J. Biochem. 119, 1157-1165 (1996)

Purification and Characterization of a Novel Hyaluronan-Binding Protein (PHBP) from Human Plasma: It Has Three EGF, a Kringle and a Serine Protease Domain, Similar to Hepatocyte Growth Factor Activator[1]

Nam-Ho Choi-Miura,*[,2] Takashi Tobe,* Jun-ichi Sumiya,* Yasuko Nakano,* Yoshihiro Sano,* Toshio Mazda,† and Motowo Tomita*

*Department of Physiological Chemistry, School of Pharmaceutical Sciences, Showa University, 1-5-8 Hatanodai, Shinagawa-ku, Tokyo 142; and †Japanese Red Cross Tokyo Blood Center, Musashino, Tokyo 180

Received for publication, January 30, 1996

A novel hyaluronan-binding protein (PHBP) was purified from human plasma by affinity chromatography on hyaluronan-conjugated Sepharose. The contaminating IgM and albumin in the partially purified preparation were removed with anti-IgG antibody-conjugated Sepharose and anti-albumin antibody-conjugated Sepharose, respectively, and no other contaminant was observed. Finally, 800 μg of PHBP was isolated from 500 ml of human plasma. PHBP gave a single 70-kDa band on SDS-PAGE under non-reducing conditions, and 50-kDa and 17-kDa bands under reducing conditions. Thus, PHBP was a heterodimer composed of 50-kDa and 17-kDa subunits, bridged by a disulfide linkage. Both subunits had novel N-terminal amino acid sequences, indicating that PHBP was a novel hyaluronan-binding protein in human plasma. The amino acid sequence deduced from the nucleotide sequence of the cloned PHBP cDNA exhibited significant homology to that of hepatocyte growth factor activator (HGFA). The results of Northern blot analysis indicated that liver, kidney, and pancreas expressed PHBP mRNA. The predicted structure of PHBP showed three epidermal growth factor (EGF) domains, a kringle domain and a serine protease domain, from its N-terminus, although HGFA has a fibronectin type II domain, an EGF domain, a fibronectin type I domain, an EGF domain, a kringle domain, and a serine protease domain, from its N-terminus.

Key words: hepatocyte growth factor, hyaluronic acid, macrophage-stimulating protein, plasma protein, serine protease.

Hyaluronic acid (HA), a high molecular weight polysaccharide composed of repeating *N*-acetyl glucosamine and D-glucuronic acid subunits, is an abundant glycosaminoglycan present in ECM, connective tissue, cartilage, bone marrow, and synovial fluid (*1*). Several proteins have HA binding properties and are found in various tissues. Thus, CD44 was found on the SV-3T3 cell surface (*2, 3*). RHAMM (receptor for HA-mediated motility) was found in the culture media of 3T3 cells and chick embryo fibroblasts (*4*). HABPs (hyaluronic acid-binding proteins) were found in chick embryo (*5*). Hyaluronectin was found in human brain (*6, 7*). Proteoglycan core proteins and link proteins were found in cartilage (*8-10*), and TSG-6 was found as a tumor necrosis factor-inducible protein in human fibroblasts (*11*).

In the catabolism of ECM, proteolysis and its control mechanism have been studied to some extent for matrix metalloproteinases (MMPs) (*12-14*) and tissue inhibitors of matrix metalloproteinases (TIMPs) (*15, 16*). The degradation of HA and its regulation mechanism, however, have not been elucidated clearly. Only limited knowledge is available on hyaluronidase (PH-20) on sperm membrane (*17*). When we searched for HA-binding proteins in human plasma using HA-conjugated Sepharose, we found that only one protein specifically bound to the column. In this paper, we describe the purification of a novel HA-binding protein (PHBP) from human plasma and the molecular cloning of its cDNA.

MATERIALS AND METHODS

Materials—Human umbilical cord HA was obtained from Sigma. EAH-Sepharose was purchased from Pharmacia. Molecular weight markers for SDS-PAGE were obtained from GIBCO BRL, and lysylendopeptidase was from Wako Pure Chemicals. Restriction enzymes were purchased from Nippon Gene, Toyobo, or Boehringer Mannheim. Recombinant Taq DNA polymerase was obtained from Takara Shuzo (Kyoto). ^{32}P-labeled nucleotides and nylon mem-

[1] This work was supported in part by a Grant-in-Aid from the Ministry of Education, Science, Sports and Culture of Japan.

[2] To whom correspondence should be addressed. Tel: +81-3-3784-8215, Fax: +81-3-3784-8216, E mail: nammiura@pharm.showa-u.ac.jp

Abbreviations: ECM, extracellular matrix; EGF, epidermal growth factor; HA, hyaluronic acid; HBP, HA-binding protein; HC, heavy chain; HGFA, hepatocyte growth factor activator; IHRP, ITI family heavy chain-related protein; ITI, inter-alpha-trypsin inhibitor; PHBP, plasma HBP; RACE, rapid amplification of cDNA ends; TSG-6, TNF-stimulated gene 6.

branes (Hybond N) were purchased from Amersham Japan. Nitrocellulose filters were from Schleicher and Schuell. Multiple Tissue Northern Blot and QUICK-Clone™ cDNA (human liver) were purchased from Clontech Laboratories. A random-primer labeling kit was obtained from Du Pont-New England Nuclear.

Purification of PHBP—HA-Sepharose was prepared according to the method of Tengblad (*18*). Briefly, EAH-Sepharose (50 ml), HA (100 mg), and *N*-ethyl-*N'*-(3-dimethylaminopropyl)-carbodiimide (5 g) were mixed in 200 ml of distilled water, and then the pH was adjusted to 4.5 with 0.1 N HCl. The mixture was rotated at room temperature overnight. The gel was washed with 0.5 M NaCl, 0.1 M sodium acetate, pH 4.0, and then 0.5 M NaCl, 0.1 M Tris-HCl, pH 8.3, three times alternately. Then, the resulting HA-Sepharose was washed with phosphate-buffered saline (PBS). Human plasma (500 ml) was applied to the HA-Sepharose column (50 ml) at the flow rate of 150 ml/h. After washing with 0.5 M NaCl containing 10 mM sodium phosphate and 0.1% NP-40, pH 7.0 (500 ml), followed by with distilled water (100 ml), the bound proteins were eluted with 0.2 M glycine-HCl, pH 2.5. The absorbance of the eluate was monitored at 280 nm, after neutralization with 1 M Tris. The contaminating IgM and albumin in the PHBP preparation were removed by passage through columns of anti-IgG antibody-conjugated Sepharose and anti-albumin antibody-conjugated Sepharose, respectively, as described previously (*19*).

Electrophoresis—SDS-PAGE was performed according to the method of Laemmli (*20*) in slab gels (10% gel or 12.5% gel) under reducing or non-reducing conditions. After electrophoresis, proteins in a gel were stained with 0.1% Coomassie Brilliant Blue G-250 in 10% acetic acid and 30% methanol for 30 min, and then the gel was destained in 10% acetic acid and 30% methanol.

