

## Mouse and Human SHPS-1: Molecular Cloning of cDNAs and Chromosomal Localization of Genes

Takuji Yamao, Takashi Matozaki,<sup>1</sup> Kenji Amano,\* Yoichi Matsuda,\* Nobuaki Takahashi,† Fukashi Ochi, Yohsuke Fujioka, and Masato Kasuga

*Second Department of Internal Medicine, Kobe University School of Medicine, Kusunoki-cho, Chuo-ku, Kobe 650, Japan;*

*\*Laboratory of Animal Genetics, School of Agricultural Sciences, Furo-cho, Chikusa-ku, Nagoya 464-01, Japan;*

*and †Kirin Brewery Co. Ltd., Central Laboratories for Key Technology, 1-13-5 Fukuura, Kanazawa-ku, Yokohama, Kanagawa 236, Japan*

Received December 24, 1996

**SHPS-1 (SHP substrate-1) is a glycosylated receptor-like protein with three immunoglobulin-like domains in its extracellular region and four YXX(L/V/I) motifs, potential tyrosine phosphorylation and SRC homology 2 (SH2) domain binding sites, in its cytoplasmic region. Various mitogens and cell adhesion induce tyrosine phosphorylation of SHPS-1 and its subsequent association with SHP-2, an SH2 domain-containing protein tyrosine phosphatase, suggesting that SHPS-1 plays a role in cell signaling in response to both growth factors and cell adhesion. The mouse and human cDNAs encoding SHPS-1 have now been isolated. The deduced amino acid sequences of rat, human, and mouse SHPS-1 show identities of 65 to 81 %. In addition to the SH2 domain binding site, a proline-rich putative SH3 domain binding site was detected in the cytoplasmic region of SHPS-1. Northern blot analysis revealed that human SHPS-1 mRNA is most abundant in brain and that the mouse mRNA is present in embryos as early as day 7. Fluorescence in situ hybridization localized the SHPS-1 gene to human chromosome 20p13 and the F3 band of mouse chromosome 2. Furthermore, interspecific backcross analysis placed the mouse SHPS-1 locus 5.0 centimorgans distal and 1.4 centimorgans proximal to the microsatellite markers *D2Mit63* and *D2Mit19*, respectively, in a region associated with the mutations coloboma (*Cm*), lethal milk (*lm*), and well-haarig (*we*). © 1997 Academic Press**

SHP-2 (also named SH-PTP2, PTP1D, Syp, PTP2C, SH-PTP3, and SAP-2) (1) is a non-transmembrane pro-

tein tyrosine phosphatase (PTPase<sup>1</sup>) that contains two SRC homology 2 (SH2) domains (2) and is thought to participate in intracellular signaling in response to various growth factors, such as platelet-derived growth factor (PDGF), epidermal growth factor (EGF), and insulin (3). SHP-2 binds to tyrosine-phosphorylated PDGF receptors in response to PDGF (4,5), as well as to tyrosine-phosphorylated insulin receptor substrate-1 in response to insulin (6). In addition, SHP-2 has been suggested to mediate the activation of RAS or mitogen-activated protein kinase, and thereby to induce DNA synthesis, in response to PDGF (7), EGF (8), insulin (9–11), and fibroblast growth factor (12). However, the mechanism by which SHP-2 mediates growth factor-induced RAS activation is not clear.

We have recently cloned a rat cDNA that encodes a glycosylated receptor-like protein of ~115 to 120 kDa termed SHPS-1 (SHP substrate-1)(13). SHPS-1 contains three immunoglobulin-like domains in its extracellular region and four YXX(L/V/I) motifs, potential tyrosine phosphorylation and SH2 domain binding sites, in its cytoplasmic region. Various mitogens, including serum, insulin, and lysophosphatidic acid, or cell adhesion induce tyrosine phosphorylation of SHPS-1 and its subsequent association with SHP-2 in cultured cells. The insulin receptor kinase or SRC kinase may directly catalyze the tyrosine phosphorylation of SHPS-1 in vivo and in vitro (13,14). Furthermore, the extent of tyrosine phosphorylation of SHPS-1 was greatly increased when a catalytically inactive SHP-2 was overexpressed in cultured cells (13–15), suggesting that SHPS-1 may be a physiological substrate for SHP-2. SHPS-1 may thus act as a docking protein and induce translocation of SHP-2 from the cytosol to the plasma membrane in response to mitogens and cell adhesion. After binding to SHPS-1, SHP-2 may catalyze its dephosphorylation and dissociate, in the process

<sup>1</sup> To whom correspondence should be addressed. Fax: 81-78-382-2080; E-mail: matozaki@med.kobe-u.ac.jp.

The abbreviations used are: PTPase, protein tyrosine phosphatase; SH<sub>2</sub> domains, src-homology-2 domains; PDGF, platelet-derived growth factor; EGF, epidermal growth factor; PCR, polymerase chain reaction; FISH, fluorescence in situ hybridization.

possibly activating an SOS-like guanine nucleotide exchange protein near the plasma membrane by also catalyzing its tyrosine dephosphorylation.

In an attempt to clarify further the physiological function of SHPS-1, we have now cloned the corresponding mouse and human cDNAs. In addition, we have determined the chromosomal locations of the human and mouse SHPS-1 genes.

