Proc. Natl. Acad. Sci. USA
Vol. 89, pp. 7417–7421, August 1992
Biochemistry

A human transmembrane protein-tyrosine-phosphatase, PTPζ, is expressed in brain and has an N-terminal receptor domain homologous to carbonic anhydrases

(cDNA/dephosphorylation/nucleotide sequence)

NEIL X. KRUEGER*† AND HARUO SAITO*‡§

*Division of Tumor Immunology, Dana–Farber Cancer Institute, and Departments of †Pathology and ‡Biological Chemistry and Molecular Pharmacology, Harvard Medical School, Boston, MA 02115

Communicated by Eugene P. Kennedy, May 18, 1992

ABSTRACT **Protein-tyrosine-phosphatases (PTPases, EC 3.1.3.48) play a crucial role in the regulation of protein tyrosine phosphorylation. Recently, it was found that the PTPase gene family exhibits a large variety of different functional domains associated with the PTPase catalytic domains. In this paper, we report the complete cDNA sequence of a human transmembrane PTPase, PTPζ, isolated from fetal brain cDNA libraries. The deduced amino acid sequence of human PTPζ is composed of a putative signal peptide of 19 amino acids, a very large extracellular domain of 1616 amino acids, a transmembrane peptide of 26 amino acids, and a cytoplasmic domain of 653 amino acids. The extracellular portion of human PTPζ contains two striking structural features: the N-terminal 280-amino acid sequence that is homologous to carbonic anhydrases (carbonate hydro-lyase, EC 4.2.1.1), and a sequence of 1048 amino acids without a cysteine residue. While it is unlikely that the carbonic anhydrase-like domain of PTPζ has any carbonic anhydrase activity, its three-dimensional structure may be quite similar to that of carbonic anhydrases, a structure that appears ideal for binding a small soluble ligand. The cytoplasmic portion of human PTPζ contains two repeated PTPase-like domains, which, when expressed in *Escherichia coli*, had PTPase activity *in vitro*. Mutational analyses indicate that only the membrane-proximal PTPase domain is catalytically active. Reverse transcription–polymerase chain reaction analyses indicate that human PTPζ is highly expressed in a glioblastoma cell line.**

The recent identification and characterization of a family of protein-tyrosine-phosphatases (PTPases, EC 3.1.3.48) suggest that dephosphorylation as well as phosphorylation of protein tyrosine residues plays an essential role in the regulation of diverse cell activities (1, 2). Given that hyperphosphorylation of protein tyrosine residues can cause cell transformation, it is possible that loss of PTPase activity may also be oncogenic. Indeed, the human gene encoding a receptor-linked PTPase, PTPγ, was mapped to the chromosomal region 3p21, which is frequently deleted in renal cell carcinomas and lung carcinomas (3).

To date, 15 human proteins containing PTPase domains have been reported. Each PTPase domain consists of ≈300 amino acid residues, of which about 40 are highly conserved. Typically, any two individual PTPase domains are 30–50% identical. Six of the cloned human PTPases are cytoplasmic enzymes, whereas the 9 other known human PTPases are transmembrane proteins with extracellular receptor-like regions connected to cytoplasmic PTPase domains (Fig. 1) (4–9). Four transmembrane PTPases have also been identified in *Drosophila* (6, 10–12). Whereas human PTPβ and its *Drosophila* homolog, DPTP10D, have only one PTPase domain, all other transmembrane PTPases contain two highly homologous PTPase domains, membrane-proximal domain 1 (D1) and C-terminal domain 2 (D2). However, our deletion and mutational analyses of LCA and LAR PTPases indicated that all of the catalytic activity was associated with D1 alone and that a conserved cysteine residue located in D1 was absolutely required for PTPase activity (6, 13, 14). Although the removal of LAR D2 by deletion did not abolish the catalytic activity of LAR D1, it did affect the substrate preference and modulator sensitivities (13, 15).

