Proc. Natl. Acad. Sci. USA
Vol. 87, pp. 2735–2739, April 1990
Biochemistry

Cloning of a cDNA for a major human protein-tyrosine-phosphatase

(tyrosine phosphorylation/dephosphorylation/human placental cDNA)

JONATHAN CHERNOFF*†, ANDREA R. SCHIEVELLA*, CHRISTINE A. JOST‡, R. L. ERIKSON*, AND BENJAMIN G. NEEL‡

*Department of Cellular and Developmental Biology, Harvard University, 16 Divinity Avenue, Cambridge, MA 02138; and ‡Molecular Medicine Unit, Beth Israel Hospital, 330 Brookline Avenue, Boston, MA 02215

Contributed by R. L. Erikson, January 3, 1990

ABSTRACT **We have isolated a cDNA clone encoding the major protein-tyrosine-phosphatase (protein-tyrosine-phosphate phosphohydrolase, EC 3.1.3.48) of human placenta. Degenerate oligonucleotides, based on the amino acid sequence of the protein, were used to amplify an internal fragment of the gene from human placental cDNA by the polymerase chain reaction. This fragment was then used to probe a human placental cDNA library. A 3.3-kilobase (kb) insert was isolated and sequenced. The insert has a single extended open reading frame that predicts a 435 amino acid protein of $M_r \approx 50{,}000$. From the amino terminus to residue 321, the deduced amino acid sequence is identical to that previously determined by peptide sequencing [Charbonneau, H., Tonks, N. K., Kumar, S., Diltz, C. D., Harrylock, M., Cool, D. E., Krebs, E. G., Fischer, E. H. & Walsh, K. A. (1989) *Proc. Natl. Acad. Sci. USA* 86, 5252–5256]; however, the sequence predicts that the protein contains an additional 114 amino acids not present in the reported peptide sequence. *In vitro* translation of the 3.3-kb insert produces a protein of M_r 56,000, in general agreement with the predicted size. The phosphatase gene appears to be present as a single copy in human genomic DNA and is transcribed into a 3.5-kb message in a variety of tissues.**

The phosphorylation of proteins at tyrosine has been recognized as an important regulatory component in signal transduction, neoplastic transformation, and, more recently, in the control of the mitotic cycle (1–4). As in systems regulated by serine or threonine phosphorylation, the phosphorylation of proteins at tyrosine is reversible, and protein-tyrosine-phosphatases (PTPases; protein-tyrosine-phosphate phosphohydrolase, EC 3.1.3.48) from a variety of tissues and organisms have been described (for review, see ref. 5). Although a general classification scheme for these enzymes does not yet exist, there are apparently a number of protein phosphatases that are highly specific for phosphotyrosyl proteins and that bear little resemblance to either the protein-serine- or protein-threonine-phosphatases or the acid and alkaline phosphatases. In most tissues, the major PTPase (as measured *in vitro* by using a variety of artificial substrates) is a metal-ion-independent, vanadate- and molybdate-sensitive species of $M_r \approx 40{,}000$. Recently, Tonks and coworkers (6, 7) have purified a form of this enzyme (PTP-1B) from human placenta and determined its amino acid sequence. The protein sequence is unrelated to those of other known phosphatases but is similar to the common leukocyte antigen (CD45) and to LAR, a homologue of the neural adhesion molecule (NCAM) (8). The PTPase activity of CD45 has been verified experimentally (9), and, recently, an additional member of the PTPase gene family has been cloned from human T cells (10). In this report, we describe the cloning of the cDNA for the major human PTPase from human placenta.§

MATERIALS AND METHODS

Oligonucleotides. Oligonucleotides were synthesized on a Milligen (Bradford, MA) apparatus by using phosphoramidite chemistry and were purified with phenyl solid-phase extraction (J. T. Baker) columns (11).

Polymerase Chain Reaction. DNA amplification was carried out on a Perkin–Elmer thermal cycler, in a 50-μl vol containing 50 mM Tris·HCl (pH 8.3), 50 mM KCl, 1.5 mM $MgCl_2$, 200 μM dNTPs, 0.01% gelatin, 0.1% Triton X-100, 1 μM of each primer, 2.5 units of *Thermus aquaticus* (*Taq*) polymerase (Promega), and 10 ng of bacteriophage DNA. Thirty cycles of amplification were performed (94°C for 40 sec, 50°C for 1 min, and 72°C for 1 min), and 10% of the product was analyzed on a 3% agarose (NuSeive; FMC) gel. The product was cut out, isolated on glass beads, and reamplified. The reamplified band was isolated as above, digested with *Eco*RI and *Bam*HI, and ligated into pBluescript SK+ (Stratagene).

