

# Primary Structure of the Pig Homologue of Human IHRP: Inter- $\alpha$ -Trypsin Inhibitor Family Heavy Chain-Related Protein

Ken Hashimoto,\* Takashi Tobe,\*<sup>1</sup> Jun-ichi Sumiya,\* Yoshihiro Sano,\*  
Nam-Ho Choi-Miura,\* Akihito Ozawa,<sup>†</sup> Hiroshi Yasue,<sup>†</sup> and Motowo Tomita\*

\*Department of Physiological Chemistry, School of Pharmaceutical Sciences, Showa University, 1-5-8 Hatanodai, Shinagawa-ku, Tokyo 142; and <sup>†</sup>Animal Genome Research Group, National Institute of Animal Industry, Ministry of Agriculture, Forestry and Fisheries, Tsukuba, Norindanchi, Ibaraki 305

Received for publication, November 8, 1995

The pig counterpart of human IHRP has been isolated from pig serum, and cDNA clones encoding this counterpart were isolated and characterized. The amino acid sequence of pig IHRP predicted from the nucleotide sequence of its cDNA shows reasonable homology to that of human IHRP. The nucleotide sequence of pig IHRP cDNA is identical to that of the mRNA partially determined to be the heavy chain of pig ITI [Buchman *et al.* (1990) *Surgery* 108, 560-566]; it is one of the major mRNAs induced in pig liver on cardiogenic shock. Thus we concluded that the reported mRNA should code for IHRP and not for the heavy chain of ITI. The pig IHRP also seems to be identical to pig-MAP, which was recently reported to be a major acute phase serum protein in pig [Gonzalez-Roman *et al.* (1995) *FEBS Lett.* 371, 227-230]. The results suggest that IHRP might be involved in acute phase reactions.

**Key words:** acute phase protein, ITI heavy chain, pig IHRP, primary structure.

We recently reported the purification and cloning of a novel human plasma 120-kDa glycoprotein that shows significant homology to the heavy chains of the inter- $\alpha$ -trypsin inhibitor family. We named the protein, IHRP (inter- $\alpha$ -trypsin inhibitor family heavy chain-related protein) (1). IHRP has been found to be identical to PK-120, which is highly sensitive to plasma kallikrein (2). IHRP was readily cleaved into 85 and 35 kDa fragments when plasma was incubated at 37°C. The nucleotide sequence of human IHRP cDNA showed the highest homology with the sequence which had been partially determined to be the heavy chain of pig ITI, although a protein corresponding to the nucleotide sequence has not been identified in pig plasma (3). This pig cDNA has been reported to encode one of the major mRNAs induced in the liver on cardiogenic shock by Buchman *et al.* (4). From these findings, we speculated that the pig cDNA might encode IHRP instead of the heavy chain of ITI, and the IHRP may be involved in acute phase reactions.

This paper deals with the purification of the IHRP-like protein from pig plasma and the cloning of the cDNA encoding pig IHRP. The present results clearly indicate that the pig cDNA reported by Buchman *et al.* (4) encodes pig IHRP.

## MATERIALS AND METHODS

**Reagents**—Polyethylene glycol (PEG) #4000 was purchased from Wako Pure Chemicals (Tokyo). Super-Q and

phenyl Toyopearl 650M were obtained from Tosoh (Tokyo). Bio Gel A-0.5 m was from Bio-Rad Laboratories (USA). UK-50 membrane filters were obtained from Toyo (Tokyo). Restriction enzymes were purchased from Nippon Gene, Toyobo, or Boehringer Mannheim, and used according to the manufacturer's instructions. Recombinant Taq DNA polymerase was obtained from Takara Shuzo (Kyoto). The <sup>32</sup>P-labeled nucleotides were purchased from Amersham Japan. Nitrocellulose filters were from Schleicher and Schuell. Agarose type II was obtained from Sigma Chemical, and low-melting-point agarose (NuSieve) from FMC. A random-primer labeling kit was obtained from Du Pont-New England.

**Purification of the Pig IHRP-Like Protein from Plasma**—Pig plasma (200 ml) was fractionated with PEG#4000. The precipitate between 5-15% (w/v) PEG#4000 was dissolved in 20 mM Tris-HCl (pH 7.4) containing 10 mM lysine-HCl. The dissolved sample was applied to a Super-Q column (3 × 20 cm) equilibrated with 20 mM Tris-HCl (pH 7.4) containing 10 mM lysine-HCl. The column was washed extensively with the same buffer to remove the unbound proteins, and then the bound proteins were eluted with a linear gradient of 0-0.25 M NaCl in the equilibration buffer. The fractions containing IHRP were detected by SDS-PAGE and pooled. Solid ammonium sulfate was added to the IHRP homologue pool to a final concentration of 2.0 M and the solution was applied to a column of phenyl Toyopearl 650 M (2.5 × 15 cm; Tosoh, Tokyo) which had been equilibrated with 2.0 M ammonium sulfate in 10 mM Tris-HCl, pH 7.0. After washing the column with the equilibrating buffer, the absorbed proteins were eluted with a linear gradient of 2-0 M ammonium sulfate in the same buffer. The fractions containing IHRP were concentrated to 5 ml by ultrafiltration with a UK-50 membrane (Toyo, Tokyo). The concen-

<sup>1</sup> To whom correspondence should be addressed. Tel: +81-3-3784-8215, Fax: +81-3-3784-8216, E-mail: ttobe@pharm.showa-u.ac.jp  
Abbreviations: ITI, inter- $\alpha$ -trypsin inhibitor; IHRP, inter- $\alpha$ -trypsin inhibitor family heavy chain-related protein; SDS-PAGE, SDS-polyacrylamide gel electrophoresis; PEG, polyethylene glycol.

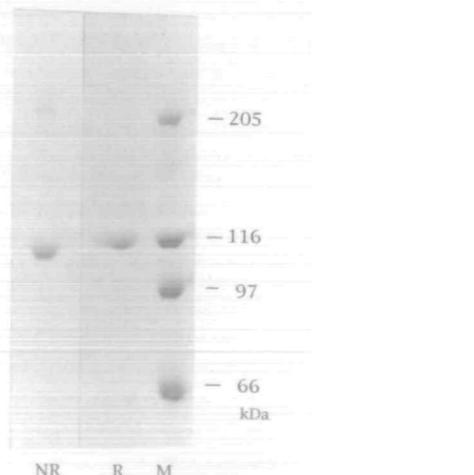
trated sample was gel-filtered on a column of Bio Gel A-0.5 m (2.5 × 95 cm; Bio-Rad) equilibrated with PBS. To remove the contaminating IgG, the fractions containing IHRP were passed through an anti-human IgG-conjugated Sepharose column (3 × 9 cm) equilibrated with PBS. Finally, 600 µg of IHRP was purified from 200 ml of pig plasma.

