Proc. Natl. Acad. Sci. USA
Vol. 87, pp. 7000–7004, September 1990
Cell Biology

Cloning of three human tyrosine phosphatases reveals a multigene family of receptor-linked protein-tyrosine-phosphatases expressed in brain

R. KAPLAN*, B. MORSE*, K. HUEBNER†, C. CROCE†, R. HOWK*, M. RAVERA*, G. RICCA*, M. JAYE*, AND J. SCHLESSINGER‡

*Rorer Biotechnology, Inc., 680 Allendale Road, King of Prussia, PA 19406; ‡Department of Pharmacology, New York University Medical Center, 550 First Avenue, New York, NY 10016; and †The Fels Institute for Cancer Research and Molecular Biology, Temple University School of Medicine, 3420 North Broad Street, Philadelphia, PA 19140

Communicated by Edmond H. Fischer, June 11, 1990

ABSTRACT **A human brainstem cDNA library in bacteriophage λgt11 was screened under conditions of reduced hybridization stringency with a leukocyte common antigen (LCA) probe that spanned both conserved cytoplasmic domains. cDNA encoding a receptor-linked protein-tyrosine-phosphatase (protein-tyrosine-phosphate phosphohydrolase, EC 3.1.3.48), RPTPase α, has been cloned and sequenced. Human RPTPase α consists of 802 amino acids. The extracellular domain of 150 residues includes a hydrophobic signal peptide and eight potential N-glycosylation sites. This is followed by a transmembrane region and two tandemly repeated conserved domains characteristic of all RPTPases identified thus far. The gene for RPTPase α has been localized to human chromosome region 20pter–20q12 by analysis of its segregation pattern in rodent–human somatic cell hybrids. Northern blot analysis revealed the presence of two major transcripts of 4.3 and 6.3 kilobases. In addition to RPTPase α, two other RPTPases (β and γ), identified in the same screen, have been partially cloned and sequenced. Analysis of sequence comparisons among LCA, the LCA-related protein LAR, and RPTPases α, β, and γ reveals the existence of a multigene family encoding different RPTPases, each containing a distinct extracellular domain, a single hydrophobic transmembrane region, and two tandemly repeated conserved cytoplasmic domains.**

The degree and pattern of phosphorylation of tyrosine residues on cellular proteins are regulated by the opposing activities of protein-tyrosine kinases (PTKases; ATP:protein-tyrosine *O*-phosphotransferase, EC 2.7.1.112) and protein-tyrosine-phosphatases (PTPases; protein-tyrosine-phosphate phosphohydrolase, EC 3.1.3.48). The structural characteristics and evolution of PTKases as well as their role in the regulation of cell growth have been considered elsewhere (1, 2). More recently, attention has also been focused on the growing family of PTPases. This family includes two types of molecules: (*i*) low molecular weight proteins such as PTPase 1B (3, 4), T-cell PTPase (5), and rat brain PTPase (6), which contain a single conserved phosphatase domain, and (*ii*) high molecular weight, receptor-linked PTPases (RPTPases) containing two tandemly repeated conserved domains separated by 56–57 amino acids. Examples of these include the leukocyte common antigen (LCA), also known as CD45 (7, 8), the LCA-related protein (LAR; ref. 9), and the LAR-related *Drosophila* proteins DLAR and DPTP (10).

Using a human LCA probe that spanned both conserved phosphatase domains, we screened a λgt11 human brainstem cDNA library under reduced stringency and identified three RPTPases, which we have designated α, β, and γ. Here we report the complete amino acid sequence of RPTPase α as well as its chromosome localization. In addition, we report the amino acid sequences of the conserved cytoplasmic domains of RPTPases β and γ. The nature of the structural homologies found among the conserved domains of these three RPTPases and previously known members of this protein family has been examined.§

MATERIALS AND METHODS

Materials. Restriction endonucleases and modifying enzymes were purchased from Boehringer Mannheim or New England Biolabs. *Taq* DNA polymerase was from Perkin–Elmer/Cetus. The λgt11 forward and reverse primers (24-mers) used in the polymerase chain reactions, as well as all sequencing primers, were synthesized on an automated DNA synthesizer (Applied Biosystems, model 380A) using either methoxy or β-cyanoethyl phosphoramidites (11). The λgt11 human brainstem cDNA library was obtained from the American Type Culture Collection (no. 37432). The LCA (CD45) clone used as a probe for screening the library was a generous gift from E. H. Fischer (University of Washington, Seattle). All sequencing reactions were performed using the Sequenase kit (United States Biochemical).