## MATERIALS AND METHODS

**Cloning of mouse and human SHPS-1 cDNAs.** A mouse brain Lambda ZAP II cDNA library (Stratagene) was screened with a 1.1-kb fragment (nt 1 to 1104) of rat SHPS-1 cDNA (13) as a probe. The probe was labeled with [ $\alpha$ - $^{32}$ P]dCTP with the use of a Random Primer DNA Labeling Kit (Takara). Hybridization was performed at 42°C in a solution containing 50% (v/v) formamide, 5× standard saline citrate (SSC), 5× Denhardt's solution, 0.1% SDS, and denatured salmon sperm DNA (100  $\mu$ g/ml), after which the nitrocellulose filters were washed four times with 2× SSC and 0.05% SDS at room temperature and once with 1× SSC and 0.1% SDS at 55°C. Eighteen positive clones isolated from the cDNA library were subcloned into the pBluescript phagemid from Lambda ZAP II with the use of the ExAssist/SORL system (Stratagene) and sequenced by dideoxy termination methods with a Sequencing Pro kit (TOYOBO). A human brain  $\lambda$ gt10 cDNA library (Clontech) was also screened as described above to isolate human SHPS-1 cDNA.

**Northern blot analysis.** A human MTN blot (Clontech) containing 2  $\mu$ g of poly(A)<sup>+</sup> RNA from various tissues was incubated consecutively with a  $^{32}$ P-labeled 1.6-kb human SHPS-1 cDNA fragment and a  $^{32}$ P-labeled human  $\beta$ -actin cDNA fragment as previously described (13). A mouse embryo MTN blot containing 2  $\mu$ g of poly(A)<sup>+</sup> RNA from various developmental stages was also incubated consecutively with a  $^{32}$ P-labeled 3.4-kb mouse SHPS-1 cDNA fragment, a  $^{32}$ P-labeled 2.0-kb human SHP-1 cDNA (16), and a  $^{32}$ P-labeled 2.5-kb human SHP-2 cDNA (9).

**Fluorescence in situ hybridization (FISH) with mouse and human chromosomes.** The direct R-banding FISH method was used for chromosomal assignment of the mouse SHPS-1 gene. R-banded chromosomes were prepared and FISH was performed as described previously (17,18). Mitogen-stimulated cultures of mouse splenocytes were synchronized by thymidine block, and were exposed to 5-bromo-deoxyuridine during the late replication stage for differential replication staining after the release from excess thymidine. R-band staining was performed by exposure of chromosome slides to ultraviolet light after staining with Hoechst 33258. The slides were incubated at 65°C for 2 h, after which chromosomes were denatured at 70°C in 2× SSC containing 70% formamide and then dehydrated in a series of ethanol solutions [70, 85, and 100% (v/v)] at 4°C.

A 3.4-kb mouse SHPS-1 cDNA fragment inserted in the Eco RI site of pBluescript was labeled by nick translation with biotin 16-dUTP (Boehringer Mannheim), precipitated together with salmon sperm DNA and *Escherichia coli* tRNA with ethanol, and denatured at 75°C for 10 min in 100% formamide. The probe was then mixed with an equal volume of hybridization solution to yield final concentrations of 50% formamide, 2× SSC, 10% dextran sulfate, and bovine serum albumin (BSA) (1 mg/ml) (Sigma), and 20  $\mu$ l of the mixture (containing 250 ng of labeled DNA) were transferred to each of the chromosome slides, which were then covered with plastic film and incubated overnight at 37°C. The slides were washed in 2× SSC containing 50% formamide for 20 min at 37°C, and then in 2× SSC and 1× SSC for 20 min each at room temperature. After rinsing in 4× SSC, the slides were incubated under coverslips for 1 h at 37°C with antibodies to biotin (Vector Laboratories) at a 1:500 dilution in

4× SSC containing 1% BSA. They were then washed with 4× SSC, 4× SSC containing 0.1% Nonidet P-40, and 4× SSC for 5 min each, incubated for 1 h at 37°C with fluorescein-conjugated antibodies to goat IgG (Nordic Immunology) at a 1:500 dilution in 4× SSC, and then washed in a shaker for 10 min each with 4× SSC, 0.1% Nonidet P-40 in 4× SSC, and 4× SSC. After rinsing with 2× SSC, the slides were finally stained with propidium iodide (0.75  $\mu$ g/ml), observed at an excitation wavelength of 450 to 490 nm (Nikon filter set B-2A) and an emission wavelength of 365 nm (UV-2A), and photographed with Kodak Ektachrome ASA100 film.

FISH localization of the human SHPS-1 gene was performed by See DNA Biotech (Downsview, Ontario, Canada) essentially as described previously (19). A 1.6-kb human SHPS-1 cDNA fragment inserted into pBluescript was used for biotinylation with dATP with the BRL BioNick labeling kit.

**Linkage mapping with interspecific backcross mice.** Recombinant mice were generated by mating males of feral-derived stocks of *M. spretus* with C57BL/6J females and backcrossing the F<sub>1</sub> females with *M. spretus* males (20). Genomic DNA prepared from the kidneys of the backcross mice was digested with restriction endonuclease, and the resulting fragments were separated by electrophoresis on a 0.8% agarose gel and transferred to a nylon membrane (Bio-Rad). The genotype of individual mice was determined by Southern hybridization with mouse SHPS-1 cDNA. Microsatellite DNA marker loci for linkage analysis were chosen on the basis of the results of cytogenetic mapping by FISH, and obtained from Research Genetics (Huntsville, AL, USA). All PCRs were performed in a final volume of 15  $\mu$ l containing 75 ng of genomic DNA and 15 pmol of each oligonucleotide primer. The amplification protocol consisted of incubation at 94°C for 5 min; 30 cycles of 94°C for 30 s, 55°C for 30 s, and 72°C for 30 s; and 72°C for 5 min. The PCR products were analyzed by agarose gel electrophoresis.

