Identification of a carbonic anhydrase-like domain in the extracellular region of RPTPγ defines a new subfamily of receptor...

Article *in* Molecular and Cellular Biology · April 1993

DOI: 10.1128/MCB.13.3.1497 · Source: PubMed

MOLECULAR AND CELLULAR BIOLOGY, Mar. 1993, p. 1497–1506
0270-7306/93/031497-10$02.00/0

Vol. 13, No. 3

Identification of a Carbonic Anhydrase-Like Domain in the Extracellular Region of RPTPγ Defines a New Subfamily of Receptor Tyrosine Phosphatases

G. BARNEA,[1] O. SILVENNOINEN,[1] B. SHAANAN,[2] A. M. HONEGGER,[1] P. D. CANOLL,[1] P. D'EUSTACHIO,[3] B. MORSE,[4] J. B. LEVY,[1] S. LAFORGIA,[5] K. HUEBNER,[5] J. M. MUSACCHIO,[1] J. SAP,[1] AND J. SCHLESSINGER[1]*

Departments of Pharmacology[1] and Biochemistry,[3] New York University Medical Center, 550 First Avenue, New York, New York 10016; Department of Biological Chemistry, The Life Sciences Institute, The Hebrew University of Jerusalem, Givat Ram, Jerusalem 91904, Israel[2]; Rhone-Poulenc Rorer, Collegeville, Pennsylvania 19426[4]; and Jefferson Cancer Institute, Thomas Jefferson Medical College, Philadelphia, Pennsylvania 19107[5]

Received 2 October 1992/Accepted 14 December 1992

The tyrosine phosphatase RPTPγ is a candidate tumor suppressor gene since it is located on human chromosome $3_p14.2$-$_p21$ in a region frequently deleted in certain types of renal and lung carcinomas. In order to evaluate its oncogenic potential and to explore its normal in vivo functions, we have isolated cDNAs and deduced the complete sequences of both human and murine RPTPγ. The murine RPTPγ gene has been localized to chromosome 14 to a region syntenic to the location of the human gene. Northern (RNA) blot analysis reveals the presence of two major transcripts of 5.5 and 8.5 kb in a variety of murine tissues. In situ hybridization analysis reveals that RPTPγ mRNA is expressed in specific regions of the brain and that the localization of RPTPγ changes during brain development. RPTPγ is composed of a putative extracellular domain, a single transmembrane domain, and a cytoplasmic portion with two tandem catalytic tyrosine phosphatase domains. The extracellular domain contains a stretch of 266 amino acids with striking homology to the zinc-containing enzyme carbonic anhydrase (CAH), indicating that RPTPγ and RPTPβ (HPTPζ) represent a subfamily of receptor tyrosine phosphatases. We have constructed a model for the CAH-like domain of RPTPγ based upon the crystal structure of CAH. It appears that 11 of the 19 residues that form the active site of CAH are conserved in RPTPγ. Yet only one of the three His residues that ligate the zinc atom and are required for catalytic activity is conserved. On the basis of this model we propose that the CAH-like domain of RPTPγ may have a function other than catalysis of hydration of metabolic CO_2.

It is now well established that the phosphorylation of tyrosine residues in cellular proteins plays an important role in the control of cell growth and differentiation (reviewed in references 5, 15, and 36). Tyrosine phosphorylation is a reversible process; protein tyrosine kinases are responsible for the incorporation of phosphate on tyrosine residues of cellular proteins, while protein tyrosine phosphatases are responsible for removal of the tyrosine-bound phosphate groups (reviewed in reference 10). Subversion of normal growth control pathways leading to oncogenesis has been shown to be caused by activation or overexpression of tyrosine kinases, which constitute a large group of dominant oncogenic proteins (reviewed in reference 15). Consistent with this view is the notion that the underexpression or inactivation of protein tyrosine phosphatases potentially results in oncogenesis. For this reason, tyrosine-specific phosphatase genes are candidate recessive oncogenes or tumor suppressor genes (21). Similarly to protein tyrosine kinases, protein tyrosine phosphatases can be divided into two classes: low-molecular-weight cytoplasmic enzymes and membrane-linked enzymes. The membrane-linked enzymes have all the hallmarks of cell surface receptors; they contain a putative extracellular domain, a single transmembrane region, and a cytoplasmic domain which usually contains two tandem catalytic domains (reviewed in references 10 and 29).

Recent studies indicate that the action of tyrosine phosphatases is not necessarily only suppressive. It was shown that members of the Src family of cytoplasmic tyrosine kinases such as $pp56^{lck}$ contain inhibitory tyrosine phosphorylation sites in their carboxy-terminal tails (reviewed in reference 14). The phosphorylation of this site by a specific tyrosine kinase, denoted CSK, leads to inhibition of tyrosine kinase activity (24). It was proposed that in T cells the dephosphorylation of such an inhibitory site in $pp56^{lck}$ by the tyrosine phosphatase CD45 leads to enhanced tyrosine phosphorylation (22). Hence, tyrosine phosphatases may function as either switch-on or switch-off signaling enzymes.

The RPTPγ gene is located in human chromosome $3_p14.2$-$_p21$, in a region found to be deleted in certain types of renal and lung carcinomas. It was therefore proposed that RPTPγ may function as a tumor suppressor gene in these cancers (21). A partial cDNA clone of the catalytic domain of RPTPγ/HPTPγ has been previously described (17, 20). In this report we describe the cDNA cloning and the complete amino acid sequence of human RPTPγ and its murine homolog. We show that RPTPγ contains a single transmembrane domain and two tandem tyrosine phosphatase domains and that the extracellular domain of RPTPγ contains a region of 266 amino acids with striking sequence similarity to the enzyme carbonic anhydrase (CAH). We also show that

* Corresponding author.

TABLE 1. DNA fragment length variant associated with the murine RPTPγ gene[a]

Allele	Size (kb)	Strains
a	2.7	DBA/2J; C3H/HeJ; AKXD-2, -3, -6, -8, -11, -12, -15, -20, -22, -23, -26, -27; BXD-5, -8, -9, -14, -15, -16, -18, -21, -22, -24, -29, -31, -32; BXH-2, -3, -4, -9, -12, -14, -19
b	2.0	C57BL/6J; 020/A; AKR/J; C57L/J; SWR/J; SJL/J; BALB/cJ; STS/A; AKXD-1, -7, -9, -10, -13, -14, -16, -18, -21, -24, -28; BXD-1, -2, -6, -11, -12, -13, -19, -20, -23, -25, -27, -28, -30; BXH-6, -7, -8, -10, -11

[a] Liver or spleen genomic DNA (10 μg) was digested with *Taq*I enzyme and analyzed by Southern blotting with a probe from the first catalytic domain of murine RPTPγ.

RPTPγ is expressed in specific regions of the rat brain and that its expression is developmentally regulated.

MATERIALS AND METHODS

Isolation of cDNA clones. (i) Human RPTPγ. The initial human RPTPγ gene clone was isolated from a λgt11 cDNA library of 1-day-old human brain stem (obtained from the American Type Culture Collection; ATCC 37432) by screening approximately 300,000 plaques with a nick-translated leukocyte common antigen probe that spanned both conserved phosphatase domains, under relaxed hybridization conditions. The rest of the cDNA clones were isolated by screening the human brain stem library with randomly primed probes from the previously isolated clones of the RPTPγ gene under high-stringency hybridization conditions.

(ii) Mouse RPTPγ. The initial mouse RPTPγ gene clone was isolated from a λgt11 mouse brain cDNA library (purchased from Clontech, Palo Alto, Calif.) by screening approximately 10^6 plaques with a randomly primed probe from human RPTPγ that spanned all of the first catalytic domain, under high-stringency hybridization conditions. The rest of the cDNA clones were isolated by screening the mouse brain library with probes from the previously isolated clones of the mouse RPTPγ gene under high-stringency hybridization conditions.

