cDNA and deduced amino acid sequence of human PK-120, a plasma kallikrein-sensitive glycoprotein

Hitoshi Nishimura[a], Ikuko Kakizaki[a], Tatsushi Muta[b], Nobumi Sasaki[a], Ping Xiao Pu[a], Toshiyuki Yamashita[a], Shigeharu Nagasawa[a,*]

[a]*Faculty of Pharmaceutical Sciences, Hokkaido University, Sapporo 060, Japan*
[b]*Department of Biology, Faculty of Science, Kyushu University 33, Fukuoka 812, Japan*

Received 9 November 1994; revised version received 28 November 1994

Abstract PK-120 is a substrate for plasma kallikrein (PK), recently purified from human plasma. Here we have established the cDNA sequence for human PK-120 mRNA. The deduced amino sequence of PK-120 revealed that it consists of 902 amino acid residues with a calculated mass of 116,423 Da. The putative cleavage sites by PK have been proposed, suggesting that PK-120 may be a precursor of a bioactive peptide. Most interestingly, PK-120 showed significant sequence identities to heavy chains (HCs) of the inter-α-trypsin inhibitor (ITI) superfamily.

Key words: cDNA cloning; Plasma kallikrein; Plasma glycoprotein; Bradykinin; Inter-α-trypsin inhibitor; Hyaluronic acid; Human

1. Introduction

Plasma kallikrein (PK) is a serine protease in plasma, and plays key roles in the activation of the blood coagulation [1], fibrinolysis [2] and complement [3] systems, the release of bradykinin from high molecular weight kininogen [4], and the cellular responses of neutrophils [5–7]. These events seem to be induced through the limited proteolysis of protein substrates by PK. Recently, we purified a novel PK-sensitive protein from human plasma and characterized it [8]. This protein, named PK-120, is a single-chain glycoprotein of 120 kDa and contains 22% (w/w) carbohydrate. PK-120 circulates in plasma at a concentration of 80 μg/ml, and it is easily degraded by a trace amount of PK. PK-120 is first cleaved by PK to yield a 100-and a 35-kDa fragments, and the resulting 100-kDa fragment is further converted to a 70-kDa fragment. However, the physiological role of PK-120 remains unclear.

In this study, we have determined the entire amino acid sequence of human PK-120 by using recombinant DNA techniques, to assess the function of PK-120. The amino sequence suggested that PK-120 is closely related to heavy chains (HCs) of the inter-α-trypsin inhibitor (ITI) superfamily. Furthermore, the putative cleavage sites by PK have been found and the results predicted that PK-120 may be a precursor of a bioactive peptide.

2. Materials and methods

2.1. Materials

Lysyl endopeptidase from *Achromobacter lyticus* M497-1 was obtained from Wako Pure Chemical Industries Ltd., Osaka. Restriction endonucleases and other DNA-modyfying enzymes were purchased from Takara Shuzo Co., Kyoto, Boehringer Mannheim GmbH, Germany, or New England Biolabs Inc., Beverly, MA. [α-^{32}P]dCTP was from DuPont-NEN Research Products, Boston, MA. Human liver poly(A)$^+$ RNA was a gift from Dr. F. Sakane (Sapporo Medical College, Japan). A human liver λgt11 cDNA library was kindly provided by Prof. M. Tomita (Showa University, Japan). A λgt10 cDNA library of human liver was also used in this study and was purchased from Clontech, Palo Alto, CA. PK was purified from human plasma as previously described [9].

2.2. Determination of partial amino acid sequences of PK-120

PK-120 was purified from human plasma as reported by Pu et al. [8]. The purified material was incubated with PK and the resulting 70- and 35-kDa fragments were isolated from the digest as previously described [8]. Both fragments were reduced, S-pyridylethylated, and then digested with lysyl endopeptidase [10]. The lysyl endopeptidase-digest was separated by reverse-phase high performance liquid chromatography on μBondasphere S-5 C8 300A (2.1 × 150 mm; Waters, Millipore Corp., MA) and TSK-GEL ODS-120T (4.6 × 250 mm; Tosoh Corp., Tokyo) columns. The purified peptides were subjected to amino-terminal sequence analysis on an Applied Biosystems model 477A gas-phase protein sequenator/Applied Biosystems 120A PTH analyzer.