**SDS-PAGE Analysis**—SDS-PAGE was performed according to Laemmli using 7.5 or 10% polyacrylamide slab gels under reducing or nonreducing conditions (5). After electrophoresis, the proteins in the gels were detected by staining with 0.1% Coomassie Brilliant Blue R-250 in 10% acetic acid and 30% methanol.

**Western Blotting**—Non-reduced samples were resolved by SDS-PAGE using the method of Laemmli, and then transferred onto nitrocellulose filters. The blotted proteins were analyzed with polyclonal antibodies against human IHRP.

**Amino Acid Sequence Analysis**—The N-terminal amino acid sequence of pig IHRP was determined with an Applied Biosystems 473A gas-phase protein sequencer. The proteins transferred onto polyvinylidene difluoride membranes were also subjected to sequence analysis.

**Cloning of the Pig IHRP cDNA**—The primers (primers N and C) for PCR were synthesized based on the nucleotide sequence reported by Buchman *et al.* (4). Both primers N and C possessed the EcoRI recognition sequence (CCGAA-TTCGG) at the 5' end. QUICK-Clone™ cDNA (pig liver, Clontech Laboratories) was used as a template for amplification of the cDNA fragments. The reaction mixture (100 µl) contained 10 mM Tris-HCl, pH 8.3, 50 mM KCl, 1.5 mM MgCl<sub>2</sub>, 0.001% (w/v) gelatin, 0.5 µg cDNA, 0.2 mM each of dNTP, 2.5 units of recombinant Taq DNA polymerase, and approximately 0.5 µM N and C primers. The reaction mixture was overlaid with one drop of light mineral oil, and the reaction was carried out for 25 cycles, followed by final extension at 72°C for 5 min using a DNA thermal cycler (Perkin Elmer Cetus). Each cycle included a heat denaturation step at 94°C for 1 min, followed by annealing of the primer to the DNA at 42°C for 1 min and DNA chain extension with Taq DNA polymerase at 72°C for



**Fig. 1. SDS-PAGE of purified pig IHRP from pig plasma.** Purified pig IHRP (1 µg) was analyzed by SDS-PAGE (7.5% gel) under non-reducing (NR) and reducing (R) conditions. After electrophoresis, proteins in the gels were stained with Coomassie Brilliant Blue R-250.

2 min. The DNA fragment produced by PCR was extracted with phenol and chloroform, and then subcloned into the EcoRI site of the pUC119 plasmid vector for nucleotide sequencing and preparation of a screening probe.

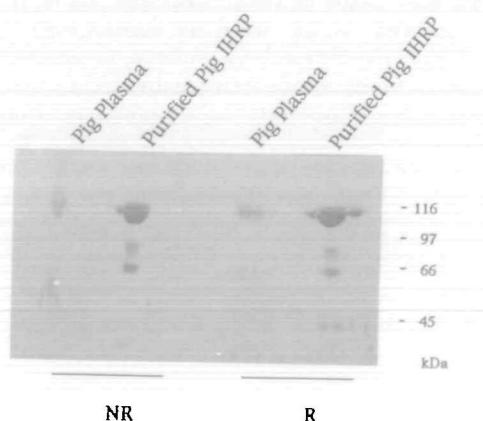
The PCR product subcloned into pUC119 was labeled with [ $\alpha$ -<sup>32</sup>P]dCTP by the random priming method (6) (Du Pont-New England Nuclear Kit) for use as a screening probe. The labeled probe was used with a conventional plaque hybridization procedure (7) to screen a normal pig liver cDNA library constructed in the  $\lambda$  ZAP phage vector. The hybridization was carried out in a solution containing 5 × saline-sodium citrate (SSC), 5 × Denhardt's, 50 mM Tris-HCl, pH 7.5, 100 µg/ml of denatured salmon sperm DNA, 0.1% SDS, 10% dextran sulfate, and 50% formamide, at 42°C, followed by washing of the membranes sequentially with 2 × SSC containing 0.1% SDS at room temperature and then with 0.1 × SSC containing 0.1% SDS at 50°C. Several rounds of screening were carried out in order to obtain positive clones.

**Characterization of Isolated Positive Clones**—Four clones ( $\lambda$ A-1,  $\lambda$ B-1,  $\lambda$ C-1, and  $\lambda$ T-I) were isolated by plaque purification and excised from the  $\lambda$  phage as cDNA inserts in the Bluescript plasmid. After retrieving the inserted DNA in the Bluescript vector, the nucleotide sequence was determined by the dideoxynucleotide chain termination method (8) with a DNA sequencer model DSQ-1000 (Shimadzu, Kyoto).

**Computer Analyses of Sequences**—A homology search, and analyses of the nucleic acid and amino acid sequences were performed using the Genetyx program (SDC, Tokyo).

## RESULTS

The IHRP-like protein from pig plasma was purified to homogeneity, as determined on SDS-PAGE (Fig. 1). The molecular mass estimated by SDS-PAGE was approximately 110 kDa. The purified protein was analyzed by Western blot analysis with a polyclonal anti-human IHRP antibody. Three immunoreactive bands of 110, 80, and 55 kDa were detected with the antibody (Fig. 2). The immunoreactive 80 and 55 kDa bands were intermediate



**Fig. 2. Western-blot analysis of purified pig IHRP with an anti-human IHRP antibody.** Purified pig IHRP (2 µg) and pig plasma were analyzed by SDS-PAGE (10% gel) under non-reducing (NR) and reducing (R) conditions. After electrophoresis, the proteins were transferred to nitrocellulose membranes, and detected with an anti-human IHRP antibody.

cleavage products, because the N-terminal amino acid sequences of the 80- and 55-kDa peptides were identical to that of the intact 110-kDa protein (Table I). The N-terminal amino acid sequence of the purified 110-kDa protein was similar to that of human IHRP, indicating this protein to be a pig homologue of human IHRP. As reported previously for human IHRP, this protein was sensitive to plasma kallikrein, also supporting that it was pig IHRP. Pig IHRP was cleaved by plasma kallikrein mainly into 55 kDa and 25 kDa peptides (Fig. 3). The N-terminal amino acid sequences of peptides generated by kallikrein were analyzed. The N-terminal amino acid sequence of the 25-kDa peptide was SVPDETS, and that of the 55 kDa one was identical to that of intact IHRP (Table I). About 600  $\mu$ g of pig IHRP was obtained from 200 ml of pig plasma.