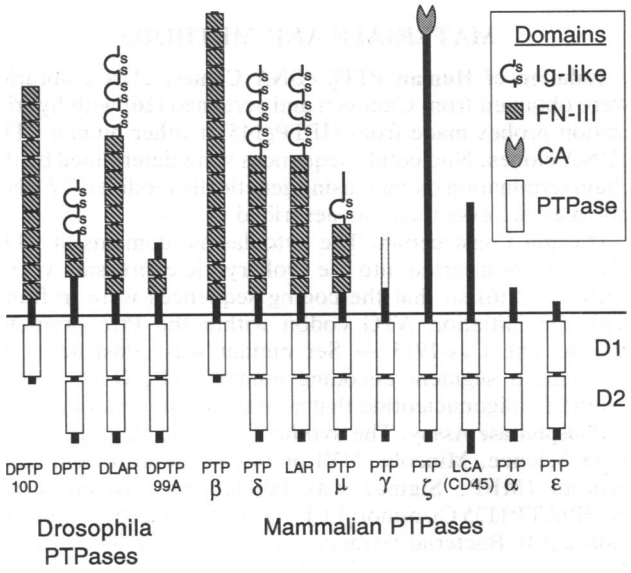

FIG. 1. Comparison of human and *Drosophila* transmembrane PTPases. The immunoglobulin (Ig)-like, fibronectin type III (FN-III), carbonic anhydrase (CA), and PTPase domains are schematically represented. Thick vertical lines indicate the segments of protein that are unique. D1 and D2, PTPase domains 1 and 2. LCA, leukocyte common antigen; LAR, LCA-related.

In contrast to the highly homologous PTPase domains, the extracellular receptor-like regions of the transmembrane PTPases are distinct in both size and structure. For example, the extracellular regions of LAR, PTPδ, PTPμ, DLAR, and DPTP contain varying numbers of immunoglobulin (Ig)-like domains and fibronectin type III (FN-III) domains (5–7, 9). The combination of Ig-like domains and FN-III domains is a

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Abbreviations: PTPase, protein-tyrosine-phosphatase; CA, carbonic anhydrase; LCA, leukocyte common antigen; LAR, LCA-related; MBP, myelin basic protein; RT–PCR, reverse transcription–polymerase chain reaction; D1, domain 1; D2, domain 2.

§To whom reprint requests should be addressed at: Division of Tumor Immunology, Dana–Farber Cancer Institute, 44 Binney Street, Boston, MA 02115.

characteristic structural motif of a family of cell adhesion molecules that includes neural cell adhesion molecule (NCAM). Thus, PTPase activities of LAR, PTPδ, PTPμ, DLAR, and DPTP may be regulated through cell-to-cell contact. In contrast, the extracellular regions of PTPβ, DPTP99A, and DPTP10D are composed only of FN-III domains (7, 10–12). LCA, PTPα, and PTPε have extracellular regions of various sizes with no obvious similarity to any known proteins (4, 7).

PTPζ and PTPγ comprise another subfamily, whose PTPase domains are 75% identical (7, 8). Here, we report the complete amino acid sequence of human PTPζ deduced from cDNA sequence[¶] and describe a type of receptor structure that is homologous to carbonic anhydrases (CAs; carbonate hydrolyase, EC 4.2.1.1).

MATERIALS AND METHODS

Isolation of Human PTPζ cDNA Clones. cDNA libraries were obtained from Clontech and screened (16) with hybridization probes made from HPTPζ145 or other human PTPζ cDNA clones. Nucleotide sequences were determined by the chain termination method using genetically modified T7 DNA polymerase, essentially as described (7, 17).

Plasmid Construction. The cytoplasmic domains of PTPζ cDNA were inserted into the prokaryotic expression vector pKKUC12 (6) so that the coding sequences were in frame with the initiation ATG codon within the Ptrc promoter region. The Cys-1913 → Ser mutant was constructed by replacing a segment encoding human PTPζ D1D2 with a synthetic oligonucleotide that contained the mutation.