Nucleotide sequence determination. DNA preparations of λgt11 RPTPγ clones were digested with *Eco*RI and subcloned into Bluescript SK+ plasmid (purchased from Strat-

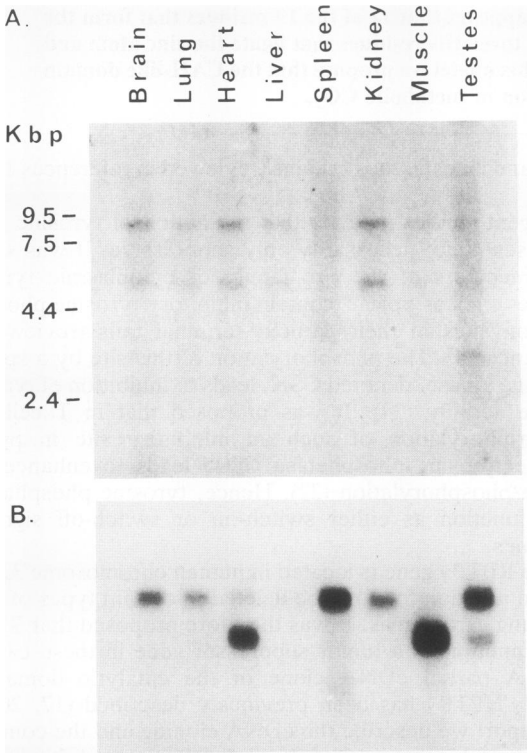


FIG. 1. Expression of RPTPγ in different murine tissues. (A) Northern blot analysis of poly(A)$^+$ RNA from different mouse tissues probed with an RPTPγ probe encompassing the first catalytic domain, the juxtamembrane domain, the transmembrane domain, and the beginning of the extracellular domain. (B) The same blot probed with a β-actin probe.

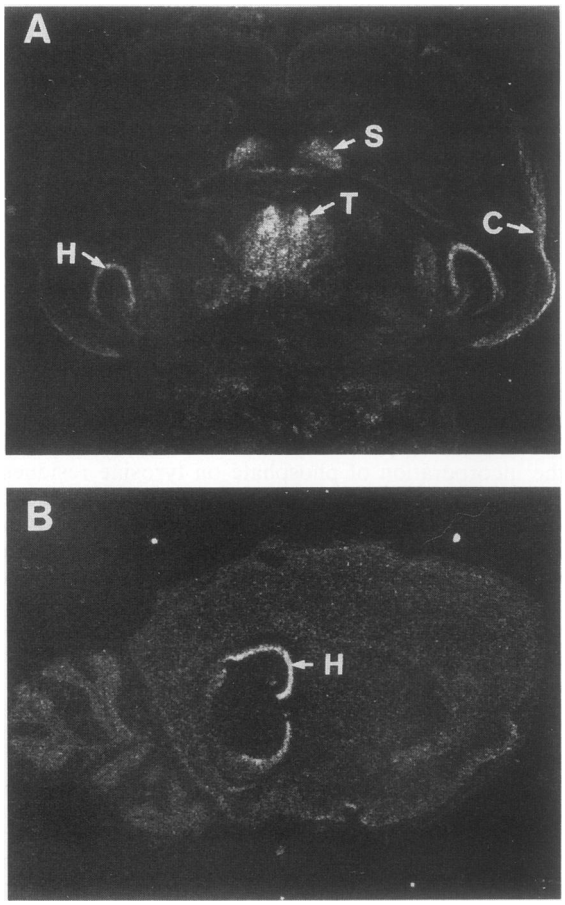


FIG. 2. In situ hybridization analysis of RPTPγ in the newborn and adult rat brain. (A) A horizontal section through the newborn rat brain shows the highest level of expression in the hippocampal formation (H), the cortex (C), the septal nuclei (S), and the midline thalamic nuclei (T). (B) A sagittal section through the adult brain shows the highest level of expression in the hippocampal formation (H).

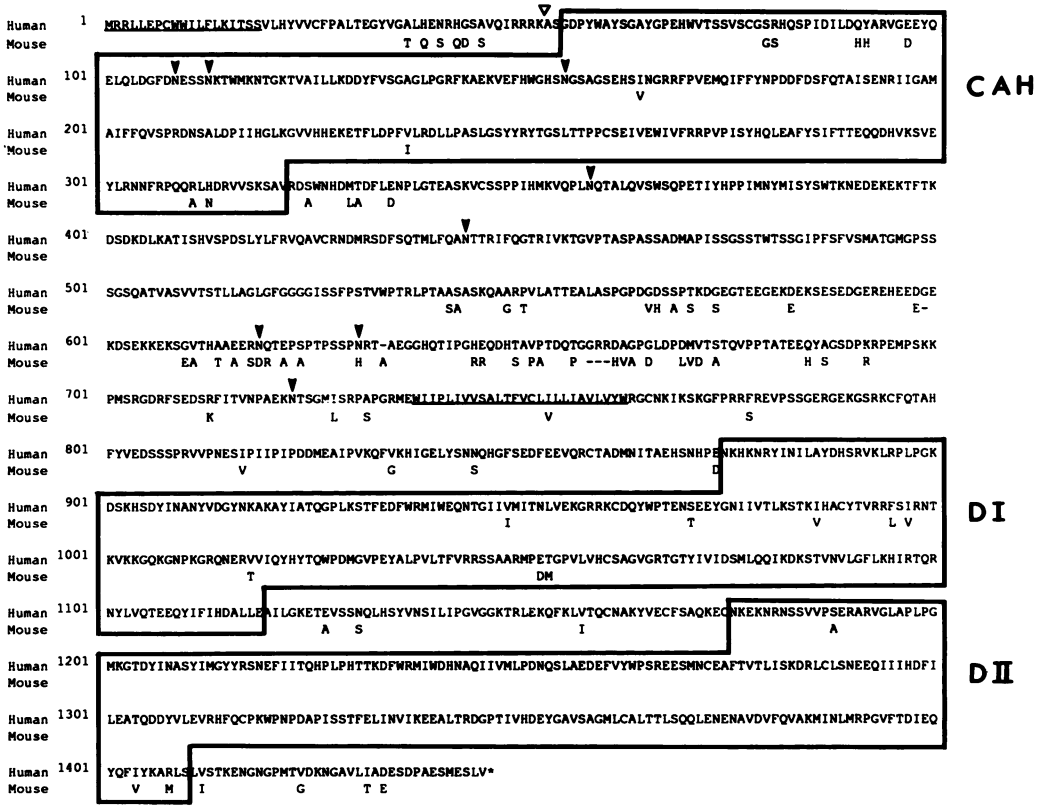


FIG. 3. Amino acid sequences of human and murine RPTPγ. The amino acid sequence of human RPTPγ is presented. Amino acids that are different in the murine gene are indicated. Dashes indicate amino acids that are not found in one of the sequences. The hydrophobic signal peptide (39) and the transmembrane domain are underlined. The potential N-glycosylation sites are indicated by arrowheads. The putative proteolytic cleavage site is indicated by an open triangle. The CAH-like domain and the phosphatase domains DI and DII are boxed.

agene, La Jolla, Calif.). Nucleotide sequences were determined by the dideoxynucleotide chain termination method (Sequenase; United States Biochemical, Cleveland, Ohio) with specific synthetic oligonucleotides as primers. All the clones were sequenced on both strands.