Buchman *et al.* determined the partial nucleotide sequence of mRNA induced by cardiogenic shock in the pig (4). Although they thought that it was derived from the pig ITI heavy chain, we speculated that the pig cDNA might encode IHRP instead of the heavy chain of ITI, because of the significant homology of the pig cDNA to that of human IHRP. Thus, utilizing primers designed from the nucleotide sequence partially determined to be the pig ITI heavy chain by Buchman *et al.* (4), we performed PCR on pig liver cDNA and obtained an amplified 900 bp DNA fragment. Nucleotide sequence analysis of the fragment confirmed that it had an identical nucleotide sequence to that reported by Buchman *et al.* (4). The nucleotide sequence of the

fragment was compared to all other sequences in the GenBank and EMBL databases, and showed the highest homology with human IHRP, as reported previously (3). After subcloning of the amplified cDNA fragment into the pUC119 plasmid vector, the PCR product was used as a probe to clone a full-length cDNA from a normal pig liver cDNA library constructed in the  $\lambda$ ZAP phage vector.

Approximately  $8 \times 10^5$  recombinant clones from the cDNA library were screened with the  $^{32}$ P-labeled PCR product, and four positive clones were isolated. These four clones,  $\lambda$ A-1,  $\lambda$ B-1,  $\lambda$ C-1, and  $\lambda$ T-1, were analyzed by PCR, to determine the length of the inserted DNA, and by digestion with several restriction enzymes. These four clones overlapped one another. One of them, clone  $\lambda$ C-1, was sequenced and ascertained to possess the full-length nucleotide sequence for pig IHRP. Figure 4 shows the restriction map and sequencing strategy for clone  $\lambda$ C-1. This clone contained an insert of 2,952 bp including a poly(A) tail of 36 bases. The open reading frame begins 6 bases downstream from the 5'-end and extends for 921 amino acids to a TAG stop codon (Fig. 5). Although the nucleotide sequence surrounding the first in-frame initia-

TABLE I. Partial amino acid sequence of pig IHRP determined by protein sequencing. 110 kDa: intact purified pig IHRP; 80 kDa peptide: immuno reactive peptide in Fig. 2; 55 kDa peptide: peptide generated by kallikrein in Fig. 3; 25 kDa peptide: peptide generated by kallikrein in Fig. 3; 22 kDa peptide: peptide generated by kallikrein in Fig. 3. All peptides, except the intact protein, were separated on a 10% polyacrylamide gel, blotted onto a PVDF membrane, and then analyzed with a gas phase protein sequencer, ABI model 473A.

Peptide or protein	N-terminal amino acid sequence
110 kDa (intact protein)	HKNDINIYS
80 kDa peptide	HKNDINIYS
55 kDa peptide	HKNDINIYS
25 kDa peptide	SVPDETS
22 kDa peptide	GATIPPPPARIQAA

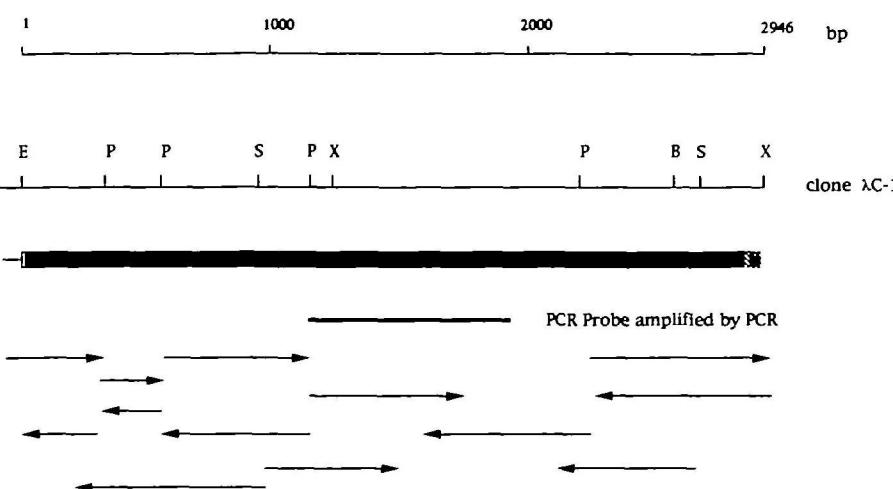


Fig. 3 SDS-PAGE analysis of plasma kallikrein digests of pig IHRP. The purified pig IHRP (2  $\mu$ g) was digested with plasma kallikrein. The digests were analyzed by SDS-PAGE (10%) under non-reducing (NR) and reducing (R) conditions, and then stained with Coomassie Brilliant Blue R-250

