Proc. Natl. Acad. Sci. USA
Vol. 87, pp. 5148–5152, July 1990
Biochemistry

Molecular cloning and chromosome mapping of the human gene encoding protein phosphotyrosyl phosphatase 1B

(tyrosine phosphorylation/dephosphorylation/human cDNA/human chromosome 20)

SHERYL BROWN-SHIMER*, KAREN A. JOHNSON*, JEANNE BENTLEY LAWRENCE†, CAROL JOHNSON†, ARTHUR BRUSKIN*, NANCY R. GREEN*, AND DAVID E. HILL*‡

*Applied bioTechnology, Inc., 80 Rogers Street, Cambridge, MA 02142; and †Department of Cell Biology, University of Massachusetts Medical School, 55 Lake Avenue North, Worcester, MA 01655

Communicated by Robert A. Weinberg, April 16, 1990

ABSTRACT **The inactivation of growth suppressor genes appears to play a major role in the malignant process. To assess whether protein phosphotyrosyl phosphatases (protein-tyrosine-phosphate phosphohydrolase, EC 3.1.3.48) function as growth suppressors, we have isolated a cDNA clone encoding human protein phosphotyrosyl phosphatase 1B for structural and functional characterization. The translation product deduced from the 1305-nucleotide open reading frame predicts a protein containing 435 amino acids and having a molecular mass of 49,966 Da. The amino-terminal 321 amino acids deduced from the cDNA sequence are identical to the empirically determined sequence of protein phosphotyrosyl phosphatase 1B [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]. A genomic clone has been isolated and used in an *in situ* hybridization to banded metaphase chromosomes to determine that the gene encoding protein phosphotyrosyl phosphatase 1B maps as a single-copy gene to the long arm of chromosome 20 in the region q13.1–q13.2.**

Changes in the phosphorylation state of protein tyrosyl residues are an important aspect of growth regulation. Ligand-induced tyrosine phosphorylation by receptor tyrosine kinases leads to changes in gene expression and alters the growth of cells (1). In addition, neoplastic transformation by oncoproteins, such as src and fps, is dependent upon intrinsic tyrosine kinase activity (2, 3). Although the overall importance of protein tyrosine phosphorylation in mediating growth regulation is appreciated, the role of dephosphorylation of phosphotyrosyl residues in regulating growth remains unclear. For example, a net increase in the phosphotyrosine content of cells due to the action of protein tyrosine kinases is well documented (4, 5), but the same net increase could easily arise due to loss or reduction in the activity of a protein phosphotyrosyl phosphatase (PTPase; protein-tyrosine-phosphate phosphohydrolase, EC 3.1.3.48). In either case, the resulting net increase in tyrosine phosphorylation could lead to oncogenic transformation. Recently, it has been postulated that PTPases could function as growth suppressors (6, 7). Support for this hypothesis comes from the observation that vanadate, a potent inhibitor of PTPases, causes a transient increase in the cellular phosphotyrosine content as well as phenotypic transformation of some fibroblast cell lines (8). A variety of cancers are associated with specific cytogenetic aberrations that may involve deletions of growth suppressor genes. One prediction for the growth suppressor hypothesis is that inactivating mutations in the genes encoding PTPases would therefore accompany some cancers.

PTPases have been identified in many different eukaryotic cell types in a broad range of sizes from both particulate and soluble cell fractions (9–11). The best characterized of these is the 37,000-Da PTPase 1B from human placenta, which has been purified to homogeneity and the amino acid sequence of which has been determined (12). A cDNA encoding a putative T cell-specific PTPase has also been isolated and sequenced (13). Recently it has been shown that the transmembrane glycoproteins CD45 (also termed L-CA or T200) and LAR share amino acid similarity with PTPase 1B and that CD45 exhibits PTPase activity (6, 12, 14–17). Thus, PTPases comprise a diverse family consisting of both transmembrane glycoproteins and cytosolic proteins. While the enzymology of intracellular PTPases is being extensively studied (10, 18), physiologically relevant substrates for these enzymes have not yet been identified, and the mode of regulation of enzymatic activity has not been elucidated. However, PTPase 1B from placenta and PTPase 5 from bovine brain have been shown to dephosphorylate the phosphotyrosyl forms of insulin receptor and v-src, respectively, *in vitro* (11, 19), and purified PTPase 1B injected into *Xenopus* oocytes delays insulin-dependent maturation (18). These results suggest that the cytoplasmic PTPases can modulate signal transduction mediated by protein tyrosine kinases.