Sequence alignments. All DNA and protein data base searches were done with the Genetic Computer Group sequence analysis software package (8). The SWISSPROT and GenBank/EMBL data bases were searched with FASTA and TFASTA, respectively (25). Proteins were aligned with the Genetics Computer Group programs LINEUP, PILEUP, PRETTY, and BESTFIT.

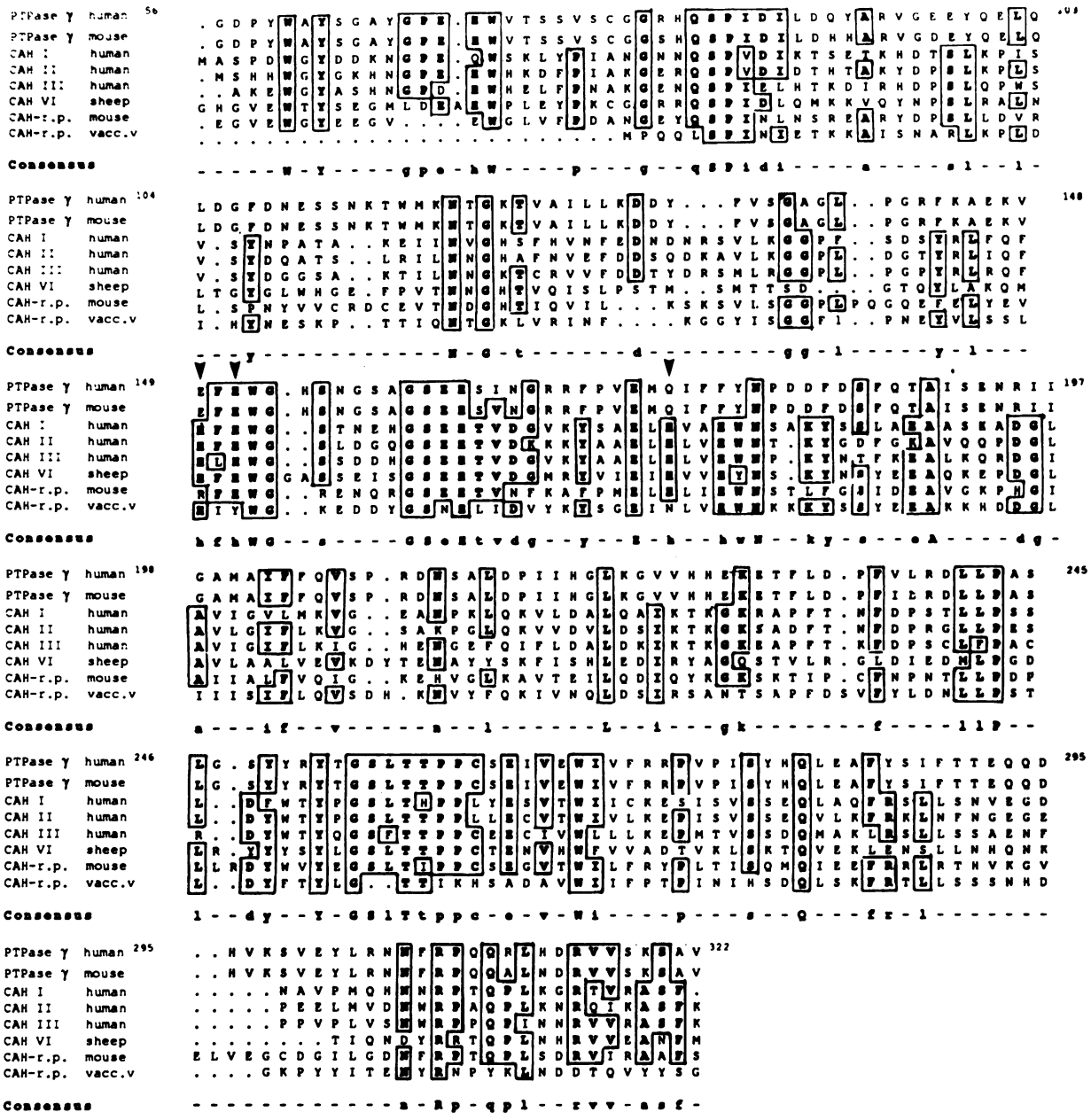
Modeling of the CAH domain and energy minimization. After alignment of the CAH domain of RPTPγ with the sequences of the soluble CAHs, the corresponding substitutions, deletions, and insertions were performed on an Evans & Sutherland and Silicon Graphic interactive displays, using the mutate options provided in program O (16). Inserted peptides and peptides flanking deleted segments were given an initial conformation which best fitted similar peptides taken from proteins with known three-dimensional structures and which are stored in the program as a data bank of conformational information. The initial side chain conformations of inserted and substituted residues were selected according to the rotamer library of Ponder and Richards (26). This rebuilt model was subjected to several cycles of energy minimization by using the program X-PLOR (4) and the energy parameter sets param19.pro and toph19.pro of the program CHARM (3). Cα atoms were constrained to remain close to their original positions in the CAH structure. The root mean square deviation between the main-chain atoms in the model thus derived and the original CAH structure is 0.7 Å (0.07 nm). All the phi/psi angles of the resulting model fall within the allowed regions of the Ramachandran plot.

Tissue expression. Poly(A)$^+$ RNA was prepared from adult mouse tissues by oligo(dT) selection as described elsewhere

```
Fbn III-7   LSPPTNLHLEANPDTGVLTVSWERSTT....PDITGYRITTTP......TNGQQGTALEEVVHADQSSCT..........FDNLSPGLEYNVSVY..TVKDDKESVPISDTIIP
Contac      PGPPGGIRIEEIRDTAV.ALTWSRGTDN..HSPISKYTIQSKTFLSEEWKDAKTEPSDIEGNMESAR..........VIDLIPWMEYEFRIIATNTLGTGEPSMPSQRI..
HLAR        SGPPRKVEVEPLNSTAV.HVYWKLPVPSKQHGQIRGYQV...TYVRLENGEPRGLPIIQDVMLAEAQW..RPEESEDYETTISGLTPETTYSVTVAAYTTKGDGARSKPKIVT..
HPTPδ       SGPPRKVEVEAVNSTSV.KVSWRSPVPNKQHGQIRGYQV...HYVRMENGEPKGQPMLKDVMLADAQW..EFDDTTEHDMIISGLQPETSYSLTVTAYTTKGDGARSKPKLVS..
RPTPγ       SSPPIHMKVQPLNQTAL.QVSWSQP.ETIYHPPIMNYMI...SYSWTKNEDEKEKTFTKD..........SDKDLKATISHVSPDSLYLFRVQAVCRNDMRSDFSQTMLF..
Consensus   --PP-----E----T-V--V-W--------H--I--Y----------------------------------------------L-P---Y---V-A----------------
```

FIG. 4. Fibronectin type III repeat in RPTPγ. The sequence of the fibronectin type III repeat of human RPTPγ is aligned with those of typical fibronectin type III repeats of the human tyrosine phosphatases LAR (32) and HPTPδ (20), domain III-7 of human fibronectin (Fbn III-7) (18), and a fibronectin domain of chicken contactin (contac) (28). Amino acid residues that are shared by four repeats or more are printed in boldface type and indicated at the bottom as the consensus sequence.

A.



B.

