 and can be regarded as alternative promoter(s). In all resulting transcripts the first exon will be missing. Moreover, those transcripts starting downstream of the start codon would give rise to a PPK molecule lacking both the signal peptide and at least 19 amino acids of the mature protein. Accordingly, because of the missing signal peptide, this molecule cannot be directed to the secretory pathway and therefore an intracellular role of the truncated PPK could be assumed. However, so far we have failed to demonstrate the existence of such a truncated PPK protein by Western blot experiments with human tissue extracts (data not shown).

Diverse transcription initiation is widespread throughout the genome, as revealed by Suzuki et al. (2001) who determined 5880 transcription start sites in 276 human genes in a 5′-full-length enriched cDNA library. In approximately 75% of these genes the most upstream and downstream transcription start sites were separated by up to 100 bp, and in ~25% the transcription starts were between 100 and >200 bp apart. Of the 276 genes, 15% had tightly clustered start sites and in only five genes did the transcription start site map to a single position. At present no analysis is available on whether different or single promoters drive gene transcription from multiple transcription start sites. However, it seems likely that a

Bereitgestellt von | Universitaetsbibliothek der LMU Muenchen
Angemeldet | 129.187.254.47
Heruntergeladen am | 18.11.13 11:20

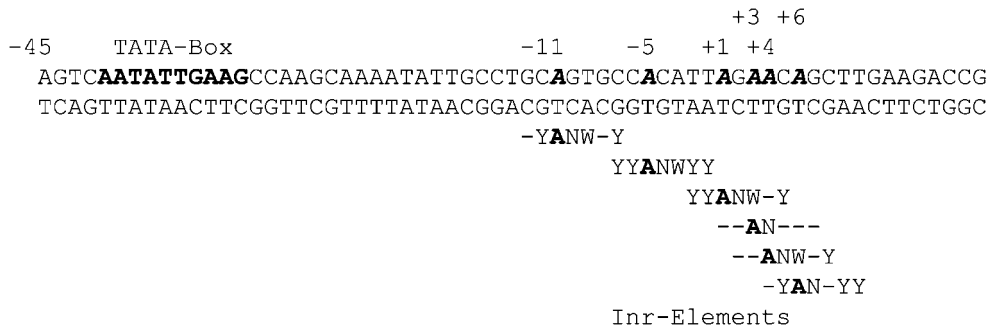

Figure 4 Transcription start sites of the *PPK* gene determined in human liver and pancreas.
A putative TATA box is shown in bold font. Experimentally determined transcription start sites are indicated by bold italics; the numbers are relative to the start site of the known mRNA. Consensus sequences of Inr elements are indicated below the sequence.

single promoter has the ability to drive transcription from different transcription start sites when they are located within an approximately 50-bp-wide region (Suzuki et al., 2001). Such use of a single promoter is certainly the case for the tightly clustered proximal transcription starts of the *PPK* gene.

The use of alternate promoters and multiple transcription start sites for a single gene has been described for multiple mammalian genes, and is regarded as part of the elaborate mechanisms that have evolved in multicellular organisms to provide a means of highly diversified temporal and spatial control of gene expression (Lemon and Tjian, 2000; Hochheimer and Tjian, 2003). In the case of *PPK*, the utilization of alternative promoters and/or alternative transcription start sites results in the expression of at least 20 different transcripts (Table 2 and Figure 2). This diversity of transcripts infers an intricate temporal and cell type-specific control of expression of the *PPK* gene and provides circumstantial evidence supporting our concept that PPK/plasma kallikrein plays hitherto unknown diverse functional roles in various tissues and cell types (Hermann et al., 1999; Neth et al., 2001). Supporting evidence for cell type-specific expression is also gained from our ongoing immunocytochemical studies (unpublished data), which for example show intense staining of PPK in distinct cells of human kidney tubules.

The experimental and data mining results demonstrate that, at least in the liver, it is the proximal region that is mainly used for transcription initiation. Therefore, for our studies on the regulation of expression of the human *PPK* gene, we chose to concentrate first on this major transcription start region. A TATA-box motive was identified at -41 to -32 (relative to the start of the known mRNA; Figure 4) which is a few bp upstream of the region most preferred for a TATA-box (Bucher, 1990), but can be expected to be a functional core promoter element. Consensus sequences of a second core promoter, the initiator element (Inr) (Smale, 1997), were identified for five of the six transcription start sites in liver, pancreas and HepG2 cells (Figure 4). No downstream promoter element (DPE) motive (G-A/T-C-G in the +30 region) (Burke and Kadonaga, 1997; Kutach and Kadonaga, 2000) was detectable. Taken together, the core promoter region of the *PPK* gene contains both a TATA-box and Inr motives in positions that can direct transcription initiation to any of the proximal transcription start sites found in liver, pancreas and HepG2 cells, whereby the action of the motives may be paired or independent.