Isolation of PTPase Clones. The amplified PTPase fragment was concatenated by treatment with T4 ligase and then labeled with [^{32}P]dCTP (New England Nuclear) to a specific activity of 5×10^8 cpm/μg by the random-priming technique (12). Two hundred thousand plaques of a human placental cDNA library in λgt11 (Clontech) were screened by standard methods (13).

DNA Sequencing. The *Eco*RI insert from λgt11 was subcloned into pBluescript, and the nucleotide sequence of the PTPase insert was obtained by the dideoxy chain-termination technique of Sanger *et al.* (14) after construction of a series of nested deletions (15). In some cases, oligonucleotides were used to obtain additional sequence information in regions not represented by the nested deletions.

Northern (RNA) Blot Analysis. Tissues were quick-frozen and powdered in liquid nitrogen, and total RNA was extracted by the guanidium isothiocyanate method (16). Ten to 15 μg was electrophoresed on a 1% agarose/6% formaldehyde gel, transferred to a Biotrans (ICN) membrane by capillary blotting with 10× standard saline citrate (SSC) (1× SSC is 0.15 M sodium chloride/0.015 M sodium citrate, pH 7), and UV cross-linked. The filter was hybridized overnight at 65°C in 0.5 M $NaPO_4$, pH 7.2/7% $NaDodSO_4$/1% bovine serum albumin with full-length [3.3 kilobases (kb)] [^{32}P]-dCTP-labeled PTPase cDNA and washed twice with 2× SSC at room temperature and twice with 0.2× SSC/0.2% $NaDodSO_4$ at 65°C. Autoradiograms were obtained by ex-

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.

Abbreviation: PTPase, protein-tyrosine-phosphatase.

†To whom reprint requests should be addressed.

§The sequence reported in this paper has been deposited in the GenBank data base (accession no. M31724).

posing the blots to Kodak XAR-5 film with intensifying screens at −70°C for 72 hr.

Southern Blot Analysis. Total genomic DNA was prepared from fresh human placenta and mouse fibroblasts by standard methods (17). Ten micrograms was digested with restriction enzymes and electrophoresed on a 0.7% agarose gel in 89 mM Tris·HCl/89 mM boric acid/2 mM EDTA. The DNA was transferred to a Hybond (Amersham) filter and UV cross-linked. Hybridization was done in 50% (vol/vol) formamide/5× standard saline phosphate with EDTA (SSPE) (1× SSPE is 0.18 M NaCl/10 mM phosphate, pH 7.4/1 mM EDTA)/1× Denhardt's solution (1× Denhardt's solution is 0.02% polyvinylpyrrolidone/0.02% Ficoll/0.02% bovine serum albumin)/1% $NaDodSO_4$/salmon sperm DNA with full-length ^{32}P-labeled PTPase probe at 42°C for 24 hr. The blot was then washed twice in 2× SSC at room temperature and twice in 2× SSC/0.1% $NaDodSO_4$ at 60°C (low-stringency) or 0.2× SSC/0.1% $NaDodSO_4$ (high-stringency) at 65°C. Autoradiograms were exposed as above for 72 hr.

RESULTS

Cloning of the Major Human Placental PTPase. Initially, we attempted to isolate PTP-1B cDNA by screening a human placental cDNA library with a 48-base-pair (bp) "guess-mer" oligonucleotide, designed to correspond to the amino acids from positions 95–110 (7). However, as this approach proved unsuccessful, we chose to use an alternative method. Degenerate oligonucleotides were constructed based on two regions of the protein, as described in Fig. 1. These regions were chosen to limit the number of possible codon combinations and also to limit the size of the DNA fragment for amplification. In addition, these regions differed from any present in CD45 or LAR to avoid amplifying fragments of those genes. By use of cDNA from a λgt11 human placental library as template, a fragment of the expected size was amplified and subsequently subcloned into pBluescript. The sequence of this fragment correctly predicted the known amino acid sequence between the two primers (data not shown).