PERCENT IDENTITY:

	RPTP γ human	RPTP γ mouse	CAH I human	CAH II human	CAH III human	CAH VI sheep	CAH-r.p mouse	CAH-r.p. vacc.virus
RPTP γ human	100.0	98.1	35.6	38.9	39.8	37.0	35.3	34.5
RPTP γ mouse		100.0	35.6	39.2	40.2	37.0	35.3	34.9
CAH I human			100.0	64.2	63.3	39.1	47.1	40.1
CAH II human				100.0	63.0	42.7	47.3	41.7
CAH III human					100.0	44.4	45.6	42.6
CAH VI sheep						100.0	40.8	36.2
CAH-r.p. mouse							100.0	33.6
CAH-r.p. vacc.v								100.0

(38), fractionated (5 μg per lane) on a formaldehyde-containing gel, transfered to Nytran (Schleicher & Schuell) by standard procedures, and probed with mouse clones MB-8 and MB-122 that encompass all of the first phosphatase domain, the juxtamembrane domain, the transmembrane domain, and the beginning of the extracellular portion.

In situ hybridization. Fresh-frozen rat tissue was cut on a cyrostat into 20-μm-thick sections and thaw mounted onto gelatin-coated slides. The sections were fixed in 4% paraformaldehyde in 0.1 M sodium phosphate (pH 7.4) for 30 min and rinsed three times for 5 min each time in 0.1 M sodium phosphate and twice for 10 min each time in 2× SSC (1× SSC is 0.15 M NaCl plus 0.015 M sodium citrate). Two different oligonucleotide probes were used in the hybridization analysis: a 51-base oligonucleotide complementary to a portion of the cytoplasmic domain and a 52-base oligonucleotide complementary to a portion of the extracellular domain. The oligonucleotides were labeled with [^{35}S]dATP (NEN DuPont) by using terminal deoxynucleotidyltransferase (Boehringer Mannheim) and purified by using Sephadex G-25 quick-spin columns (Boehringer Mannheim). The specific activity of the labeled probes was between 5×10^8 and 1×10^9 cpm/μg. Prehybridization and hybridization were carried out in a buffer containing 50% deionized formamid, 4× SSC, 1× Denhardt's solution, 500 μg of denatured salmon sperm DNA per ml, 250 μg of yeast tRNA per ml, and 10% dextran sulfate. The tissue was incubated for 12 h at 45°C in hybridization solution containing the labeled probe (10^6 cpm per section) and 10 mM dithiothreitol. Controls for specificity were performed on adjacent sections by adding a 30-fold concentration of the unlabeled oligonucleotide or by hybridization with the sense probe. After hybridization the sections were washed in two changes of 2× SSC at room temperature for 1 h, 1× SSC at 55°C for 30 min, 0.5× SSC at 55°C for 30 min, and 0.5× SSC at room temperature for 15 min and dehydrated in 60, 80, and 100% ethanol. After air drying, the sections were exposed to X-ray film for 5 to 10 days.

Chromosomal localization of murine RPTPγ. To define the genetic locus encoding mouse RPTPγ, we used a probe that contains the first catalytic domain of mouse RPTPγ to search for a restriction fragment length variant among inbred strains of mice. Southern blotting analysis of *Taq*I-digested DNA revealed two DNA fragments, of 2.9 and 1.8 kb, shared by all strains examined plus a fragment of either 2.7 kb (C3H/HeJ and DBA/2J) or 2.0 kb (all other strains examined). The inheritance of this DNA variant in the AKXD, BXD, and BXH recombinant inbred strains of mice defined a genetic locus linked to *Odc-9* (8 recombinants among 61 strains) and *Plau* (11 recombinants among 49 strains) near the centromeric end of the linkage map of chromosome 14 (Table 1). We propose *Ptpg* (phosphotyrosine phosphatase gamma) as the symbol for the locus, consistent with the symbol *Ptpa* previously assigned for mouse RPTPα (30).

Nucleotide sequence accession numbers. The sequence data for human and murine RPTPγ have been deposited in the GenBank data base under accession numbers L09247 and L09562, respectively.

RESULTS AND DISCUSSION

Resolving the issue of whether RPTPγ may function as a tumor suppressor gene (21) will require a detailed screening of tumors for genomic rearrangements and point mutations and reintroduction of wild-type RPTPγ into tumor cells. Since so far the genomic analysis of RPTPγ has been performed only with the partial sequence that has been described previously (17), we have cloned the full-length human cDNA. In addition, we have cloned the murine homolog of the RPTPγ gene to facilitate an analysis of its tissue expression, as well as its normal in vivo function. The murine RPTPγ gene was localized to chromosome 14 (Table 1) to a region syntenic to the chromosomal localization of the human RPTPγ gene (21). Northern (RNA) blot analysis shows that RPTPγ is widely expressed in different murine tissues (Fig. 1). Two major RPTPγ transcripts of 5.5 and 8.5 kb were detected in the murine brain, lung, kidney, heart, skeletal muscle, liver, spleen, and testes. An additional shorter transcript of approximately 3.0 kb was detected in testes.

The two RPTPases most related to RPTPγ are RPTPβ/HPTPζ (19, 22a) and a *Drosophila* phosphatase, DPTP99A (13, 34, 40). Since these phosphatases are specifically expressed in the central nervous system, we have analyzed the expression pattern of RPTPγ in the brains of newborn and adult rats by in situ hybridization. In newborn rats the highest level of expression was detected in the hippocampal formation, in the septal and midline thalamic nuclei, and in the cortex (Fig. 2A). However, in the adult rat brain RPTPγ is highly expressed in the hippocampal formation (Fig. 2B) but not in the septal and midline thalamic nuclei or in the cortex (data not shown). Hybridization with probes derived from either the cytoplasmic or the extracellular domain gave similar results. The addition of a 30-fold concentration of unlabeled oligonucleotides completely blocked the labeling in all areas. Furthermore, no signal was observed in adjacent sections hybridized with the sense probe. These results demonstrate that the two probes hybridize to mRNA in a sequence-specific manner. The transient expression in the septal and midline thalamic nuclei and in the cortex of the newborn rat brain indicates that the expression of RPTPγ is developmentally regulated and may play a role in the development of these regions. It is noteworthy that the patterns of expression of RPTPγ and the closely related RPTPβ are totally different. In the embryo, RPTPβ is expressed in the ventricular and subventricular zones of the brain and spinal cord. In the adult brain, RPTPβ is expressed in the Purkinje cell layer of the cerebellum, the dentate gyrus, and the subependymal layer of the anterior horn of the lateral ventricle (22a).

The complete amino acid sequences of the human and murine RPTPγ deduced from cDNA clones are presented in Fig. 3. Translation of the cDNA sequences reveals the

FIG. 5. Alignment of the CAH-like domain in RPTPγ with different forms of CAH. (A) The sequences of the CAH-like domains of human and murine RPTPγ are aligned with representative sequences of the CAH family: human CAH-1, CAH-2, and CAH-3 (sequences can be found in SWISSPROT accession numbers P00915, P00918, and P07451, respectively), sheep CAH-6 and vaccinia virus (vacc.v) CAH-like protein (CAH-r.p.) (SWISSPROT accession numbers P0860 and P04195, respectively), and mouse CAH-related protein (CAH-r.p.) (GenBank accession number X6197). Residues conserved in at least five of the eight sequences are boxed. The positions in human RPTPγ of the first and last amino acids in each line are indicated. The three His residues involved in Zn binding in CAH are indicated with arrowheads. (B) A matrix of the percent identity between the CAH-like domains in human and murine RPTPγ and the CAH sequences shown above derived from the alignments shown in panel A.