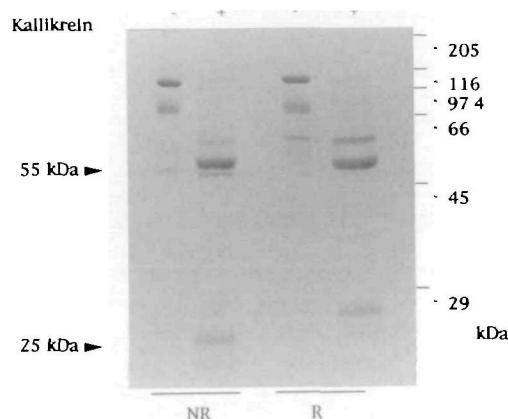
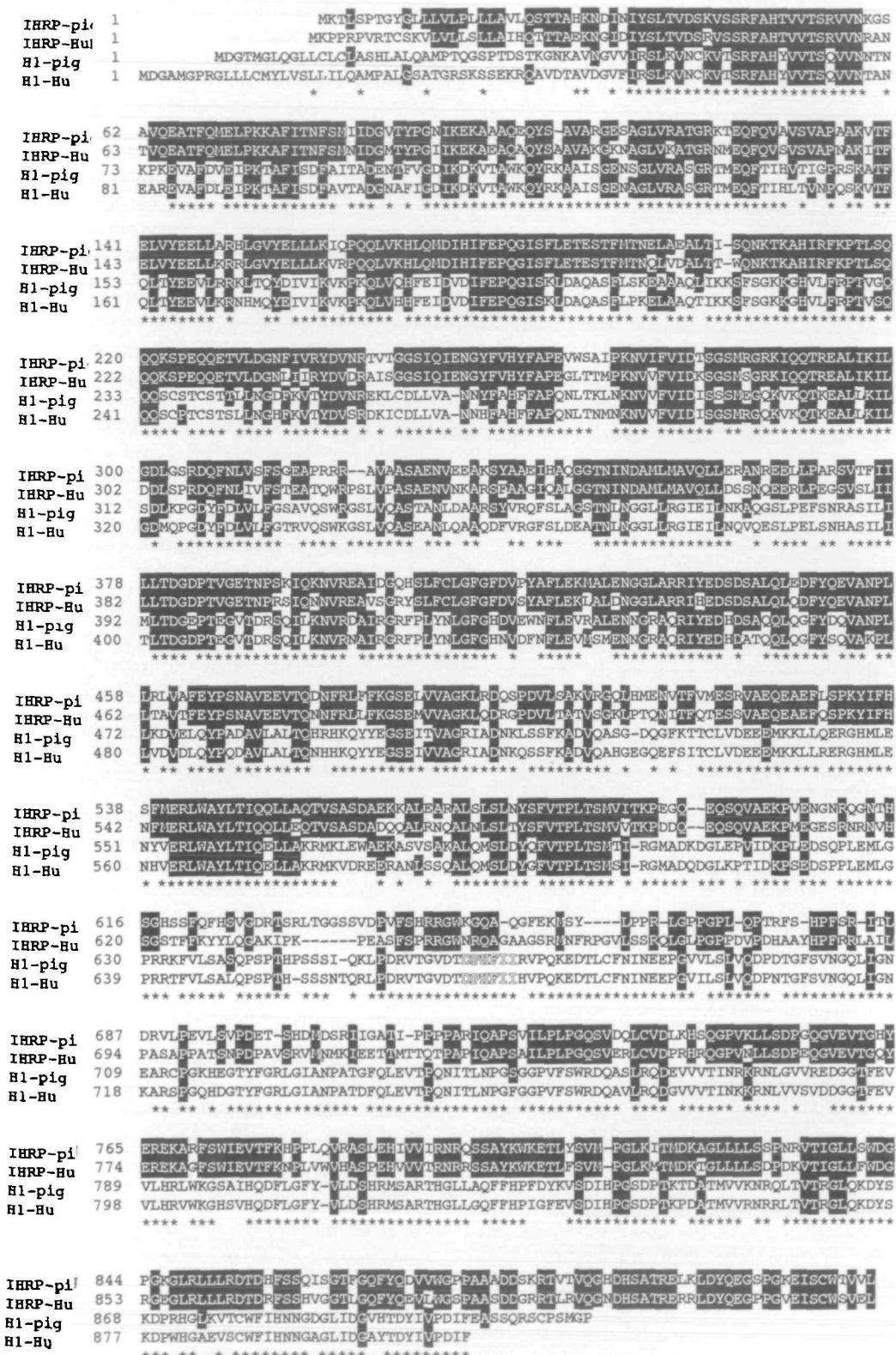


Fig. 4. Schematic representation of the restriction map, PCR probe, cDNA clones, and sequence strategy for pig IHRP cDNA. The closed bar designates the IHRP-coding sequence oriented in the 5' to the 3' direction. The nucleotide scale is indicated at the top. The PCR product which was used for screening the pig cDNA library and the cDNA clone isolated from the pig liver cDNA library are aligned with the restriction map. The sequence strategy used for determination of the pig IHRP cDNA sequence is shown as arrows. Horizontal arrows indicate the direction and extent of the sequencing. The shaded box indicates the poly(A) tail.