To obtain the complete cDNA for this enzyme, the PTPase gene fragment was used to probe the human placental cDNA library. A single plaque, with a 3.3-kb insert, hybridized to the probe. The plaque was purified, and the insert was subcloned into pBluescript and sequenced (Fig. 2). The sequence contains a single long open translational reading frame of 1305 bp, which encodes a protein of 435 amino acids. The putative translation inhibition codon at nucleotide 91 conforms to the translation start consensus sequence (A/G)NNATGG (18). The open reading frame is terminated by a stop codon (TAG) at nucleotide 1396. A polyadenylylation signal is present at nucleotides 3200–3205, followed by a polyadenylate tract. Over the first 321 residues, the predicted amino acid sequence is in complete agreement with the published amino acid sequence as determined by protein-sequencing techniques. However, the predicted sequence does not terminate at this point but, rather, extends an additional 114 residues before a termination codon is reached. The predicted molecular weight for the phosphatase becomes, then, 49,996 rather than the reported 37,354. We believe that this clone specifies the full-length protein, as the predicted initiating methionine has been found to be blocked by N-acetylation in the authentic protein (7). It remains possible that alternative forms of the gene are transcribed or that the transcript contains a longer 5′ leader sequence, as human PTP-1B mRNA appears somewhat larger than 3.3 kb (see below).

Unlike CD45 and LAR, PTP-1B has no repeated domains. The carboxyl terminus is rich in hydrophobic residues and is unrelated to the carboxyl terminus of the T-cell PTPase. Our deduced amino acid sequence was compared with other reported sequences by using the FASTA algorithm (19). This analysis revealed weak similarities between the carboxyl terminus of PTP-1B and that of the yeast protein KAR1 (20), as well as with a short stretch of the hypothetical Epstein–Barr virus protein BALF (21) (data not shown).

***In Vitro* Translation.** mRNA was synthesized by using T7 polymerase from a vector containing the 3.3-kb insert, followed by *in vitro* translation in the presence of [^{35}S]methionine (Fig. 3). The translated protein migrates as an M_r 56,000 protein, 6000 larger than expected. The 5′ region of the PTPase clone lacks any ATG codons upstream of the predicted initiating methionine, making it unlikely that the translated protein contains any extra amino acids at the amino terminus. At the carboxyl terminus, translation should stop at the TAG codon at position 1396–1399. Although our nucleotide sequence could be in error, this is extremely doubtful, as the sequence of both strands is unambiguous in this region. Posttranslational modifications, which may alter protein mobility on $NaDodSO_4$/acrylamide gels, have not been reported under these conditions in a rabbit reticulocyte lysate. The amino acid composition of PTP-1B does not suggest an obvious explanation for the anomalous migration seen in this experiment. When this cDNA is expressed in *Escherichia coli*, a highly active PTPase of apparent M_r 56,000 is produced (A.S., unpublished data).

Tissue Distribution of PTP-1B mRNA. Northern blot analysis of RNAs from a variety of human tissues demonstrates that the predominant phosphatase transcript is present as a 3.5-kb species. A second, larger band of ≈4 kb is also