Amino Acid Sequence Analysis—After electrophoresis, proteins in the gel were transferred to a Pro Blot membrane (Applied Biosystems) in 20 mM Tris containing 150 mM glycine and 20% methanol electrically. The proteins on the membrane were stained with 0.1% Coomassie Brilliant Blue G-250 in 1% acetic acid and 40% methanol for 1 min, and then the membrane was destained with 50% methanol. The protein bands were cut off and their N-terminal amino acid sequences were analyzed with a 473A Protein Sequencer (Applied Biosystems). For intra-molecular amino acid sequence analysis, the purified PHBP (70 μg) was reduced and carboxymethylated according to the method of Cleland (*21*). The carboxymethylated PHBP was digested with 1 μg of lysylendopeptidase at 37°C overnight in 10 mM Tris-HCl, pH 7.0. The peptides were separated by HPLC on a Syn Chropak RP-P column (2.1×250 mm; Syn Chrom) as follows. The digested sample was applied to the column, which had been equilibrated with 5% acetonitrile in 0.1% trifluoroacetic acid (TFA). The peptides were eluted with a linear gradient of acetonitrile (5–50%, 45 min) in 0.1% TFA at the flow rate of 0.5 ml/min. The peptide peaks were monitored by measuring the absorbance at 214 nm and appropriate peaks were collected. The N-terminal amino acid sequence of each material was analyzed as above.

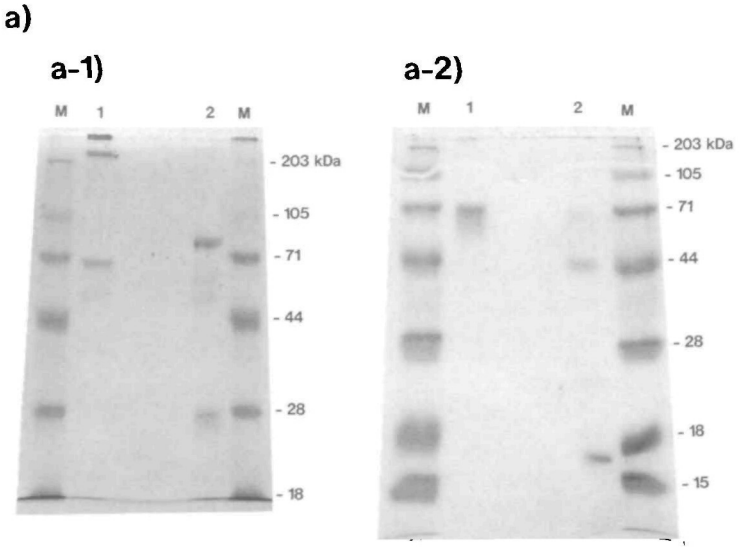
Preparation of a Specific Screening Probe for PHBP cDNA—Several sense and antisense mixed oligonucleotide primers were synthesized based on the amino acid sequences of the peptides generated on digestion of PHBP with lysylendopeptidase. Various combinations of the primers were used for PCR with QUICK-Clone™ cDNA (human liver; Clontech Laboratories) as a template for amplification of a PHBP cDNA fragment. The reaction mixture (100 μl) comprised 10 mM Tris-HCl, pH 8.3, 50 mM KCl, 1.5 mM $MgCl_2$, 0.001% gelatin, 0.5 μg cDNA, 0.2 mM of each dNTP, 2.5 units of recombinant Taq DNA polymerase, and 0.5 μM primer. The reaction mixture was overlaid with one drop of light mineral oil. The reactions were carried out for 5 cycles (94°C: 30 s, 37°C: 30 s, 72°C: 1 min), and 30 cycles (94°C: 30 s, 55°C: 30 s, 72°C: 1 min), followed by final extension at 72°C for 5 min, using a DNA thermal cycler (Perkin Elmer Cetus). When the primers (N1 and C1) based on the amino acid sequences, TVCLPD-GSFPSECHISGWGVTETGK (primer N1, TLCLPD, sense, 64-fold degeneracy) and PGQDTCQGDSGGPLTCEK (primer C1, GPLTCEK, antisense, 256-fold degeneracy) were used for PCR, an about 250-bp DNA fragment was amplified. To confirm the PCR product, it was reamplified using primers N2 (CHISGW) and C2 (QDTCQGD), which were based on other amino acid sequences within the same peptides.

Screening of Human Liver cDNA—The PCR product subcloned into pUC119 was labeled with [α-^{32}P]dCTP by the random priming method (*22*) (Du Pont-New England Nuclear Kit) for use as a screening probe. The labeled probe was used in a conventional plaque hybridization procedure (*23*) to screen a human liver cDNA library constructed in the λgt11 phage vector (Clontech Laboratories). The hybridization was carried out in a solution comprising 5× 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. The membranes were washed 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 to obtain positive clones.

Characterization of the Isolated Positive Clones and Construction of Full Length cDNA—The inserts of two independent positive clones (C7-1 and C8-1) were mapped using several restriction enzymes and then subcloned into the pUC119 vector. After subcloning of the inserted DNA, the nucleotide sequence was determined by the dideoxynucleotide chain termination method (*24*) with a Shimadzu DNA sequencer model DSQ1000. The sequence of clone C8-1 was all included in that of clone C7-1, which did not contain the full length cDNA of PHBP. In order to construct the full length cDNA, the 5′-end of the PHBP transcript was amplified by the PCR technique according to the method described by Edwards *et al.* with a slight modification (*25*). To obtain the (−)-strand of PHBP cDNA, the human liver cDNA (QUICK Clone™ cDNA; Clontech) was amplified using primer PHBP-anti-900, which is complementary to the nucleotide sequence just downstream of the 5′-end of clone C7-1. The reaction was performed as follows: 7 min at 96°C for initial denaturation, 25 cycles of amplification using a step program (96°C, 1 min; 55°C, 1 min; 72°C, 2 min), and final extension at 72°C for 15 min. The excess amounts of primers and dNTP were removed with a Centricon-30 (Amicon, USA). The sample was concentrated to 23 μl by centrifugation under reduced pressure. The (−)-strand cDNA was then tailed with poly(C) at its 3′-end by terminal deoxynucleotidyltrans-

ferase with 1 mM dCTP. After heat treatment of the reaction mixture at 65°C for 15 min, the mixture was diluted to 300 μl with 10 mM Tris-HCl, 1 mM EDTA, pH 8.0. The diluted mixture (10 μl) was amplified with primer amp-dG-20 [5′-CCGAATTCGG$(G)_{20}$-3′] and primer PHBP-anti-900. Both primer amp-dG-20 and primer PHBP-anti-900 contain *Eco*RI recognition sites at their 5′-ends for subcloning. The PCR reaction was carried out as follows; the reaction mixture was denatured at 95°C for 5 min and then cooled to 72°C, followed by the addition of 2.5 units of Taq DNA polymerase, and then was overlaid with one drop of mineral oil. For initial extension, the reaction mixture was annealed at 50°C for 2 min and then extended at 72°C for 30 min. The following amplification was carried out with a step program (94°C, 1 min; 50°C, 1 min; 72°C, 2 min, 35 cycles) and final extension at 72°C for 15 min. The PCR product (5′-RACE) was extracted with phenol/chloroform, digested with *Eco*RI, and then separated on a 1.5% agarose gel. The resulting PCR product was subcloned into the *Eco*RI-digested pUC119 plasmid vector.