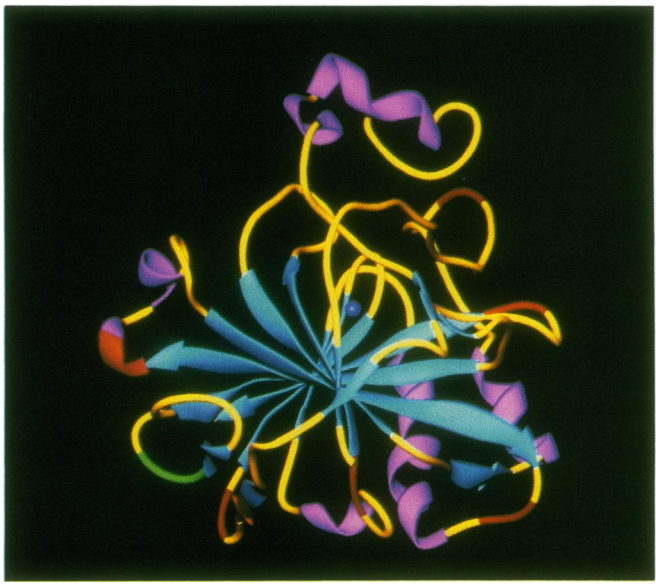


FIG. 6. Deletions and insertions in the RPTPγ sequence mapped onto the CAH structure. A ribbon diagram (6) of human CAH-2 with regions corresponding to deletions and insertions in the sequence of RPTPγ shown in green and red, respectively. Helices are shown in magenta, β strands are in blue, turns are in yellow, and the zinc atom is in dark blue. Note that all the deletions and insertions fall in exposed regions of CAH.

existence of an open reading frame of 1,445 amino acids for human RPTPγ and 1,442 amino acids for the murine homolog. The initiation Met in both human RPTPγ and the murine homolog are preceded by in-frame stop codons. Both proteins contain a signal sequence (underlined in Fig. 3) and a putative extracellular domain of 736 amino acids for human RPTPγ and 733 amino acids for the murine gene. The extracellular domain of human RPTPγ contains eight putative N-linked glycosylation sites, six of which are conserved in the murine homolog (marked with arrowheads in Fig. 3). The amino-terminal region of the extracellular domain of RPTPγ (residues 56 to 322) has a striking sequence similarity to the enzyme CAH. This CAH-like domain is preceded by 4 basic amino acids (ArgArgArgLys) that resemble the cleavage site in the extracellular domain of the insulin receptor (35) separating the α and β subunits of the insulin receptor. This sequence motif may function as a cleavage site for proteolytic enzymes (1). It is of note that a similar cleavage site has been identified in the extracellular domain of another receptor-type phosphotyrosine phosphatase, denoted LAR (31, 41).

The CAH-like domain is followed by one fibronectin type III repeat, a motif found in many cell surface proteins. Alignment of the fibronectin type III sequence of RPTPγ with typical fibronectin type III repeats of other proteins is presented in Fig. 4. The remaining 293 amino acids of the extracellular domain are devoid of any Cys residues and can be subdivided into three regions: a Ser/Thr-rich region in which 32% of the amino acids are Ser and Thr (residues 442 to 560), followed by a region composed of 90% charged and polar amino acids (residues 561 to 662) and a region with no similarity to any known sequence motif. Hence, the region downstream from the fibronectin repeat may function as a spacer separating the CAH-like domain and the fibronectin type III repeat from the transmembrane region.

The extracellular domain of RPTPγ is followed by a typical transmembrane domain of 23 amino acid residues. As in most known receptor tyrosine phosphatases, the intracellular domain of RPTPγ contains two tandem phosphatase domains (10). It is noteworthy that the second phosphatase domain of RPTPγ has an Asp residue at position 1351 instead of a conserved Cys residue thought to be essential for catalytic activity (11, 27). Interestingly, an Asp residue is also found in a similar position in RPTPβ (17) and in the *Drosophila* phosphatase 99A (13, 34, 40). Moreover, the second catalytic domain of RPTPγ contains an insert of 15 amino acids (residues 1299 to 1313) identical to the insert found in corresponding position in RPTPβ (17) and so far

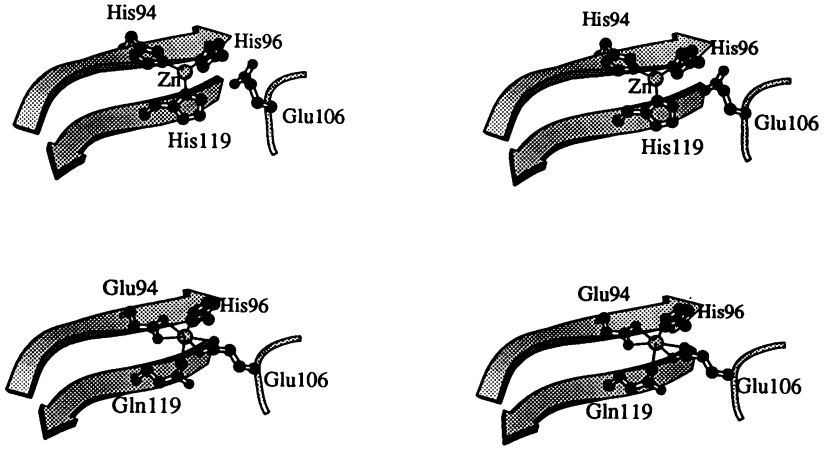


FIG. 7. Comparison between the zinc-binding site of CAH and the putative metal-binding site in the RPTPγ model. (Top) Stereo view of the zinc-binding site in human CAH-2 viewed roughly from the direction of the water molecule which is the fourth zinc ligand in this enzyme (OHH263 [9]). Note the tetrahedral coordination of the zinc. Bonds from the zinc to ligand atoms are shown in thin lines. (Bottom) Stereo view of the putative metal-binding site in the RPTPγ domain (same view as top drawing), after applying the substitutions His-94 to Glu and His-119 to Gln and energy minimization (see text). Glu-106, which is part of the active site in human CAH-2, has been rotated to become one of the potential ligands. Note the planar constellation of atoms around the putative metal. Bidentate coordination is shown for Glu-94 and Glu-106. All the numbers of residues are according to CAH-2.

appears to be unique to these two phosphatases. This insert contains a Tyr residue in position 1308 which is preceded by two Asp residues and followed by a Val residue, suggesting that it may function as a phosphorylation site for tyrosine kinases.

Human RPTPγ and its murine homolog are 90% identical at the nucleotide level and 95% identical at the amino acid level, and most of the differences are conservative substitutions (Fig. 3). This strikingly high conservation is found in both the cytoplasmic and extracellular domains. This degree of conservation between human and mouse RPTPγ is higher than in some of the other phosphatases, such as the extracellular portion of RPTPα (17, 20, 23, 30). The most conserved parts are the fibronectin repeat (100% identity), the CAH-like domain (97% identity), and the Ser/Thr-rich domain (97% identity). The rest of the extracellular domain is the least conserved (70% identity). It is noteworthy that the cluster of charged and polar amino acids that follows the Ser/Thr-rich domain is found in both human RPTPγ and the murine homolog in spite of the lower degree of sequence conservation. The high degree of identity between the human and murine sequences in the CAH-like domain, the fibronectin type III repeat, and the Ser/Thr-rich domain may suggest that these regions are required for a conserved biological functions.

As already mentioned, the amino-terminal region of the extracellular domain of RPTPγ contains a region of 266 amino acids with a striking sequence similarity to the enzyme CAH. CAH catalyzes the hydration of metabolic CO_2 or the dehydration of HCO_3^- in the following reaction: $CO_2 + H_2O \leftrightarrows H^+ + HCO_3^-$. CAHs are ubiquitously expressed enzymes with extremely efficient turnover rate of 10^6 s^{-1} for CO_2 hydration. It has also been shown that CAHs are able to hydrolyze certain esters and to hydrate specific aldehydes. All CAHs are zinc metalloenzymes in which the zinc atom is required for the catalytic activity. Seven types of CAH have been identified so far (reviewed in reference 33). CAH-1, CAH-2, and CAH-3 are cytoplasmic enzymes, CAH-4 is an extracellular glycoprotein, CAH-5 is a mitochondrial enzyme, CAH-6 is a secreted enzyme, and CAH-7 is a membrane-bound enzyme. In addition, vaccinia virus also contains a transmembrane protein with a CAH-like domain in its extracellular portion.