Isolation of cDNA Clones. Approximately 300,000 plaques from a λgt11 cDNA library of 1-day-old human infant brainstem were screened on duplicate nitrocellulose filters under conditions of reduced stringency with a nick-translated LCA probe that spanned both conserved phosphatase domains (3). Hybridization was carried out at 55°C overnight in a solution of 5× SSPE (SSPE is 10 mM NaH_2PO_4, pH 7.4/0.18 M NaCl/1 mM EDTA) containing 0.25% nonfat dry milk, 0.1% SDS, and ^{32}P-labeled LCA probe at 10^6 cpm/ml. The filters were washed three times for 20 min at 55°C in 2× SSPE/0.2% SDS and then processed for autoradiography. This screen yielded 79 duplicate positives; 12 of these, showing varying degrees of hybridization to the LCA probe, were plaque-purified by repetitious screening with the same probe. The polymerase chain reaction (12) was then used to determine the sizes of the cDNA inserts. The DNA templates consisted of portions of the eluates from each pure plaque, heated at 75°C for 15 min to release the DNA. The templates were primed with the λgt11 forward and reverse primers. The reaction mixtures (0.1 ml) were prepared as described (13). Amplification was achieved by performing 30 cycles, each

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: LCA, leukocyte common antigen; LAR, LCA-related protein; PTKase, protein-tyrosine kinase; PTPase, protein-tyrosine-phosphatase; RPTPase, receptor-linked PTPase.

§The nucleotide sequence of human RPTPase α has been deposited in the GenBank data base (accession no. M34668). The sequences of human RPTPases β and γ will be deposited when they have been completed.

including 1.5 min of denaturation at 94°C, 2 min of annealing at 65°C, and 4 min of extension at 72°C, in an automated Perkin–Elmer/Cetus DNA thermal cycler. A portion of each sample (15 μl) was analyzed by electrophoresis through a 1% agarose gel containing ethidium bromide at 1 μg/ml (14). DNA was prepared from the 4 largest clones by using LambdaSorb (Promega) and then digested with *Eco*RI. The fragments were subcloned separately into the *Eco*RI site of M13mp18 for sequencing. Nucleotide sequences were determined by the dideoxynucleotide chain-termination method (15) using modified T7 polymerase (16).

Analyses of Nucleotide and Amino Acid Sequences. All computer analyses of sequence data were performed on a Micro VAX II using programs written by IntelliGenetics. DNA sequences were analyzed and assembled using the GEL program. Hydrophobic analyses of proteins were based on the algorithm of Kyte and Doolittle (17), as implemented in the PEP program. Protein sequence alignments were done using the GENALIGN program (18–20). Initial alignments were done using the Jimenez-Montano reduced protein alphabet.¶

Northern Blots. Samples containing either 20 μg of total RNA or 2 μg of poly(A)$^+$ RNA were resolved in a formaldehyde/agarose gel and transferred to nitrocellulose. RPTPase α and β-actin probes were labeled by random priming (14). Hybridizations and washes were carried out at 65°C as described (22). Blots hybridized with the RPTPase α probe were exposed to XAR-2 x-ray film (Kodak) with an intensifying screen for 72 hr at −80°C. Results were obtained from the actin-probed blots after 15 hr under the same conditions.

Chromosome Localization of Human RPTPase α. Isolation, propagation, and characterization of parental and somatic cell hybrids used in this study have been described (23–25). Presence of specific human chromosomes or regions of chromosomes has been confirmed by DNA hybridization using probes for genes assigned to specific chromosome regions. Hybrid DNAs were digested with an excess of restriction endonuclease *Hin*dIII or *Eco*RI, sized by electrophoresis in 0.8% agarose gels, transferred to nylon filters, and hybridized as described (23). The RPTPase α probe consisted of the 3′-most 0.8 kilobase (kb) of clone 31-4 (see Fig. 1*B*).

¶Jimenez-Montano, M. & Zamora-Cortina, L., Proceedings 7th International Biophysics Congress, 1981, Mexico City.