## RESULTS AND DISCUSSION

### Cloning of Mouse and Human SHPS-1 cDNAs

Screening of mouse brain cDNA library with a 1.1-kb fragment of rat SHPS-1 cDNA yielded 10 overlapping clones that encompassed a mouse full-length SHPS-1 cDNA. The mouse cDNA comprised 3979 bp and contained a single open reading frame of 1527 bp [nucleotides (nt) 410 to 1936] that encoded a protein of 509 amino acids (Fig. 1A). The first ATG codon (nt 410 to 412) matched the Kozak consensus sequence for translation initiation and presumably represents the start codon. The 3' noncoding region of the cDNA contains a typical polyadenylation signal (AATAAA). The NH<sub>2</sub>-terminal 28 amino acids are hydrophobic and likely constitute a signal peptide, and residues 344 to 368 correspond to a transmembrane domain. The extracellular region of mouse SHPS-1 contains three homologous immunoglobulin (Ig)-like domains (V, C1, and C1, respectively), as does that of rat SHPS-1 (13), and 14 potential N-linked glycosylation sites (NXS or NXT, where X is any amino acid). The cytoplasmic region of mouse SHPS-1 contains four potential tyrosine phosphorylation sites (Y408ADL, Y432ASI, Y449ADL, and Y473ASV), as does that of the rat protein (13). Thus, the overall structure of mouse SHPS-1 is similar to that of rat SHPS-1. The cytoplasmic region of mouse

## A

```

GGGCCGCCCTCGCGCTGGCCTCGCGACGGCTCCGCACAGCCCGCACTGCTCTGCGAGCT      ----- 300 bp ----- 360
CCGAGCGCGCTCACCGCCGCTCTCCCTCTGCTCTGTCAGCCGCGGCCCATGGAGCCCGCGGCCCGGCCCTGGCCGCTAGGGCCGCT 450
                                     M E P A G P A P G R L G P L
GCTGCTCTGCTGCTGCTCTCCGCGCTCTGTTTCTGTACAGGAGTACACGGGAAAGAACTGAAGGTGACTCAGCCTGAGAAATCAGTGT 540
L L C L L L S A S C F C T G V T G K E L K V T Q P E K S V S 16
TGTTGCTGCTGGGATTGACCGTTCTGAAGTGCACCTTTGACCTCCTTGTGCGGTGGGACCCATTAGTGTTACAGAGGATAGGGCA 630
V A A G D S T V L N C T L T S L L P V G P I K W Y R G V G Q
AAGCCGGCTGTTGATCTACAGTTTCACAGGAGAACACTTTCTCGAGTTACAAATGTTTCAGATGCTACTAAGAGAAACAATATGGACT 720
S R L L L L Y S F T G E H F P R V T N V S D A T K R N N M D F
TTCCATCCGTATCAGTAATGTACCCCGAAGATGCCGGTACCTACTACTGTGTGAAGTTCAGAAAGGACCATCAGAGCCTGACACAGA 810
S I R I S N V T P E D A G T Y Y C V K F Q K G P S E P D T E
AATACAATCTGGAGGGGAACAGAGGTCTATCTACGCCAAACCTTCTCCACGGAGGTATCCGGCCAGACAGAGGGGCATACCTGA 900
I Q S G G G Y T E V Y V L A K P S P P E V S G P A D R G I P D
CCGAAAGTGAACCTTACCTGCAAGTCTCTATGGCTTCTCTCCCGGAATATCACCTGAAGTGGTTCAAAGATGGGCAAGAACCTCACCA 990
Q K V N F T C K S H G F S P R N I T L K W F K D G Q E L H H
CTTGGAGACCCCGTGAACCTAGTGGAAAGATGTCTCTACAACATCTCCAGCACAGTCAGGGTGGTACTAACTCCATGATGTTGA 1080
L E T T V N P S G K N V S Y N I S S T V R V V L N S M D V H
TTCTAAGGTCATCTGCGAGGTAGCCACATACCTTGGATAGAAGCCCTTTCGTGGGATTGTAACTGTCTAACTTCATCCGAGTTTC 1170
S K V I C E V A H I T L D R S P L R G I A N L S N F I R V S
ACCCACCGTGAAGGTACCCCAACAGTCCCGACGTCAATGAACAGGTGAACCTCACCTGCCGGGTGAGAGGTTTACCCCGAGGATCT 1260
P T V K V T Q Q S P T S M N Q V N L T C R A E R F Y P E D L
CCAGCTGATCTGGCTGGAGAATGGAAACGTATCACGGAATGACACGCCCAAGAATCTCACAAAGAACACGGATGGGACCTATAATTACAC 1350
Q L I W L E N G N V S R N D T P K N L T K N T D G T Y N Y T
AAGCTTGTCTCGTGAAGTCATCTGCTCATAGAGAGGACGTGGTGTTCACGTGCCAGGTGAAGCACGACCAACAGCCAGCGATCACCCG 1440
S L F L V N S S A H R E D V V F T C Q V K H D Q Q P A I T R
AAACATCCCGTGGGACTTGGCCACTCGAGTGATCAAGGGAGCATCAAACTTCCCTGGTAATAATGCTACCCACAAGTGAATGT 1530
N H T V L D G L A H S D Q G S M Q T F P G N N A T H N W N V
CTTTCATCGGTGGGCGTGGCGTGTGCTTGTGCTGCTGATGGCTGCTCTACCTCCTCCGGATCAACAGAGAAGGCCAA 1620
F I G V G V A C A L L V V L L M A A L Y L L R I K Q K K A K 376