The CAH-like domains in human and murine RPTPγ were aligned with the amino acid sequences of the different forms of CAH (Fig. 5A). Comparison of the sequences revealed clusters of identical amino acids that usually match regions of high conservation of sequences between the different forms of CAH. Very few insertions and deletions were required in order to align the CAH domain in RPTPγ to that of the different forms of CAH. It is noteworthy that 11 of the 19 residues that form the active site of CAH (9) are also found in RPTPγ. Of the three His residues that ligate the zinc atom in CAH (indicated with arrowheads in Fig. 5A), only one, in position 151, is conserved in RPTPγ, whereas the other two are replaced by Glu and Gln residues (positions 149 and 175, respectively). The CAH domain of RPTPγ shares 35 to 40% sequence identity with all known CAHs (Fig. 5B).

On the basis of the striking sequence similarity of this region with CAHs, we have constructed a model for the structure of this region in RPTPγ by substitution of amino acids from the RPTPγ sequence into equivalent positions in the known three-dimensional crystal structure of CAH (9) (entry lca2 in the Brookhaven Protein Data Bank). This was followed by energy minimization (4) in which the Cα atoms were restrained to their positions in the original CAH structure. In the absence of direct structural evidence it is difficult to assess the correctness of the resulting model, but three features point strongly to its relevance to the actual fold of the RPTPγ domain. First, as all the deletions and insertions relative to the CAH sequence map to surface regions and all but one fall in stretches of the polypeptide chain lacking any well-defined secondary structure (Fig. 6), they could all be accommodated without distorting the original model. Second, substitutions of internal residues in the RPTPγ domain follow a pattern that preserves the tight packing of the two hydrophobic cores in CAH (9), as is often encountered in families of homologous proteins (2). Although replacement of some aromatic residues by aliphatic ones in the RPTPγ domain (Phe-66 in CAH to Val-124 in RPTPγ, Phe-70 to Leu-128, Phe-93 to Val-148, and Phe-176 to Leu-233) reduces somewhat the aromatic character of the hydrophobic core proximal to the zinc-binding site (9), the tight packing is maintained through the substitutions Ile-59 in CAH to Trp-116 in RPTPγ, Val-160 to Ile-217, and Ser-56 to Asn. Third, inspection of the sequence of the CAH domain of mouse RPTPγ reveals that substitutions relative to the human gene also follow the pattern of conservation of a tightly packed hydrophobic core (e.g., Ile-165 in human RPTPγ to Val in the murine protein and Val-237 in the human protein to Ile in the murine homolog).

Particularly intriguing is the fate of the zinc-binding site and the active site of CAH in the RPTPγ domain. Sequence alignment shows that two of the conserved histidines ligating the zinc in all known CAHs have been replaced in RPTPγ (His-94 in CAH to Glu-149 in RPTPγ and His-119 to Gln-175). Inspection of zinc-binding sites in proteins whose three-dimensional structure is known (37) reveals that none of them contains Gln. Furthermore, loss of zinc-binding capability as a result of His-to-Gln mutation was reported in the growth hormone family (7). Moreover, inspection of the energy-minimized model of the CAH domain of RPTPγ suggests that the residues Glu-149, His-151, Glu-162, and Gln-175 can form a planar constellation of atoms (Fig. 7) that is often found in the binding sites of octahedrally coordinated transition metals such as manganese (e.g., see reference 12).

A CAH-like domain was also found in the amino-terminal region of RPTPβ/HPTPζ (19, 22a). It appears that the salient features of the CAH-like domain in RPTPγ, such as the tightly packed hydrophobic core and the replacement of two of the three conserved His residues, are also observed in the CAH domain of RPTPβ. This similarity may reflect functional parallels between these two domains. The biological role of the CAH domains of RPTPγ and RPTPβ is not known. In view of the fact that only one of the three His residues that ligate zinc and are crucial for CAH activity is conserved, it is conceivable that the CAH domains of RPTPγ and RPTPβ may lack the ability to bind zinc. It is possible, however, that the site occupied by zinc in CAH will be capable of binding other transition metals and thus have a function other than hydration of metabolic CO_2. Interestingly, the homology between RPTPγ and RPTPβ extends into the downstream fibronectin type III repeat which is flanked in both proteins by two conserved Cys residues. An alignment of the sequences of these two domains in RPTPγ and RPTPβ (Fig. 8A) shows that they share approximately 37% sequence identity. The fibronectin type III repeat in RPTPγ is followed by a stretch of 293 amino acids that is characterized by a lack of Cys residues and thus may serve as a spacer that separates the CAH-like domain