RESULTS

Isolation and Characterization of Human RPTPase cDNA Clones. In an effort to identify new members of the PTPase family, we screened 300,000 plaques from a human infant brainstem cDNA library in λgt11 under nonstringent conditions using a nick-translated LCA probe that spanned both conserved phosphatase domains. Four of our initial 79 duplicate positives were sequenced in their entirety. Two clones, 31-4 and 27-1, contained overlapping portions of the entire coding region of a RPTPase that we designated α (Fig. 1*B*). The combined lengths of clones 31-4 and 27-1 equaled 3615 bp (Fig. 1*A*), encoding a protein of 802 amino acids (Fig. 1*C*) and containing an additional 695 bp and 510 bp, respec-

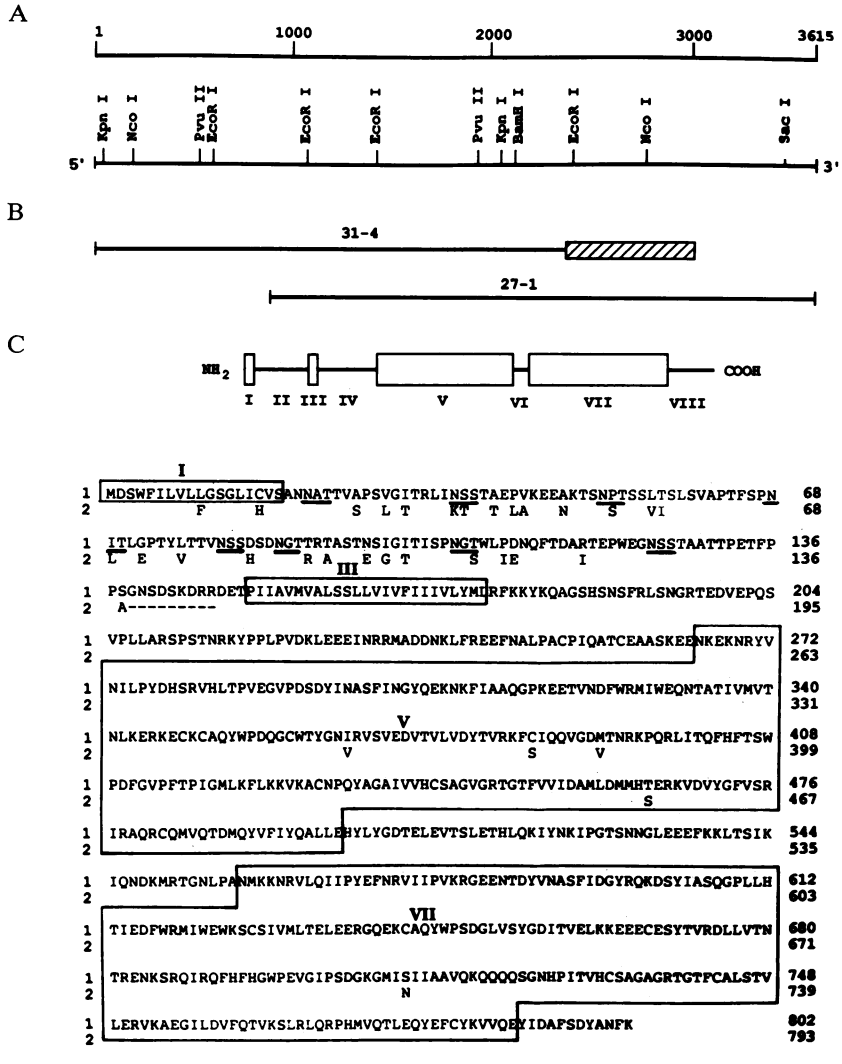

FIG. 1. Structure of human RPTPase α deduced from the sequence of cDNA clones. (*A*) Composite restriction map [3615 base pairs (bp)] of overlapping clones 31-4 and 27-1, which together contain the entire coding region of human RPTPase α. (*B*) Relative positions of clones 31-4 and 27-1. Both strands of each clone were sequenced in their entirety by using a series of oligonucleotide primers. The hatched region in clone 31-4 corresponds to the fragment used as probe for the Northern blot (Fig. 3) as well as for the chromosome assignment. (*C*) Comparison of the amino acid sequences of human (line 1) and mouse (line 2) RPTPase α. The single-letter amino acid code is used. Only the differences are shown. The dashed line indicates a stretch of amino acids not present in the mouse sequence. The coding portion of human RPTPase α, and its position relative to clones 31-4 and 27-1 (*B*), is shown at the top. The following regions are designated: signal peptide (I), extracellular domain with potential N-glycosylation sites for the human protein underlined (II), transmembrane (III), juxtamembrane (IV), first phosphatase domain (V), interdomain (VI), second phosphatase domain (VII), C terminus (VIII).