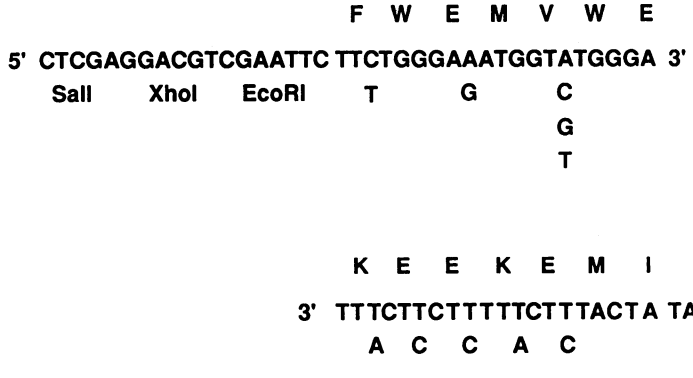

FIG. 1. Oligonucleotide primers used for polymerase chain reaction amplification of internal PTP-1B fragment. Amino acids at positions 95–101 and 128–134 (7) were reverse translated into all possible codons that could potentially specify these residues. The indicated multicloning sites were synthesized at the 5′ end of each primer to facilitate ligating the amplified gene fragment into Bluescript vectors. The nucleotide sequence of each primer is presented below the amino acid sequence in single-letter code, and positions of degeneracy are indicated.

```
1   GGGCGGGCCTCGGGGCTAAGAGCGCGACGCCTAGAGCGGCAGACGGCGCAGTGGGCCGAGAAGGAGGCGCAGCAGCCGCCCTGGCCCGTCATGGAGATGG  100
                                                                                                  M  E  M  E
5   AAAAGGAGTTCGAGCAGATCGACAAGTCCGGGAGCTGGGCGGCCATTTACCAGGATATCCGACATGAAGCCAGTGACTTCCCATGTAGAGTGGCCAAGCT  200
      K  E  F  E  Q  I  D  K  S  G  S  W  A  A  I  Y  Q  D  I  R  H  E  A  S  D  F  P  C  R  V  A  K  L
38  TCCTAAGAACAAAAACCGAAATAGGTACAGAGACGTCAGTCCCTTTGACCATAGTCGGATTAAACTACATCAAGAAGATAATGACTATATCAACGCTAGT  300
      P  K  N  K  N  R  N  R  Y  R  D  V  S  P  F  D  H  S  R  I  K  L  H  Q  E  D  N  D  Y  I  N  A  S
71  TTGATAAAAATGGAAGAAGCCCAAAGGAGTTACATTCTTACCCAGGGCCCTTTGCCTAACACATGCGGTCACTTTTGGGAGATGGTGTGGGAGCAGAAAA  400
      L  I  K  M  E  E  A  Q  R  S  Y  I  L  T  Q  G  P  L  P  N  T  C  G  H [F  W  E  M  V  W  E] Q  K  S
105 GCAGGGGTGTCGTCATGCTCAACAGAGTGATGGAGAAAGGTTCGTTAAAATGCGCACAATACTGGCCACAAAAAGAAGAAAAAGAGATGATCTTTGAAGA  500
      R  G  V  V  M  L  N  R  V  M  E  K  G  S  L  K  C  A  Q  Y  W  P  Q [K  E  E  K  E  M  I] F  E  D
138 CACAAATTTGAAATTAACATTGATCTCTGAAGATATCAAGTCATATTATACAGTGCGACAGCTAGAATTGGAAAACCTTACAACCCAAGAAACTCGAGAG  600
      T  N  L  K  L  T  L  I  S  E  D  I  K  S  Y  Y  T  V  R  Q  L  E  L  E  N  L  T  T  Q  E  T  R  E
171 ATCTTACATTTCCACTATACCACATGGCCTGACTTTGGAGTCCCTGAATCACCAGCCTCATTCTTGAACTTTCTTTTCAAAGTCCGAGAGTCAGGGTCAC  700
      I  L  H  F  H  Y  T  T  W  P  D  F  G  V  P  E  S  P  A  S  F  L  N  F  L  F  K  V  R  E  S  G  S  L
205 TCAGCCCGGAGCACGGGCCCGTTGTGGTGCACTGCAGTGCAGGCATCGGCAGGTCTGGAACCTTCTGTCTGGCTGATACCTGCCTCCTGCTGATGGACAA  800
      S  P  E  H  G  P  V  V  V  H  C  S  A  G  I  G  R  S  G  T  F  C  L  A  D  T  C  L  L  L  M  D  K
238 GAGGAAAGACCCTTCTTCCGTTGATATCAAGAAAGTGCTGTTAGAAATGAGGAAGTTTCGGATGGGGTTGATCCAGACAGCCGACCAGCTGCGCTTCTCC  900
      R  K  D  P  S  S  V  D  I  K  K  V  L  L  E  M  R  K  F  R  M  G  L  I  Q  T  A  D  Q  L  R  F  S
271 TACCTGGCTGTGATCGAAGGTGCCAAATTCATCATGGGGGACTCTTCCGTGCAGGATCAGTGGAAGGAGCTTTCCCACGAGGACCTGGAGCCCCCACCCG 1000
      Y  L  A  V  I  E  G  A  K  F  I  M  G  D  S  S  V  Q  D  Q  W  K  E  L  S  H  E  D  L  E  P  P  P  E
305 AGCATATCCCCCCACCTCCCCGGCCACCCAAACGAATCCTGGAGCCACACAATGGGAAATGCAGGGAGTTCTTCCCAAATCACCAGTGGGTGAAGGAAGA 1100