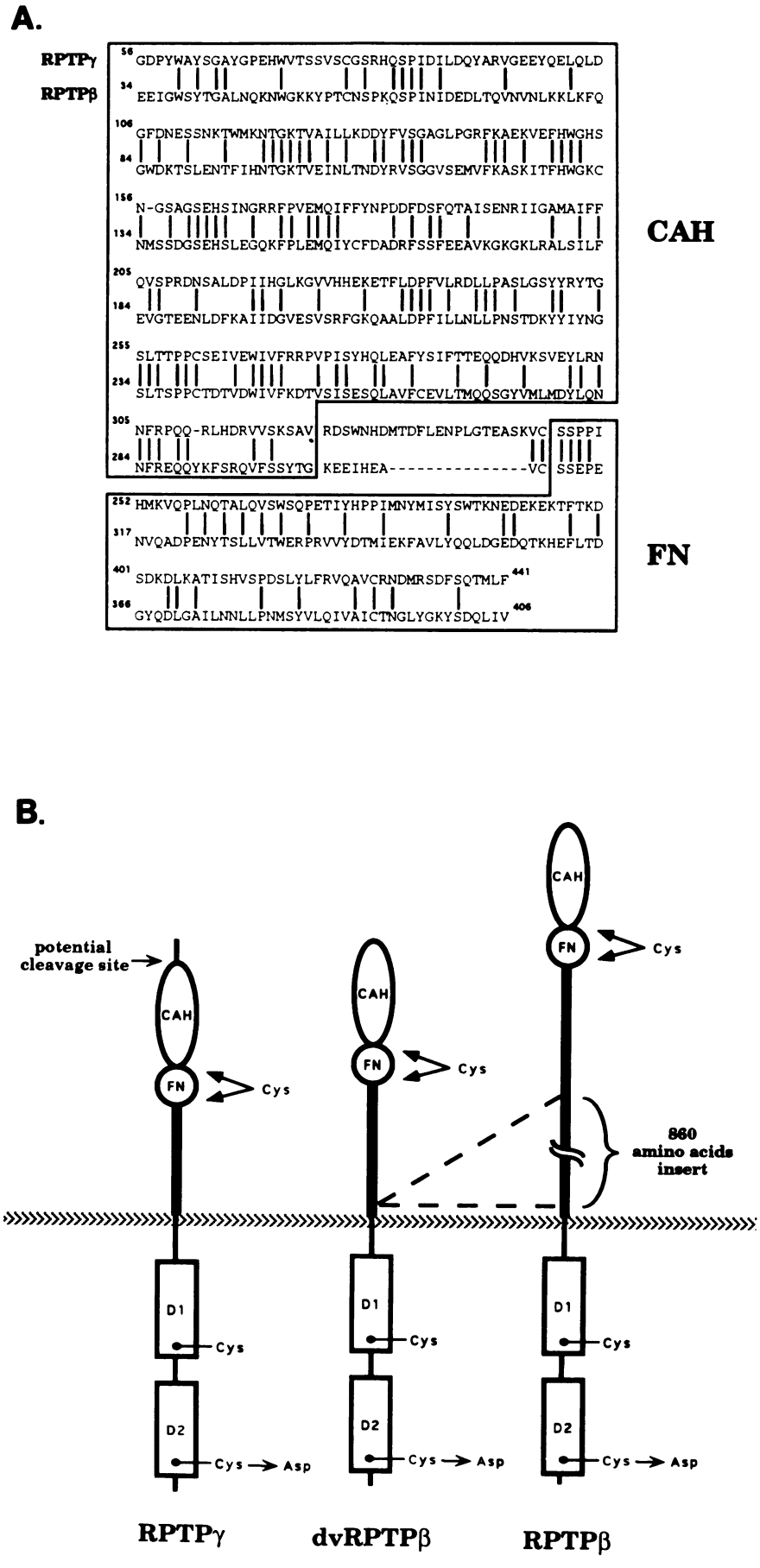

A.
RPTPγ
RPTPβ
GDPYWAYSGAYGPEHWVTSSVSCGSRHQSPIDILDQYARVGEEYQELQLD
EEIGWSYTGALNQKNWGKKYPTCNSPKQSPINIDEDLTQVNVNLKKLKFQ
GFDNESSNKTWMKNTGKTVAILLKDDYFVSGAGLPGRFKAEKVEFHWGHS
GWDKTSLENTFIHNTGKTVEINLTNDYRVSGGVSEMVFKASKITFHWGKC
N-GSAGSEHSINGRRFPVEMQIFFYNPDDFDSFQTAISENRIIGAMAIFF
NMSSDGSEHSLEGQKFPLEMQIYCFDADRFSSFEEAVKGKGKLRALSILF
CAH
QVSPRDNSALDPIIHGLKGVVHHEKETFLDPFVLRDLLPASLGSYYRYTG
EVGTEENLDFKAIIDGVESVSRFGKQAALDPFILLNLLPNSTDKYYIYNG
SLTTPPCSEIVEWIVFRRPVPISYHQLEAFYSIFTTEQQDHVKSVEYLRN
SLTSPPCTDTVDWIVFKDTVSISESQLAVFCEVLTMQQSGYVMLMDYLQN
NFRPQQ-RLHDRVVSKSAV RDSWNHDMTDFLENPLGTEASKVC SSPPI
NFREQQYKFSRQVFSSYTG KEEIHEA--------------VC SSEPE
HMKVQPLNQTALQVSWSQPETIYHPPIMNYMISYSWTKNEDEKEKTFTKD
NVQADPENYTSLLVTWERPRVVYDTMIEKFAVLYQQLDGEDQTKHEFLTD
SDKDLKATISHVSPDSLYLFRVQAVCRNDMRSDFSQTMLF
GYQDLGAILNNLLPNMSYVLQIVAICTNGLYGKYSDQLIV
FN
B.
potential cleavage site
CAH
FN
Cys
860 amino acids insert
D1
D2
Cys → Asp
RPTPγ
dvRPTPβ
RPTPβ

FIG. 8. RPTPγ and RPTPβ define a new subfamily of receptor tyrosine phosphatases. (A) Alignment of the CAH-like domains and the fibronectin type III repeats of RPTPγ and RPTPβ (22a). The sequence in this region of RPTPβ is identical to that reported by Krueger and Saito (19) for RPTPζ. The aligned CAH domains and fibronectin type III repeats (FN) are boxed. Identical amino acids are indicated by a connecting line. (B) Schematic diagram summarizing the conserved features that define the subfamily of RPTPγ and RPTPβ. The extracellular regions of RPTPγ and the two forms of RPTPβ (RPTPβ and the deletion variant [dvRPTPβ]) contain CAH-like domains (CAH), fibronectin type III repeats (FN), and spacers of variable length that are characterized by a very low content of cysteine residues (indicated by thick lines). The conserved cysteine residues that flank the fibronectin type III repeat are marked. The cytoplasmic regions of RPTPγ and RPTPβ contain two typical phosphatase domains (D1 and D2). A conserved Cys residue in the first phosphatase domain and an Asp residue that replaces the Cys residue in the second phosphatase domain are indicated. The potential cleavage site in RPTPγ and the 860-amino-acid insert in RPTPβ which is not present in the deletion variant are denoted.

and the fibronectin repeat from the transmembrane domain. Two forms of RPTPβ have been identified (22a); both contain intact CAH-like domains and fibronectin type III repeats followed by a Cys-free region. The two forms of RPTPβ differ in the length of the Cys-free spacer, which contains 1,048 amino acids in the long form and 384 amino acids in the deletion variant. As mentioned earlier, the high similarity between RPTPγ and RPTPβ extends to the cytoplasmic domains of RPTPγ and RPTPβ.

On the basis of these similarities we suggest that these two phosphatases define a new subfamily of receptor tyrosine phosphatases (Fig. 8B). The elucidation of the biological function of RPTPγ and RPTPβ and the role of the CAH-like domain in these proteins may require the identification of their putative natural ligands and the binding region of these orphan receptors.

ACKNOWLEDGMENTS

This work was supported by grants from SUGEN, Inc. (to J. Schlessinger), and from HFSPO (to J. Schlessinger); J. Sap is a recipient of a postdoctoral fellowship from HFSPO. Computing was supported by NSF under grant DIR-8908095.

We thank Kiki Nelson for synthesis of peptides and oligonucleotides.