	GCCAA	-1
ATGAAGACCCCTCCTCCCTACTGGCTACGGCCTTCTGGTCTGCCCTGCTGGCTGCTTCAGAGCACCACGGCCCACAAGAATGACATC	96	
<u>M K T L S P T G Y G L L L V L P L L A V L Q S T T A</u> <u>E K N D I</u>	32	
AAACATCTACAGTCTACCGTGGACTCCAAGGTCTCGCCGATTTGCCACACAGTCGTTACAGGCCAGTGGTCAACAAGGGCAGTGTGCAAG	192	
<u>N I Y S L T V D S K V S S R F A H T V V T S R V V N K G S A V Q</u>	64	
GAGGCCACCTCCAGATGGAGCTGCCAAGAAGGCTTCATCACCAACTCTCCATGATCATGATGGTGTGACCTACCCAGGTAACATCAAGGAG	288	
<u>E A T F Q M E L P K K A F I T N F S M I I D G V T Y P G N I K E</u>	96	
AAAGGCTGAGCCCAGGAGCAGTACAGGCCGTGCCAGGGAGAGAGTCGCTGGCCTTGTCAAGGGCACTGGGAGAAAGACAGAGCAGTCCAGGTG	384	
<u>K A A A Q E Q Y S A V A R G E S A G L V R A T G R K T E Q F Q V</u>	128	
GCAGTCAGCGTGGCTCTGCTGCCAAGGTACCCCTCGAGCTGGTGTATGAGGAGCTCTGGCAGGGCATCTGGAGTTATGAGCTGCTGTTGAAA	480	
<u>A V S V A P A A K V T F F E L V Y E E L L A R H L G V Y E L L L K</u>	160	
ATCCAGGCCAGCAGCTGGTCAAGCACCTGCAGATGGACATTACATCTCGAGGCCCTAGGGCATCAGCTTCTGGAGACAGAGGACACCTTATG	576	
<u>I Q P Q Q L V K H L Q M D I H I F E P Q G I S F L E T E S T F M</u>	192	
ACCAATGAACGGCAGAGGCCCTCACCATATCACAGAACAGACTAACAGCTCACATCCGATTCAAGGCCACACTCTCCAGCAGCAGAACAGTCCCCA	672	
<u>T N E L A E A L T I S Q N K T K A H I R F K P T L S Q Q Q K S P</u>	224	
GAGCAGCAGGAAACAGTCCTGGATGGCAACTTCATGTCGCTATGATGTGAAACGGCAGACTCACTGGGGTTCCATTAGATCGAGAACATGGCTAC	768	
<u>E Q Q E T V L D G N F I V R Y D V N R T V T G G S I Q I E N G Y</u>	256	
TTTGTGCAATTACTTTGCCCGGGAGGCTGGTCTGCAATACCAAGAACGTGATTTGTCAATTGACACGGAGGGCTCCATGAGGGCAGGAAATC	864	
<u>F V H Y F A P E V W S A I P K N V I F V I D T S G S M R G R K I</u>	288	
CAGCAGAGCCGGGAAGGCCATAATTAGATCTGGTGAACCTCGGCTCCCGCAGGGCTCACCTTGTCAAGCTTCAGTGGGAGAACACCCAGGAGA	960	
<u>Q Q T R E A L I K I L G D L G S R D Q F N L V S F S G E A P R R</u>	320	
AGGGCTGTGCAAGCTCAGCTGAGAACGAGCTATGCTGCCGAATTCATGCCAGGGAGGACCAATATAATGATGATGATG	1056	
<u>R A V A A S A E N V E E A K S Y A A E I H A Q G G T N I N D A M</u>	352	
CTGATGGCGTGCAGCTGGAAAAGAGCCAACCGGGAGGAGCTGCTGCCCGCAGGGAGCTTACCTCATCATCCCTCACCGATGGCGACCC	1152	
<u>L M A V Q L L E R A N R E E L L P A R S V T F I I L L T D G D P</u>	384	
ACTGTGGGGAGACCAACCCCTCGAAGATCCAGAACAGTCGGGAAGCCATAGAGGCCAGCACGCCCTTCTGCCCTGGCTTCGGCTTGTAC	1248	
<u>T V G E T N P S K I Q K N V R E A I D G Q H S L F C L G F G F D</u>	416	
GTCCCCCTACGCCCTTCTGGAGAACAGTGGCACTGGAGAACATGGCGCTGGCCGGCATCTATGAGGACTCTGACTCTGCCCTGCACTCGAGGAC	1344	
<u>V P Y A F L E K M A L E N G G L A R R I Y E D D S D A L Q L E D</u>	448	
TTCTTACCAAGGGTGGCCAACCCACTGCTGAGGTGGCTTGAGTACCCAAGCAATGCCGTGGAGGAGTCACCGAGAACACTCCGGCTG	1440	
<u>F Y Q E V A N P L L R L V A F E Y P S N A V E E V T Q D N F R L</u>	480	
TTCTTCAAAAGGCTCTGAGTTGGTGGCCGGAAAGCTCCGGGACAGAGCCCCGATGTGCTCTCCGCCAAAGTCAGGGGCAACTGCACATGGAG	1536	
<u>F F K G S E L V V A G K L R D Q S P D V L S A K V R G Q L H M E</u>	512	
AAATGTCACCTTCGTAATGGAGTCCAGGGTAGCAGAGCAGGAGGGCTGAGCCCCAAGTACATCTTCCACAGCTCATGGAGAGACTCTGG	1632	
<u>N V T F V M E S R V A E Q E A E F L S P K Y I F H S F M E R L W</u>	544	
GCATACCTGACCATCCAGCAACTGTTGGCGCAAACAGTCCTGCGCAGATGCTGAGAACAGGCCCTGAAGGCCAGCCCTGAGCTGTCACTC	1728	
<u>A Y L T I Q Q L L A Q T V S A S D A E K K A L E A R A L S S L</u>	576	
AACTACACCTTGTCAACCCCTCTCACATCTATGCTGATCACCAACCTGAGGCCAAAGAACAGTCAGGTTGCTGAGAACAGCCGTGAAAATGGA	1824	
<u>N Y S F V T P L M V I T K P E G Q E Q S Q V A E K P V E N G</u>	608	
AAACGACAGGGAAACACCCACTCAGGTCACTTCTCTCAATTCTCATGGAGAACAGAACATCCAGACTAACAGGAGGAGCAGTGTAGAC	1920	
<u>N R Q G N T H S G H S F Q F H S V G D R T S R L T G G S S V D</u>	640	
CCTGTTTTCTCACAGAAGGGCTGGAAAGGACAAGGCCAAAGGTTTGAGAAAATGCTTACCTCCACCGAGATTAGGACCCCCCGGACCTCTT	2016	
<u>P V F S H R R G W K G Q A Q G F E K M S Y L P P R L G P P G P L</u>	672	
CAGCCTACTCGTTTCTCATCCGTTCCGTATAACCTGGACCGGGTGTGCCAGAGGTGCTATCCGTTCTGATGAAACAGTCACATGACATG	2112	
<u>Q P T R F S H P F S R I T L D R V L P E V L S V P D E T S H D M</u>	704	
GATTCAAGAACATAGGAGCCACGATACCTCCACCCCTGCCGATCCAGGCTCTCCGTATCTGCCACTGCCAGGGTGGACCAG	2208	
<u>D S R I I G A T I P P P P A R I Q A P S V I L P L P G Q S V D Q</u>	736	
CTCTGTGTGGACCTCAAGCACTCCCAGGGCCAGTGAAGCTGCTCTGAGACCCCTGCCAAGGGGTTGAAGTGAATGCCACTATGAGAGGGAGAAG	2304	
<u>L C V D L K H S Q G P V K L L S D P G Q G V E V T G H Y E R E K</u>	768	
GCCCCGCTCTCATGGATGAGTCACCTCAAGCACCCGCCACTGCAAGGTTGCTGCATCCCTGGACCATAGTAGTAGTATTGGAAACGCCAAAGC	2400	
<u>A A R F S W I E V T F K H P P L Q V R A S L E H I V V I R N R Q S</u>	800	
TCTGCGCTAACAGTGGAGGAAACACTCTACTCACTGATGCCGGCTCAAGATAACCATGGACAAGGCCACTTCTCTGCTCAGCAGGCCAAAC	2496	
<u>S A Y K W K E T L Y S V M P G L K I T M D K A G L L L S S P N</u>	832	
AGAGTGACCATCGGCCCTGCTCTGGATGCCAGGGCTCCGACTCTCTGCGGGACACTGACCAACTCTCCAGCCAGATCAGTGGG	2592	
<u>R V T I G L L S W D G P G K G L R L L L R D T D H F S S Q I S G</u>	864	
ACCTTGGCCAGTTTACCAAGGACGTTGGCTGGGGACCCCCAGCAGCAGATGACAGCAAGCGAACAGTGCACAGTCAGGGACATGACACTCT	2688	
<u>T F G Q F Y Q D V V W G P P A A A D D S K R T V T V Q G H D H S</u>	896	
GCCACCAAGAGAGCTAACAGCTGGATTACCAAGGGATCCCCGGGAAAGAGATTCTCTGCTGGACTGTGGTGTAGTTCTGATGGGAGGAGTTA	2784	
<u>A T R E L K L D Y Q E G S P G K E I S C W T V V L ***</u>	921	
CACCCGCCCTCATGCTGCCCTTGGCAGATGGCTGCCACACTGTAACACAGGTCAAGCTGTGGGCCCTGGAAACATCATGGGAGATGTATT	2880	
<u>TCACTCATAAAATAAGAGAGGTGATGTGAAAAA</u>	2946	

Fig. 5. Nucleotide sequence of IHRP cDNA and the amino acid sequence predicted from the nucleotide sequence. The nucleotide sequence of the constructed cDNA for IHRP is shown on the first line, and the amino acid sequence predicted from the nucleotide sequence is shown on the second line. The amino acid sequences which were confirmed chemically are double underlined. Nucleotides and amino

acids are numbered on the right. Nucleotide residues are numbered from +1 at the first base of the putative initiation codon (methionine). A polyadenylation signal (AATAAA) is indicated by a double underline. The potential *N*-glycosylation sites (-Asn-Xaa-Ser/Thr-) are shown by a single underline. Histidine at the N-terminus of the mature protein is shown as an outlined character.