```

----- 1900 bp -----

```

GGGGTCAACATCTTCCACACGGTTCACGAGCCCGAGAAGAACGCCAGGGAATAACCCAGATCCAGGACACAAATGACATCAACGACAT 1710
G S T S S T R L H E P E K N A R E I T Q I Q D T N D I N D I 406
CACATACGACACCTGAATCTGCCCCAAGAGAAGAGCCCGCACCCTGGGCCCTGAGCCTAACACCCACAGAAATGCAAGCATTGA 1800
T Y A D L N L P K E K K P A P R A P E P N N H T E Y A S I E 436
GACAGGCAAGTGCCTAGGCCAGAGGATACCCCTACCTATGCTGACCTGGACATGGTCCACCTCAGCGGGGCACAGCCAGCCCCCAAGCC 1890
T G K V P R P E D T L T Y A D L D M V H L S R A Q P A P K P 466
TGAGCCATCTTCTCAGAGTATGCTAGTGTCAGGTCCAGAGGAAGTGAATGGGCTGTGCTGTACTAGGCCCATCCCCACAAGTTT 1980
E P S F S E Y A S V Q V Q R K * 481
TCTTGCTACATGAGTGG ----- CAGGCCCTGTTTGTACATAGTTGCAACTTAAATTTT 3940
GCTTGCAAAATATTTTGTAAATAAGATTCTCGGTAAC 3979

```

## B

GTACAGTCTTTG  
V Q S L

**FIG. 1.** Nucleotide and deduced amino acid sequences of mouse SHPS-1 cDNA. A. The putative signal peptide is indicated by a solid underline, and the transmembrane region by a bold underline. The AATAAA box near to the polyadenylated 3' end of the cDNA is indicated by dots. The three Ig-like domains are indicated by dotted (the V-like domain) and dashed (the C1-like domains) lines; cysteine residues that potentially form disulfide bonds in these domains are boxed. The 14 potential N-linked glycosylation sites (NXS or NXT) are in boldface. The putative binding sites for SH2 domains of SHP-2 are in boldface with underlines and dots. A potential binding site for SH3 domains is indicated by a double underline. The numbers to the right of each row refer to nucleotide position (upper) or to amino acid position in the predicted mature SHPS-1 protein (lower). B. The 12-bp insertion between nt 1681 and 1682 that is likely the result of alternative splicing and its deduced amino acid sequence. The insertion site is indicated by an arrowhead in (A).

SHPS-1 also contains a polyproline motif, XP $\phi$ PXXP ( $\phi$ , hydrophobic residue) (residues 418 to 424), that represents a potential SH3 domain binding site (21), suggesting that SHPS-1 may interact not only with an SH2 domain-containing PTPase but also with an SH3 domain-containing signaling molecule such as a SRC family kinase or GRB2 (21).

One of the mouse SHPS-1 cDNA clones (clone C) contained a 12-bp insertion (GTACAGTCTTTG) between

nt 1681 and 1682 (Fig. 1B). This insertion does not result in a frame shift, and examination of the genomic organization of the mouse gene revealed a splice junction (22) between nt 1681 and 1682 (data not shown), suggesting that the additional 12-bp insertion in clone C is due to alternative splicing.

Screening of a human brain cDNA library with a <sup>32</sup>P-labeled rat SHPS-1 cDNA probe yielded 1625-bp human SHPS-1 cDNA that contained a single open

Rat	: MEPA GPAPGR LGPLLFCLLLSASCFCAGASGK-ELKVTQADKSVSVAAGDSATLN <b>ETVSS</b>	31
Mouse	: .....L.....T.VT...-.....PE.....TV.....LT.	31
Human	: .....-.....A...AWS.VA.EE..Q.I.P.....E..I.H...T.	32
Rat	: LTPVGPIKWEFKGEGONRSFIYSFIGGEHFPRI <b>TVNS</b> DATKRNNMDFESICIS <b>SVT</b> PEDAGT	91
Mouse	: .L.....YR.V...S.LL...T-.....V.....R.....	90
Human	: .I.....Q..R.A.PA.EL..NQKE-.....V.T..ES...E.....S...I..A....	91
Rat	: YY <b>Q</b> VKFQKGIVEPDTEIKSGGGTTLTVLAKPSSPEVSGPDSRGSPGQTV <b>NET</b> QKSYGFSP	151
Mouse	: .....PS.....Q.....EV.....P.....AD..I.D.K.....H....	150
Human	: .....R...-S-...F...A..E.S.R...A.V...AA.AT.QH..S...E.H....	149
Rat	: <b>RNIT</b> LKWLKDGKELSHLETTIS-SKS <b>SVNSYNIS</b> TSVSVKLSPEDIHSRV <b>IDE</b> VAHVTLLEG	210
Mouse	: .....F...Q...H.....VNP.GK.....R.V.NSM.V..K.....I..DR	210
Human	: .D....F.N.N..SDFQ.N.DPV.ES...S.H..AK.V.TR..V..Q.....Q.	209
Rat	: RPL <b>NGTAN</b> FSNIIRVSPTLKITQQPLTPASOV <b>NLT</b> QVOVKFYPKALQINWLENG <b>NLS</b> RTD	270
Mouse	: S..R.I..L..F.....V.V...SP.SMN.....RAER...ED..I.....V....	270
Human	: D..R....L.ET...P...EV...VRAEN...V...R...QR...T.....V...E	269
Rat	: KPEHFTDNDRGTY <b>NYT</b> SLFLVN <b>SSA</b> HREDVV <b>ET</b> QVEHDSQPAITEN <b>NHT</b> VRAFAHSSSGG	330
Mouse	: T.K <b>NL</b> .K.T.....K..Q....R...LGL...DQ.	330
Human	: TASTV.E.K....WM.WL...V....D..KL.....G...VSKS.DLKVS..PKEQ.	329
Rat	: SMETIPDNNAYYN <b>WN</b> VFIGVGVACALLV <b>VLLMA</b> LYLLRIKQKKAKGSTSSTRLHEPEKN	390
Mouse	: ..Q.F.G...TH.....	390
Human	: .N-.AAE.TGSNER.IY.V...V.T...A.....V..R...Q.....	388
Rat	: AREITQIQDTNDINDITY <b>AD</b> LNL <b>PKEK</b> K <b>PAP</b> RVPEPN <b>NHTE</b> Y <b>ASI</b> ETGK <b>LP</b> EDTLTY <b>YA</b>	450
Mouse	: .....A.....V.....	450
Human	: .....- - - - -G...QAA.....Q.SPQ.AS.....	443
Rat	: <b>DL</b> DMVHLNRA--QTPK <b>PE</b> PSFSEY <b>YASV</b> QVQRK	481
Mouse	: .....S...- - -A.....	481
Human	: .....TPK..A.....P..	476