Phosphatase Assay. The synthetic peptide Raytide (Oncogene Science, Mineola, NY) or bovine brain myelin basic protein (MBP; Sigma) was labeled at tyrosine using [γ-^{32}P]ATP (125 Ci/mmol; 1 Ci = 37 GBq) and p56^{lck} tyrosine kinase (14). Bacterial extracts were prepared as described (7). The phosphatase assay (50 μl) contains 25 mM Hepes (pH 7.3), 5 mM EDTA, 50 mM NaCl, 2.5 μg of bovine serum albumin, 10 mM dithiothreitol, ^{32}P-labeled substrate (40 nCi), and 5 μl of appropriately diluted sample to be assayed. After 30 min at 37°C, the reaction was terminated by addition of 750 μl of charcoal mixture [0.9 M HCl/90 mM sodium pyrophosphate/2 mM NaH_2PO_4/4% (vol/vol) Norit A]. Samples were microcentrifuged, and the radioactivity in 400 μl of supernatant was measured.[1]

Detection of Human PTPζ mRNA from Cell Lines by Reverse Transcription–Polymerase Chain Reaction (RT–PCR). The T-lymphocyte line REX is from our collection. Other cell lines were obtained from the American Type Culture Collection. RNA was prepared from cultured cells by using 4 M guanidinium isothiocyanate (18). cDNA was synthesized using Superscript RNase H$^-$ reverse transcriptase (Bethesda Research Laboratories) (19). PCR was done in a 50-μl mixture that contained 2.5 units of *Taq* DNA polymerase, 30 pmol of sense and antisense primers, 0.25 mM dNTPs, 10 mM Tris·HCl (pH 8.3), 50 mM KCl, 1.25 mM $MgCl_2$, and 0.5 μg of gelatin. Samples were amplified using the following protocol: 1 min at 92°C, 2 min at 60°C, and 2 min at 72°C, for 35 cycles. After amplification, 4 μl of each sample was electrophoresed through a 1% agarose gel in 40 mM Tris acetate buffer (pH 7.8) containing 2 mM EDTA and 1 μg of ethidium bromide per ml.

The following primers were used for PCR: PTPζ sense (positions 5716–5736), 5′-AGTGTGCAAGTGCTTGCCTAT-3′; PTPζ antisense (positions 6270–6250), 5′-TGCAGAATAGTCACTCTGCTG-3′; β-actin sense (positions 214–233), 5′-ATGGTGGGCATGGGTCAGAA-3′; β-actin antisense (positions 1039–1019), 5′-CAGTGATCTCCTTCTGCATCC-3′.

[¶]The sequence reported in this paper has been deposited in the GenBank data base (accession no. M93426).

RESULTS AND DISCUSSION

Isolation of Human PTPζ cDNA Clones from Fetal Brain cDNA Libraries. Previously, we isolated two human PTPζ cDNA clones from a fetal liver cDNA library (7). These clones, HPTPζ-145 and -147, encode only the cytoplasmic PTPase domains and transmembrane segment of human PTPζ. To obtain further coding sequences, several human cDNA libraries were screened for PTPζ clones. Although no additional PTPζ clones were isolated from tonsil, placenta, fetal kidney, and fetal lung libraries, we identified 74 PTPζ cDNA clones from two fetal brain cDNA libraries (Clontech). Fig. 2 shows the relative sizes of representative human PTPζ cDNA clones.

Structure of PTPζ. The complete nucleotide sequences of HPTPζ-145, -246, and -256, and substantial sequences of all the other cDNA clones shown in Fig. 2, were determined. By combining these data, the entire protein-coding region and 3′ noncoding sequence including poly(A) was obtained (Fig. 3). The human PTPζ cDNA sequence contains a long open reading frame, nucleotides 148–7089, that encodes a protein of 2314 amino acids.

The N terminus of the deduced human PTPζ amino acid sequence has a hydrophobic amino acid stretch characteristic of a signal peptide. The most likely cleavage site predicted by the method of von Heijne (20) is between the 19th and 20th amino acids from the N terminus. Thus, the predicted structure of the mature human PTPζ protein consists of a 1616-amino acid extracellular region, a 26-amino acid transmembrane segment, and a 653-amino acid cytoplasmic region. The cytoplasmic region is composed of duplicated PTPase-like domains (D1 and D2) that are highly homologous to human PTPγ D1 and D2, respectively: amino acid identity between PTPζ D1 and PTPγ D1 is 75.1% and that between PTPζ D2 and PTPγ D2 is 74.9%. The amino acid sequence identity of PTPζ to other PTPases is between 30% and 50% for both D1 and D2, with the notable exception of the D1 region of *Drosophila* DPTP99A, which is 60.0% identical to human PTPζ D1 (10–12). In contrast, DPTP99A D2 is only 35.8% identical to human PTPζ D2.