      H  I  P  P  P  P  R  P  P  K  R  I  L  E  P  H  N  G  K  C  R  E  F  F  P  N  H  Q  W  V  K  E  E
338 GACCCAGGAGGATAAAGACTGCCCCATCAAGGAAGAAAAAGGAAGCCCCTTAAATGCCGCACCCTACGGCATCGAAAGCATGAGTCAAGACACTGAAGTT 1200
      T  Q  E  D  K  D  C  P  I  K  E  E  K  G  S  P  L  N  A  A  P  Y  G  I  E  S  M  S  Q  D  T  E  V
371 AGAAGTCGGGTCGTGGGGGGAAGTCTTCGAGGTGCCCAGGCTGCCTCCCCAGCCAAAGGGGAGCCGTCACTGCCCGAGAAGGACGAGGACCATGCACTGA 1300
      R  S  R  V  V  G  G  S  L  R  G  A  Q  A  A  S  P  A  K  G  E  P  S  L  P  E  K  D  E  D  H  A  L  S
405 GTTACTGGAAGCCCTTCCTGGTCAACATGTGCGTGGCTACGGTCCTCACGGCCGGCGCTTACCTCTGCTACAGGTTCCTGTTCAACAGCAACACATAGCC 1400
      Y  W  K  P  F  L  V  N  M  C  V  A  T  V  L  T  A  G  A  Y  L  C  Y  R  F  L  F  N  S  N  T  ***
    TGACCCTCCTCCACTCCACCTCCACCCACTGTCCGCCTCTGCCCGCAGAGCCCACGCCCGACTAGCAGGCATGCCGCGGTAGGTAAGGGCCGCCGGACCG 1500
    CGTAGAGAGCCGGGCCCCGGACGGACGTTGGTTCTGCACTAAAACCCATCTTCCCCGGATGTGTGTCTCACCCCTCATCCTTTTACTTTTTGCCCCTTCC 1600
    ACTTTGAGTACCAAATCCACAAGCCATTTTTTGAGGAGAGTGAAAGAGAGTACCATGCTGGCGGCGCAGAGGGAAGGGGCCTACACCCGTCTTGGGGCTC 1700
    GCCCCACCCAGGGCTCCCTCCTGGAGCATCCCAGGCGGCGCACGCCAACAGCCCCCCCCTTGAATCTGCAGGGAGCAACTCTCCACTCCATATTTATTTA 1800
    AACAATTTTTTCCCCAAAGGCATCCATAGTGCACTAGCATTTTCTTGAACCAATAATGTATTAAAATTTTTTGATGTCAGCCTTGCATCAAGGGCTTTAT 1900
    CAAAAAGTACAATAATAAATCCTCAGGTAGTACTGGGAATGGAAGGCTTTGCCATGGGCCTGCTGCGTCAGACCAGTACTGGGAAGGAGGACGGTTGTAA 2000
    GCAGTTGTTATTTAGTGATATTGTGGGTAACGTGAGAAGATAGAACAATGCTATAATATATAATGAACACGTGGGTATTTAATAAGAAACATGATGTGAG 2100
    ATTACTTTGTCCCGCTTATTCTCCTCCCTGTTATCTGCTAGATCTAGTTCTCAATCACTGCTCCCCCGTGTGTATTAGAATGCATGTAAGGTCTTCTTGT 2200
    GTCCTGATGAAAAATATGTGCTTGAAATGAGAAACTTTGATCTCTGCTTACTAATGTGCCCCATGTCCAAGTCCAACCTGCCTGTGCATGACCTGATCAT 2300
    TACATGGCTGTGGTTCCTAAGCCTGTTGCTGAAGTCATTGTCGCTCAGCAATAGGGTGCAGTTTTCCAGGAATAGGCATTTGCTAATTCCTGGCATGACA 2400
    CTCTAGTGACTTCCTGGTGAGGCCCAGCCTGTCCTGGTACAGCAGGGTCTTTGCTGTAACTCAGACATTCCAAGGGTATGGGAAGCCATATTCACACCTCA 2500
    CGCTCTGGACATGATTTAGGGAAGCAGGACACCCCCCGCCCCCCACCTTTGGGATCAGCCTCCGCCATTCCAAGTCAACACTCTTCTTGAGCAGACCGTG 2600
    ATTTGGAAGAGAGGCACCTGCTGGAAACCACACTTCTTGAAACAGCCTGGGTGACGGTCCTTTAGGCAGCCTGCCGCCGTCTCTGTCCCGGTTCACCTTG 2700
    CCGAGAGAGGCGCGTCTGCCCCACCCTCAAACCCTGTGGGGCCTGATGGTGCTCACGACTCTTCCTGCAAAGGGAACTGAAGACCTCCACATTAAGTGGC 2800
    TTTTTAACATGAAAAACACGGCAGCTGTAGCTCCCGAGCTACTCTCTTGCCAGCATTTTCACATTTTGCCTTTCTCGTGGTAGAAGCCAGTACAGAGAAA 2900
    TTCTGTGGTGGGAACATTCGAGGTGTCACCCTGCAGAGCTATGGTGAGGTGTGGATAAGGCTTAGGTGCCAGGCTGTAAGCATTCTGAGCTGGCTTGTTG 3000
    TTTTTAAGTCCTGTATATGTATGTAGTAGTTTGGGTGTGTATATATAGTAGCATTTCAAAATGGACGTACTGGTTTAACCTCCTATCCTTGGAGAGCAGC 3100
    TGGCTCTCCACCTTGTTACACATTATGTTAGAGAGGTAGCGAGCTGCTCTGCTATATGCCTTAAGCCAATATTTACTCATCAGGTCATTATTTTTTACAA 3200
    TGGCCATGGAATAAACCATTTTTACAAAAATAAAAACAAAAAAAGCAAAAAAAAAA
```