**FIG. 2.** Alignment of the deduced amino acid sequences of rat, mouse, and human SHPS-1. Dots within the sequences of mouse or human SHPS-1 represent amino acids that are identical to those in the rat protein. Gaps introduced for optimal alignment are indicated by hyphens. Other features of the three proteins are indicated as in Fig. 1.

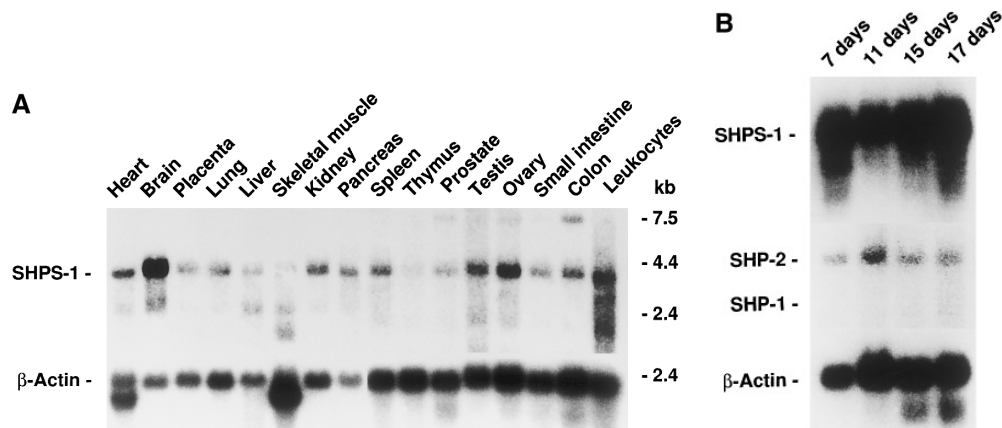
reading frame encoding a protein of 503 amino acids (Fig. 2). The amino acid sequences of rat, human, and mouse SHPS-1 share overall identities and similarities of ~65 to 80% and >90%, respectively (Fig. 2, Table 1). The homology between mouse and rat proteins is greater than that of either protein to human SHPS-1, with regard to both the overall structure and specific regions. The four potential tyrosine phosphorylation and SH2 domain binding sites of SHPS-1 (YADL, YASI, YADL, and YASV) are completely conserved among all three species. Whereas residues 418 to 424 of rat

SHPS-1 (KPAPRV) and mouse SHPS-1 (KPAPRAP) completely matched the SH3 domain binding site (XPφPXXP) motif, residues 411 to 417 (KPAPQAA) of human SHPS-1 only partially match this consensus sequence. The identity of the extracellular domain of SHPS-1 among the three species, is less than that of the cytoplasmic region (Table 1). Although the three Ig-like structures are conserved among the three species, only 4 of the 15 potential N-linked glycosylation sites in rat SHPS-1 are also present in the human protein.

**TABLE 1**  
Sequence Homology of Rat, Mouse, and Human SHPS-1

Comparison	Overall		Extracellular domain		Transmembrane domain		Cytoplasmic domain	
	I	S	I	S	I	S	I	S
Rat-mouse	80.7	95.2	73.8	92.7	100	100	96.5	100
Rat-human	65.1	91.1	60.2	91.2	68.0	100	79.1	89.6
Mouse-human	66.2	91.5	61.9	91.9	68.0	100	80.0	89.6

The amino acid sequences of rat, mouse, and human SHPS-1 were compared with the GENETYX computer program. The percentages of amino acid residues that are identical (I) or similar (S) for the full length proteins or the indicated domains were calculated for each pair of species. Amino acids considered similar are R and K; E and D; T and S; V, I, L, and A; and Y and F.



**FIG. 3.** Northern blot analysis of SHPS-1 mRNA in various human tissues and during mouse embryogenesis. A. Poly(A)<sup>+</sup> RNA from various human tissues was subjected to Northern blot analysis with <sup>32</sup>P-labeled human SHPS-1 (upper panel) and, subsequently,  $\beta$ -actin (lower panel) cDNAs as probes. The positions of molecular size standards (in kilobases) are shown on the right. B. Poly(A)<sup>+</sup> RNA from mouse embryos at various stages of development was subjected to Northern blot analysis with <sup>32</sup>P-labeled mouse SHPS-1 (upper panel), human SHP-1 and human SHP-2 (middle panel), and human  $\beta$ -actin (lower panel) cDNAs as probes.