Northern Blot Analysis—A Multiple Tissue Northern Blot (Clontech Laboratories), which had been blotted with polyadenylated RNA from various human tissues, was used to analyze the expression of PHBP mRNA. The filter was



b)

50 kDa

Sequence 1 (Minor): FSLMSLLESLDPDWT

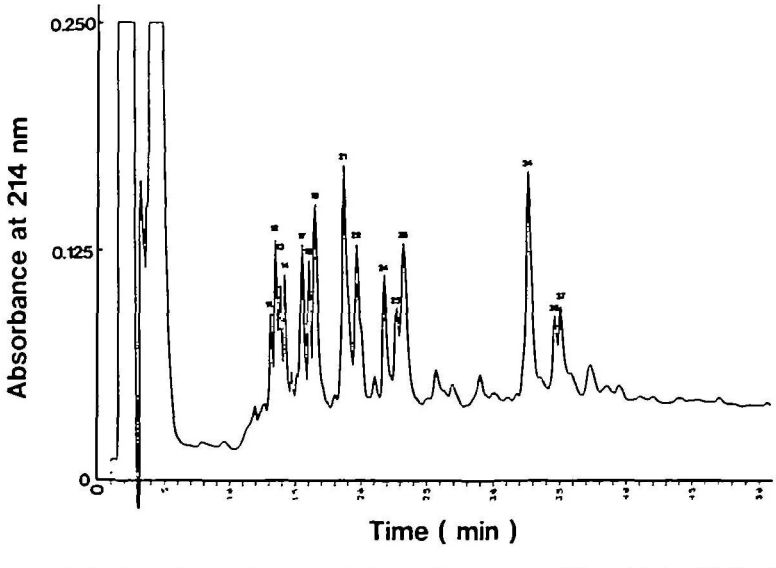
Sequence 2 (Major): SLLESLDPDWTPDQY

17 kDa

Sequence 1 (Minor): IYGGFKSTAGKHPWQ

Sequence 2 (Major): STAGKHPWQASLQSS

Fig. 1. **Purification of PHBP from human plasma.** (a) SDS-PAGE. (a-1) The eluate from a HA-Sepharose column was analyzed by SDS-PAGE (10% gel) under non-reducing (1) or reducing (2) conditions. After electrophoresis, proteins in the gel were stained with Coomassie Brilliant Blue G-250. (a-2) The fraction (PHBP) passing through anti-IgG antibody-Sepharose and anti-albumin antibody-Sepharose columns was analyzed by SDS-PAGE (12.5% gel) as described above (2 μg/lane). M: molecular weight marker proteins. (b) The N-terminal amino acid sequences of the 50- and 17-kDa materials of PHBP. After SDS-PAGE under reducing conditions, the 50- and 17-kDa bands were transferred to a Pro Blot membrane (Applied Biosystems) and then stained with Coomassie Brilliant Blue G-250. Each band was cut out and subjected to sequence analysis with a 473A Protein Sequencer.



b)

Peak 13: PGQDTCQGDSGGPLTCEK

Peak 14: LKPVDGHCALESK

Peak 18: FTCACPDQFK

Peak 19: HPYTGPSCSQVVPVCRPNPCQNGATCSRHK

Peak 24: PWCFIK

Peak 25: FLNWIK

Peak 26: TVCLPDGSFPSGSECHISGWGVTETGK

Peak 34: MXRTVNQHACLYWNSHLLLQENYNMFMEDA

Fig. 2. **Intra-molecular amino acid sequence analyses of PHBP.** (a) HPLC of the digest of reduced and carboxymethylated PHBP with lysylendopeptidase. Reduced and carboxymethylated PHBP (70 μg) was digested with lysylendopeptidase (1 μg) at 37°C overnight in 10 mM Tris-HCl, pH 7.0. The digested sample was applied to a column of Syn Chropak RP-P (2.1×150 mm; Syn Chrom), which had been equilibrated with 5% acetonitrile, 0.1% TFA. After washing with the same solution, peptides were eluted with a linear gradient of acetonitrile (5–50%, 45 min) in 0.1% TFA at the flow rate of 0.5 ml/min. (b) The N-terminal amino acid sequences of the PHBP peptides. The N-terminal amino acid sequences of the PHBP peptides obtained as above were analyzed with a 473A Protein Sequencer.

a)

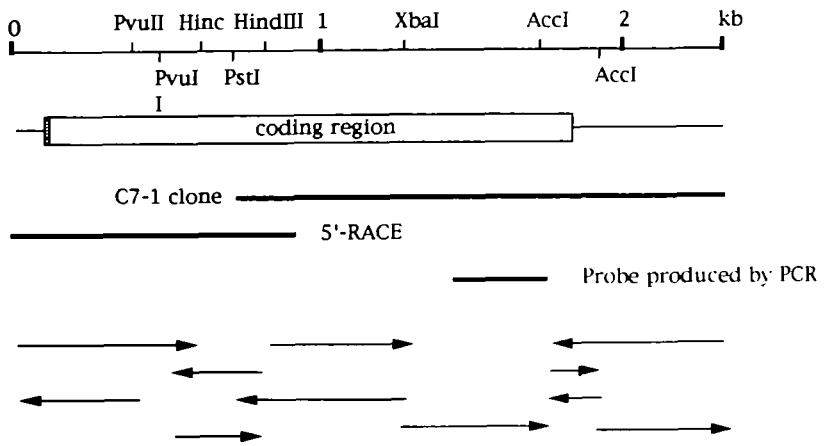


Fig. 3. **Molecular cloning of PHBP cDNA.** (a) Schematic representation of the restriction map, the PCR probe, the cDNA clones and the sequence strategy for the PHBP cDNA. The boxed bar denotes the PHBP coding sequence oriented in the 5′ to 3′ direction. The nucleotide scale is given at the top. The PCR product which was used for screening of the human liver cDNA library, the cDNA clone (C7-1) isolated from the human liver cDNA library, and the 5′-RACE cDNA are aligned with the restriction map. Horizontal arrows indicate the direction and extent of the sequencing. (b) Nucleotide sequence of PHBP cDNA and the deduced amino acid sequence. The nucleotide sequence of the constructed cDNA for PHBP is shown on the first line and the amino acid sequence deduced from the nucleotide sequence is shown on the second line. The amino acid sequences which were confirmed protein chemically are underlined. Nucleotide residues are numbered from +1 at the first base of the putative initiation codon. A polyadenylation signal is indicated by a double underline. The potential *N*-glycosylation sites are shown by thick underlines.