A computer search of the Protein Identification Resource protein sequence data base (Release 28.0) detected significant levels of sequence similarity between the N-terminal 280 amino acids of human PTPζ with various CAs (Fig. 4). CAs catalyze reversible hydration of CO_2 to bicarbonate ion, and

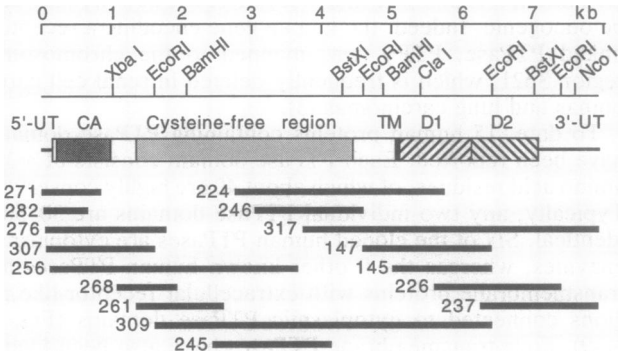

FIG. 2. Restriction map of HPTPζ cDNA and schematic model of the coding region. Shaded boxes represent functional domains, whereas thin lines at the both ends represent 5′ and 3′ untranslated (UT) regions. Below the restriction maps, thick lines indicate the sizes and extents of representative HPTPζ cDNA clones. CA, CA-like region; TM, transmembrane segment; kb, kilobases.