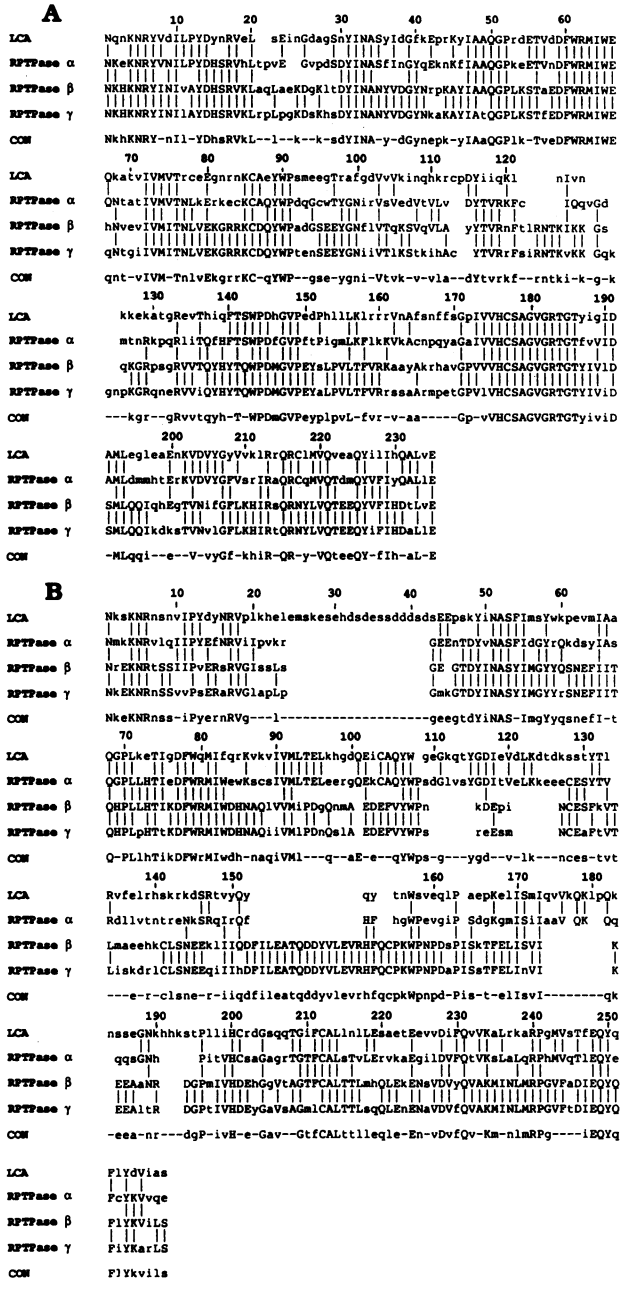

FIG. 2. Comparison of the amino acid sequences of the first (*A*) and second (*B*) conserved phosphatase domains of human RPTPases LCA, α, β and γ. CON is the consensus sequence: a capital letter indicates complete agreement, while a small letter indicates agreement among two or three of the four sequences. A dash indicates no consensus.

tively, of 5′ and 3′ untranslated region. Two of the four clones contained portions of genes coding for two additional RPTPases, designated β and γ (Fig. 2). Like RPTPase α, these two proteins contain typical hydrophobic transmembrane regions and distinct extracellular domains, indicating that they also represent separate RPTPases (unpublished results).

Simultaneous with the cloning and sequencing of human RPTPase α, the murine counterpart of this gene was also cloned (26). A comparison of the mouse and human protein sequences is seen in Fig. 1*C*. With the exception of the extracellular domain, where some variability exists, only 5 residues are found that differ between the two proteins. An examination of the structure of human RPTPase α reveals the following features: a relatively short extracellular domain consisting of 150 residues that includes a hydrophobic signal