b)

```
   CCTGAATCCTTGGAGACTGACATTTTTCCCCCCTAAAGGCATAGACAACAAAAGAAATTTTATTGAGAGGAAAACACAAGTCCTTAAACTGCAAAG   -1
ATGTTTGCCAGGATGTCTGATCTCCATGTTCTGCTGTTAATGGCTCTGGTGGGAAAGACAGCCTGTGGGTTCTCCCTGATGTCTTTATTGGAAAGCCTG   99
M  F  A  R  M  S  D  L  H  V  L  L  L  M  A  L  V  G  K  T  A  C  G  F  S  L  M  S  L  L  E  S  L    33
GACCCAGACTGGACCCCTGACCAGTATGATTACAGCTACGAGGATTATAATCAGGAAGAGAACACCAGTAGCACACTTACCCATGCTGAGAATCCTGAC  198
D  P  D  W  T  P  D  Q  Y  D  Y  S  Y  E  D  Y  N  Q  E  E  N  T  S  S  T  L  T  H  A  E  N  P  D    66
TGGTACTACACTGAGGACCAAGCTGATCCATGCCAGCCCAACCCCTGTGAACACGGTGGGGACTGCCTCGTCCATGGGAGCACCTTCACATGCAGCTGC  297
W  Y  Y  T  E  D  Q  A  D  P  C  Q  P  N  P  C  E  H  G  G  D  C  L  V  H  G  S  T  F  T  C  S  C    99
CTGGCTCCTTTCTCTGGGAATAAGTGTCAGAAAGTGCAAAATACGTGCAAGGACAACCCATGTGGCCGGGGCCAATGTCTCATTACCCAGAGTCCTCCC  396
L  A  P  F  S  G  N  K  C  Q  K  V  Q  N  T  C  K  D  N  P  C  G  R  G  Q  C  L  I  T  Q  S  P  P   132
TACTACCGCTGTGTCTGTAAACACCCTTACACAGGTCCCAGCTGCTCCCAAGTGGTTCCTGTATGCAGGCCAAACCCCTGCCAGAATGGGGCTACCTGC  495
Y  Y  R  C  V  C  K  H  P  Y  T  G  P  S  C  S  Q  V  V  P  V  C  R  P  N  P  C  Q  N  G  A  T  C   165
TCCCGGCATAAGCGGAGATCCAAGTTCACCTGTGCCTGTCCCGACCAGTTCAAGGGGAAATTCTGTGAAATAGGTTCTGATGACTGCTATGTTGGCGAT  594
S  R  H  K  R  R  S  K  F  T  C  A  C  P  D  Q  F  K  G  K  F  C  E  I  G  S  D  D  C  Y  V  G  D   198
GGCTACTCTTACCGAGGGAAAATGAATAGGACAGTCAACCAGCATGCGTGCCTTTACTGGAACTCCCACCTCCTCTTGCAGGAGAATTACAACATGTTT  693
G  Y  S  Y  R  G  K  M  N  R  T  V  N  Q  H  A  C  L  Y  W  N  S  H  L  L  Q  E  N  Y  N  M  F   231
ATGGAGGATGCTGAAACCCATGGGATTGGGGAACACAATTTCTGCAGAAACCCAGATGCGGACGAAAAGCCCTGGTGCTTTATTAAAGTTACCAATGAC  792
M  E  D  A  E  T  H  G  I  G  E  H  N  F  C  R  N  P  D  A  D  E  K  P  W  C  F  I  K  V  T  N  D   264
AAGGTGAAATGGGAATACTGTGATGTCTCAGCCTGCTCAGCCCAGGACGTTGCCTACCCAGAGGAAAGCCCCACTGAGCCATCAACCAAGCTTCCGGGG  891
K  V  K  W  E  Y  C  D  V  S  A  C  S  A  Q  D  V  A  Y  P  E  E  S  P  T  E  P  S  T  K  L  P  G   297
TTTGACTCCTGTGGAAAGACTGAGATAGCAGAGAGGAAGATCAAGAGAATCTATGGAGGCTTTAAGAGCACGGCGGGCAAGCACCCATGGCAGGCGTCC  990
F  D  S  C  G  K  T  E  I  A  E  R  K  I  K  R  I  Y  G  G  F  K  S  T  A  G  K  H  P  W  Q  A  S   330
CTCCAGTCCTCGCTGCCTCTGACCATCTCCATGCCCCAGGGCCACTTCTGTGGTGGGGCGCTGATCCCACCCCTGCTGGGTGCTCACTGCTGCCCACTGC 1089
L  Q  S  S  L  P  L  T  I  S  M  P  Q  G  H  F  C  G  G  A  L  I  H  P  C  W  V  L  T  A  A  H  C   363
ACCGACATAAAAACCAGACATCTAAAGGTGGTGCTAGGGGACCAGGACCTGAAGAAAGAAGAATTTCATGAGCAGAGCTTTAGGGTGGAGAAGATATTC 1188
T  D  I  K  T  R  H  L  K  V  V  L  G  D  Q  D  L  K  K  E  E  F  H  E  Q  S  F  R  V  E  K  I  F   396
AAGTACAGCCACTACAATGAAAGAGATGAGATTCCCCACAATGATATTGCATTGCTCAAGTTAAAGCCAGTGGATGGTCACTGTGCTCTAGAATCCAAA 1287
K  Y  S  H  Y  N  E  R  D  E  I  P  H  N  D  I  A  L  L  K  L  K  P  V  D  G  H  C  A  L  E  S  K   429
TACGTGAAGACTGTGTGCTTGCCTGATGGGTCCTTTCCCTCTGGGAGTGAGTGCCACATCTCTGGCTGGGGTGTTACAGAAACAGGAAAAGGGTCCCGC 1386
Y  V  K  T  V  C  L  P  D  G  S  F  P  S  G  S  E  C  H  I  S  G  W  G  V  T  E  T  G  K  G  S  R   462
CAGCTCCTGGATGCCAAAGTCAAGCTGATTGCCAACACTTTGTGCAACTCCCGCCAACTCTATGACCACATGATTGATGACAGTATGATCTGTGCAGGA 1485
Q  L  L  D  A  K  V  K  L  I  A  N  T  L  C  N  S  R  Q  L  Y  D  H  M  I  D  D  S  M  I  C  A  G   495
AATCTTCAGAAACCTGGGCAAGACACCTGCCAGGGTGACTCTGGAGGCCCCCTGACCTGTGAGAAGGACGGCACCTACTACGTCTATGGGATAGTGAGC 1584
N  L  Q  K  P  G  Q  D  T  C  Q  G  D  S  G  G  P  L  T  C  E  K  D  G  T  Y  Y  V  Y  G  I  V  S   528
TGGGGCCTGGAGTGTGGGAAGAGGCCAGGGGTCTACACCCAAGTTACCAAATTCCTGAATTGGATCAAAGCCACCATCAAAAGTGAAAGTGGCTTCTAA 1683
W  G  L  E  C  G  K  R  P  G  V  Y  T  Q  V  T  K  F  L  N  W  I  K  A  T  I  K  S  E  S  G  F  ***  560
GGTACTGTCTTCTGGACCTCAGAGCCCACTCTCCTTGGCACCCTGACACCGGGAGGCCTCATGGCCAACAATGGACACCTCCAGAGCCTCCAGGGGACC 1782
ACACAGTAGACTATCCTACTCTAAGCAGAGACAACTGCCACCCAGCCTGGGCCTTCCCAGACCAGCATTTGCACAATATCACCAGGCTTCTTCTGCCTC 1881
CCTTGGTAACCCAAGGAATGATGGAATCAACACAACATAGTATGTTTGCTTTCCTTACCCAATTGTACCTTCTAGAAAATCAGTGTTCACAGAGACTGC 1980
CTCCACCACAGGCATCCTGCAAATGCAGACTCCAGAATCCCCAGCATCAGCGGGAACCACCATCACATCTTTATTCCTCAGCCCAGACACTCGAGGCAC 2079
TCAACAGAATCAGCCATCCACGTCTAGGTATCAGAGAGGACCACAAATACAACATTCTCCATCTGCTTTCAGAGTTATTATTTTAATAAAGGAAGATCT 2178
GGGATGGGCTGGTGGGCCATTCCAGCTTGCCGAAATCAAGCCATCTGAAGCCTGTCTCTGGTGAACAAACTTCCTCTCTGGCCTCTCAGGAATCAGGGT 2277
GGCATGGCTCACAACAGCAGGGCCTTCTTCTTTTT  2312
```

prehybridized for 2 h at 42°C and then hybridized at 42°C for 16 h in a solution comprising 50% formamide, 10× Denhardt's, 5×SSPE, 2% SDS, and 100 μg/ml of denatured salmon sperm DNA. The cDNA probe was labeled with [α-^{32}P]dCTP by the random-primer method (*22*), and then added to the hybridization solution to give a final concentration of 10^6 cpm/ml. The filter was washed twice with 2× SSC containing 0.05% SDS at room temperature for 20 min, and then washed twice with 0.1×SSC containing 0.1% SDS at 50°C for 20 min, and then exposed for 20 h with an intensifying screen at −70°C.