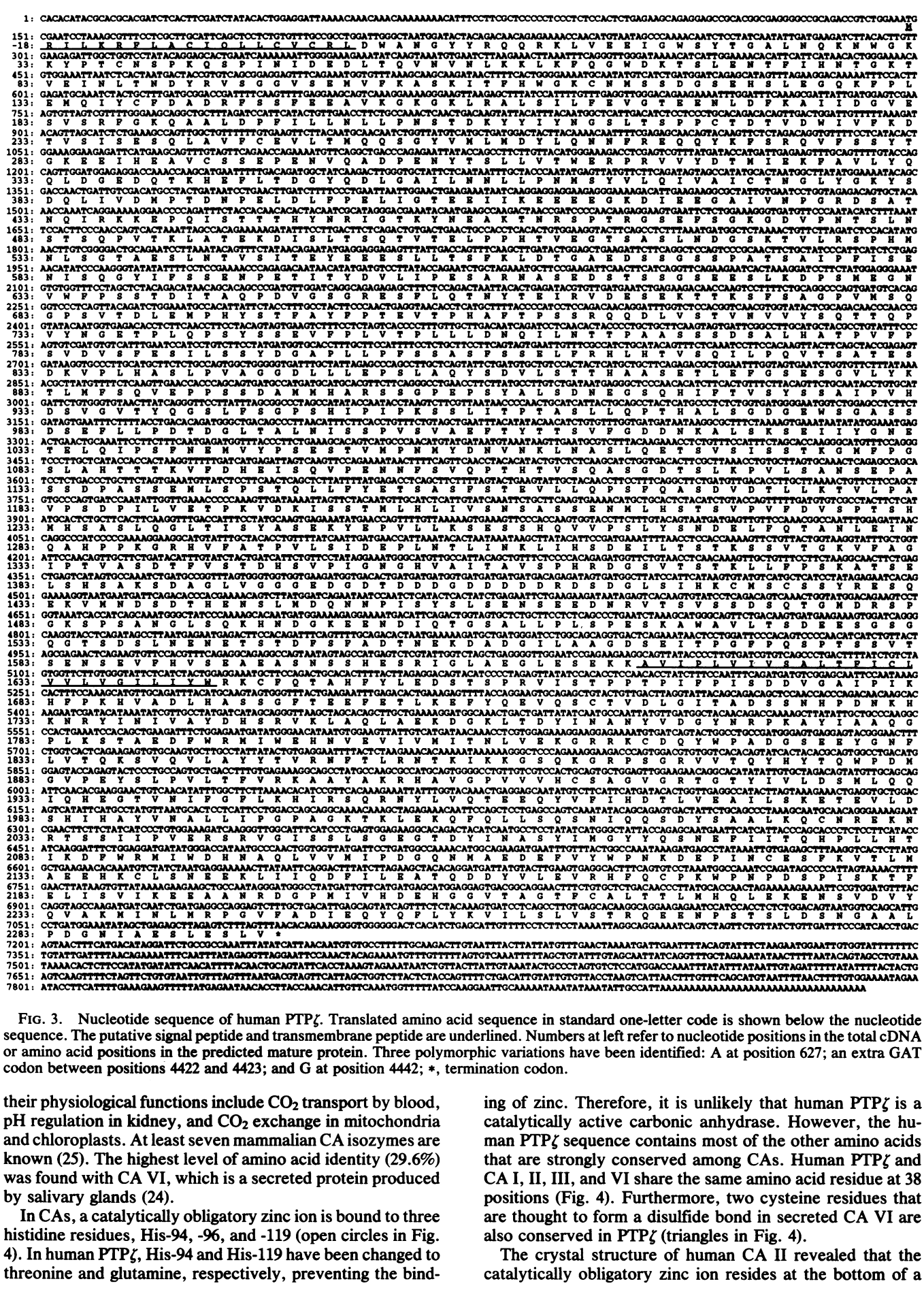

FIG. 3. Nucleotide sequence of human PTPζ. Translated amino acid sequence in standard one-letter code is shown below the nucleotide sequence. The putative signal peptide and transmembrane peptide are underlined. Numbers at left refer to nucleotide positions in the total cDNA or amino acid positions in the predicted mature protein. Three polymorphic variations have been identified: A at position 627; an extra GAT codon between positions 4422 and 4423; and G at position 4442; *, termination codon.

their physiological functions include CO_2 transport by blood, pH regulation in kidney, and CO_2 exchange in mitochondria and chloroplasts. At least seven mammalian CA isozymes are known (25). The highest level of amino acid identity (29.6%) was found with CA VI, which is a secreted protein produced by salivary glands (24).

In CAs, a catalytically obligatory zinc ion is bound to three histidine residues, His-94, -96, and -119 (open circles in Fig. 4). In human PTPζ, His-94 and His-119 have been changed to threonine and glutamine, respectively, preventing the binding of zinc. Therefore, it is unlikely that human PTPζ is a catalytically active carbonic anhydrase. However, the human PTPζ sequence contains most of the other amino acids that are strongly conserved among CAs. Human PTPζ and CA I, II, III, and VI share the same amino acid residue at 38 positions (Fig. 4). Furthermore, two cysteine residues that are thought to form a disulfide bond in secreted CA VI are also conserved in PTPζ (triangles in Fig. 4).

The crystal structure of human CA II revealed that the catalytically obligatory zinc ion resides at the bottom of a

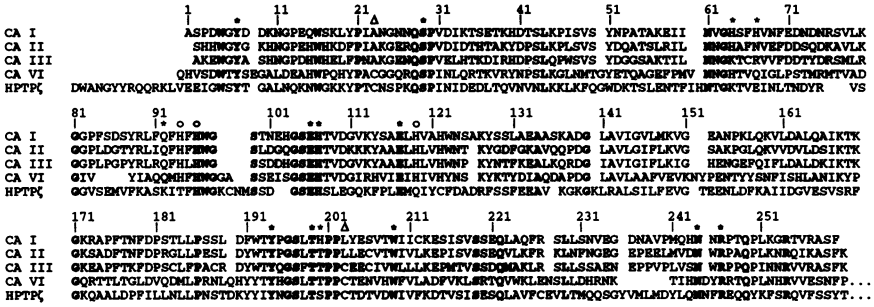

FIG. 4. Alignment of the sequences of human (H) PTPζ with human CA I, II, III, and VI (21–24). Positions where all sequences share an identical amino acid are highlighted by boldface type. Zinc-bound histidine residues are labeled by open circles, and amino acid residues that form the hydrogen-bond network in CAs are marked by asterisks. Cysteine residues, which in CA VI form an intramolecular disulfide bond, are indicated by triangles. The numbers above the amino acid sequences refer to the positions in CA I.