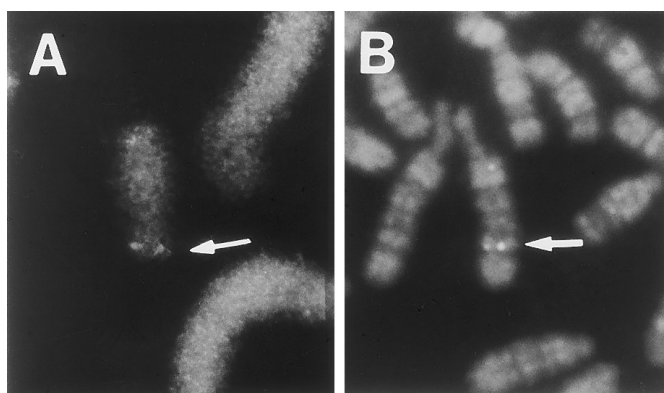
#### Northern Blot Analysis

Northern blot analysis of polyadenylated [poly(A)<sup>+</sup>] RNA revealed that the major human SHPS-1 mRNA (~4.2 kb) was present in all tissues examined, being most abundant in brain and with substantial amounts also observed in heart, spleen, testis, ovary, and peripheral blood leukocytes (Fig. 3A). SHP-2 is enriched in synaptic membranes, indicating that it may be important in neuronal function (23). Furthermore, immunohistochemical staining of human brain cortex showed that SHPS-1 is localized in neuronal cells (Fukunaga, K., Matozaki, T. and Kasuga M., unpublished data), suggesting that both SHPS-1 and SHP-2 may play a role in neurons. In addition to the 4.2-kb SHPS-1 mRNA, transcripts of smaller size (~2.4 kb or <2.4 kb) were also observed in heart, muscle, testis, and leukocytes, suggesting that multiple alternatively spliced forms of the primary SHPS-1 transcript may exist.

Because SHP-2 mRNA is present in the early mouse embryo (4), we also determined the abundance of SHPS-1 mRNA in the mouse embryo. SHPS-1 mRNA was detected together with SHP-2 mRNA in 7-day embryos, and both were present through 17 days (Fig. 3B). SHP-1, another SH2 domain-containing PTPase (1), forms a complex with tyrosine-phosphorylated SHPS-1 in v-Src-transformed rat fibroblasts (13,24). SHP-1 is predominantly expressed in hematopoietic cells and is thought to play an inhibitory role in cytokine-stimulated proliferation of these cells (1,3). In contrast to SHP-2 mRNA, SHP-1 transcripts were virtually undetectable in 7- to 17-day-stage mouse embryos (Fig. 3B). Homozygous SHP-2 knockout mice die before day 10.5 of embryonic development (25). Thus, these results suggest that SHPS-1 together with SHP-2 may play an important role in early embryogenesis.

#### FISH Mapping of Human and Mouse SHPS-1 Genes and Interspecific Linkage Mapping in Mouse

The chromosomal location of the human SHPS-1 gene was determined by FISH. A total of 100 metaphase cells was examined, 73% of which exhibited specific hybridization signals at chromosome 20p13 (Fig. 4A). Thus, these data localize the human SHPS-1 gene to chromosome 20p13. The human SHPS-1 cDNA appears to be the same as the CCA53 cDNA and the CCA12 cDNA, both of which were recently cloned by direct screening of a human brain cDNA library with an oligonucleotide containing CCA repeats (26). However, the nucleotide sequence of the CCA53 cDNA sub-

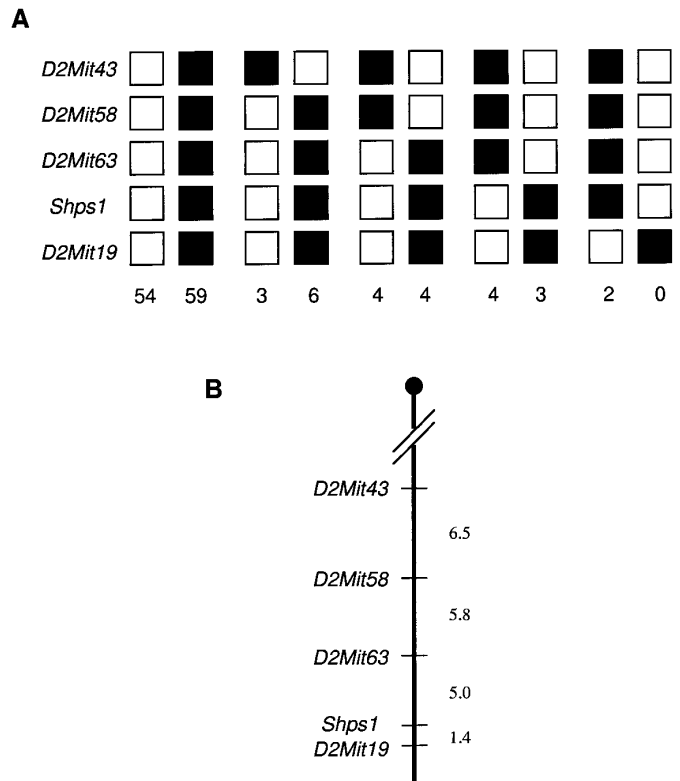


**FIG. 4.** FISH localization of the human and mouse SHPS-1 gene. A. FISH was performed as described in Materials and Methods. Double fluorescence spots were detected on R-banded chromosome 20 at the p13 region (arrow). B. The mouse SHPS-1 gene was localized on mouse R-banded chromosomes by FISH with a biotinylated 3.4-kb cDNA fragment as probe. The hybridization signals (arrows) are located in the F3 band of chromosome 2. The metaphase spreads were photographed with a Nikon B-2A.

mitted to the database appears to contain multiple errors and the function of its encoded protein was not identified. We also detected 17 CCA repeats ~400 bp downstream from the stop codon in the 3' untranslated region of our human SHPS-1 cDNA clone (data not shown). Expansion of trinucleotide repeats has been suggested to be responsible for several hereditary neurological disorders, including myotonic dystrophy, fragile X syndrome, spinal and bulbar atrophy, Huntington's disease, spinocerebellar ataxia, Machado-Joseph disease, and dentatorubral-pallidoluysian atrophy (26,27). According to the OMIM database, no hereditary neuronal disorder has so far been mapped to chromosome 20p13. However, given that human SHPS-1 mRNA is most abundant in the brain, expansion of the triplet repeats in the SHPS-1 gene may well underlie some neurological disease.