FIG. 2. Nucleotide and predicted amino acid sequence of PTP-1B. The DNA sequence is numbered at right, and amino acids are numbered at left; amino acids appear in one-letter code beneath nucleotides. The termination codon is marked with stars, the polyadenylylation signal (AATAAA) is underlined, and the regions used to prime amplification of the internal fragment are boxed.

apparent, which could represent the transcript of a related gene or alternately spliced versions of a single primary PTPase message. The strongest hybridizations are to samples from kidney and liver. This pattern of distribution is similar to that observed by direct measurement of PTPase activity (22, 23).

Gene Structure. Genomic DNA from human placenta and mouse fibroblasts was digested with three restriction enzymes and subjected to Southern blot analysis with the PTPase probe (Fig. 4). The human placental PTPase gene appears as a single band ≈7–12 kb in size, unlike the complex pattern for the human T-cell PTPase gene (≈70 kb) described by Cool *et al.* (10). The pattern for mouse DNA is similar to that seen for human DNA. The simplicity of the hybridization pattern suggests that the gene is present as a single copy in human genomic DNA.

DISCUSSION

In this communication, we describe the molecular cloning of a cDNA corresponding to the major human PTPase previously purified and characterized by Fischer and colleagues (7). In contrast to their protein sequence analysis, however, the cDNA sequence predicts a protein of M_r 50,000, which contains, in addition to the reported 321 amino acids, an additional 114 (mostly hydrophobic) residues at the carboxyl terminus. We have previously reported the purification of a similar enzyme from bovine brain, termed PTPase-5, which

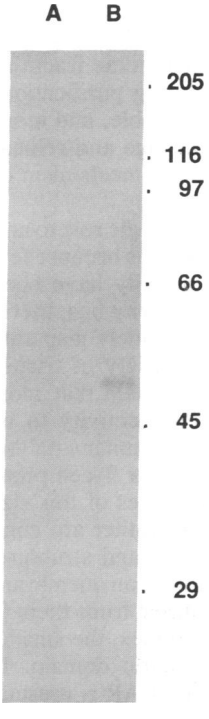

FIG. 3. *In vitro* translation of PTP-1B mRNA. Bluescript plasmids with and without the 3.3-kb PTP insert were transcribed from the T7 promoter. One microgram of mRNA was translated with micrococcal-nuclease-treated rabbit reticulocyte lysate (Promega) under conditions recommended by the manufacturer with [^{35}S]methionine in a 50-μl vol. Ten microliters of the extract was analyzed by electrophoresis on a 10% NaDodSO$_4$/polyacrylamide gel. Proteins were detected by autoradiography after 24 hr. Lanes: A, control mRNA; B, mRNA from Bluescript plasmid containing the 3.3-kb PTPase insert. M_r markers ($\times 10^{-3}$) are indicated by numbers.