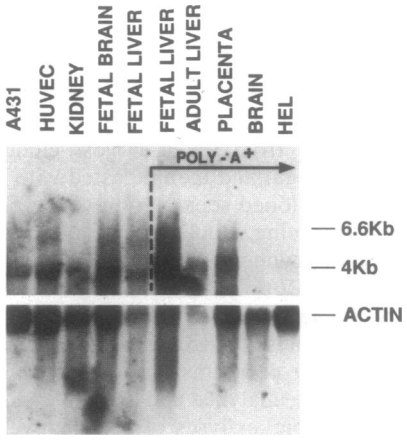

FIG. 3. Relative expression of human RPTPase α in various tissues and cell lines determined by Northern blot hybridization with RPTPase α probe (*Upper*) and B-actin probe (*Lower*). Total (five left lanes) or poly(A)$^+$ (five right lanes) RNA samples from the indicated human cell lines or tissues were analyzed. A431 and HEL are epidermoid carcinoma and erythroleukemia cell lines, respectively. All other lanes represent flash-frozen tissue samples. HUVEC is an acronym for human umbilical-vein endothelial cell tissue.

peptide containing the only cysteine in this region. There are eight potential N-glycosylation sites, and since this domain is also very rich in serine and threonine, there exist a number of potential O-glycosylation sites as well. The extracellular domains of RPTPase α, LCA, and LAR appear to be structurally unrelated. There is a hydrophobic transmembrane region anchored on both sides by charged residues. This is followed by the two tandemly repeated conserved phosphatase domains of about 235 residues each, which are separated by 57 amino acids, typical of RPTPases such as LCA (7, 8), LAR (9), and the two *Drosophila* PTPases DLAR and DPTP (10).

Fig. 2 *A* and *B* show the alignments of the amino acids within the first and second conserved phosphatase domains, respectively, of LCA and RPTPases α, β, and γ. It is readily apparent that among the four RPTPases, β and γ share the greatest sequence similarity. It was reported (10) that among the sequences of the conserved phosphatase domains of PTPase 1B, LCA, LAR, DLAR, and DPTP there are 29 invariant residues. While many of these residues are also present in both phosphatase domains of RPTPases α, β, and γ, it is interesting that the second conserved phosphatase domains of both β and γ lack a number of these amino acids, including the two cysteines at positions 104 and 201 in phosphatase domain 2 of LCA (see Fig. 2*B*).

Northern Blot Analysis. RPTPase α expression was examined in various cell lines and tissues (Fig. 3). The results indicate the presence of two major RNA transcripts of approximately 4.3 and 6.3 kb, respectively. The larger of the two species appears to be more prevalent in fetal tissues and is particularly prominent in the poly(A)$^+$ fetal liver sample, where there is also the highest relative amount of the 4.3-kb transcript. It is possible that the differential expression of the two transcripts is developmentally regulated and/or a result of alternative splicing mechanisms, a feature seen with LCA (7). The adult brain shows the least relative expression of RPTPase α. The results suggest that RPTPase α is expressed to some degree throughout many tissues. Murine RPTPase α was shown to be expressed in many tissues and cell lines and most abundantly in brain and kidney (26).

Chromosome Localization of the Human RPTPase α Gene. DNAs from 17 rodent–human somatic cell hybrids carrying overlapping subsets of human chromosome regions representing the entire human genome were tested for presence of

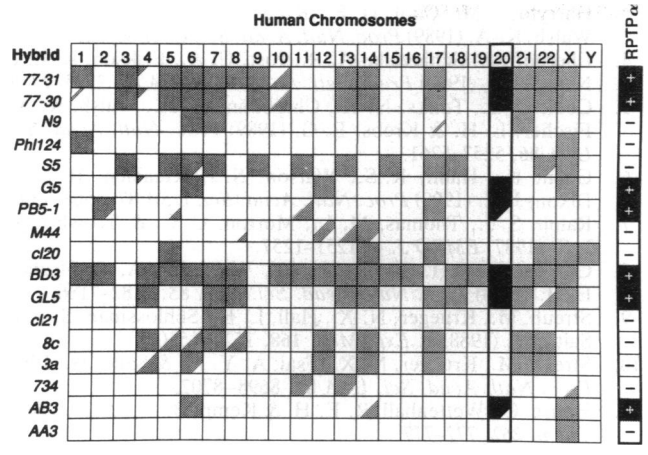

FIG. 4. Chromosomal localization of human RPTPase α by analysis of a panel of 17 rodent–human somatic cell hybrids. A completely stippled box indicates that the hybrid contained the chromosome indicated in the upper row; lower-right stippling indicates presence of the long arm (or part of the long arm, indicated by a smaller fraction of stippling) of the chromosome; upper-left stippling indicates presence of the short arm (or partial short arm) of the chromosome; an open box indicates absence of the chromosome. The column for chromosome 20 is boldly outlined and stippled to highlight correlation of presence of this chromosome (or chromosome region) with the presence of the RPTPase α gene. The pattern of retention of the human RPTPase α sequences in the hybrids is shown at right (RPTPα), where presence of the gene is indicated by a stippled box with a plus sign and absence of the gene is indicated by an open box enclosing a minus sign.