**Fig. 6 Comparison of the amino acid sequence of pig IHRP with that of human IHRP or pig ITI heavy chain.** Sequences were aligned using the Genetyx program (SDC, Tokyo). Amino acid residues identical with ones in pig IHRP are shown by black boxes. The amino acid sequence shown by the outlined letters in H1 represents the signal for processing to form a complex with a bikunin. The

asterisks represent the identical amino acid residues between the pig and human ITI heavy chains. IHRP-pi, pig inter- $\alpha$ -trypsin inhibitor family heavy chain-related protein; IHRP-Hu, human inter- $\alpha$ -trypsin inhibitor family heavy chain-related protein, H1-pig, pig inter- $\alpha$ -trypsin inhibitor heavy chain 1.

TABLE II. Homologies of nucleotide and amino acid sequences. IHRP: inter- $\alpha$ -trypsin inhibitor family heavy chain-related protein (3); H1: inter- $\alpha$ -trypsin inhibitor family heavy chain 1 (23); H2: inter- $\alpha$ -trypsin inhibitor family heavy chain 2 (24); H3: inter- $\alpha$ -trypsin inhibitor family heavy chain 3 (25).

		Human				Pig		Amino Acid Sequence Homology (%)
		IHRP	H1	H2	H3	IHRP	H1	
Human	IHRP	100	42	42	52	71	34	
	H1	55	100	39	55	34	80	
	H2	49	52	100	40	34	38	
	H3	57	59	48	100	39	52	
Pig	IHRP	78	55	50	58	100	35	
	H1	55	83	52	64	55	100	

Nucleotide Sequence Homology

tion codon at nucleotides 1-3, CCAAAATGA, does not match well the consensus sequence, CCACCATGG, described by Kozak (9, 10), this codon appears to be the most likely candidate for the translation start site because the methionine was followed by 27 hydrophobic amino acid residues which may be a signal peptide sequence. The length of the hydrophobic amino acid stretch was suitable for a signal peptide. In the 3' non-coding region, a consensus polyadenylation signal, AATAAA (11), and a 3' poly(A) tail are positioned 19 nucleotides upstream of the first A of the poly(A) tract and 3' end, respectively. The open reading frame encodes the putative signal sequence of 27 amino acid residues and a mature protein of 894 amino acid residues whose N-terminal 9 residues were confirmed by protein sequencing (Fig. 5). The amino acid sequences determined by protein sequencing were all contained in this reading frame, indicating that the  $\lambda$  T-1 clone codes pig IHRP (Table I and Fig. 6). The molecular mass of the mature protein without carbohydrate was calculated to be 99,348 Da.

The  $\lambda$  T-1 clone contained an extra nucleotide sequence in its 3' region, as compared with the other clones. This extra nucleotide sequence in this clone may be derived from an intron, because the sequences are in good agreement with the GT-AG rule (12), and all three reading frames predicted from this extra nucleotide sequence were found to contain several stop codons (data not shown).

The nucleotide sequence of the full length pig IHRP cDNA contains the nucleotide sequence partially determined to be the pig ITI heavy chain, and exhibits considerable homology with those of pig ITI heavy chain and human IHRP; the nucleotide sequence of pig IHRP cDNA showed 78, 55, 55, 50, and 58% homology to those of human IHRP, pig ITI-H1, and human ITI-H1, -H2, and -H3, respectively. These homologies are summarized in Table II.

## DISCUSSION

In this study, we purified the pig counterpart of human inter- $\alpha$ -trypsin inhibitor family heavy chain-related protein

(IHRP) from pig plasma, and isolated four independent cDNA clones encoding pig IHRP from a normal pig liver cDNA library constructed in the  $\lambda$  ZAP phage vector. The purified protein was immunoreactive to an anti-human IHRP antibody and sensitive to plasma kallikrein, suggesting that the protein is the pig counterpart of human IHRP. In addition, the molecular weight of this protein estimated by SDS-PAGE was similar to that of human IHRP (120 kDa) (1, 2). The molecular weight of pig IHRP calculated from the amino acid sequence predicted from its cDNA is slightly smaller than that estimated by SDS-PAGE, the difference perhaps being caused by glycosylation. There are five potential N-glycosylation sites, namely Asn<sub>60</sub>, Asn<sub>205</sub>, Asn<sub>242</sub>, Asn<sub>513</sub>, and Asn<sub>577</sub>, which are conserved in human IHRP with the exception of Asn<sub>242</sub> (Fig. 5). The N-terminal amino acid sequence of this protein, HKNDINIYSL, is highly homologous to that of human IHRP, EKNGIDIYSL (3). We therefore concluded that this protein was pig IHRP. Like human IHRP, this protein was highly sensitive to plasma kallikrein, which cleaved it into 55 kDa and 25 kDa peptides (Fig. 3). Only one putative cleavage site for plasma kallikrein (Phe<sub>449</sub>-Arg<sub>450</sub>-Leu<sub>451</sub>) is present in the amino acid sequence predicted from the nucleotide sequence of pig IHRP cDNA (Fig. 5). This site corresponds to the cleavage site closest to the N terminus among the three putative cleavage sites of human IHRP. The molecular masses of the two peptides predicted on the cleavage at this site were calculated to be 50,359 and 49,000 Da, respectively. The 50,359 Da peptide should correspond to the 55 kDa peptide generated from pig IHRP with plasma kallikrein (Fig. 3), because its N-terminal amino acid sequence was identical to that of the intact protein. The 25 kDa peptide, which was also generated on kallikrein digestion (Fig. 3), is considerably smaller than the expected molecular mass of 49 kDa for the remaining C-terminal fragment. This may have been caused by proteolytic trimming of the N-terminal region of the 49 kDa peptide, because the N-terminal amino acid sequence of the 25 kDa peptide corresponds to the sequence, Ser<sub>695</sub>-Val<sub>696</sub>-Pro<sub>697</sub>-Asp<sub>698</sub>-Glu<sub>699</sub>-Thr<sub>700</sub>-, which is about 200-amino-acid downstream from the putative cleavage site. The putative cleavage site for kallikrein could not be found in the N-terminal region of the 25 kDa peptide. Therefore, further work is needed to determine whether or not this site is a real plasma kallikrein cleavage site.