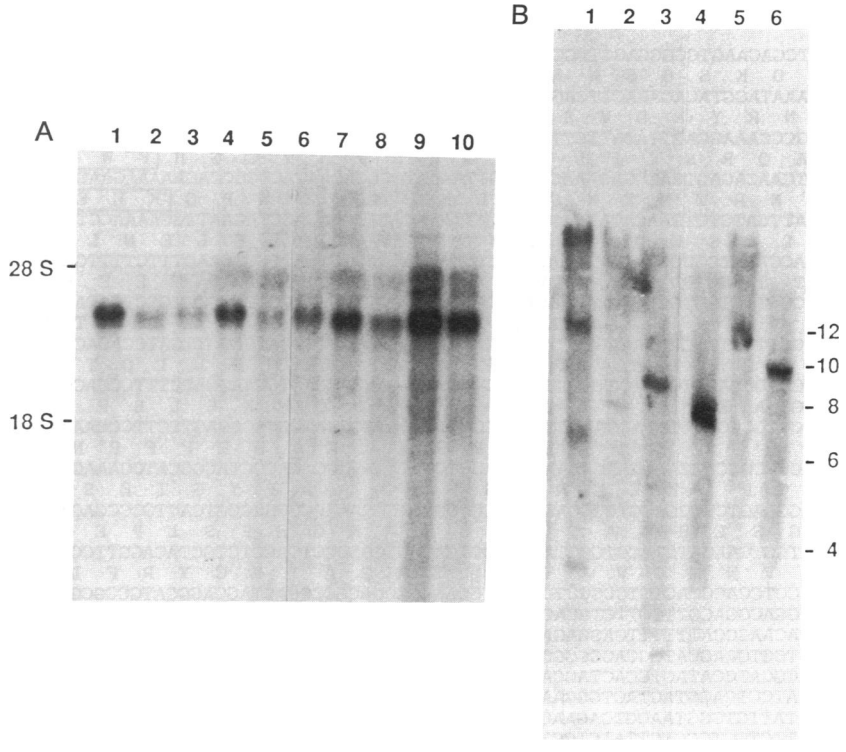

FIG. 4. Hybridization analysis of PTP-1B mRNA and genomic sequences. (*A*) Blot hybridization of PTP-1B mRNA. Ten micrograms of total RNA was electrophoresed per lane: human liver (lane 1), heart (atrium) (lane 2), heart (ventricle) (lane 3), HG-1 fibroblast cells (lane 4), lung (lane 5), EJ bladder carcinoma cell line (lane 6), Calu-6 lung cancer line (lane 7), placenta (lane 8), kidney (medulla) (lane 9), and kidney (cortex) (lane 10). The filter was exposed overnight. The positions of the 18S and 28S RNAs are marked. (*B*) Southern blot analysis of human placental and mouse fibroblast genomic DNA. Ten micrograms of murine (lanes 1–3) or human (lanes 4–6) DNA was digested with *Hin*dIII (lanes 1 and 4), *Eco*RI (lanes 2 and 5), or *Bam*HI (lanes 3 and 6). Human DNA was washed under high-stringency conditions, and mouse DNA was washed under low-stringency conditions. Numbers indicate the size of DNA markers in kb.

shares many properties with the human placental enzyme but has a M_r of 48,000 (24). Swarup and Subrahmanyam (25) purified a particulate PTPase from rat spleen that migrated at $M_r \approx 50{,}000$ when protease inhibitors were present in homogenization and extraction buffers but only at $M_r \approx 36{,}000$ when such precautions were not taken. This reduction in size was accompanied by an increase in enzymatic activity. Similar data have also been reported by others (26). A plausible explanation for these findings is that the enzyme is normally found in association with the membrane fraction, but that its hydrophobic tail is cleaved during purification, rendering the phosphatase smaller, more soluble, and more active. Conceivably, such proteolytic cleavage and relocation to the cytosol might represent a normal mechanism of enzyme regulation *in vivo*.

At present, it is difficult to ascribe a physiologic role to any of the particular PTPases. This situation obtains because few key protein targets for tyrosine kinase activity have been identified, despite intense study. PTPase activity has, therefore, been measured by using a variety of widely disparate phosphotyrosyl proteins. In view of the variety of tissues examined and substrates used, it is remarkable that most investigators have found the major PTPase activity to be represented by a protein of $M_r \approx 40{,}000$ with similar catalytic properties. Based both on the sequence of the T-cell phosphatase and on the present study, the PTPases of this size appear not to represent a single entity but rather are comprised of a group of proteins of similar size and structure. These enzymes are also closely related to the transmembrane PTPases, such as CD45 and LAR, but differ from them in several vital aspects. Unlike the latter enzymes, the smaller PTPases appear to contain a single catalytic domain. In addition, whereas the activity of CD45 and LAR is presumably regulated by binding of extracellular ligand, different mechanisms must exist for regulating the smaller phosphatases. PTP-1B and the T-cell PTPase are nearly identical over the first ≈300 amino acids but differ completely at their carboxyl termini. It would not be surprising if regulation of enzymatic activity or substrate specificity were dictated by this region.

Despite the lack of known physiologically significant substrates for tyrosine kinase activity, the PTPase described in this and previous studies is likely to have an important function, as it represents the major activity in most tissues examined. In addition, the markedly hydrophobic character of the carboxyl-terminus sequence of this enzyme suggests that it is membrane-associated, as are the majority of tyrosine kinases (1) and, presumably, at least some of their important substrates. In this context, we note that Tonks and coworkers (7) found that approximately half the PTP-1B activity was found in the particulate fraction.

The availability of PTPase clones may provide an avenue for analyzing the role of tyrosine phosphorylation in various cellular processes. The PTP-1B species may be particularly valuable in this regard, as the *in vitro* activity of the M_r 38,000 form of this enzyme is ≈100 times that of CD45 (in the absence of ligand) (9). This highly active form of the enzyme could be introduced into cells by engineering truncated versions of the PTP-1B cDNA in appropriate expression vectors.