2.3. Polymerase chain reaction (PCR)

Based on the amino acid sequence (TGLLLLSDPDKVTIGLLFWDGRGEGL) of the 35-kDa fragment-derived peptide, two degenerate oligonucleotides, 5′-GA(T/C)CC(G/A/T/C)GA(T/C)AA(G/A)GT(G/A/T/C)AC(G/A/T/C)AT-3′ and 5′-TC(G/A/T/C)CC(G/A/T/C)CG(G/A)CC(G/A)TCCCA(G/A)AA-3′, were designed as PCR primers. First strand cDNA was synthesized from 0.1 μg of poly(A)$^+$ RNA of human liver with random primers, using a GeneAmp RNA PCR kit (Perkin-Elmer). The reaction mixture was used as a template for the following reaction. PCR (40 cycles) was performed using an AmpliTaq (Perkin-Elmer), with denaturation at 94°C for 1 min, annealing at 50°C for 2 min, and extension at 72°C for 2 min. The PCR product (50 bp) was subcloned into a pCRII vector (Invitrogen, San Diego, CA) and sequenced with an Applied Biosystems 373A DNA sequencer using the Dyedeoxy terminator cycle sequencing kit (Applied Biosystems).

2.4. cDNA cloning of PK-120

The cloned 50-bp DNA was labeled with [α-^{32}P]dCTP by PCR as described above, and used to probe the human liver λgt11 cDNA library. The λgt10 cDNA library of human liver was secondly screened with the insert of the positive clone λPK5, after labeling with [α-^{32}P]dCTP using a Ready-To-Go DNA Labelling Kit (Pharmacia, Uppsala, Sweden). Hybridization was carried out at 60°C or at 65°C in a solution of 7% (w/v) polyethylene glycol 6000, 10% SDS and 100 μg/ml yeast tRNA, and washed with 2× SSC (1× SSC = 15 mM

*Corresponding author. Fax: (81) (11) 706-4990.

Abbreviations: PK, plasma kallikrein; ITI, inter-α-trypsin inhibitor; HC, heavy chain of the ITI superfamily; PCR, polymerase chain reaction; HA, hyaluronic acid.

The nucleotide sequence data reported here will be available in the GenBank/EMBL nucleotide sequence databases with the accession number D38535.