About 600  $\mu$ g of purified pig IHRP was obtained from 200 ml of pig plasma. The yield of pig IHRP was thus very low, the concentration of human IHRP being about 80  $\mu$ g/ml of plasma (2). The low yield may have resulted from the degradation of IHRP during the purification, because IHRP is highly susceptible to proteolysis. As shown in Fig. 2, the purified pig IHRP preparation contained 80- and 55-kDa immunoreactive peptides, which possessed an identical N-terminal amino acid sequence to that of intact pig IHRP, while only the 110-kDa protein was detected on Western blot analysis in pig plasma.

The nucleotide sequence of pig IHRP cDNA contains an identical nucleotide sequence to that partially determined by Buchman *et al.* to be the pig ITI heavy chain (4). It exhibits considerably higher homology with that of human IHRP than with that of either pig ITI family heavy chain or any of the human ITI family heavy chains (Table II). As shown in Fig. 5, the nucleotide sequence of the cDNA

contained a full open reading frame, starting at nucleotides 1-3 of an initiation codon and terminating at nucleotides 2764-2766 of a stop codon (TAG). The 2,763-bp open reading frame codes for a polypeptide of 921 amino acid residues. The amino acid sequence of pig IHRP predicted from this nucleotide sequence shows reasonable homology to that of human IHRP (Fig. 6). A comparison of the pig IHRP amino acid sequence with that of pig ITI heavy chain is also shown in Fig. 6. It is evident that there is a similarity between pig IHRP and pig ITI heavy chain, but the amino acid sequence of pig IHRP exhibits much greater homology with that of human IHRP than with that of pig ITI heavy chain reported by Suzuki *et al.* (unpublished; accession number, D38754). As we previously reported for human IHRP, similarity between pig IHRP and pig ITI heavy chain is more noticeable in the N-terminal two-thirds of the protein. The sequences conserved in this region of human ITI-heavy chains are well conserved in the same regions of pig IHRP and pig ITI-heavy chain. In the conserved region, the heavy chains of human ITI have been reported to contain three homologous sequences, a calcium binding site (13), the reactive site of human  $\alpha_2$ -thiol-protease inhibitor (13), and von Willebrand factor type A (14), which correspond to residues Asp<sub>381</sub>-Gly<sub>387</sub>, Gly<sub>484</sub>-Gly<sub>491</sub>, and Thr<sub>244</sub>-Asn<sub>468</sub>, respectively, in pig IHRP. As we reported for human IHRP (3), pig IHRP also exhibits no significant homology with pig ITI heavy chain or any of the heavy chains of the human ITI family in the C-terminal 300 amino acids. In contrast, pig IHRP revealed high homology throughout the entire sequence with human IHRP, with the exception of the region from Val<sub>505</sub> to Pro<sub>718</sub>. This region, interestingly, is rich in proline and relatively hydrophobic, suggesting that it may be exposed to the environment in a random coil structure, so that many proteases can readily cleave this region. Indeed, no large fragments containing this region were obtained on kallikrein digestion.

Three members of the inter- $\alpha$ -trypsin inhibitor (ITI) family (serine protease inhibitors) have been purified from human plasma. One of them is called inter- $\alpha$ -trypsin inhibitor, consisting of heavy chain 1 (HC1), heavy chain 2 (HC2), and bikunin (15). Another is called pre- $\alpha$ -trypsin inhibitor, consisting of heavy chain 3 (HC3) and bikunin (16). The other is a complex of HC2 and bikunin (17). In all of them, the peptide chains (HC1, HC2, HC3, and bikunin) are covalently linked to each other through an unusual structure. The interchain linkage is mediated through esterification of the  $\alpha$ -carbon of the C-terminal Asp of a heavy chain with C6 of an internal N-acetylgalactosamine of the glycosaminoglycan chain O-linked to Ser<sub>10</sub> of the bikunin (18). All heavy chains are synthesized as precursor forms, and processed to remove the C-terminal portion. The resulting C-terminal aspartic acid contributes to the formation of a complex with bikunin.

Before or during the linkage processing, a C-terminal peptide with a calculated molecular weight of about 25 kDa is released from the heavy chain precursor (17-19). Although the N-terminal two-thirds of the amino acid sequence of IHRP shows significant homology to those of ITI heavy chains, there is no significant homology in the C-terminal one-third of the amino acid sequence. In particular, IHRP does not contain the sequence, DPHFII, which is thought to be a signal for the linkage processing, therefore, IHRP seems not to form a complex with bikunin.

The functions of the various proteins belonging to the ITI family are not precisely known yet. In particular, the function of heavy chains of ITI has not been elucidated yet. A 150 kDa protein, containing a partial peptide sequence of bikunin, has been reported to stabilize the cumulus extracellular matrix (20, 21). The members of the ITI family described above comprise complexes with bikunin or bikunin itself. The extracellular matrix-stabilizing activity therefore may be due to bikunin in the complex or free bikunin. Possibly, the heavy chains of the ITI family regulate the protease inhibitor activity of bikunin. However, IHRP is likely to have a unique function by itself, because it does not form a complex with bikunin. Furthermore, no ITI heavy chains have been reported to be free in the plasma.

The present report clearly indicates that one of the major mRNAs induced in the pig liver on cardiogenic shock (4) codes for IHRP and not for the heavy chain of ITI. Thus, IHRP may be involved in acute phase reactions. Acute phase reactants, *e.g.*, procoagulant fibrinogen and plasma proteinase inhibitors, are proteins synthesized in the liver and secreted into the bloodstream, which, as a teleological group, protect the organism against either the cause of the shock or the consequences of resuscitation. Production in the liver, secretion into the bloodstream, and possession of the consensus sequence of thiolprotease inhibitors also support that IHRP may be an acute phase reactant. The homology of the nucleotide sequence of human IHRP to that of human ATP-dependent protease also implies that IHRPs may be involved in acute phase reactions, because the ATP-dependent protease is induced by heat shock.

During the preparation of this manuscript, Gonzalez-Ramon *et al.* reported the isolation of pig-MAP, which is a major acute phase serum protein in the pig (22). A partial N-terminal amino acid sequence of pig-MAP is identical to that of pig IHRP reported here with the exception of Ser<sub>22</sub> and Lys<sub>23</sub>. Although the reason for this discrepancy is not known, pig-MAP may be pig IHRP.

Hence, IHRP seems to be the major protein induced in the acute phase in pigs, and the function of IHRP as an acute phase reactant remains to be elucidated.