The chromosomal localization of the mouse SHPS-1 gene was also determined by direct R-banding FISH with a mouse cDNA fragment as the probe. The mouse gene was localized to the R-positive F3 band of chromosome 2 (17,28) (Fig. 4B), a locus that corresponds well to human chromosome 20p13 (29), the side of the human gene.

To characterize further the mouse SHPS-1 locus (*Shps1*), we performed fine linkage mapping by interspecific backcross analysis with progeny derived from the mating of C57BL/6 mice and *Mus spretus*. Genomic DNAs from C57BL/6 × *M. spretus* F<sub>1</sub> mice was digested separately with each of six different restriction endonucleases (Apa I, Bam HI, Eco RI, Hind III, Kpn I, and Pst I) and analyzed by Southern blot hybridization for informative restriction fragment length variants (RFLVs) with a 238-bp mouse SHPS-1 cDNA fragment as probe. RFLVs between C57BL/6 and *M. spretus* were detected by digestion with Eco RI or Kpn I (data not shown). The RFLVs apparent with Kpn I digestion in either C57BL/6 or *M. spretus* were used to follow the segregation of *Shps1* in a total of 139 interspecific backcross mice. We examined the concordance of the segregation of RFLVs identified by Southern blot hybridization with the segregation of four microsatellite marker loci: *D2Mit43*, *D2Mit58*, *D2Mit63*, and *D2Mit19*. Gene order was determined by minimizing the number of recombination events required to explain the allele distribution patterns. Comparative pairwise loci analysis showed that the gene order and recombination frequency, expressed as genetic distance (+SE) in centimorgans (cM) and the number of recombinants, are as follows: centromere (Hc2) – *D2Mit43* – [ $6.5 \pm 2.1$  (9/139)] – *D2Mit58* – [ $5.8 \pm 2.0$  (8/139)] – *D2Mit63* – [ $5.0 \pm 1.9$  (7/139)] – *Shps1* – [ $1.4 \pm 1.0$  (2/139)] – *D2Mit19* – telomere (Fig. 5). We have compared our interspecific linkage map of mouse chromosome 2 with a recent composite mouse linkage map showing the location of uncloned mouse mutations (29). The com-



**FIG. 5.** Linkage mapping of mouse *Shps1* on chromosome 2. **A.** Segregation patterns of mouse *Shps1* and flanking microsatellite markers (*D2Mit43*, *D2Mit58*, *D2Mit63*, and *D2Mit19*) in backcross mice. Each column represents the chromosome identified in the backcross progeny that was inherited from the C57BL/6 × *M. spretus* F<sub>1</sub> parent. The open squares represent the presence of the *M. spretus* allele, and the closed squares the C57BL/6 allele. The number of offspring inheriting each type of chromosome is given at the bottom of each column. **B.** Partial linkage map of chromosome 2 showing the location of *Shps1* in relation to the flanking DNA markers. Recombination distances between loci are shown in centimorgans to the right of the chromosome.

parison revealed that three spontaneous mutations, wellhaarig (*we*), lethal milk (*lm*), and coloboma (*Cm*), map close to *Shps1*. Mice heterozygous for the semi-dominant mutation *Cm* exhibit ocular dysmorphology and neuronal abnormalities (30). The size of the eye is reduced at birth, and a ventral segment of the choroid, constituting ~25% of the whole, is absent. In adults, the eyeballs are rotated ventrally so that the pupil may be partly hidden by the lower lid. The animals also exhibit abnormal posture as well as head shaking and circling (head bobbing). In addition, *Cm* mice are extremely hyperactive, with spontaneous locomotor activity three times that of normal mice (31). *Cm* mice carry a 1- to 2-cM deletion on chromosome 2 that encompasses the genes encoding the nerve terminal protein SNAP-25 and phospholipase C- $\beta$ 1 (31,32). The incorporation of a transgene encoding SNAP-25 into the *Cm*

strain rescued the hyperactivity phenotype but not the ocular dysmorphology and head bobbing (33). Given that *D2Mit19*, a microsatellite marker located near to *Shps1*, is included in the *Cm* deletion (32), and that SHPS1 mRNA is abundant in the nervous system, the deletion or mutation of the SHPS-1 gene might contribute to the *Cm* phenotype.

## ACKNOWLEDGMENTS

Supported by a grant-in-aid for Cancer Research and a grant-in-aid for Scientific Research from the Ministry of Education, Science, and Culture of Japan, and by a grant from the Ciba-Geigy Foundation (Japan) for the Promotion of Science. The nucleotide sequence data reported in this paper have been deposited in the DDBJ, EMBL, and GenBank databases under the accession numbers D86043 (human SHPS-1 cDNA), D87967 (mouse SHPS-1 cDNA), and D87968 [mouse SHPS-1 cDNA with the 12-bp insertion (long type)].