conical substrate binding cleft, and many amino acids surrounding the cleft are involved in an extensive hydrogen-bond network (asterisks in Fig. 4) (26). Preliminary molecular modeling suggested that the amino acid sequence of human PTPζ is superimposable onto the structure of CA II without major distortion (data not shown). When the conserved amino acids in human PTPζ were located on the CA II crystal structure, many were found to line the inner wall of the cleft and correspond to the positions involved in the hydrogen-bond network. Therefore, the CA-like domain of PTPζ probably forms a deep and moderately wide hydrophobic pocket that would be suitable as a ligand binding site (27, 28). Because CAs bind a variety of compounds such as sulfonamides and aldehydes as well as carbon dioxide and bicarbonate, the ligand of PTPζ may have structural similarity to these compounds. However, the hydrophobic pocket of CA is large enough to accommodate a short peptide (27, 28). Except for the N-terminal CA-like domain, the remainder of the extracellular region had no significant sequence similarity to any known protein sequences. A possibly significant feature, however, is that a 1048-amino acid stretch, amino acids 374–1421, lacks any cysteine (see Fig. 3).

PTPζ D1 Expressed in *Escherichia coli* Is Catalytically Active. To examine whether both D1 and D2 of human PTPζ have PTPase activity, we expressed the whole cytoplasmic region (HPTPζ-D1D2) or only D1 (HPTPζ-D1) by using the prokaryotic expression vector pKKUC12. We also made a mutant of HPTPζ-D1D2 that contained a Cys → Ser mutation at the catalytically obligatory cysteine residue at position 1913 in D1 (HPTPζ-D1D2-C1913S) (Fig. 5 *Upper*). The alkaline phosphatase-defective *E. coli* strain DHB4 was transformed with these expression plasmids, and the production of PTPζ peptides was induced by isopropyl β-D-thiogalactoside. The extracts prepared from the induced *E. coli* were serially diluted and assayed for PTPase activities using [^{32}P]phosphotyrosine-labeled MBP or oligopeptide Raytide as substrate. HPTPζ-D1D2 and HPTPζ-D1 expressed in *E. coli* had PTPase activity but the C1913S mutation completely abolished the PTPase activity of HPTPζ-D1D2 (Fig. 5 *Lower*), suggesting that D2 lacks PTPase activity in our *in vitro* assay system. These results are consistent with the fact that human PTPζ D2 contains a naturally occurring amino acid substitution that changes the obligatory cysteine residue to aspartate. Therefore, these data and our previous results on LCA and LAR (6, 13) indicate that the role of D2 in these PTPases is regulatory rather than catalytic, in spite of the strong similarity of the amino acid sequences between D1 and D2. However, D2 of PTPα has been reported to have a weak catalytic activity (29). PTPζ dephosphorylated Raytide substantially better than MBP. Previously, we have