SSDI 0014-5793(94)01364-0

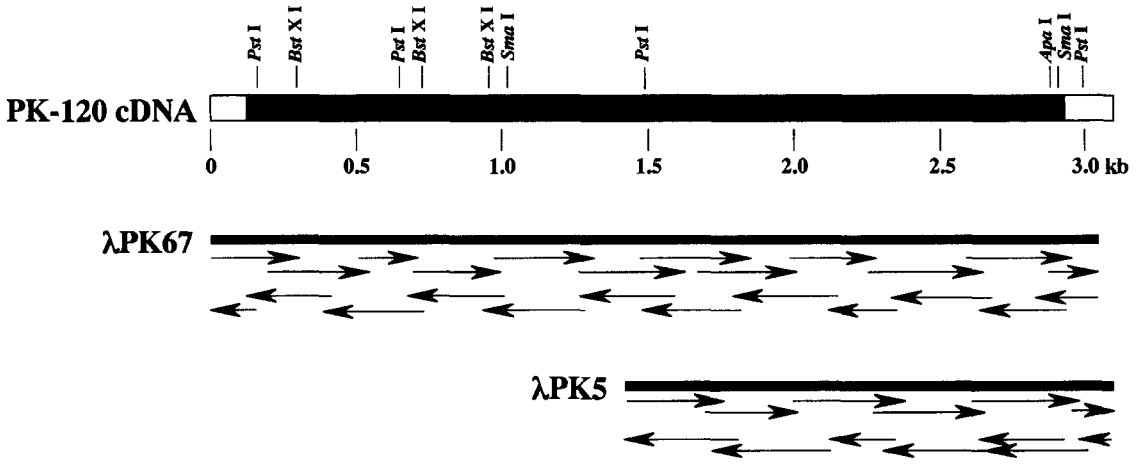


Fig. 1. Restriction map and sequence strategy. The black box at the top shows the open reading frame. Each arrow indicates a direction and an analyzed area on a sequencing.

sodium citrate and 150 mM NaCl, pH 7.0) containing 0.1% SDS at 60°C or at 65°C. Inserts of the positive clones were subcloned into a pBluescript II SK− vector (Stratagene) and then sequenced in both directions by sequence-primer walking.

2.5. Northern blotting

A Multiple Tissue Northern Blot (Clontech), to which poly(A)$^+$ RNAs from various human organs were transferred, was hybridized with a ^{32}P-labeled *Bst*XI fragment of λPK67 in 5× SSC, 10× Denhardt's solution (1× Denhardt's solution = 0.02% polyvinylpyrrolidone, 0.02% bovine serum albumin and 0.02% Ficoll), 50 mM sodium phosphate (pH 6.8), 50% formamide and 100 μg/ml yeast tRNA at 42°C for 16 h and then washed in 0.1× SSC containing 0.1% SDS at 65 °C. The washed filter was subjected to autoradiography at −80 °C.

2.6. Homology search

The nucleotide and the deduced amino acid sequences were compared with all entries in the GenBank/EMBL Databases and the SWISS-PROT Protein Database, respectively, with the GENETYX-MAC system (Software Development Co., Tokyo).

3. Results

3.1. cDNA cloning of PK-120

The amino acid sequences of several lysyl endopeptidase peptides derived from the 70- and the 35-kDa fragments were determined. A pair of oligonucleotide primers corresponding to the amino acid sequence of the 35-kDa fragment-derived peptide were used to amplify a partial cDNA fragment from human liver poly(A)$^+$ RNA. PCR yielded a 50-bp cDNA fragment, and the deduced amino acid sequence of the fragment was identical to the peptide sequence. With the cDNA fragment, 6.5×10^5 independent clones of the human liver λgt11 cDNA library were screened. Of 11 positive clones, the longest positive clone, λPK5, was sequenced. This clone (1.7 kb) was shown to encode a carboxy-terminal portion of PK-120, a 3′ non-coding region, and a poly(A) tail (Fig. 1). To obtain clone(s) encoding the remaining part of PK-120, 1×10^6 independent clones of the human liver λgt10 cDNA library were also screened with the ^{32}P-labeled λPK5 insert. Of 106 positive clones, the clone λPK67 (3.0 kb) was sequenced. λPK67 encodes a 5′ non-coding region and an amino-terminal portion of PK-120 including a putative signal sequence (Fig. 1). The composite nucleotide sequence obtained from these two clones spans 3,058 bp (not including a poly(A) tail) (Fig. 2). The ATG codon at nucleotide 130 begins an open reading frame of 2,790 bp followed by a TAG termination codon. In the 3′ non-coding region, there is one polyadenylation signal with the AATAAA sequence. The open reading frame encodes a putative signal sequence of 28 amino acid residues and a mature protein of 902 amino acid residues, of which 211 residues were confirmed by the protein sequencing. The molecular mass of the mature protein without carbohydrate was calculated to be 116,423 Da. There are four potential N-glycosylation sites (-Asn-Xaa-Ser/ Thr-; Asn^{53}, Asn^{179}, Asn^{489} and Asn^{549}). Putative cleavage sites by PK (-Phe-Arg-Xaa-) [11] have also been proposed (Arg^{455}, Arg^{633} and Arg^{660}).