In our promoter activity study the high transcriptional activity of the -1675/+22 reporter gene construct demonstrated that the 1.7-kbp sequence upstream of hE1 contains *cis*-acting elements important for the regulation of *PPK* gene transcription. Deletion mutation analyses (Figure 3) identified two regions with high stimulatory effects on transcription. The first one, the putative promoter, was found within the proximal sequence from -668 bp to the transcription start site, with the *cis*-acting elements mainly localized in the segments -102/-152 and -205/-452. The second one was detected in the 189-bp region -1488 to -1675 and is presumably an enhancer region. Transcription-repressing activity was found in the region -668 to -1488 between the promoter and the enhancer, mainly localized within the segments -668 to -1176 and -1298 to -1488. *In silico* searches for transcription factor binding sites in the promoter region were carried out using two different methods (Quandt et al., 1995; Lenhard et al., 2003). However, even with the most stringent parameters, too many potential transcription factor binding sites were found to allow for a conclusive decision on which sites are biologically functional and which are 'innocent bystanders'. Therefore, further experiments have been designed to identify the critical binding sites.

PPK mRNA is expressed in many human tissues at different levels (Neth et al., 2001) and PPK participates as a protease in various biological processes. Therefore, in order to comply with the spatially and temporally varying demands for the functional activity of PPK, distinct mechanisms for fine-tuned transcriptional regulation of the *PPK* gene may be expected. In accordance, our experimental and data mining investigations show that at least four promoter regions and diverse transcription start sites can be utilized, and tissue-dependent alternative use of promoters and transcription starts is evident from a comparison of the start sites in liver and kidney.

Materials and methods

Determination of transcriptional start sites of *PPK*

Transcriptional start sites were determined by performing RNA ligase-mediated rapid amplification of 5′-cDNA ends (RLM-

Bereitgestellt von | Universitaetsbibliothek der LMU Muenchen
Angemeldet | 129.187.254.47
Heruntergeladen am | 18.11.13 11:20

Table 3 Gene-specific primers for RLM-RACE.

Designation		Position	Sequence (5′–3′)
Human			
hP-r139	Exon 2	19560916–19560895	TAGCAAACAAGGAAATGAAATA
hP-r124-Nhe	Exon 2	19560901–19560876	GCTAGC-TGAAATAAGTTGCTTGCTTGAATAAA
hE1a-or71	hE1a	19541931–19541910	GCCAGCAGAGAAGGGCACGTAG
hE1a-ir47	hE1a	19541907–1941884	ATATGGATGGCGCCCTTGTGCATT
Mouse			
mP-or9146	mE1a	75–49	GAGGTCCAGCAGAAAATGGCACATAGG
mP-ir9126	mE1a	55–29	ACATAGGCATAGGGATGGCGTCCTTGG
mP-or22799	mE1c	605–583	CCTGGAAGGATGGTCACGTTGTG
mP-ir22771	mE1c	577–550	GGAGGAGAGGGAGTCTTCACATGAAGAT

The position numbers refer to human NT_022792.15 and mouse PPK mRNA NM_008455.1. Underlined nucleotides indicate an NheI recognition sequence.

RACE; Invitrogen, Karlsruhe, Germany) with 250 ng of poly(A)$^+$ mRNA from human liver, pancreas, and kidney and from mouse liver (Clontech, Heidelberg, Germany). 5′-RLM-RACE ensures that only full-length 5′-ends of mRNA molecules with an intact Cap structure are selectively reverse transcribed and amplified by nested PCR. For reverse transcription, random hexamer priming was employed. The gene-specific primers used for primary and nested PCR are listed in Table 3. The PCR products were cloned into the plasmids pCRII-TOPO or pCR4-TOPO (Invitrogen) and the DNA sequences of the inserts were determined by a commercial service (MediGenomix, Martinsried, Germany).

Isolation of the prekallikrein promoter region

The upstream region of the human prekallikrein gene was cloned using the Genome Walker Kit (Clontech). Primary and nested PCRs were carried out with the Advantage Genomic Polymerase Mix (Clontech) using forward primers complementary to the adapter sequences and the gene-specific reverse primers 5′-AAT TGC TTC TTG GAG GTG AGT-3′ and 5′-GCT AGC TGA ACG GTC TTC AAG CTG TTC T-3′ (NheI recognition site underlined) which correspond to nucleotides 60–40 and 22–1 of the prekallikrein cDNA sequence. A 1.7-kbp product was obtained, which was cloned into the pCRII-TOPO vector and sequenced on both strands. The plasmid was cut with MluI and NheI and the resulting fragment was purified by agarose gel electrophoresis and cloned into the plasmid pSEAP2-Basic (Clontech), which was designated PPK/pSEAP2-Basic. Plasmid DNA was isolated from 100-ml overnight cultures of single clones by the Mobius 1000 Kit (Novagen, Bad Soden, Germany) and stored at -20°C.