## REFERENCES

- Adachi, M., Fischer, E. H., Ihle, J., Imai, K., Jirik, F., Neel, B., Pawson, T., Shen, S. H., Thomas, M., Ullrich, A., and Zhao, Z. (1996) *Cell* **85**, 15.
- Koch, C. A., Anderson, D., Moran, M. F., Ellis, C., and Pawson, T. (1991) *Science* **252**, 668–674.
- Matozaki, T., and Kasuga, M. (1996) *Cell Signal* **8**, 13–19.
- Feng, G. S., Hui, C. C., and Pawson, T. (1993) *Science* **259**, 1607–1611.
- Kazlauskas, A., Feng, G. S., Pawson, T., and Valius, M. (1993) *Proc. Natl. Acad. Sci. USA* **90**, 6939–6942.
- Kuhné, M. R., Pawson, T., Lienhard, G. E., and Feng, G. S. (1993) *J. Biol. Chem.* **268**, 11479–11481.
- Valius, M., and Kazlauskas, A. (1993) *Cell* **70**, 225–236.
- Bennett, A. M., Hausdorff, S. F., O'Reilly, A. M., Freeman, R. M., and Neel, B. G. (1996) *Mol. Cell. Biol.* **16**, 1189–1202.
- Noguchi, T., Matozaki, T., Horita, K., Fujioka, Y., and Kasuga, M. (1994) *Mol. Cell. Biol.* **14**, 6674–6682.
- Milarski, K. L., and Saltiel, A. R. (1994) *J. Biol. Chem.* **269**, 21239–21243.
- Xiao, S., Rose, D. W., Sasaoka, T., Maegawa, H., Burke, T. R., Roller, P. P., Shoelson, S. E., and Olefsky, J. M. (1994) *J. Biol. Chem.* **269**, 21244–21248.
- Tang, T. L., Freeman, R. M., O'Reilly, A. M., Neel, B. G., and Sokol, S. Y. (1995) *Cell* **80**, 473–483.
- Fujioka, Y., Matozaki, T., Noguchi, T., Iwamatsu, A., Yamao, T., Takahashi, N., Tsuda, M., Takada, T., and Kasuga, M. (1996) *Mol. Cell. Biol.* **16**, 6887–6899.
- Noguchi, T., Matozaki, T., Fujioka, Y., Yamao, T., Tsuda, M., Takada, T., and Kasuga, M. (1996) *J. Biol. Chem.* **271**, 27652–27658.
- Matozaki, T., Noguchi, T., Suzuki, T., and Kasuga, M. (1995) *Adv. Prot. Phosphatases* **9**, 319–338.
- Uchida, T., Matozaki, T., Matsuda, K., Suzuki, T., Matozaki, S., Nakano, O., Wada, K., Konda, Y., Sakamoto, C., and Kasuga, M. (1993) *J. Biol. Chem.* **268**, 11845–11850.
- Matsuda, Y., Harada, Y. N., Natsuume-Sakai, S., Lee, K., Shiomi, T., and Chapman, V. M. (1992) *Cytogenet. Cell Genet.* **61**, 282–285.
- Matsuda, Y., and Chapman, V. M. (1995) *Electrophoresis* **16**, 261–272.
- Heng, H. H. Q., and Tsui, L. C. (1994) In situ hybridization protocols: Methods in Molecular Biology (Choo, K. H. A., Ed.) pp. 109–122, Humana Press, Clifton, NJ.
- Matsuda, Y., Imai, T., Shiomi, T., Saito, T., Yamauchi, M., Fukao, T., Akao, Y., Seki, N., Ito, T., and Hori, T. (1996) *Genomics* **34**, 347–357.
- Yu, H., Chen, J. K., Feng, S., Dalgarno, D. C., Brauer, A. W., and Schreiber, S. L. (1994) *Cell* **76**, 933–945.
- Mount, S. M. (1982) *Nucleic Acids Res.* **10**, 459–472.
- Suzuki, T., Matozaki, T., Mizoguchi, A., and Kasuga, M. (1995) *Biochem. Biophys. Res. Commun.* **211**, 950–959.
- Matozaki, T., Uchida, T., Fujioka, Y., and Kasuga, M. (1994) *Biochem. Biophys. Res. Commun.* **204**, 874–881.
- Arrandale, J. M., Gore-Willse A., Rocks, S., Ren, J. M., Zhu, J., Davis, A., Livingston, J. N., and Rabin, D. U. (1996) *J. Biol. Chem.* **271**, 21353–21358.
- Margolis, R. L., Breschel, T. S., Li, S. H., Kidwai, A. S., Antonarakis, S. E., McInnis, M. G., and Ross, C. A. (1995) *Somat. Cell Mol. Genet.* **21**, 279–284.
- Zoghbi, H. Y. (1996) *Nature Genet.* **14**, 237–238.
- Evans, E. P. (1996) in Standard G-band karyotype: Genetic Variants and Strains of the Laboratory Mouse (Lyon, M. F., Rastan, S., and Brown, S. D. M., Eds.), pp. 1446–1449, Oxford University Press, Oxford.
- Siracusa, L. D., Morgan, J. L., Fisher, J. K., Abott, C. M., and Peters, J. (1996) *Mammal. Genome* **6**, S51–S63.
- Doolittle, D. P., Davisson, M. T., Guidi, J. N., and Green, M. C. (1996) Catalog of mutant genes and polymorphic loci: Genetic Variants and Strains of the Laboratory Mouse (Lyon, M. F., Rastan, S., and Brown, S. D. M., Eds.), pp. 17–854, Oxford University Press, Oxford.
- Hess, E. J., Jinnah, H. A., Kozak, A., and Wilson, M. C. (1992) *J. Neurosci.* **12**, 2865–2874.
- Hess, E. J., Collins, K. A., Copeland, N. G., Jenkins, N. A., and Wilson, M. C. (1994) *Genomics* **21**, 257–261.
- Hess, E. J., Collins, K. A., and Wilson, M. C. (1996) *J. Neurosci.* **16**, 3104–3111.