3.2. Northern Blotting

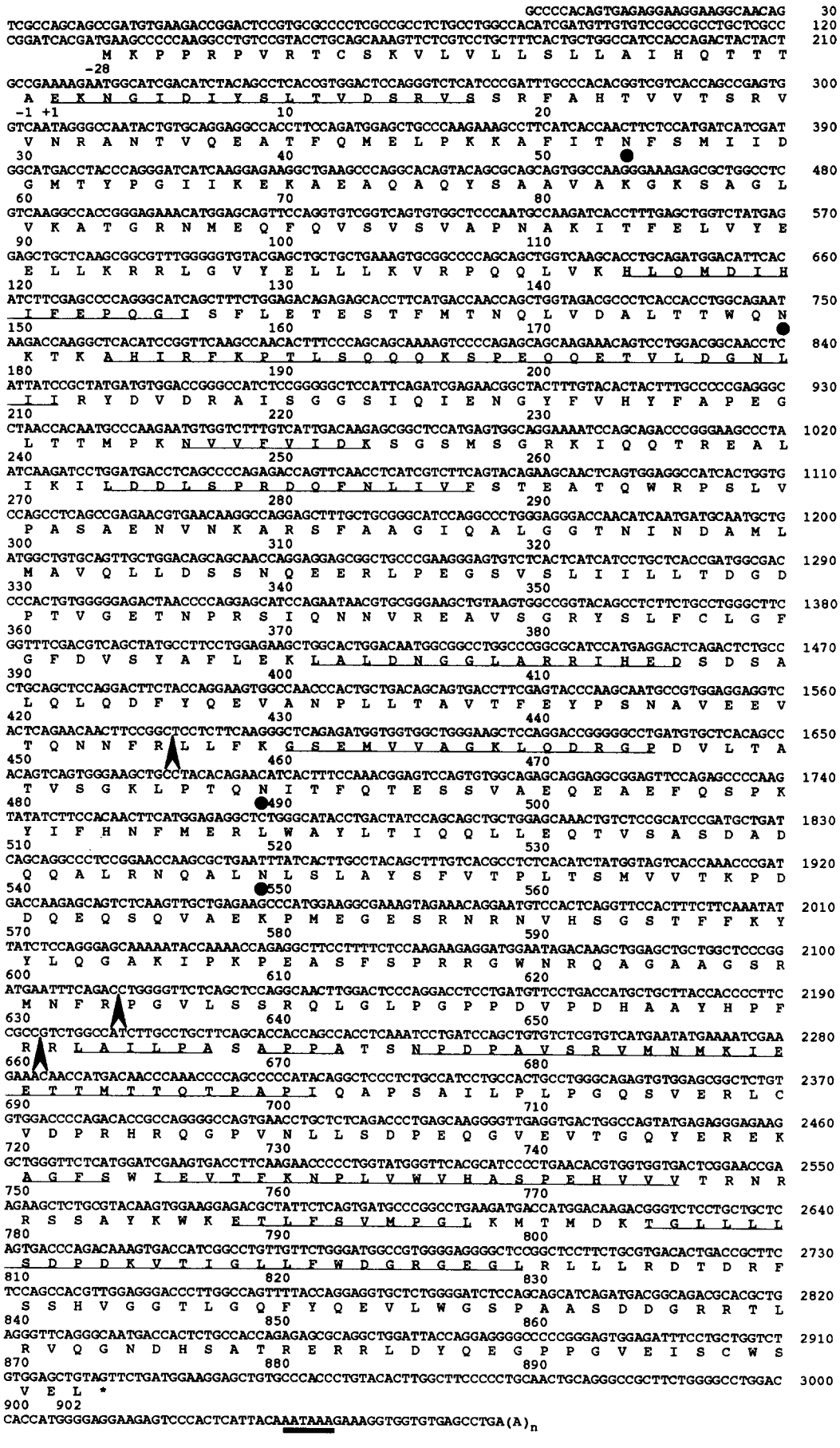
Northern blot analysis of poly(A)$^+$ RNAs of various human organs was carried out. As shown in Fig. 3, a band at 3.3 kb, in addition to a faint 4.0-kb band, was observed only in the poly(A)$^+$ RNA of liver (lane 5).