Reporter plasmids

Prekallikrein promoter segments of various lengths were generated by PCR with the PPK/pSEAP2 plasmid as a template. PCR was performed with Thermozyme (Invitrogen) using the primers shown in Table 4. The PCR products were cloned into the plasmid pCRII-TOPO and sequenced on both strands. Cloning into pSEAP-Basic and subsequent plasmid preparation was achieved as described above.

Cell culture, transfection, and SEAP assay

HepG2 cells were purchased from DSMZ (Braunschweig, Germany) and maintained at 37°C in RPMI medium (PAA, Pasching, Austria) supplemented with 10% fetal calf serum (FCS) (PAN Biotech, Aidenbach, Germany). Cells grown to confluence in 75-cm^2 flasks were harvested by trypsination (PAN Biotech) and replated in 12-well plates at approximately 40% confluence (2 ml of medium per well). The cells were grown overnight at 37°C in 5% CO_2 to 50–70% confluence. The medium was replaced by 0.9 ml of fresh RPMI medium with 10% FCS, and 0.1 ml of transfection mixture containing 3 μl of FuGENE 6 transfection reagent (Roche, Mannheim, Germany) and 1 μg of DNA in RPMI medium was added.

Secretory alkaline phosphatase (SEAP) activity was determined using the chemiluminescent SEAP assay system (Clontech). After 24 and/or 48 h, 15-μl aliquots from each cell culture supernatant were transferred into the wells of a 96-well microtest plate; 45 μl of dilution buffer was added and the plate was incubated for 30 min at 65°C to inactivate endogenous alkaline phosphatase. After cooling, 60 μl of assay buffer and, after 5 min at room temperature, 60 μl of CSPD substrate solution

Table 4 Primers used for the construction of reporter plasmids.

Promoter mutant	Position	Nucleotide sequence (5′–3′)
-23/+22	-23/-2	ACGCGT-AATATTGCCTGCAGTGCCACAT
-102/+22	-102/-81	ACGCGT-GATGTTCATGGAATATGTTGAC
-152/+22	-152/-131	ACGCGT-GACATCACTCCCTGAAATAGTT
-205/+22	-205/-184	ACGCGT-GGGATGAGCGCTAGAAACTCCT
-452/+22	-452/-431	ACGCGT-GTGGTGAAGTTAGGTTCTATCT
-668/+22	-668/-647	ACGCGT-GCTAACGTGGAGGCTAGATAGA
-1176/+22	-1176/-1155	ACGCGT-AGTGTGGCCTTGCACTGAAGTA
-1298/+22	-1298/-1277	ACGCGT-TCTTAAGGCAATTATTTTTAGC
-1488/+22	-1488/-1467	ACGCGT-CTAGTCTGCTGCTGAAGCTGTT
-1675/+22	+22/+1	GCTAGC-TGAACGGTCTTCAAGCTGTTCT

The numbers indicate the position relative to the 5′-end of the known mRNA. Underlined nucleotides represent MluI and NheI recognition sequences.

Bereitgestellt von | Universitaetsbibliothek der LMU Muenchen
Angemeldet | 129.187.254.47
Heruntergeladen am | 18.11.13 11:20

(Clontech) were added. The luminescence was quantified after 10 min.

BLAST analyses and search for a TATA box

For the alignment of sequences, the NCBI program BLAST 2 Sequences (Tatusova and Madden, 1999) was employed. The settings were: match, 1; mismatch, -2; gap open, 1; gap extension, 1; x_dropoff, 50; expect, 10; wordsize, 10; and filter, yes. Accession numbers of the PPK mRNA sequences are: human, NM_000892.2; mouse, NM_008455.1; and rat, NM_012725.1. Accession numbers of the genomic sequences are: human, NT_022792.15; mouse, NW_000342.1; and rat, NW_043042.1. For the TATA-box search, the program accessible at www.itb.cnr.it/webgene was used.

Acknowledgments

This work was supported in part by the Deutsche Forschungsgemeinschaft, grant FI204/6-1.