To assess the role of the cytoplasmic PTPases in growth regulation and to address the question of whether PTPases can specifically function as growth suppressors, we have used degenerate oligonucleotides deduced from the PTPase 1B amino acid sequence to isolate a cDNA clone.§ The 3.2-kilobase cDNA has a continuous open reading frame that predicts a protein of 435 amino acids. By using the cDNA as a probe, a genomic clone has been isolated and used in a nonisotopic *in situ* hybridization reaction to demonstrate that the human PTP1B gene, which encodes PTPase 1B, is located on the long arm of chromosome 20.

MATERIALS AND METHODS

Degenerate Oligonucleotides Used for cDNA Isolation. Two minimally degenerate oligonucleotides were used to screen a human placental cDNA library (Clontech). The oligopeptide sequences [amino acid numbering based on Charbonneau *et al.* (12)] and their corresponding degenerate oligonucleotides (degenerate positions are shown within parentheses) used were (*i*) Trp^{96}–Gln^{102} and the antisense oligonucleotide 5′-TG(T/C)TCCCA(G/A/T/C)ACCAT(T/C)TCCCA and (*ii*)

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 phosphotyrosyl phosphatase.

‡To whom reprint requests should be addressed.

§The sequences reported in this paper have been deposited in the GenBank data base [accession nos. M33689 (PTP1B), M33688 (exon A), M33687 (exon B), M33686 (exon C), M33685 (exon D), and M33684 (exon E)].

Glu^{132}–Thr^{138} and the sense oligonucleotide GA(G/A)ATG-AT(A/C/T)TT(T/C)GA(G/A)GA(T/C)AC. The antisense oligonucleotide designed from oligopeptide 1 was chosen for further use as a sequencing primer to identify the 5′ end of any open reading frame in putative cDNA clones.

Isolation and Characterization of PTP1B Clones. Recombinant bacteriophage (500,000) from a human placental cDNA library were screened for sequences encoding PTPase 1B by transferring bacteriophage DNA to nitrocellulose and challenging for hybridization with the two degenerate oligonucleotides. The final hybridization wash conditions were 2× SSC (1× SSC = 0.15 M NaCl/0.015 M sodium citrate, pH 7)/0.1% SDS at 42°C for 30 min. Radiolabeling of oligonucleotides with T4 polynucleotide kinase, oligonucleotide purification, filter hybridization conditions, DNA preparations, restriction digestions, and Southern analysis were as described (20). Approximately 2×10^6 recombinant bacteriophage from a human genomic library (Clontech) were screened as described (20) for DNA sequences hybridizing to a radiolabeled (21) PTPase 1B cDNA clone. DNA sequence determination was by the chain-termination method (22) using the Sequenase reaction (United States Biochemical).

Chromosome Mapping. A recombinant λ phage genomic clone was used for gene mapping by fluorescence detection of *in situ* hybridization. The basic procedure for localizing single-copy genes with fluorescence detection has been described (23, 24), and a detailed analysis of banding procedures will be described elsewhere (45). The denatured probe, previously labeled by nick-translation with biotin dUTP, was present at a final concentration of 5 μg/ml in the hybridization solution. An excess of nonspecific competitors including total human placental DNA was also included to block hybridization to repetitive sequence elements. Specific hybridization was detected by staining with avidin conjugated to fluorescein.

Metaphase chromosomes were identified by diaminophenylindole banding, enhanced by prior incorporation of 5-bromodeoxyuridine into chromosomal DNA. In addition, Giemsa/trypsin banding prior to hybridization was done to confirm these results. The localization of the gene and interpretation of banding patterns in 10-12 metaphases were performed by two independent observers. The localization was also analyzed by simple measurement of signal position relative to the total length of the chromosome in 27 metaphase figures.