the human RPTPase α locus by Southern blot analysis (23–25). The results (Fig. 4) show that presence of the human RPTPase α locus in hybrid cells correlates only with presence of a partial human chromosome 20. The data also allow a regional localization for the RPTPase α locus, since hybrids PB5-1 and AB3 are each missing a part of the long arm of chromosome 20 and yet retain the RPTPase α locus. Thus, the human RPTPase α gene maps to 20pter–20q12.

DISCUSSION

The protein phosphatases are composed of at least two separate and distinct families (27), the protein-serine/threonine-phosphatases and the protein-tyrosine-phosphatases. This is in contrast to protein kinases, which show clear sequence similarity between serine/threonine- and tyrosine-specific enzymes. There appear to be two varieties of PTPase molecules: the first group is comprised of small, soluble enzymes that contain a single catalytic domain, such as placental PTPase 1B (3, 4), T-cell PTPase (5), and rat brain PTPase (6). The second group is made up of the more complex, receptor-linked enzymes, RPTPases such as LCA (7, 8), LAR (9), and the LAR-related *Drosophila* proteins DLAR and DPTP (10).

The sequences of the conserved phosphatase domains of the three human RPTPases identified here (α, β, and γ) have been compared with one another as well as with those of LCA, LAR, and two soluble PTPases, placental phosphatase 1B and T-cell PTPase (Table 1). The two soluble enzymes have a sequence identity of 70%; however, when each is compared with the RPTPases (phosphatase domains PD1 or PD2), this number drops to 29–42%. In all cases, the soluble PTPases showed a greater identity with PD1 than with PD2 of the RPTPases. RPTPase α appears to be most related to LAR, since their PD1 sequences are 56% identical and their PD2 sequences are 52% identical. The conserved domains of RPTPases β and γ are most related to each other, even more so than are the two soluble PTPases, since β and γ are 75% identical in both PD1 and PD2. Whether or not this degree of similarity is functionally significant remains to be seen. It is interesting that the relationship between PD1 and PD2 within any RPTPase appears, in general, to be no closer than that seen between different members of the family; i.e., the identities between PD1 and PD2 range from a high of 47% for LAR to a low of 29% for RPTPase γ. Every RPTPase characterized thus far contains two phosphatase domains, but the reason for this remains unclear. Several possibilities come to mind: (*i*) both domains are catalytically active and behave in a cooperative manner; (*ii*) both domains are catalytically active but have different substrate specificities; (*iii*) only one domain has enzymatic activity, while the other is regulatory in function. The data in Fig. 2 and Table 1 indicate that there is somewhat more constraint against sequence divergence in PD1 than in PD2. This may be consistent with the notion that PD1 possesses the catalytic activity, while PD2 may serve a different function, perhaps regulatory in nature.

The gene for human RPTPase α has been mapped to chromosome region 20pter–20q12. It was previously proposed (28) that all human chromosome 20 loci mapped thus far in mouse are found on mouse chromosome 2. This appears to be true for RPTPase α as well (26). The long arm of human chromosome 20 is involved in translocations and deletions in myeloid disorders and neoplasms (29). It will be useful to determine whether the human RPTPase α locus is specifically involved in deletions on 20q, thus making it a candidate tumor-suppressor gene, or antioncogene. Along the same lines, deletion of mouse chromosome 2 may be involved in

Table 1. Percent identities between conserved phosphatase domains

	PTPase 1B	T-cell PTPase	LCA		LAR		RPTPase α		RPTPase β		RPTPase γ	
			PD1	PD2	PD1	PD2	PD1	PD2	PD1	PD2	PD1	PD2
PTPase 1B	100											
T-cell PTPase	70	100										
LCA PD1	37	36	100									
LCA PD2	30	26	31	100								
LAR PD1	39	42	50	28	100							
LAR PD2	29	33	42	34	47	100						
RPTPase α PD1	36	38	50	32	56	45	100					
RPTPase α PD2	33	34	40	32	41	52	43	100				
RPTPase β PD1	35	39	41	31	33	41	47	33	100			
RPTPase β PD2	29	30	31	30	31	34	31	37	30	100		
RPTPase γ PD1	35	34	32	29	39	36	34	32	75	27	100	
RPTPase γ PD2	29	29	30	28	32	36	31	34	33	75	29	100

Alignments of the conserved phosphatase domains were carried out as described in *Materials and Methods*. The regions compared are designated in Fig. 1*C* and Fig. 2. PD, phosphatase domain.

the development of radiation-induced myeloid leukemia in SJL/J mice (30).