RESULTS

Purification of PHBP—When human plasma was applied to a column of HA-Sepharose, a small fraction (0.02%) of plasma proteins bound to the column and were eluted with 0.2 M glycine-HCl, pH 2.5. These HA-binding proteins gave >203-, 70-, and 50-kDa bands on SDS-PAGE under non-reducing conditions, and 80-, 55-, 50-, and 28-kDa ones under reducing conditions (Fig. 1, a-1). From these results, we assumed that IgM and albumin contaminated the eluate. Then, the HA-Sepharose column eluate was applied to columns of anti-IgG antibody-conjugated Sepharose and anti-albumin antibody-conjugated Sepharose. The resulting eluate gave a single band of 70 kDa on SDS-PAGE under non-reducing conditions (Fig. 1, a-2, lane 1). Finally, 800 μg of HA-binding protein (PHBP) was purified from 500 ml of human plasma. The purified PHBP bound to newly prepared HA-Sepharose, indicating the direct binding of PHBP to HA (data not shown). N-Terminal amino acid sequence analysis of PHBP revealed several sequences, suggesting that PHBP was composed of several peptides. PHBP gave 50- and 17-kDa bands on SDS-PAGE under reducing conditions (Fig. 1, a-2, lane 2), and both materials exhibited two N-terminal amino acid sequences, respectively (Fig. 1b). Sequence 2 (major sequence) started from the fifth amino acid of sequence 1 (minor sequence) in the 50-kDa material, and sequence 2 (major sequence) started from the seventh amino acid of sequence 1 (minor sequence) in the 17-kDa material. These results suggested the trimming of four and six amino acids from the N-terminals of the 50- and 17-kDa materials, respectively. A homology search of these sequences with the published databanks [Genetyx CD30, NBRF-PDB (PIR) Rel. 44 and SWISS-PROT Rel. 31] indicated that PHBP was a novel protein and the sequence of the 17-kDa material showed partial homology to that of hepatocyte growth factor activator (HGFA), which is known as a serine protease (data not shown). From these results, we concluded that PHBP was a novel HA-binding heterodimer, in which the two subunits were bridged by disulfide linkage(s).

Intra-Molecular Amino Acid Sequence Analyses of PHBP—PHBP (70 μg) was reduced and carboxymethylated, and then digested with lysylendopeptidase (1 μg). The digested sample was applied to a reversed-phase HPLC column and the peptides were separated (Fig. 2a). The N-terminal amino acid sequence of each peptide supported that PHBP was a novel protein (Fig. 2b), and peak materials 13, 18, and 24 showed significant homologies to HGFA (data not shown). Peak materials 11, 12, 17, 21, and 22 gave two or three amino acids at every step, while peak materials 36 and 37 showed no detectable sequences.

Isolation and Sequencing of PHBP cDNA Clones—PCR was applied to amplify the cDNA fragment encoding PHBP from human liver cDNA. The PCR product amplified with the two mixed oligonucleotide primers (N1 and C1) was about 250 bp in size and was confirmed to encode the amino acid sequences of the peptides which were used for the design of the PCR primers, the nested PCR and nucleotide sequence analysis (data not shown). After subcloning of the amplified cDNA fragment into the pUC119 plasmid vector, it was used as a probe to clone the full-length PHBP cDNA from a human liver cDNA library. Approximately 8×10^5 clones of the cDNA library were screened with the ^{32}P-labeled probe, two positive clones being isolated. These two clones (clones C7-1 and C8-1) were sequenced and found to be truncated at their 5′-ends. The full length cDNA of PHBP was constructed as described under "MATERIALS AND METHODS." Figure 3a shows a restriction map of clone C7-1, the region of the probe produced by PCR, the 5′-RACE, and the sequencing strategy for PHBP cDNA. The constructed cDNA had an overall length of 2,312 bp and contained an open reading frame encoding 560 amino acid residues. The molecular weight of the precursor of PHBP was calculated to be 62,670 from the sequence. The complete nucleotide sequence of PHBP cDNA and the amino acid sequence deduced from it are shown in Fig. 3b. Although the nucleotide sequence surrounding the first in-frame initiation codon at nucleotides 1-3, GCCTGATGT, does not match well the consensus sequence, CCACCATGG, described by Kozak (*26, 27*), this codon appears to be the most likely candidate for the translation start site because this methionine is followed by 23 hydrophobic amino acid residues, which may be in a signal peptide sequence (*28*). The second in-frame codon, ATG, 9 bp downstream from the first in-frame one, is also possibly the initiation codon. The third in-frame codon, ATG, 24 bp further downstream from the second in-frame one, is unlikely to be the initiation codon because the putative signal peptide is too short. As shown in Fig. 3b, the C7-1 clone included a 3′-poly(A) tail and the consensus polyadenylation signal, AATAAA (*29*), positioned at 46 nucleo-

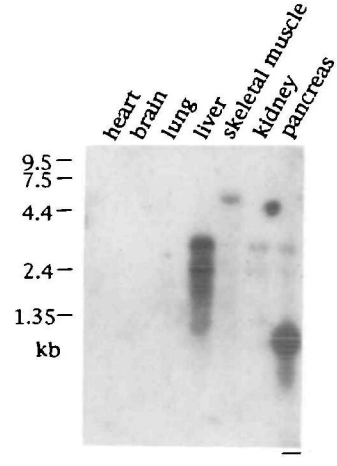


Fig. 4. **Northern blot analysis of PHBP mRNA in human tissues.** A Multiple Tissue Northern Blot obtained from Clontech was used for hybridization with the PHBP cDNA. The filter was hybridized with the ^{32}P-labeled cDNA probe and visualized by autoradiography.

tides upstream of the first A of the poly(A) tract.

Distribution of the PHBP Message in Adult Human Tissues—The expression of PHBP mRNA in human adult tissues was examined by Northern blot hybridization. A Multiple Tissue Northern Blot (Clontech Laboratories), which had been blotted with polyadenylated RNAs from human adult heart, brain, placenta, lung, liver, skeletal muscle, kidney, and pancreas, was probed with the ^{32}P-labeled cDNA (clone C7-1). Two hybridized signals (3 and 2.3 kb) were detected in the liver, kidney and pancreas (Fig. 4). A small band of 1.3 kb was observed for the pancreas besides the signals of 3 and 2.3 kb. In skeletal muscle, only a 6 kb band was detected. The significance of the appearance of these 1.3- and 6-kb bands is not clear at present.