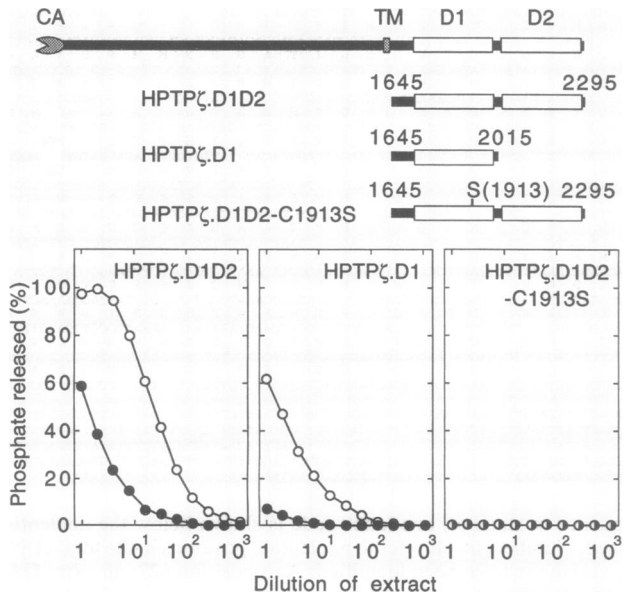

FIG. 5. (*Upper*) Cytoplasmic domains expressed in *E. coli*. The whole PTPζ molecule is shown for comparison. Open boxes indicate PTPase D1 and D2. Numbers above the boxes indicate the amino acid positions. CA, CA-like region; TM, transmembrane segment. (*Lower*) PTPase activity in extracts of *E. coli* DHB4 cells transformed with expression plasmids. The amounts of the expressed PTPζ peptide in each extract were comparable as visualized by SDS/PAGE (data not shown). PTPase activities in serially diluted extracts were assayed using [^{32}P]phosphotyrosine-labeled Raytide (○) or MBP (●) at 37°C for 30 min. The amount of radioactivity released as inorganic phosphate was expressed as the percentage of the total input radioactivity.

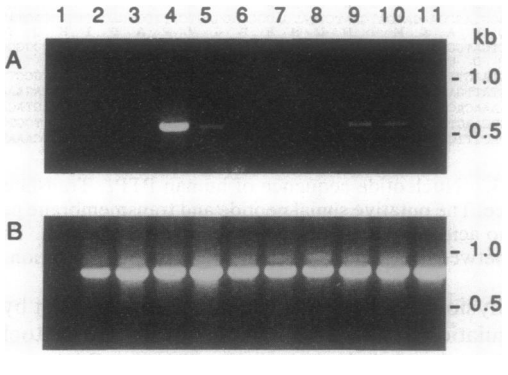

FIG. 6. RT–PCR detection of PTPζ mRNA (*A*) or control β-actin mRNA (*B*) in cultured human cell lines. Expected sizes of the amplified fragments are 555 base pairs for PTPζ and 826 base pairs for β-actin. Lanes: 1, no RNA; 2, T lymphoma REX; 3, B lymphoma Raji; 4, glioblastoma U-373MG; 5, neuroglioma H4; 6, neuroblastoma SK-N-SH; 7, neuroblastoma IMR-32; 8, kidney adenocarcinoma A-704; 9, hepatoblastoma HepG2; 10, normal fetal lung cell line FHs738; 11, intestinal smooth muscle cell line HISM.

demonstrated that LCA PTPase also has higher activity with Raytide than with MBP, whereas LAR, PTPα, PTPβ, and PTPδ PTPases have higher activity with MBP than with Raytide, and that such substrate specificity is conferred, at least partly, by D2 (7, 13). The catalytically inactive PTPζ D2 may also influence substrate specificity.

Expression of PTPζ in Human Cell Lines. To identify cell types that express PTPζ, we examined cultured human cell lines for the presence of PTPζ mRNA by RT-PCR. cDNA was synthesized from mRNA isolated from several cell lines, and PTPζ sequence and control β-actin sequence were amplified using *Taq* DNA polymerase. Fig. 6 shows that a glioblastoma cell line, U373MG, expressed PTPζ strongly; very weak PTPζ expression was observed in several cell lines. The large number of PTPζ cDNA clones found in brain libraries and the high level of expression in a glioblastoma cell line indicate that PTPζ is expressed preferentially in brain. Elucidation of the physiological role of PTPζ in brain function and/or development requires further study.

We thank Drs. Stan Tabor and Charles C. Richardson for T7 DNA polymerase, Dr. Christopher T. Walsh for $p56^{lck}$ kinase, Dr. David W. Christianson for molecular modeling and suggestions on CA structures, Dr. Michel Streuli for valuable advice, Dr. Tim Ernst for synthetic oligonucleotides, and Ms. Tran Thai for technical assistance. This work was supported in part by National Institutes of Health Grants CA51132 and AI26598, and by funds from Hoffmann–LaRoche, Inc.