Finally, the technique of DNA amplification by using mixed primers may prove useful in identifying additional members of the PTPase gene family. Inspection of the published amino acid sequences of various PTPases reveals several apparent consensus domains, to which degenerate oligonucleotides could be made. Such an approach has many

inherent advantages over alternate methods, such as using mixed or guess-mer oligonucleotides to probe cDNA libraries under low-stringency conditions (for discussion, see ref. 27). Indeed, we have already detected an apparent cDNA homologue to PTP-1B in murine tissues by using this approach (data not shown). It remains to be seen whether this method can be extended to more distantly related organisms.

Note. While this work was in progress, we learned that a similar clone has been isolated by David Hill (personal communication).

We thank David Hill of Applied Biotechnology for communication of his unpublished results on a PTP-1B cDNA clone. We are grateful to Drs. S. Izumo, S. Alper, F. Flier, and J. Gebert of Beth Israel Hospital for human RNA samples. This work was supported by grants from the National Institutes of Health (CA 49152, CA 08563, and CA 42580). R.L.E. is an American Cancer Society Professor of Cellular and Developmental Biology.

1. Hunter, T. & Cooper, J. A. (1985) *Annu. Rev. Biochem.* **54,** 897–930.
2. Dunphy, W. G. & Newport, J. W. (1989) *Cell* **58,** 181–191.
3. Morla, A. O., Draetta, G., Beach, D. & Wang, J. Y. J. (1989) *Cell* **58,** 193–203.
4. Gould, K. L. & Nurse, P. (1989) *Nature (London)* **342,** 39–45.
5. Lau, K.-H. W., Farley, J. R. & Baylink, D. J. (1989) *Biochem. J.* **257,** 23–36.
6. Tonks, N. K., Diltz, C. D. & Fischer, E. H. (1988) *J. Biol. Chem.* **263,** 6722–6730.
7. Charbonneau, H., Tonks, N. K., Kumar, S., Diltz, C. D., Harrylock, M., Cool, D. E., Krebs, E. G., Fischer, E. H. & Walsh, K. A. (1989) *Proc. Natl. Acad. Sci. USA* **86,** 5252–5256.
8. Strueli, M., Krueger, N. X., Hall, L. R., Schlossman, S. F. & Saito, H. (1988) *J. Exp. Med.* **168,** 1523–1530.
9. Tonks, N. K., Charbonneau, H., Diltz, C. D., Fischer, E. H. & Walsh, K. A. (1988) *Biochemistry* **27,** 8695–8700.
10. 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.
11. Horn, T. & Urdea, M. S. (1989) *Nucleic Acids Res.* **16,** 11559–11571.
12. Feinberg, A. P. & Vogelstein, B. (1983) *Anal. Biochem.* **132,** 6–13.
13. Maniatis, T., Fritsch, E. F. & Sambrook, J. (1982) *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Lab., Cold Spring Harbor, NY).
14. Sanger, F., Nicklen, S. & Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. USA* **74,** 5463–5467.
15. Heinikoff, S. (1984) *Gene* **28,** 351–359.
16. Chirgwin, J. M., Przybyla, A. E., MacDonald, R. J. & Rutter, W. J. (1979) *Biochemistry* **24,** 5294–5303.
17. Gattoni, S., Kirschmeier, P., Weinstein, I. B., Escobeda, J. & Dina, D. (1982) *Mol. Cell. Biol.* **2,** 42–51.
18. Kozak, M. (1986) *Cell* **44,** 283–292.
19. Pearson, W. P. & Lipman, D. J. (1988) *Proc. Natl. Acad. Sci. USA* **85,** 2444–2448.
20. Rose, M. D. & Fink, G. R. (1987) *Cell* **48,** 1047–1060.
21. Baer, R., Bankier, A. T., Biggen, M. D., Deininger, P. L., Farrell, P. J., Gibson, T. J., Hatfull, G., Hudson, G. S., Satchwell, S. C., Seguin, C., Tuffnell, P. S. & Barrell, B. G. (1984) *Nature (London)* **310,** 207–211.
22. Shriner, C. L. & Brautigan, D. L. (1984) *J. Biol. Chem.* **259,** 11383–11390.
23. Chernoff, J. & Li, H.-C. (1985) *Arch. Biochem. Biophys.* **240,** 135–145.
24. Jones, S. W., Erikson, R. L., Ingebritsen, V. M. & Ingebritsen, T. S. (1989) *J. Biol. Chem.* **264,** 7747–7753.
25. Swarup, G. & Subrahmanyam, G. (1989) *J. Biol. Chem.* **264,** 7801–7808.
26. Rottenberg, S. A. & Brautigan, D. L. (1987) *Biochem. J.* **243,** 747–754.
27. Gould, S. J., Subramani, S. & Scheffler, I. E. (1989) *Proc. Natl. Acad. Sci. USA* **86,** 1934–1938.