DISCUSSION

HA-binding proteins have been found in several tissues, although HA-binding proteins in plasma have not been reported so far. In this paper, we described the purification and characterization of a novel HA-binding protein (PHBP) in human plasma. The HA-Sepharose column eluate contained only PHBP, IgM, and albumin, suggesting that PHBP was a major HA-binding protein in human plasma (Fig. 1b), and the contaminating IgM and albumin could be removed with anti-IgG antibody-Sepharose and anti-albumin antibody-Sepharose, respectively. Our results indicated that PHBP was a disulfide-bonded heterodimer. Sequence 2 of the 50-kDa material showed cleavage on the C-terminal side of methionine, and sequence 2 of the 17-kDa material showed cleavage on the C-terminal side of lysine, therefore, the cleavage of both N-termini might be caused by different proteolytic enzymes.

We isolated two independent cDNA clones (C7-1 and C8-1) encoding PHBP from a human liver cDNA library. Clone C7-1 was found to contain the full sequence of clone C8-1 insert cDNA and to be truncated at the 5′-end. In order to obtain the full length cDNA, the 5′-end of the human PHBP transcript was amplified by the PCR technique. The nucleotide sequence of the constructed cDNA (nucleotides −96 to 2312) contained the full open reading frame, starting at nucleotides 1-3 of the initiation codon and terminating at nucleotides 1681-1683 of a stop codon (TAA). The 1,680 bp open reading frame coded for a polypeptide of 560 amino acid residues. All amino acid sequences of the peptides generated on digestion of PHBP with lysylendopeptidase are included in this reading frame, supporting that this cDNA encodes PHBP (Fig. 3b).

A homology search of the nucleotide sequence of PHBP cDNA revealed that it was identical to that of hepatocyte growth factor activator (HGFA)-like mRNA registered in DDBJ by Kitamura *et al.* (accession number, D49742). The cDNA of HGFA-like mRNA is 3,008 bp in length, while PHBP cDNA is 2,312 bp in length. This discrepancy is due to the utilization of a different polyadenylation signal in the 3′-region of the PHBP transcript. The precursor of PHBP consists of 560 amino acid residues. The overall amino acid sequence of the PHBP precursor shows significant sequence homology (about 31%) to that of HGFA (*30*). Interestingly, the PHBP precursor is about 100 amino acid residues shorter than the HGFA precursor. HGFA consists of characteristic structural domain, a type I and a type II fibronectin domains, two EGF domains, a kringle domain and a serine protease domain, while PHBP consists of three EGF domains, a kringle domain and a serine protease

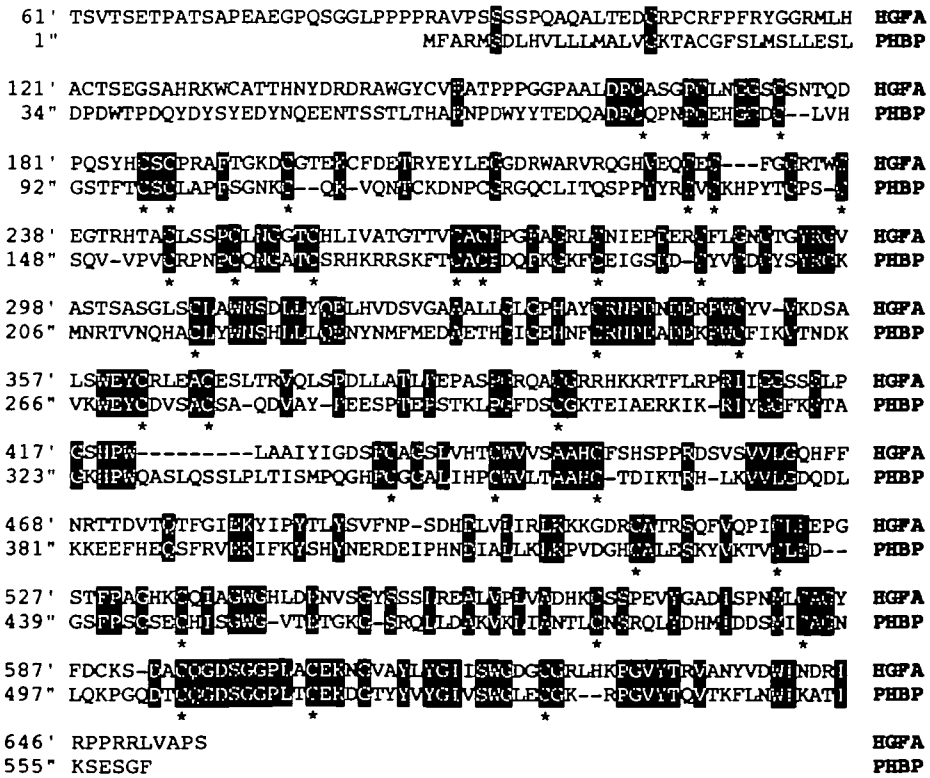


Fig. 5. **Similarity of the amino acid sequences of PHBP and HGFA.** Sequences were aligned using the Genetyx program (SDC, Tokyo). Identical amino acid residues are shown by black boxes. Asterisks indicate cysteine residues conserved in the two sequences.

domain. The N-terminal amino acids of the 50- and 17-kDa fragments of PHBP correspond to Phe_1 and Ile_{291} in the deduced amino acid sequence, respectively. Thus, a mature form of PHBP may be produced on proteolytic cleavage of the peptide bond between Arg_{290}-Ile_{291} of a single peptide precursor. However, the 17-kDa fragment is much smaller than the molecular mass calculated from the amino acid sequence from Ile_{291} to Phe_{537} of the C-terminus, suggesting further proteolytic degradation in the C-terminal region of the fragment. The 17-kDa fragment is unlikely to contain Ser_{486}, which is essential for serine protease activity as an active site. It is notable that the peptide corresponding to Phe_{546}-Lys_{551} was recovered on digestion of PHBP with lysylendopeptidase. Therefore, PHBP should have the C-terminal region which is absent from the 17-kDa fragment, although we could not identify the C-terminal region as a band on SDS-PAGE. We detected weak proteolytic activity in the PHBP preparation with ITI as a substrate (data not shown). The weak protease activity of PHBP might be explained by the finding that a major portion of our PHBP preparation is degraded in the C-terminal catalytic domain; an active form of PHBP as a serine protease would consist of a 50-kDa fragment and an about 25-kDa fragment corresponding to Ile_{291}-Phe_{537}, which contains an intact catalytic domain.

All cysteine residues contained in three-fourths of the C-terminal region of the PHBP precursor are located virtually identically to those of HGFA with the exception of Cys_{589} in HGFA (Fig. 5). Therefore, PHBP is predicted to have a similar secondary structure to that of HGFA in this region. The predicted structural domains in the PHBP precursor are shown in Fig. 6. PHBP as well as HGFA is synthesized as a single peptide precursor, and then may be converted to a two-chain form of 50- and 25-kDa fragments linked by a disulfide bond. In the case of HGFA activation, the HGFA precursor is truncated at its N-terminal region, resulting in the 34-kDa active form containing a catalytic domain. On the other hand, active PHBP derived from the precursor may still retain the N-terminal domains through disulfide-bonded linkages.