3.3. Homology search

The nucleotide and the deduced amino acid sequences of PK-120 cDNA were compared with all entries in the databases. These results indicated that PK-120 is highly homologous to HCs of the ITI superfamily. Fig. 4 shows an alignment of the amino acid sequences of PK-120 and the HCs. PK-120 exhibited significant sequence identities to HC1 (34%) and HC2 (31%) of human ITI and HC3 (38%) of human pre-α-trypsin inhibitor. Furthermore, an amino-terminal portion of PK-120

→

Fig. 2. The composite nucleotide sequence (upper) and the deduced amino acid sequence (lower) of PK-120. A polyadenylation signal (AATAAA) is indicated by a thick underline. The amino acid residues confirmed by the protein sequencing are underlined. The putative cleavage sites by PK (-Phe-Arg-Xaa-) are marked by *arrowheads*. The potential N-glycosylation sites (-Asn-Xaa-Ser/Thr-) are shown by black circles.



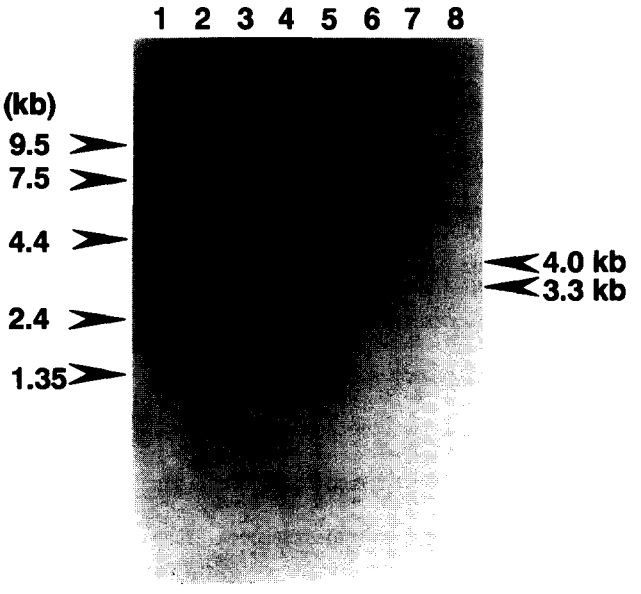


Fig. 3. Northern blot analysis of PK-120 mRNA. A Multiple Tissue Northern Blot (Clontech), to which poly(A)$^+$ RNAs from various human organs were transferred, was hybridized with a ^{32}P-labeled *Bst*XI fragment of λPK67. The organs analyzed here are as follows: lane 1, heart; lane 2, brain; lane 3, placenta; lane 4, lung; lane 5, liver; lane 6, skeletal muscle; lane 7, kidney; and lane 8, pancreas. For details see section 2.

(residues 1–660) showed higher homologies to these proteins (HC1 40%, HC2 39%, and HC3 47%). ITI consists of two homologous HCs and one light chain including two Kunitz-type trypsin inhibitor domains (bikunin), and these were covalently linked to a chondroitin sulfate. The amino acid sequences of the HCs around the processing site for the chondroitin sulfate binding are well conserved. Interestingly, the amino acid sequence of PK-120 corresponding to the processing site was completely different from those of the HCs.

4. Discussion

Here we have obtained two cDNA clones (λPK5 and λPK67) from human liver cDNA libraries and sequenced them. As shown in Fig. 3, an intense band at 3.3 kb was observed only in the poly(A)$^+$ RNA of liver by using a ^{32}P-labeled *Bst*XI fragment of λPK67. A faint band at 4.0 kb, which may correspond to a PK-120-related molecule, was also observed in liver. These results suggested that the composite sequence of 3,058 bp (not considering a poly(A) tail) obtained from these two clones covers almost the full-length PK-120 cDNA, and that PK-120 is mainly synthesized in liver.