References

Akita, K., Okuno, M., Enya, M., Imai, S., Moriwaki, H., Kawada, N., Suzuki, Y., and Kojima, S. (2002). Impaired liver regeneration in mice by lipopolysaccharide via TNF-α/kallikrein-mediated activation of latent TGF-β. Gastroenterology *12*, 35–36.

Beaubien, G., Rosinski-Chupin, I., Mattei, M.G., Mbikay, M., Chrétien, M., and Seidah, N.G. (1991). Gene structure and chromosomal localization of plasma kallikrein. Biochemistry *30*, 1628–1635.

Bhoola, K.D., Figueroa, C.D., and Worthy, K. (1992). Bioregulation of kinins: kallikreins, kininogens, and kininases. Pharmacol. Rev. *44*, 1–80.

Blais, C., Marceau, F., Rouleau, J.L., and Adam, A. (2000). The kallikrein-kininogen-kinin system: lessons from the quantification of endogenous kinins. Peptides *21*, 1903–1940.

Blaukat, A. (2003). Structure and signalling pathways of kinin receptors. Andrologia *35*, 17–23.

Bucher, P. (1990). Weight matrix descriptions of four eukaryotic RNA polymerase II promoter elements derived from 502 unrelated promoter sequences. J. Mol. Biol. *212*, 563–578.

Burke, T.W., and Kadonaga, J.T. (1997). The downstream core promoter element, DPE, is conserved from *Drosophila* to humans and is recognized by TAFII60 of *Drosophila*. Genes Dev. *11*, 3020–3031.

Chung, D.W., Fujikawa, K., McMullen, B.A., and Davie, E.W. (1986). Human plasma prekallikrein, a zymogen to a serine protease that contains four tandem repeats. Biochemistry *25*, 2410–2417.

Colman, R.W., and Schmaier, A.H. (1997). Contact system: a vascular biology modulator with anticoagulant, profibrinolytic, antiadhesive, and proinflammatory attributes. Blood *90*, 3819–3843.

Colman, R.W., and Wong, P.Y. (1979). Kallikrein-kinin system in pathologic conditions. In: Handbook of Experimental Pharmacology, Vol. XXV (Supplement), E.G. Erdös, ed. (Berlin, Germany: Springer-Verlag), pp. 569–607.

Davie, E.W. (2003). A brief historical review of the waterfall/cascade of blood coagulation. J. Biol. Chem. *278*, 50891–50832.

DiScipio, R.G. (1982). The activation of the alternative pathway C3 convertase by human plasma kallikrein. Immunology *45*, 587–595.

Hauert, J., Nicoloso, G., Schleuning, W.D., Bachmann, F., and Schapira, M. (1989). Plasminogen activators in dextran sulfate-activated euglobulin fractions: a molecular analysis of factor XII- and prekallikrein-dependent fibrinolysis. Blood *73*, 994–999.

Hecquet, C., Becker, R.P., Tan, F.L., and Erdös, E.G. (2002). Kallikreins when activating bradykinin B-2 receptor induce its redistribution on plasma membrane. Int. Immunopharmacol. *2*, 1795–1806.

Hermann, A., Arnhold, M., Kresse, H., Neth, P., and Fink, E. (1999). Expression of plasma prekallikrein mRNA in human nonhepatic tissues and cell lineages suggests special local functions of the enzyme. Biol. Chem. *380*, 1097–1102.

Hochheimer, A., and Tjian, R. (2003). Diversified transcription initiation complexes expand promoter selectivity and tissue-specific gene expression. Genes Dev. *1*, 130–132.

Ichinose, A., Fujikawa, K., and Suyama, T. (1986). The activation of pro-urokinase by plasma kallikrein and its inactivation by thrombin. J. Biol. Chem. *261*, 3486–3489.

Joseph, K., Tholanikunnel, B.G., and Kaplan, A.P. (2002). Heat shock protein 90 catalyzes activation of the prekallikrein-kininogen complex in the absence of factor XII. Proc. Natl. Acad. Sci. USA *99*, 896–900.

Kutach, A.K., and Kadonaga, J.T. (2000). The downstream promoter element DPE appears to be as widely used as the TATA box in *Drosophila* core promoters. Mol. Cell. Biol. *20*, 4754–4764.

Lemon, B., and Tjian, R. (2000). Orchestrated response: a symphony of transcription factors for gene control. Genes Dev. *10*, 2551–2569.

Lenhard, B., Sandelin, A., Mendoza, L., Engström, P., Jareborg, N., and Wasserman, W.W. (2003). Identification of conserved regulatory elements by comparative genome analysis. J. Biol. *2*, 13.

Motta, G., Rojkjaer, R., Hasan, A.A., Cines, D.B., and Schmaier, A.H. (1998). High molecular weight kininogen regulates prekallikrein assembly and activation on endothelial cells: a novel mechanism for contact activation. Blood *91*, 516–528.