The HGF-converting enzyme reported by Mizuno *et al.* is also a heterodimeric protein, which has similar enzymatic properties to HGFA (*31*). But the molecular mass (90 kDa) of the HGF-converting enzyme consisting of 65- and

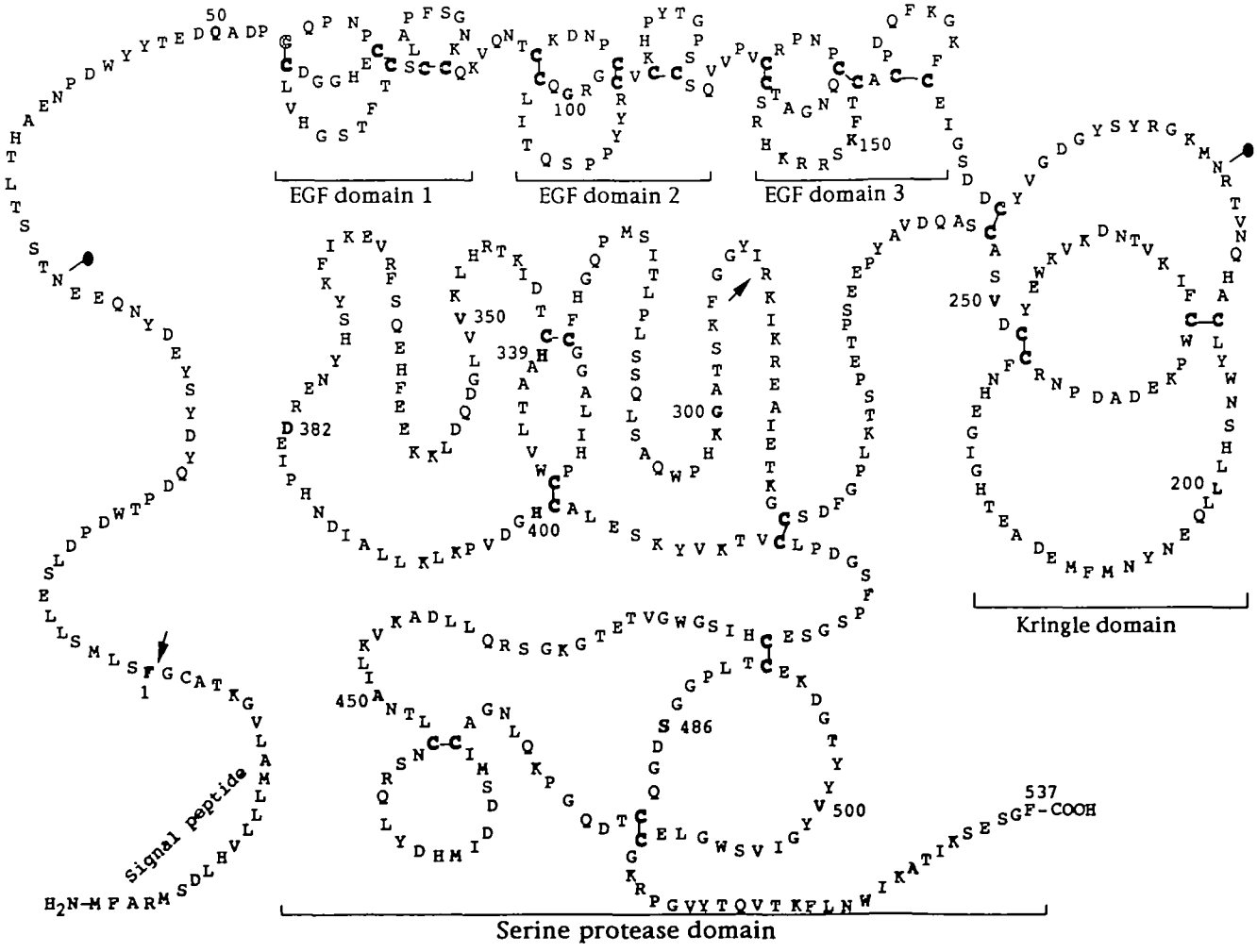


Fig. 6. **Schematic representation of the putative structure of PHBP.** Possible disulfide linkages are based on the sequence similarity to HGFA. The putative sites for proteolytic processing to generate active PHBP are indicated by arrows. Amino acids for the potential catalytic triad of the serine protease active site are indicated by the bold characters. The cysteines linked by disulfide bonds are indicated by the outlined characters.

32-kDa fragments clearly differs from that of PHBP. Although the function of PHBP has not been clarified yet, PHBP may exhibit its biological functions on binding to HA in ECM, connective tissue, cartilage, bone marrow, and synovial fluid. The protease which activates the PHBP precursor and the natural substrate of active PHBP should be elucidated.

The results of Northern blot analysis showed that PHBP was expressed in liver, skeletal muscle, kidney, and pancreas. The sizes of the signals, however, varied with the tissue. In liver, two signals were detected. The band of 3 kb may correspond to the HGFA-like mRNA registered by Kitamura *et al.* The other band, 2.3 kb, should correspond to PHBP mRNA. Besides the signals detected in liver, a strong hybridization signal (1.3 kb) was detected in pancreas. This signal is too short to encode PHBP. A large band (6 kb) was also detected in skeletal muscle. At present, it is unknown whether these bands reflect an alternatively spliced form of PHBP mRNA or a cross-hybridized mRNA.

Regarding the ECM-stabilizing activity of the ITI family, the heavy chains of the family are cleaved and bind to HA firmly (*32-37*). However, the enzyme involved in the cleavage of the heavy chains of the ITI family and the mechanism by which the heavy chains bind to HA tightly have not been elucidated. Although several proteases could cleave ITI (*38*), and human leukocyte elastase cleaved HC2 of ITI preferentially (*39*), PHBP might also be a candidate of the enzyme concerned in this process, because PHBP exhibited HA binding properties and cleaved HC2 of ITI weakly. Further work is needed to elucidate the natural substrate and the nature of PHBP.

We are grateful to Ms. Kiyomi Saito for her skillful technical assistance.