The deduced amino acid sequence of PK-120 indicated that it consists of 902 amino acid residues, of which 23% were confirmed at the protein level. Through the protein sequencing, Ser^{668}, Thr^{673}, and Ser^{674} could not be identified as phen-

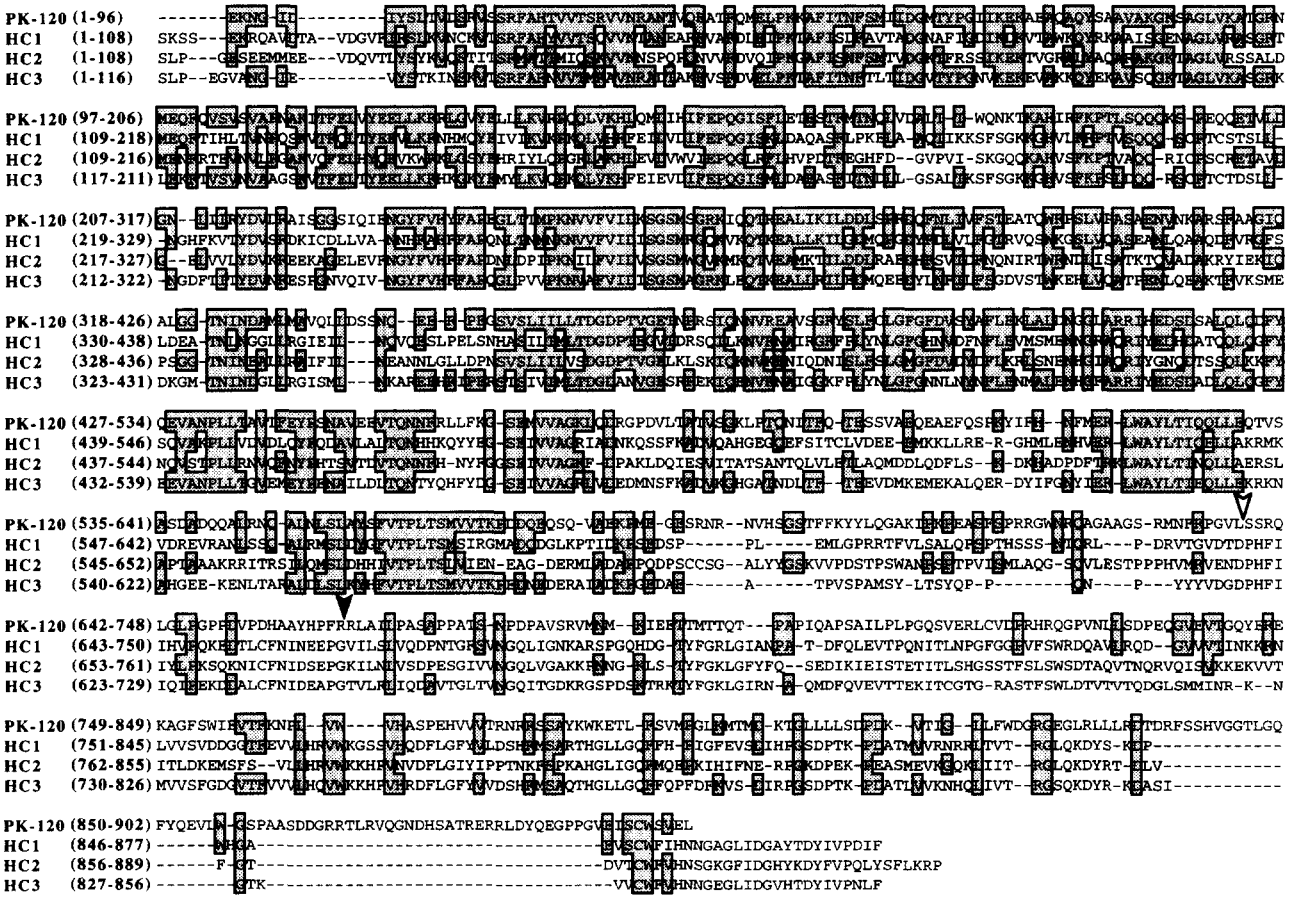


Fig. 4. Alignment of amino acid sequences of PK-120 and HCs of the ITI superfamily. Identical residues to those of PK-120 are boxed and shadowed. HC1, heavy chain 1 of human ITI [15,16]; HC2, heavy chain 2 of human ITI [16,17]; and HC3, a heavy chain of human pre-α-trypsin inhibitor [18]. A black arrowhead shows the PK-cleavage site of PK-120. An open arrowhead shows the processing site of the HCs for the chondroitin sulfate-binding.

ylthiohydantoin amino acids (Fig. 2). Since Pu et al. have suggested the existence of N- and O-linked sugar chains in PK-120 [8], these amino acid residues are probably O-glycosylated. Furthermore, the potential N-glycosylation sites were identified (Asn^{53}, Asn^{179}, Asn^{489} and Asn^{549}): therefore at least one of them seems to be *N*-glycosylated.

From the sequence data, it was found that PK-120 contains only three cysteine residues (Cys^{386}, Cys^{719} and Cys^{897}) (Fig. 2). We have carried out the experiments of the titration for the SH-group on PK-120, 70- and 35-kDa fragments (data not shown). The results suggested that Cys^{719} and Cys^{897} form a disulfide bond within the 35-kDa fragment, and that Cys^{386} is present as a free cysteine residue.