1. Saito, H. & Streuli, M. (1991) *Cell Growth Differ.* **2**, 59–65.
2. Fischer, E. H., Charbonneau, H. & Tonks, N. K. (1991) *Science* **253**, 401–406.
3. LaForgia, S., Morse, B., Levy, J., Barnea, G., Cannizzaro, L. A., Li, F., Nowell, P. C., Boghosian-Sell, L., Glick, J., Weston, A., Harris, C. C., Drabkin, H., Patterson, D., Croce, C. M., Schlessinger, J. & Huebner, K. (1991) *Proc. Natl. Acad. Sci. USA* **88**, 5036–5040.
4. Streuli, M., Hall, L. R., Saga, Y., Schlossman, S. F. & Saito, H. (1987) *J. Exp. Med.* **166**, 1548–1566.
5. Streuli, M., Krueger, N. X., Hall, L. R., Schlossman, S. F. & Saito, H. (1988) *J. Exp. Med.* **168**, 1523–1530.
6. Streuli, M., Krueger, N. X., Tsai, A. Y. M. & Saito, H. (1989) *Proc. Natl. Acad. Sci. USA* **86**, 8698–8702.
7. Krueger, N. X., Streuli, M. & Saito, H. (1990) *EMBO J.* **9**, 3241–3252.
8. Kaplan, R., Morse, B., Huebner, K., Croce, C., Howk, R., Ravera, M., Ricca, G., Jaye, M. & Schlessinger, J. (1990) *Proc. Natl. Acad. Sci. USA* **87**, 7000–7004.
9. Gibbink, M. F. B. G., van Etten, I., Hateboer, G., Suijkerbuijk, R., Beijersbergen, R. L., van Kessel, A. G. & Moolenaar, W. H. (1991) *FEBS Lett.* **290**, 123–130.
10. Tian, S. S., Tsoulfas, P. & Zinn, K. (1991) *Cell* **67**, 675–685.
11. Yang, X., Seow, K. T., Bahri, S. M., Oon, S. H. & Chia, W. (1991) *Cell* **67**, 661–673.
12. Hariharan, I. K., Chuang, P. T. & Rubin, G. M. (1991) *Proc. Natl. Acad. Sci. USA* **88**, 11266–11270.
13. Streuli, M., Krueger, N. X., Thai, T., Tang, M. & Saito, H. (1990) *EMBO J.* **9**, 2399–2407.
14. Tsai, A. Y. M., Itoh, M., Streuli, M., Thai, T. & Saito, H. (1991) *J. Biol. Chem.* **266**, 10534–10543.
15. Itoh, M., Streuli, M., Krueger, N. X. & Saito, H. (1992) *J. Biol. Chem.* **267**, 12356–12363.
16. Benton, W. & Davis, R. (1977) *Science* **196**, 180–182.
17. Tabor, S. & Richardson, C. C. (1989) *J. Biol. Chem.* **264**, 6447–6458.
18. Chomczynski, P. & Sacchi, N. (1987) *Anal. Biochem.* **162**, 156–159.
19. Rothstein, D. M., Saito, H., Streuli, M., Scholossman, S. F. & Morimoto, C. (1992) *J. Biol. Chem.* **267**, 7139–7147.
20. von Heijne, G. (1986) *Nucleic Acids Res.* **14**, 4683–4690.
21. Barlow, J. H., Lowe, N., Edwards, Y. H. & Butterworth, P. H. W. (1987) *Nucleic Acids Res.* **15**, 2386–2386.
22. Montgomery, J. C., Venta, P. J., Tashian, R. E. & Hewett-Emmett, D. (1987) *Nucleic Acids Res.* **15**, 4687.
23. Wade, R., Gunning, P., Eddy, R., Show, T. & Kedes, L. (1986) *Proc. Natl. Acad. Sci. USA* **83**, 9571–9575.
24. Aldred, P., Fu, P., Barrett, G., Penschow, J. D., Wright, R. D., Coghlan, J. P. & Fernley, R. T. (1991) *Biochemistry* **30**, 569–575.
25. Tashian, R. E. (1990) *BioEssays* **10**, 186–192.
26. Eriksson, A. E., Jones, T. A. & Liljas, A. (1988) *Proteins Struct. Funct. Genet.* **4**, 274–282.
27. Nair, S. K., Tiffany, L. C., Christianson, D. W. & Fierke, C. A. (1991) *J. Biol. Chem.* **266**, 17320–17325.
28. Alexander, R. S., Nair, S. K. & Christianson, D. W. (1991) *Biochemistry* **30**, 11064–11072.
29. Wang, Y. & Pallen, J. C. (1991) *EMBO J.* **10**, 3231–3237.