Cloning of the cDNA for a Mouse Homologue of Human PHBP: a Novel Hyaluronan-Binding Protein

Ken HASHIMOTO, Takashi TOBE,* Jun-ichi SUMIYA, Ken-ichi SAGUCHI, Yoshihiro SANO, Yasuko NAKANO, Nam-Ho CHOI-MIURA, and Motowo TOMITA

Department of Physiological Chemistry, School of Pharmaceutical Sciences, Showa University, 1-5-8, Hatanodai, Shinagawa-ku, Tokyo 142, Japan. Received June 9, 1997; accepted August 2, 1997

The cDNA which encodes the mouse counterpart of human plasma hyaluronan-binding protein (PHBP) was isolated and characterized. The clone contained an insert of 2153 bp, which contained the 1674-bp open reading frame coding for a polypeptide of 558 amino acid residues. The amino acid sequence of mouse PHBP predicted from the nucleotide sequence of cDNA shows reasonable homology to that of human PHBP. Like human PHBP, the amino acid sequence predicted from the nucleotide sequence of mouse PHBP cDNA exhibited significant homology to that of human hepatocyte growth factor activator (HGFA).

Key words: mouse PHBP; primary structure; HGFA-like mRNA

We have recently reported the purification and cloning of a novel human plasma hyaluronan-binding protein that showed significant homology to the hepatocyte growth factor activator (HGFA). We named the protein PHBP (plasma hyaluronan-binding protein).1) Human PHBP is a hetero-dimer protein composed of 50 and 25 kDa subunits which are bridged by disulfide linkage. Human PHBP consists of 560 amino acid residues, including a possible signal peptide at the N-terminus. Both subunits are encoded by the same mRNA and are produced from a common translation product by proteolytic processing. The large and small subunits are located at the N-terminal and C-terminal regions of the precursor, respectively. The large subunit contains three epidermal growth factor (EGF) domains and a kringle domain, and the small subunit is similar in amino acid sequence to the serine protease domain of HGFA. Thus, although human PHBP exhibits a characteristic feature which is different from hepatocyte growth factor activator (HGFA),2) the physiological function of PHBP may be thought to act as a serine protease in vivo. The nucleotide sequence of human PHBP cDNA was identical to the sequence which had been determined to be hepatocyte growth factor activator like (HGFA like) mRNA, although a protein corresponding to the nucleotide sequence has not been identified in human plasma. The function of PHBP is unknown to date. To elucidate the function of this protein, we tried to isolate the cDNA which encodes the mouse counterpart of PHBP.

MATERIALS AND METHODS

Reagents ISOGEN-LS was purchased from Wako Pure Chemicals (Osaka, Japan). Restriction enzymes were purchased from Nippon Gene, Toyoobo, or Boehringer Mannheim and used according to the manufacturer's instructions. RecombinantTaq DNA polymerase was obtained from Takara Shuzo (Tokyo, Japan). 32P-Labeled nucleotide was purchased from Amersham Japan Corp. (Tokyo, Japan). Nitrocellulose filters were from Schleicher and Schuell. Agarose type II was obtained from Sigma Chemical Co. (St. Louis, MO, U.S.A.), and low-melting-point agarose (NuSieve) was from FMC (Rockland, ME, U.S.A.). A random-primer labeling kit was obtained from Du Pont-New England (Boston, MA, U.S.A.).

Cloning of the Mouse PHBP cDNA The human PHBP cDNA was labeled with [α-32P]dCTP by a random priming method3) (Du Pont-New England Nuclear Kit) to use as a screening probe. The labeled probe was used in a conventional plaque hybridization procedure4) to screen a normal mouse liver cDNA library constructed in a λZIP phage vector by cross hybridization. The hybridization was carried out in a solution containing 5 × saline-sodium citrate (SSC), 5 × Denhardt's solution, 50 mM Tris-HCl, pH 7.5, 100 mg/ml of denatured salmon sperm DNA, 0.1% sodium dodecyl sulfate (SDS), 10% dextran sulfate, and 50% formamide at 42°C, then the membranes were washed sequentially with 2 × SSC containing 0.1% SDS at room temperature, followed by 0.1 × SSC containing 0.1% SDS at 45°C. Several rounds of screening were carried out in order to obtain positive clones.