On the protein-based study of PK-120, we could not identify a 30-kDa fragment, which should be released from the carboxy-terminal portion of the 100-kDa fragment, suggesting that the *C*-terminal 30-kDa region is degraded into small fragments [8]. The amino acid sequence of PK-120 predicted the putative cleavage sites of PK (-Phe-Arg-Xaa-; Arg^{455}, Arg^{633} and Arg^{660}) (Fig. 2). As for Arg^{455}, it is unclear whether this residue is a PK cleavage site or not. It is reasonable to consider that Arg^{660} is one of the real PK cleavage sites, because the deduced amino acid sequence from Leu^{662} is identical to the N-terminal sequence of the 35-kDa fragment. In a previous study, the N-terminal residue of the 35-kDa fragment (corresponding to residue 661) was identified to be Asp [8], while the cDNA sequence data suggested that this residue is Arg. The reason for this discrepancy is unknown. Since the yield for a phenylthiohydantoin-Arg is very low on Edman degradation, residue 661 may be Arg, rather than Asp. From these results, it is likely that a peptide of 27 amino acid residues (Pro^{634}–Arg^{660}) is released from PK-120 by the PK digestion. The amino acid sequence of the carboxy-terminal portion of this peptide is similar to that of bradykinin released from high molecular weight kininogen by PK, suggesting that PK-120 may be a precursor of a bioactive molecule such as bradykinin. In fact, we could isolate this peptide from the PK digest of PK-120 (data not shown). These results suggested that Arg^{633} and Arg^{660} are the PK cleavage sites and that this peptide is released from PK-120 by the PK digestion. The functional analysis of the peptide is actively ongoing in our laboratory.