REFERENCES

1. **Barr, P. J.** 1991. Mammalian subtilisins: the long-sought dibasic processing endoproteases. Cell **66**:1–3.
2. **Bordo, D., and P. Argos.** 1990. Evolution of protein cores constraints in point mutations as observed in globin tertiary structures. J. Mol. Biol. **211**:975–988.
3. **Brooks, B. R., R. E. Bruccoleri, B. D. Olfason, D. J. States, S. Swaminathan, and M. Karplus.** 1983. CHARM: a program for macromolecular energy minimization and dynamics calculations. J. Comput. Chem. **4**:187–217.
4. **Brunger, A. T.** 1992. X-PLOR (version 3.0) manual. Yale University, New Haven, Conn.
5. **Cantley, L. C., K. R. Auger, C. Carpenter, B. Duckworth, A. Graziani, R. Kapeller, and S. Soltoff.** 1991. Oncogenes and signal transduction. Cell **64**:281–302.
6. **Carson, M.** 1987. Ribbon models of macromolecules. J. Mol. Graphics **5**:103–106.
7. **Cunningham, B. C., S. Bass, G. Fuh, and J. A. Wells.** 1990. Zinc mediation of the binding of human growth hormone to the human prolactin receptor. Science **250**:1709–1712.
8. **Devereux, J., P. Haeberli, and O. Smithies.** 1989. A comprehensive set of sequence analysis programs in the VAX. Nucleic Acids Res. **12**:387–396.
9. **Ericksson, A. E., T. A. Jones, and A. Liljas.** 1988. Refined structure of human carbonic anhydrase II at 2.0 A resolution. Proteins **4**:274–282.
10. **Fischer, E. H., H. Charbonneau, and N. K. Tonks.** 1991. Protein tyrosine phosphatases: a diverse family of intracellular and transmembrane enzymes. Science **253**:401–406.
11. **Guan, K., and J. E. Dixon.** 1991. Evidence for protein-tyrosine-phosphatase catalysis proceeding via a cysteine-phosphate intermediate. J. Biol. Chem. **266**:17026–17030.
12. **Hardman, K. D., R. C. Agarwal, and M. J. Freiser.** 1982. Manganese and calcium binding sites of concanavalin A. J. Mol. Biol. **157**:69–86.
13. **Hariharan, I. K., P. T. Chuang, and G. M. Rubin.** 1991. Cloning and characterization of a receptor-class phosphotyrosine phosphatase gene expressed on central nervous system axons in *Drosophila melanogaster*. Proc. Natl. Acad. Sci. USA **88**: 11266–11270.
14. **Hunter, T.** 1987. A tail of two src's mutatis mutandis. Cell **49**:1–4.
15. **Hunter, T.** 1991. Cooperation between oncogenes. Cell **64**:249–270.
16. **Jones, T. A., J. Y. Zou, S. W. Cowan, and M. Kjeldgaard.** 1991. Improved methods for binding protein models in electron density maps and the location of errors in these models. Acta Crystallogr. Sect. A**47**:110–119.
17. **Kaplan, R., B. Morse, K. Huebner, C. Croce, R. Howk, M. Ravera, G. Ricca, M. Jaye, and J. Schlessinger.** 1990. Cloning of three human tyrosine phosphatases reveals a multigene family of receptor-linked protein-tyrosine-phosphatases expressed in brain. Proc. Natl. Acad. Sci. USA **87**:7000–7004.
18. **Kornblihtt, A. R., H. K. Umezawa, K. Vibe-Pedersen, and F. E. Baralle.** 1985. Primary structure of human fibronectin: differential splicing may generate at least 10 polypeptides from a single gene. EMBO J. **4**:1755–1759.
19. **Krueger, N. X., and H. Saito.** 1992. A human transmembrane protein-tyrosine phosphatase, RPTPζ, is expressed in brain and has an N-terminal receptor domain homologous to carbonic anhydrase. Proc. Natl. Acad. Sci. USA **89**:7417–7421.
20. **Krueger, N. X., M. Streuli, and H. Saito.** 1990. Structural diversity and evolution of human receptor-like protein tyrosine phosphatases. EMBO J. **9**:3241–3252.
21. **LaForgia, S., B. Morse, J. Levy, G. Barnea, L. A. Cannizzaro, F. Li, P. C. Nowell, L. Boghosian-Sell, J. Glick, A. Weston, C. C. Harris, H. Drabkin, D. Patterson, C. M. Croce, J. Schlessinger, and K. Huebner.** 1991. Receptor protein-tyrosine phosphatase γ is a candidate tumor suppressor gene at human chromosome region 3p21. Proc. Natl. Acad. Sci. USA **88**:5036–5040.
22. **Ledbetter, J. A., N. K. Tonks, E. H. Fischer, and E. A. Clark.** 1988. CD45 regulates signal transduction and lymphocyte activation by specific association with receptor molecules on T or B cells. Proc. Natl. Acad. Sci. USA **85**:8628–8632.

22a. **Levy, J. B., et al.** Submitted for publication.

23. **Matthews, R. J., E. D. Cahir, and M. L. Thomas.** 1990. Identification of an additional member of the protein-tyrosine-phosphatase family: evidence for alternative splicing in the tyrosine-phosphatase family: evidence for alternative splicing in the tyrosine phosphatase domain. Proc. Natl. Acad. Sci. USA **87**:4444–4448.
24. **Nada, S., M. Okada, A. MacAuley, J. A. Cooper, and H. Nakagawa.** 1991. Cloning of a complementary DNA for a protein-tyrosine kinase that specifically phosphorylates a negative regulatory site of $p60^{c\text{-}src}$. Nature (London) **351**:69–72.
25. **Pearson, W. R., and D. J. Lipman.** 1988. Improved tools for biological sequence comparison. Proc. Natl. Acad. Sci. USA **85**:2444–2448.
26. **Ponder, J. W., and F. M. Richards.** 1987. Tertiary templates for proteins: use of packing criteria in the enumeration of allowed sequences for different structural classes. J. Mol. Biol. **193**:775–791.

27. **Pot, D. A., and J. E. Dixon.** 1992. Active site labeling of a receptor-like protein tyrosine phosphatase. J. Biol. Chem. **267:** 140–143.
28. **Ranscht, B., and M. T. Dours.** 1988. Sequence of contactin, a 130-kD glycoprotein concentrated in areas of interneuronal contact, defines a new member of the immunoglobulin supergene family in the nervous system. J. Cell Biol. **107:**1561–1573.
29. **Saito, H., and M. Streuli.** 1991. Molecular characterization of protein tyrosine phosphatases. Cell Growth Differ. **2:**59–65.
30. **Sap, J., P. D'Eustachio, D. Givol, and J. Schlessinger.** 1990. Cloning and expression of a widely expressed receptor tyrosine phosphatase. Proc. Natl. Acad. Sci. USA **87:**6112–6116.
31. **Streuli, M., N. X. Kreuger, P. Dariniello, M. Tang, J. M. Munro, W. A. Blattler, D. A. Adler, C. M. Disteche, and H. Saito.** 1992. Expression of the receptor-linked protein tyrosine phosphatase LAR: proteolytic cleavage and shedding of the CAM-like extracellular region. EMBO J. **11:**897–907.
32. **Streuli, M., N. X. Kreuger, L. R. Hall, S. F. Schlossman, and H. Saito.** 1988. A new member of the immunoglobulin superfamily that has a cytoplasmic region homologous to the leukocyte common antigen. J. Exp. Med. **168:**1523–1530.
33. **Tashian, R. E.** 1989. The carbonic anhydrases: widening perspectives on their evolution expression and function. Bioessays **10:**186–192.
34. **Tian, S. S., P. Tsoulfas, and K. Zinn.** 1991. Three receptor-linked protein-tyrosine phosphatases are selectively expressed on central nervous system axons in the drosophila embryo. Cell **67:**675–685.
35. **Ullrich, A., J. R. Bell, E. Y. Chen, R. Herrera, L. M. Petruzzelli, T. J. Dull, A. Gray, L. Coussens, Y. C. Liao, M. Tsubokawa, A. Mason, P. H. Seeburg, C. Grunfeld, O. M. Rosen, and J. Ramachandran.** 1985. Human insulin receptor and its relationship to the tyrosine kinase family of oncogenes. Nature (London) **313:**756–761.
36. **Ullrich, A., and J. Schlessinger.** 1990. Signal transduction by receptors with tyrosine kinase activity. Cell **61:**203–212.
37. **Vallee, B. L., and D. S. Auld.** 1990. Zinc coordination, function, and structure of zinc enzymes and other proteins. Biochemistry **29:**5647–5659.
38. **Vennstrom, B., and J. M. Bishop.** 1982. Isolation and characterization of chicken DNA homologous to the two putative oncogenes of avian erythroblastosis virus. Cell **28:**135–143.
39. **von Heijne, G.** 1986. A new method for predicting signal sequence cleavage sites. Nucleic Acids Res. **14:**4683–4690.
40. **Yang, X., K. T. Seow, S. M. Bahri, S. H. Oon, and W. Chia.** 1991. Two drosophila receptor-like tyrosine phosphatase genes are expressed in a subset of developing axons and pioneer neurons in the embryonic CNS. Cell **67:**661–673.
41. **Yu, Q., T. Lenardo, and R. A. Weinberg.** 1992. The N-terminal and C-terminal domains of a receptor tyrosine phosphatase are associated by non-covalent linkage. Oncogene **7:**1051–1057.