REFERENCES

1. Laurent, T.C. and Fraser, J.R.E. (1992) Hyaluronan. *FASEB J.* **6**, 2397-2404
2. Underhill, C.B. and Toole, B.P. (1979) Binding of hyaluronate to the surface of cultured cells. *J. Cell Biol.* **82**, 475-484
3. Culty, M., Miyake, K., Kincade, P.W., Sikorski, E., Butcher, E.C., and Underhill, C. (1990) The hyaluronate receptor is a member of the CD44 (H-CAM) family of cell surface glycoproteins. *J. Cell Biol.* **111**, 2765-2774
4. Yang, B., Zhang, L., and Turley, E.A. (1993) Identification of two hyaluronan-binding domains in the hyaluronan receptor, RHAMM. *J. Biol. Chem.* **268**, 8617-8623
5. Banerjee, S.D. and Toole, B.P. (1991) Monoclonal antibody to chick embryo hyaluronan-binding protein: Changes in the distribution of the binding protein during early brain development. *Dev. Biol.* **146**, 186-197
6. Delpech, B. and Halavent, C. (1981) Characterization and purification from human brain of a hyaluronic acid-binding glycoprotein, hyaluronectin. *J. Neurochem.* **36**, 855-859
7. Delpech, B. (1982) Immunochemical characterization of the hyaluronic acid-hyaluronectin interaction. *J. Neurochem.* **38**, 978-984
8. Doege, K., Hassell, J.R., Caterson, B., and Yamada, Y. (1986) The link protein cDNA sequence reveals a tandemly repeated protein structure. *Proc. Natl. Acad. Sci. USA* **83**, 3761-3765
9. Neame, P.J., Christner, J.E., and Baker, J.R. (1987) Cartilage proteoglycan aggregates: The link protein and proteoglycan amino-terminal globular domains have similar structures. *J. Biol. Chem.* **262**, 17768-17778
10. Doege, K.J., Sasaki, M., Kimura, T., and Yamada, Y. (1991) Complete coding sequence and deduced primary structure of the human cartilage large aggregating proteoglycan, aggrecan: Human-specific repeats, and additional alternatively spliced forms. *J. Biol. Chem.* **266**, 894-902
11. Wisniewski, H.-G., Burgess, W.H., Oppenheim, J.D., and Vilcek, J. (1994) TSG-6, an arthritis-associated hyaluronan-binding protein, forms a stable complex with the serum protein inter-alpha-inhibitor. *Biochemistry* **33**, 7423-7429
12. Gross, J. and Lappiere, C.M. (1962) Collagenolytic activity in amphibian tissues: A tissue culture assay. *Proc. Natl. Acad. Sci. USA* **48**, 1014-1022
13. Woessner, J.F., Jr. (1991) Matrix metalloproteinases and their inhibitors in connective tissue remodeling. *FASEB J.* **5**, 2145-2154
14. Basset, P., Bellocq, J.P., Wolf, C., Stoll, I., Hutin, P., Limacher, J.M., Podhajcer, O.L., Chenard, M.P., Rio, M.C., and Chambon, P. (1990) A novel metalloproteinase gene specifically expressed in stromal cells of breast carcinomas. *Nature* **348**, 699-704
15. Wilhelm, S.M., Collier, I.E., Marmer, B.L., Eisen, A.Z., Grant, G.A., and Goldberg, G.I. (1989) SV40-transformed human lung fibroblasts secrete a 92-kDa type IV collagenase which is identical to that secreted by normal human macrophages. *J. Biol. Chem.* **264**, 17213-17221
16. Mast, A.E., Enghild, J.J., Nagase, H., Suzuki, K., Pizzo, S.V., and Salvesen, G. (1991) Kinetics and physiologic relevance of the inactivation of alpha1-proteinase inhibitor, alpha1-antichymotrypsin, and antithrombin III by matrix metalloproteinases-1 (tissue collagenase), -2 (72-kDa gelatinase/type IV collagenase), and -3 (stromelysin). *J. Biol. Chem.* **266**, 15810-15816
17. Gmachl, M., Sagan, S., Ketter, S., and Kreil, G. (1993) The human sperm protein PH-20 has hyaluronidase activity. *FEBS Lett.* **336**, 545-548
18. Tengblad, A. (1979) Affinity chromatography on immobilized hyaluronate and its application to the isolation of hyaluronate-binding proteins from cartilage. *Biochim. Biophys. Acta* **578**, 281-289
19. 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 the inter-alpha-trypsin inhibitor family from human plasma. *J. Biochem.* **117**, 400-407
20. Laemmli, U.K. (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **227**, 680-685
21. Cleland, W.W. (1964) Dithiothreitol, a new protective reagent for SH groups. *Biochemistry* **3**, 480-482
22. Feiberg, A.P. and Vogelstein, B. (1983) A technique for radiolabeling DNA restriction endonuclease fragments to high specific activity. *Anal. Biochem.* **132**, 6-13
23. Maniatis, T., Fritsch, E.F., and Sambrook, J. (1982) *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY
24. Sanger, F., Nicklen, S., and Coulson, A.R. (1977) DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. USA* **74**, 5463-5467
25. Edwards, J.B.D.M., Delort, J., and Mallet, J. (1991) Oligodeoxyribonucleotide ligation to single-stranded cDNAs: A new tool for cloning 5′ ends of mRNAs and for constructing cDNA libraries by in vitro amplification. *Nucleic Acids Res.* **19**, 5227-5232
26. Kozak, M. (1986) Point mutations define a sequence flanking the AUG initiator codon that modulates translation by eukaryotic ribosomes. *Cell* **44**, 283-292
27. Kozak, M. (1987) An analysis of 5′-noncoding sequences from 699 vertebrate messenger RNAs. *Nucleic Acids Res.* **15**, 8125-8148
28. von Heijne, G. (1982) Signal sequences are not uniformly hydrophobic. *J. Mol. Biol.* **159**, 537-541
29. Wickens, M.D. and Stephenson, P. (1984) Role of the conserved AAUAAA sequence: Four AAUAAA point mutants prevent messenger RNA 3′ end formation. *Science* **226**, 1045-1051
30. Miyazawa, K., Shimomura, T., Kitamura, A., Kondo, J., Morimoto, Y., and Kitamura, N. (1993) Molecular cloning and sequence analysis of the cDNA for a human serine protease

responsible for activation of hepatocyte growth factor: Structural similarity of the protease precursor to blood coagulation factor XII. *J. Biol. Chem.* **268**, 10024-10028
31. Mizuno, K., Tanoue, Y., Okano, I., Harano, T., Takada, K., and Nakamura, T. (1994) Purification and characterization of hepatocyte growth factor (HGF)-converting enzyme: Activation of pro-HGF. *Biochem. Biophys. Res. Commun.* **198**, 1161-1169
32. Jessen, T.E., Faarvang, K.L., and Ploug, M. (1988) Carbohydrate as a covalent crosslink in human inter-alpha-trypsin inhibitor: A novel plasma protein structure. *FEBS Lett.* **230**, 195-200
33. Balduyck, M., Laroui, S., Mizon, C., and Mizon, J. (1989) A proteoglycan related to the urinary trypsin inhibitor (UTI) links the two heavy chains of inter-alpha-trypsin inhibitor. *Biol. Chem. Hoppe-Seyler* **370**, 329-336
34. 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
35. Yoneda, M., Suzuki, S., and Kimata, K. (1990) Hyaluronic acid associated with the surfaces of cultured fibroblasts is linked to a serum-derived 85-kDa protein. *J. Biol. Chem.* **265**, 5247-5257
36. Huang, L., Yoneda, M., and Kimata, K. (1993) A serum-derived hyaluronan-associated protein (SHAP) is the heavy chain of the inter-alpha-trypsin inhibitor. *J. Biol. Chem.* **268**, 26725-26730
37. Jessen, T.E., Odum, L., and Johnsen, A.H. (1994) In vivo binding of human inter-alpha-trypsin inhibitor free heavy chains to hyaluronic acid. *Biol. Chem. Hoppe-Seyler* **375**, 521-526
38. Dietl, T., Dobrinski, W., and Hochstrasser, K. (1979) Human inter-alpha-trypsin inhibitor: Limited proteolysis by trypsin, plasmin, kallikrein and granulocytic elastase, and inhibitory properties of the cleavage products. *Hoppe-Seyler's Z. Physiol. Chem.* **360**, 1313-1318
39. Balduyck, M., Piva, F., Mizon, C., Maes, P., Malki, N., Gressier, B., Michalski, C., and Mizon, J. (1993) Human leucocyte elastase (HLE) preferentially cleaves heavy chain H2 of inter-alpha-trypsin inhibitor (ITI). *Biol. Chem. Hoppe-Seyler* **374**, 895-901