Characterization of Isolated Positive Clones Only one clone (221) was isolated by plaque purification and excised from the λ phage as a cDNA insert in Bluescript plasmid. After retrieving the inserted DNA in the Bluescript vector, the nucleotide sequence was determined by the dideoxynucleotide chain termination method4) with a DNA sequencer model DQA-1000 (Shimadzu, Kyoto).

Southern Blot Analysis DNA was extracted from mouse liver. The liver was homogenized, then lysed with SDS (1% w/v) and proteinase K (100 mg/ml) at 37°C for 16 h in a lysis buffer (100 mM NaCl, 25 mM EDTA). Proteins were removed by phenol/chloroform extraction and DNA was precipitated with ethanol. After the precipitated DNA was dried, 20 mg of the DNA was dissolved in TE, and completely digested at 37°C with EcoRI, BamHI, HindIII or PstI. The DNA fragments were separated on a 0.7% agarose gel by electrophoresis. After depurinating, denaturing and neutralizing the gel, the DNA fragments were transferred onto a nylon membrane (Hybond N). The membrane was probed with 32P-labeled cDNA.

Northern Blot Analysis The total RNAs were isolated

© 1997 Pharmaceutical Society of Japan
from mouse liver, kidney, heart, spleen, and lung using ISOGEN-LS. The RNAs were separated on an agarose gel containing formaldehyde, then transferred onto a nylon filter. The filter was prehybridized for 2 h at 42 °C and then hybridized at 42 °C for 16 h in a solution containing 50% formamide, 10 × Denhardt's solution, 5 × SSPE solution, 2% SDS, and 100 mg/ml of denatured salmon sperm DNA. The cDNA probe was labeled with [z-32P]-dCTP by the random-primer method and added to the hybridization solution to give a final concentration of 106 cpm/ml. The filter was washed with 2 × SSC containing 0.05% SDS at room temperature for 20 min × 2 times and washed with 0.1 × SSC containing 0.1% SDS at 50 °C for 20 min × 2 times, and then exposed for 20 h with an intensifying screen at −70 °C.

RESULTS

Approximately 8 × 105 recombinant clones from the cDNA library were screened with the 32P-labeled human PHBP cDNA, and only one positive clone was isolated. This clone, 221, was analyzed by PCR to determine the length of inserted DNA and by digestion with several restriction enzymes, then sequenced. The clone was ascertained to possess the full-length nucleotide sequence for mouse PHBP. Figure 1 shows the restriction map and sequencing strategies for this clone. This clone contained an insert of 2153 bp containing a poly(A) tail of 14 bases. The open reading frame begins 10 bases downstream from the 5'-end and extends for 558 amino acids to a TGA stop codon (Fig. 2). Although the nucleotide sequence surrounding the first in-frame initiation codon at nucleotides 1–3, CACAGATGT, does not match well with the consensus sequence, CCACCATGG, described by Kozak, this codon appears to be the most likely candidate for the translation start site because this methionine was followed by 22 hydrophobic amino acid residues which may act as a signal peptide. The length of hydrophobic amino acid stretch was suitable to act as a signal peptide. In the 3' non-coding region, a consensus polyadenylation signal, AATAAA, and a 3' poly(A) tail are positioned 24 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 23 amino acid residues and a mature protein of 553 amino acid residues (Fig. 2). The molecular mass of the mature protein without carbohydrate was calculated to be 59798.11 from the amino acid sequence.

The expression of PHBP mRNA in adult mouse tissues was examined by Northern blot hybridization. The filter, which had been blotted with total RNAs from mouse adult liver, kidney, heart, spleen, and lung, was probed with the 32P-labeled mouse PHBP cDNA. A strong signal was detected in the liver and kidney, but not in the heart, spleen and lung (Fig. 4).

To determine whether mouse PHBP was encoded by a single or multiple gene(s), Southern blot hybridization was performed with DNA from mouse liver. The hybridization pattern with cDNA encoding this protein was rather simple (Fig. 5). It is likely that the gene encoding PHBP is a single copy.