## REFERENCES

1. Choi-Miura, N.-H., Sano, Y., Oda, E., Nakano, Y., Tobe, T., Yanagishita, T., Taniyama, M., Katagiri, T., and Tomita, M. (1995) Purification and characterization of a novel glycoprotein which has significant homology to heavy chains of inter- $\alpha$ -trypsin inhibitor family from human plasma. *J. Biochem.* 117, 400-407
2. Nishimura, H., Kakizaki, I., Muta, T., Sasaki, N., Pu, X.P., Yamashita, T., and Nagasawa, S. (1995) cDNA and deduced amino acid sequence of human PK-120, a plasma kallikrein-sensitive glycoprotein. *FEBS Lett.* 357, 207-211
3. Saguchi, K., Tobe, T., Hashimoto, K., Sano, Y., Nakano, Y., Choi-Miura, N.-H., and Tomita, M. (1995) Cloning and characterization of cDNA for inter- $\alpha$ -trypsin inhibitor family heavy chain-related protein (IHRP), a novel human plasma glycoprotein. *J. Biochem.* 117, 14-18
4. Buchman, T.G., Cabin, D.E., Vickers, S., Deutszman, C.S., Delgado, E., Sussman, M.M., and Bulkley, G.B. (1990) Molecular biology of circulatory shock. Part II. Expression of four groups of hepatic genes is enhanced after resuscitation from cardiogenic shock. *Surgery* 108, 560-566
5. Laemmli, U.K. (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* 227, 680-685

6. Feinberg, A.P. and Vogelstein, B. (1983) A technique for radio-labeling DNA restriction endonuclease fragments to high specific activity. *Anal. Biochem.* **132**, 6-13
7. Maniatis, T., Fritsch, E.F., and Sambrook, J. (1989) *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY
8. Sanger, F., Nicklen, S., and Coulson, A.R. (1977) DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. USA* **77**, 5463-5467
9. Kozak, M. (1986) Point mutations define a sequence flanking the AUG initiation codon that modulates translation by eukaryotic ribosomes. *Cell* **44**, 283-292
10. Kozak, M. (1987) An analysis of 5'-noncoding sequence from 699 vertebrate messenger RNAs. *Nucleic Acids Res.* **15**, 8125-8133
11. Wickens, M. and Stephenson, P. (1984) Role of the conserved AAUAAA sequence: Four AAUAAA point mutants prevent messenger RNA 3' end formation. *Science* **226**, 1045-1051
12. Stephen, M.M. (1982) A catalogue of splice junction sequences. *Nucleic Acids Res.* **10**, 459-472
13. Salier, J.P., Diarra-Mehrpor, M., Sesboue, R., Bourguignon, J., Benarous, R., Ohkubo, I., Kurachi, K., and Martin, J.P. (1987) Isolation and characterization of cDNAs encoding the heavy chain of human inter-alpha-trypsin inhibitor (I $\alpha$ TI): Unambiguous evidence for multipolypeptide chain structure of I $\alpha$ TI. *Proc. Natl. Acad. Sci. USA* **84**, 8272-8276
14. Perkins, S.J., Smith, K.F., Williams, S.C., Haris, P.I., Chapman, D., and Sim, R.B. (1994) The secondary structure of the von Willebrand factor type A domain in factor B of human complement by Fourier transform infrared spectroscopy. *J. Mol. Biol.* **238**, 104-119
15. Jessen, T., Faarvagn, K.L., and Plong, M. (1988) Carbohydrate as covalent crosslink in human inter-alpha-trypsin inhibitor: A novel plasma protein structure. *FEBS Lett.* **230**, 195-200
16. Enghild, J.J., Thogersen, I.B., Pizzo, S.V., and Salvesen, G. (1989) Analysis of inter-alpha-trypsin inhibitor and a novel trypsin inhibitor, pre-alpha-trypsin inhibitor, from human plasma. *J. Biol. Chem.* **264**, 15975-15981
17. Enghild, J.J., Salvesen, G., Thogersen, I.B., Valnickova, Z., Pizzo, S.V., and Hefta, S.A. (1993) Presence of the protein-glycosaminoglycan-protein covalent cross-link in the inter-alpha-inhibitor-related proteinase inhibitor heavy chain2/bikunin. *J. Biol. Chem.* **268**, 8711-8716
18. Enghild, J.J., Salvesen, G., Hefta, S.A., Thogersen, I.B., Rutherford, S., and Pizzo, S.V. (1991) Chondroitin 4-sulfate covalently cross-links the chain of the human blood protein pre- $\alpha$ -inhibitor. *J. Biol. Chem.* **266**, 747-751
19. Morelle, W., Capon, C., Balduyck, M., Sautiere, P., Kouch, M., Michalski, C., Fournet, B., and Mizon, J. (1994) Chondroitin sulfate covalently cross-links the three polypeptide chains of inter-alpha-trypsin inhibitor. *Eur. J. Biochem.* **221**, 881-888
20. Chen, L., Mao, S.J.T., and Larsen, W.J. (1992) Identification of a factor in fetal bovine serum that stabilizes the cumulus extracellular matrix: A role for a member of the inter-alpha-trypsin inhibitor family. *J. Biol. Chem.* **267**, 12380-12386
21. Castillo, G.M. and Templeton, D.M. (1993) Subunit structure of bovine ESF (extracellular-matrix stabilizing factors). *FEBS Lett.* **318**, 292-296
22. Gonzalez-Ramon, N., Alava, M.A., Sarsa, J.A., Pineiro, M., Escartin, A., Garcia-Gil, A., Lampreave, F., and Pineiro, A. (1995) The major acute phase serum protein in pigs is homologous to human plasma kallikrein sensitive PK-120. *FEBS Lett.* **371**, 227-230
23. Diarra-Mehrpor, M., Bourguignon, J., Bost, F., Sesboue, R., Muschio, F., Sarafan, N., and Martin, J.-P. (1992) Human inter- $\alpha$ -trypsin inhibitor: Full-length cDNA sequence of the heavy chain H1. *Biochim. Biophys. Acta* **1132**, 114-118
24. Gebhard, W., Schreitmuller, T., Hochstrasser, K., and Wachter, E. (1988) Complementary DNA and derived amino acid sequence of the precursor of one of the three protein components of the inter- $\alpha$ -trypsin inhibitor complex. *FEBS Lett.* **229**, 63-67
25. Bourguignon, J., Diarra-Mehrpor, M., Thiberville, L., Bost, F., Sesboue, R., and Martin, J.-P. (1993) Human pre- $\alpha$ -trypsin inhibitor-precursor heavy chain cDNA and deduced amino-acid sequence. *Eur. J. Biochem.* **212**, 771-776