While we are beginning to understand more about the structure and diversity of the PTPases, much remains to be learned about their cellular functions. Tonks *et al.* (31) suggested that the small, soluble enzymes may have a "housekeeping" function. On the other hand, the RPTPases would be more restricted in their activities because they are located in the cell membrane and potentially regulated by extracellular ligands. Some progress has been made with regard to understanding the role of LCA (CD45) in T cells. T-cell clones deficient in their expression of surface LCA failed to proliferate in response to antigen or crosslinked CD3 (32). It was reported (33, 34) that the PTPase activity of LCA plays a role in the activation of $pp56^{lck}$, a lymphocyte-specific PTKase. It was hypothesized that the phosphatase activity of LCA activates $pp56^{lck}$ by dephosphorylation of a C-terminal tyrosine residue. This event may in turn be related to T-cell activation.

Streuli *et al.* (10) used site-directed mutagenesis to determine whether or not any of four conserved cysteines in LCA (two in each phosphatase domain) was required for activity toward artificial substrates. They concluded that only one of the cysteine residues is essential for PTPase activity (residue 177 of PD1 of LCA, Fig. 2*A*) and that most likely only the first phosphatase domain has enzymatic activity. However, they did not rule out the possibility that the second domain is also able to dephosphorylate a different substrate. More recently, Streuli *et. al.* (21) determined that the second conserved domains of LCA and LAR had no detectable phosphatase activity but that sequences within these domains could influence substrate specificity.

While the cytoplasmic domains of RPTPase α, β, and γ are highly conserved, the extracellular domains of these receptors are unrelated to one another as well as to those of LAR and LCA (data not shown). This suggests that each of these receptors has its own distinct ligand. It is likely that the binding of such ligands to the RPTPases plays a crucial role, together with growth factor receptors exhibiting PTKase activity, in the regulation of the level of tyrosine phosphorylation of target proteins involved in signal transduction. The diversity of the RPTPases described in this report reveals the existence of a multigene family. Further analysis of structure–function relationships among these membrane receptors will provide important insights into the mechanisms involved in cell growth, differentiation, and oncogenesis.

This study was initiated as a collaboration with the laboratory of Dr. E. H. Fischer. We acknowledge fruitful discussion with Dr. Fischer and members of his laboratory including Dr. N. Tonks, Dr. G. Daum, and Dr. D. Cool. We gratefully acknowledge technical assistance by M. Nathan, E. Murray, and T. Druck and secretarial assistance by R. Ratkiewicz. We acknowledge the critical review of the manuscript by Dr. D. Hurwitz and Dr. A. Zilberstein. This study was supported in part by National Institutes of Health Grants CA21124 and CA39860 (to C.C. and K.H.).