As shown in Fig. 4, a computer-assisted homology search indicated that PK-120 is closely related to HCs of the ITI superfamily. PK-120 showed high sequence homologies to HC1 (34%) and HC2 (31%) of human ITI and HC3 (38%) of human pre-α-trypsin inhibitor. Interestingly, the 100- and the 70-kDa fragments are more homologous to the HCs than PK-120. Recently, it has been reported that bovine pre-α-trypsin inhibitor (or ITI) stabilizes cumulus extracellular matrix, probably due to the binding of the protein to hyaluronic acid (HA) [12], and that the HCs of human and bovine ITIs act as HA-associated proteins [13,14]. HA is a major component of extracellular matrix, which plays modulatory roles in tissue morphogenesis, cell migration, and cell proliferation. It is likely that the 100- and/or the 70-kDa fragments, rather than PK-120, bind to HA, leading to the stabilization of extracellular matrix. In conclusion, the digestion of PK-120 by PK may be a important reaction for many fundamental biological processes as described above.

Acknowledgements: We express our thanks to Prof. M. Tomita (Showa University) for providing a human liver λgt11 cDNA library, and to Dr. F. Sakane (Sapporo Medical College) for providing poly(A)$^+$ RNA of human liver. We are also grateful to Prof. S. Iwanaga (Kyushu University) for valuable discussions, and to Drs. M. Tada (Hokkaido University) and Y. Miura (Kyushu University) for homology search.

References

[1] Furie, B. and Furie, B.C. (1988) Cell 53, 505–518.
[2] Ichinose, A., Kisiel, W. and Fujikawa, K. (1984) FEBS Lett. 175, 412–418.
[3] Discipio, R.G. (1982) Immunology 45, 587–595.
[4] Kozin, F. and Cochrane, C.G. (1988) in: Inflammation: Basic Principles and Clinical Correlates (Gallin, J.I., Goldstein, I.M. and Synderman, R. eds) pp. 101–120, Raven Press, New York.
[5] Kaplan, A.P., Kay, A.B. and Austen, K.F. (1972) J. Exp. Med. 135, 81–97.
[6] Schapira, M., Despland, E., Scott, C.F., Boxer, L.A. and Colman, R.W. (1982) J. Clin. Invest. 69, 1199–1202.
[7] Wachtfogel, Y.T., Kucich, U., James, H.L., Scott, C.F., Schapira, M., Zimmerman, M., Cohen, A.B. and Colman, R.W. (1983) J. Clin. Invest. 72, 1672–1677.
[8] Pu, P.X., Iwamoto, A., Nishimura, H. and Nagasawa, S. (1994) Biochim. Biophys. Acta (in press).
[9] Nagase, H. and Barrett, A.J. (1981) Biochem. J. 193, 187–192.
[10] Nishimura, H., Takeya, H., Miyata, T., Suehiro, K., Okamura, Y., Niho, Y. and Iwanaga, S. (1993) J. Biol. Chem. 268, 24041–24046.
[11] Kawabata, S., Miura, T., Morita, T., Kato, H., Fujikawa, K., Iwanaga, S., Takada, K., Kimura, T. and Sakakibara, S. (1988) Eur. J. Biochem. 172, 17–25.
[12] Chen, L., Mao, S.J.T. and Larsen, W.J. (1992) J. Biol. Chem. 267, 12380–12386.
[13] Yoneda, M., Suzuki, S. and Kimata, K. (1990) J. Biol. Chem. 265, 5247–5257.
[14] Huang, L., Yoneda, M. and Kimata, K. (1993) J. Biol. Chem. 268, 26725–26730.
[15] Diarra-Mehrpour, M., Bourguignon, J., Bost, F., Sesboüé, Muschio, F., Sarafan, N., and Martin, J. -P. (1992) Biochim. Biophys. Acta 1132, 114–118.
[16] Gebhard, W., Schreitmüller, T., Hochstrasser, K. and Wachter, E. (1989) Eur. J. Biochem. 181, 571–576.
[17] Gebhard, W., Schreitmüller, T., Hochstrasser, K. and Wachter, E. (1988) FEBS Lett. 229, 63–67.
[18] Bourguignon, J., Diarra-Mehrpour, M., Thiberville, L., Bost, F., Sesboüé, R. and Martin, J.-P. (1993) Eur. J. Biochem. 212, 771–776.