Motta, G., Shariat-Madar, Z., Mahdi, F., Sampaio, C.A.M., and Schmaier, A.H. (2001). Assembly of high molecular weight kininogen and activation of prekallikrein on cell matrix. Thromb. Haemost. *86*, 840–847.

Neth, P., Arnhold, M., Nitschko, H., and Fink, E. (2001). The mRNAs of prekallikrein, factors XI and XII, and kininogen, components of the contact phase cascade are differentially expressed in multiple non-hepatic human tissues. Thromb. Haemost. *85*, 1043–1047.

Peek, M., Moran, P., Mendoza, N., Wickramasinghe, D., and Kirchhofer, D. (2002). Unusual proteolytic activation of pro-hepatocyte growth factor by plasma kallikrein and coagulation factor XIa. J. Biol. Chem. *277*, 47804–47809.

Quandt, K., Frech, K., Karas, H., Wingender, E., and Werner, T. (1995). MatInd and MatInspector: new fast and versatile tools for detection of consensus matches in nucleotide sequence data. Nucleic Acids Res. *23*, 4878–4884.

Schapira, M., Despland, E., Scott, C.F., Boxer, L.A., and Colman, R.W. (1982). Purified human plasma kallikrein aggregates human blood neutrophils. J. Clin. Invest. *69*, 1199–1202.

Schapira, M., Scott, C.F., Boxer, L.A., and Colman, R.W. (1983). Activation of human polymorphonuclear leukocytes by purified human plasma kallikrein. Adv. Exp. Med. Biol. *156*, 747–753.

Seidah, N.G., Ladenheim, R., Mbikay, M., Hamelin, J., Lutfalla, G., Rougeon, F., Lazure, C., and Chretien, M. (1989). The cDNA structure of rat plasma kallikrein. DNA *8*, 563–574.

Seidah, N.G., Sawyer, N., Hamelin, J., Mion, P., Beaubien, G., Brachpapa, L., Rochemont, J., Mbikay, M., and Chretien, M. (1990). Mouse plasma kallikrein. cDNA structure, enzyme characterization, and comparison of protein and mRNA levels among species. DNA Cell Biol. *9*, 737–748.

Selvarajan, S., Lund, L.R., Takeuchi, T., Craik, C.S., and Werb, Z. (2001). A plasma kallikrein-dependent plasminogen cascade required for adipocyte differentiation. Nat. Cell Biol. *3*, 267–275.

Bereitgestellt von | Universitaetsbibliothek der LMU Muenchen
Angemeldet | 129.187.254.47
Heruntergeladen am | 18.11.13 11:20

Shariat-Madar, Z., Mahdi, F., and Schmaier, A.H. (2002). Assembly and activation of the plasma kallikrein/kinin system: a new interpretation. Int. Immunopharmacol. *2*, 1841–1849.

Smale, S.T. (1997). Transcription initiation from TATA-less promoters within eukaryotic protein-coding genes. Biochim. Biophys. Acta *1351*, 73–88.

Suzuki, Y., Taira, H., Tsunoda, T., Mizushima-Sugano, J., Sese, J., Hata, H., Ota, T., Isogai, T., Tanaka, T., Morishita, S., et al. (2001). Diverse transcriptional initiation revealed by fine, large-scale mapping of mRNA start sites. EMBO Rep. *2*, 388–393.

Suzuki, Y., Yamashita, R., Nakai, K., and Sugano, S. (2002). DBTSS: database of human transcriptional start sites and full-length cDNAs. Nucleic Acids Res. *30*, 328–331.

Tatusova, T.A., and Madden, T.L. (1999). Blast 2 Sequences, a new tool for comparing protein and nucleotide sequences. FEMS Microbiol. Lett. *174*, 247–250.

Wachtfogel, Y.T., Kucich, U., James, H.L., Scott, C.F., Schapira, M., Zimmerman, M., Cohen, A.B., and Colman, R.W. (1983). Human plasma kallikrein releases neutrophil elastase during blood coagulation. J. Clin. Invest. *72*, 1672–1677.

Yu, H.R., Anderson, P.J., Freedman, B.I., Rich, S.S., and Bowden, D.W. (2000). Genomic structure of the human plasma prekallikrein gene, identification of allelic variants, and analysis in end-stage renal disease. Genomics *69*, 225–234.

Received September 28, 2004; accepted December 6, 2004

Bereitgestellt von | Universitaetsbibliothek der LMU Muenchen
Angemeldet | 129.187.254.47
Heruntergeladen am | 18.11.13 11:20