RESULTS

Isolation and Structural Analysis of cDNA Clones Encoding PTPase 1B. One recombinant bacteriophage from a human placental cDNA library was shown by Southern analysis to contain a DNA fragment of 3.2 kilobases that hybridized to both of the degenerate oligonucleotides (data not shown). The fragment was subcloned into pGEM3Z, and the DNA sequence was determined. The cDNA sequence of the PTP1B gene, the deduced translation product, and sequences corresponding to the degenerate oligonucleotides used in the cDNA cloning are shown in Fig. 1. The observed 1305-nucleotide open reading frame predicts a protein of 435 amino

```
GCGCGACGCGGCCTAGAGCGGCAGACGGCGCAGTGGGCCGAGAAGGAGGCGCAGCAGCCGCCCTGGCCCGTCATGGAGATGGAAAAGGAGTTCGAGCAGATCGACAAGTCCGGGAGCTGG   48
                                                                                    MetGluMetGluLysGluPheGluGlnIleAspLysSerGlySerTrp

GCGGCCATTTACCAGGATATCCGACATGAAGCCAGTGACTTCCCATGTAGAGTGGCCAAGCTTCCTAAGAACAAAAACCGAAATAGGTACAGAGACGTCAGTCCCTTTGACCATAGTCGG   168
AlaAlaIleTyrGlnAspIleArgHisGluAlaSerAspPheProCysArgValAlaLysLeuProLysAsnLysAsnArgAsnArgTyrArgAspValSerProPheAspHisSerArg

ATTAAACTACATCAAGAAGATAATGACTATATCAACGCTAGTTTGATAAAAATGGAAGAAGCCCAAAGGAGTTACATTCTTACCCAGGGCCCTTTGCCTAACACATGCGGTCACTTTTGG   288
IleLysLeuHisGlnGluAspAsnAspTyrIleAsnAlaSerLeuIleLysMetGluGluAlaGlnArgSerTyrIleLeuThrGlnGlyProLeuProAsnThrCysGlyHisPheTrp

GAGATGGTGTGGGAGCAGAAAAGCAGGGGTGTCGTCATGCTCAACAGAGTGATGGAGAAAGGTTCGTTAAAATGCGCACAATACTGGCCACAAAAAGAAGAAAAAGAGATGATCTTTGAA   408
GluMetValTrpGluGlnLysSerArgGlyValValMetLeuAsnArgValMetGluLysGlySerLeuLysCysAlaGlnTyrTrpProGlnLysGluGluLysGluMetIlePheGlu

GACACAAATTTGAAATTAACATTGATCTCTGAAGATATCAAGTCATATTATACAGTGCGACAGCTAGAATTGGAAAACCTTACAACCCAAGAAACTCGAGAGATCTTACATTTCCACTAT   528
AspThrAsnLeuLysLeuThrLeuIleSerGluAspIleLysSerTyrTyrThrValArgGlnLeuGluLeuGluAsnLeuThrThrGlnGluThrArgGluIleLeuHisPheHisTyr

ACCACATGGCCTGACTTTGGAGTCCCTGAATCACCAGCCTCATTCTTGAACTTTCTTTTCAAAGTCCGAGAGTCAGGGTCACTCAGCCCGGAGCACGGGCCCGTTGTGGTGCACTGCAGT   648
ThrThrTrpProAspPheGlyValProGluSerProAlaSerPheLeuAsnPheLeuPheLysValArgGluSerGlySerLeuSerProGluHisGlyProValValValHisCysSer

GCAGGCATCGGCAGGTCTGGAACCTTCTGTCTGGCTGATACCTGCCTCTTGCTGATGGACAAGAGGAAAGACCCTTCTTCCGTTGATATCAAGAAAGTGCTGTTAGAAATGAGGAAGTTT   768
AlaGlyIleGlyArgSerGlyThrPheCysLeuAlaAspThrCysLeuLeuLeuMetAspLysArgLysAspProSerSerValAspIleLysLysValLeuLeuGluMetArgLysPhe

CGGATGGGGCTGATCCAGACAGCCGACCAGCTGCGCTTCTCCTACCTGGCTGTGATCGAAGGTGCCAAATTCATCATGGGGGACTCTTCCGTGCAGGATCAGTGGAAGGAGCTTTCCCAC   888
ArgMetGlyLeuIleGlnThrAlaAspGlnLeuArgPheSerTyrLeuAlaValIleGluGlyAlaLysPheIleMetGlyAspSerSerValGlnAspGlnTrpLysGluLeuSerHis