DISCUSSION

In this study, we have isolated a cDNA clone which encodes mouse plasma hyaluronan-binding protein (PHBP) from a normal mouse liver cDNA library constructed in the azap phage vector. As shown in Fig. 2, 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 1675–1677 of a stop codon (TGA). The 1674-bp open reading frame codes for a polypeptide of 558 amino acid residues, which is two residues shorter than that of human PHBP. The amino acid sequence of mouse PHBP which was predicted from the nucleotide sequence shows reasonable homology to that of human PHBP (Fig. 3). The overall amino acid sequence of the mouse PHBP precursor shows significant sequence homology (about 78%) to that of human PHBP, and especially, in the putative small subunit containing

---

Fig. 1. Schematic Representation of the Restriction Map and Sequence Strategy of Mouse PHBP cDNA

The open box designates the PHBP-coding sequence oriented in the 5' to the 3' direction. The shaded box indicates the putative signal peptide. The nucleotide scale is indicated at the top. The sequence strategy used for the determination of mouse PHBP cDNA sequence is shown as arrows. Horizontal arrows indicate the direction and extent of the sequencing. E, EcoRI; P, PvuI; S, SalI.

---

NII-Electronic Library Service
Fig. 2. Nucleotide Sequence of Mouse PHBP cDNA and Its Predicted Amino Acid Sequence from the Nucleotide Sequence

The nucleotide sequence of the cDNA for PHBP is shown on the first line, and the amino acid sequence predicted from the nucleotide sequence is shown on the second line. 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 single underline.

Fig. 3. Comparison of the Amino Acid Sequence of Mouse PHBP with That of Human PHBP

Sequences were aligned using the Genetyx program (SDC, Tokyo). Amino acid residues identical with human PHBP are shown by black boxes. Asterisks indicate cysteine residues conserved in the two sequences. The vertical arrow indicates the putative site for proteolytic cleavage. The arrow heads indicate the consensus amino acid residues for serine protease.

serine protease domain the homology is more than 88%, and three amino acid residues, His360, Asp403, and Ser507, which are essential for serine protease, are completely conserved in this region. All cysteine residues contained in mature human PHBP were completely conserved in mouse PHBP as well, suggesting that mouse PHBP is predicted to have a similar secondary structure to that of human PHBP. Therefore, although the N-
mRNA registered in DDBJ by Kitamura et al. (accession number, D49742), except for human PHBP. We therefore concluded that this protein was mouse PHBP. Also, the nucleotide sequences of five cDNA clones (accession numbers AA268125, AA237499, AA108230, AA108229 and AA217892) are completely included in our nucleotide sequence. These nucleotide sequences have been registered as EST in GenBank, and were cloned from the house mouse musculus cDNA library by Marra et al. Like human PHBP, this protein may be produced as a single peptide, which is cleaved into 50 and 25 kDa peptides. The putative cleavage site to form a hetero-dimer (—Lys310–Arg311–Ile312–Tyr313–Gly314–Gly315–Phe316–Lys317–Ser318—), which was identified in human PHBP, is well conserved in the amino acid sequence predicted from the nucleotide sequence of mouse PHBP cDNA (Fig. 3). The molecular masses of two peptides predicted from the cleavage at this site are calculated to be 32 607.65 and 27 208.48 Da, respectively.

The results of Northern blot analysis showed that PHBP was expressed in the liver and kidney, and not in the heart, spleen and lung (Fig. 4). The signal of the kidney was stronger than that of the liver. The expression of this protein was also detected in human adult liver and kidney, but in human, the signal of liver was stronger than that of kidney.15 And like human PHBP, the two hybridization signals were detected in liver, indicating the presence of transcripts of two different sizes.

Southern blot analysis indicates that the gene which codes mouse PHBP seems to be a single copy, and that the gene may be spread over 13 kb (Fig. 5).

Although the function of PHBP has not yet been clarified, as previously reported, PHBP may exhibit its biological functions as a serine protease. The protease which activates the PHBP precursor to form a hetero-dimer and the natural substrate of active PHBP should be elucidated. For this purpose, mouse PHBP cDNA may come to be a good tool, for example, to examine the induction of this protein in mouse under the various experimental conditions or to create a recombinant PHBP precursor to search for the activator of PHBP.

REFERENCES