1. Hunter, T. & Cooper, J. A. (1985) *Annu. Rev. Biochem.* **54,** 897–930.
2. Ullrich, A. & Schlessinger, J. (1990) *Cell* **61,** 203–212.
3. Charbonneau, H., Tonks, N. K., Kumar, S., Diltz, D. C., Harrylock, M., Cool, D. E., Krebs, E. G., Fischer, E. H. & Walsh, K. A. (1989) *Proc. Natl. Acad. Sci. USA* **86,** 5252–5256.
4. Chernoff, J., Schievella, A. R., Jost, C. A., Erikson, R. L. & Neel, B. G. (1990) *Proc. Natl. Acad. Sci. USA* **87,** 2735–2739.
5. Cool, D. E., Tonks, N. K., Charbonneau, H., Walsh, K. A., Fischer, E. H. & Krebs, E. G. (1989) *Proc. Natl. Acad. Sci. USA* **86,** 5257–5261.
6. Guan, K., Haun, R. S., Watson, S. J., Geahlen, R. L. & Dixon, J. E. (1990) *Proc. Natl. Acad. Sci. USA* **87,** 1501–1505.
7. Ralph, S. J., Thomas, M. L., Morton, C. C. & Trowbridge, I. S. (1987) *EMBO J.* **6,** 1251–1257.
8. Charbonneau, H., Tonks, N. K., Walsh, K. A. & Fischer, E. H. (1988) *Proc. Natl. Acad. Sci. USA* **85,** 7182–7186.
9. Streuli, M., Krueger, N. X., Hall, L. R., Schlossman, S. F. & Saito, H. (1988) *J. Exp. Med.* **168,** 1523–1530.
10. Streuli, M., Krueger, N. X., Tsai, A. Y. M. & Saito, H. (1989) *Proc. Natl. Acad. Sci. USA* **86,** 8698–8702.
11. House, C., Wettenhall, R. E. H. & Kemp, B. E. (1987) *J. Biol. Chem.* **262,** 772–777.
12. Saiki, R. K., Scharf, S., Faloona, F., Mullis, K. B., Horn, G. T., Erlich, H. A. & Arnhemim, N. (1985) *Science* **230,** 1350–1354.
13. Dionne, C. A., Kaplan, R., Seuanez, H., O'Brien, S. J. & Jaye, M. (1990) *Biotechniques* **8,** 190–194.
14. Maniatis, T., Fritsch, E. F. & Sambrook, J. (1982) *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Lab., Cold Spring Harbor, NY).
15. Sanger, F., Nicklen, S. & Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. USA* **74,** 5463–5467.
16. Tabor, S. & Richardson, C. C. (1987) *Proc. Natl. Acad. Sci. USA* **84,** 4767–4771.
17. Kyte, J. & Doolittle, R. F. (1982) *J. Mol. Biol.* **157,** 105–132.
18. Sobel, E. & Martinez, H. M. (1985) *Nucleic Acids Res.* **14,** 363–374.
19. Karlin, S., Ghandour, G., Foulser, D. E. & Korn, L. J. (1984) *Mol. Biol. Evol.* **1,** 357–370.
20. Needleman, S. B. & Wunsch, C. D. (1970) *J. Mol. Biol.* **48,** 443–453.
21. Streuli, M., Krueger, N. X., Thai, T., Tang, M. & Saito, H. (1990) *EMBO J.* **9,** 2399–2407.
22. Church, G. & Gilbert, W. (1984) *Proc. Natl. Acad. Sci. USA* **81,** 1991–1995.
23. Durst, M., Croce, C. M., Gissman, L., Schwarz, E. & Huebner, K. (1987) *Proc. Natl. Acad. Sci. USA* **84,** 1070–1074.
24. Ku, D.-H., Travali, S., Calabretta, B., Huebner, K. & Baserga, R. (1989) *Somatic Cell Mol. Genet.* **15,** 297–307.
25. Juan, C.-C., Hwang, J., Liu, A. A., Whang-Peng, J., Knutsen, T., Huebner, K., Croce, C. M., Zhang, H., Wang, J. C. & Liu, L. F. (1988) *Proc. Natl. Acad. Sci. USA* **85,** 8910–8913.
26. Sap, J., D'Eustachio, P., Givol, D. & Schlessinger, J. (1990) *Proc. Natl. Acad. Sci. USA* **87,** 6112–6116.
27. Hunter, T. (1989) *Cell* **58,** 1013–1016.
28. Lalley, P. A., Davisson, M. T., Graves, J. A. M., O'Brien, S. J., Womack, J. E., Roderick, T. H., Creau-Goldberg, N., Hillyard, A. I., Doolittle, D. P. & Rogers, J. A. (1989) *Cytogenet. Cell Genet.* **51,** 503–532.
29. Trent, J. M., Kaneko, Y. & Mitleman, F. (1989) *Cytogenet. Cell Genet.* **51,** 533–562.
30. Trakhtenbrot, L., Krauthgamer, R., Resnitzky, P. & Haran-Ghera, N. (1988) *Leukemia* **2,** 545–550.
31. Tonks, N. K., Charbonneau, H., Diltz, C. D., Fischer, E. H. & Walsh, K. A. (1988) *Biochemistry* **27,** 8695–8701.
32. Pingel, J. T. & Thomas, M. L. (1989) *Cell* **58,** 1055–1065.
33. Mustelin, T., Coggeshall, K. M. & Altman, A. (1989) *Proc. Natl. Acad. Sci. USA* **86,** 6302–6306.
34. Ostergaard, H. L., Shackelford, D. A., Hurley, T. R., Johnson, P., Hyman, R., Sefton, B. M. & Trowbridge, I. S. (1989) *Proc. Natl. Acad. Sci. USA* **86,** 8959–8963.