GAGGACCTGGAGCCCCCACCCGAGCATATCCCCCCACCTCCCCGGCCACCCAAACGAATCCTGGAGCCACACAATGGGAAATGCAGGGAGTTCTTCCCAAATCACCAGTGGGTGAAGGAA   1008
GluAspLeuGluProProProGluHisIleProProProProArgProProLysArgIleLeuGluProHisAsnGlyLysCysArgGluPhePheProAsnHisGlnTrpValLysGlu

GAGACCCAGGAGGATAAAGACTGCCCCATCAAGGAAGAAAAAGGAAGCCCCTTAAATGCCGCACCCTACGGCATCGAAAGCATGAGTCAAGACACTGAAGTTAGAAGTCGGGTCGTGGGG   1128
GluThrGlnGluAspLysAspCysProIleLysGluGluLysGlySerProLeuAsnAlaAlaProTyrGlyIleGluSerMetSerGlnAspThrGluValArgSerArgValValGly

GGAAGTCTTCGAGGTGCCCAGGCTGCCTCCCCAGCCAAAGGGGAGCCGTCACTGCCCGAGAAGGACGAGGACCATGCACTGAGTTACTGGAAGCCCTTCCTGGTCAACATGTGCGTGGCT   1248
GlySerLeuArgGlyAlaGlnAlaAlaSerProAlaLysGlyGluProSerLeuProGluLysAspGluAspHisAlaLeuSerTyrTrpLysProPheLeuValAsnMetCysValAla

ACGGTCCTCACGGCCGGCGCTTACCTCTGCTACAGGTTCCTGTTCAACAGCAACACATAGCCTGACCCTCCTCCACTCCACCTCCACCCACTGTCCGCCTCTGCCCGCAGAGCCCACGCC   1368
ThrValLeuThrAlaGlyAlaTyrLeuCysTyrArgPheLeuPheAsnSerAsnThr X

CGACTAGCAGGCATGCCGCGGTAGGTAAGGGCCGCCGGACCGCGTAGAGAGCCGGGCCCCGGACGGACGTTGGTTCTGCACTAAAACCCATCTTCCCCGGATGTGTGTCTCACCCCTCAT   1488
CCTTTTACTTTTTGCCCCTTCCACTTTGAGTACCAAATCCACAAGCCATTTTTTGAGGAGAGTGAAAGAGAGTACCATGCTGGCGGCGCAGAGGGAAGGGGCCTACACCCGTCTTGGGGC   1608
TCGCCCCACCCAGGGCTCCCTCCTGGAGCATCCCAGGCGGGCGGCACGCCAGACAGCCCCCCCCTTGAATCTGCAGGGAGCAACTCTCCACTCCATATTTATTTAAACAATTTTTTCCCC   1728
AAAGGCATCCATAGTGCACTAGCATTTTCTTGAACCAATAATGTATTAAAATTTTTTGATGTCAGCCTTGCATCAAGGGCTTTATCAAAAAGTACAATAATAAATCCTCAGGTAGTACTG   1848
GGAATGGAAGGCTTTGCCATGGGCCTGCTGCGTCAGACCAGTACTGGGAAGGAGGACGGTTGTAAGCAGTTGTTATTTAGTGATATTGTGGGTAACGTGAGAAGATAGAACAATGCTATA   1968
ATATATAATGAACACGTGGGTATTTAATAAGAAACATGATGTGAGATTACTTTGTCCCGCTTATTCTGCTCCCTGTTATCTGCTAGATCTAGTTCTCAATCACTGCTCCCCCGTGTGTAT   2088
TAGAATGCATGTAAGGTCTTCTTGTGTCCTGATGAAAAATATGTGCTTGAAATGAGAAACTTTGATCTCTGCTTACTAATGTGCCCCATGTCCAAGTCCAACCTGCCTGTGCATGACCTG   2208
ATCATTACATGGCTGTGGTTCCTAAGCCTGTTGCTGAAGTCATTGTCGCTCAGCAATAGGGTGCAGTTTTCCAGGAATAGGCATTTGCCTAATTCCTGGCATGACACTCTAGTGACTTCC   2328
TGGTGAGGCCCAGCCTGTCCTGGTACAGCAGGGTCTTGCTGTAACTCAGACATTCCAAGGGTATGGGAAGCCATATTCACACCTCACGCTCTGGACATGATTTAGGGAAGCAGGGACACC   2448
CCCCGCCCCCCACCTTTGGGATCAGCCTCCGCCATTCCAAGTCGACACTCTTCTTGAGCAGACCGTGATTTGGAAGAGAGGCACCTGCTGGAAACCACACTTCTTGAAACAGCCTGGGTG   2568
ACGGTCCTTTAGGCAGCCTGCCGCCGTCTCTGTCCCGGTTCACCTTGCCGAGAGAGGCGCGTCTGCCCCACCCTCAAACCCTGTGGGGCCTGATGGTGCTCACGACTCTTCCTGCAAAGG   2688
GAACTGAAGACCTCCACATTAAGTGGCTTTTTAACATGAAAAACACGGCAGCTGTAGCTCCCGAGCTACTCTCTTGCCAGCATTTTCACATTTTGCCCTTTCTCGTGGTAGAAGCCAGTAC   2808
AGAGAAATTCTGTGGTGGGAACATTCGAGGTGTCACCCTGCAGAGCTATGGTGAGGTGTGGATAAGGCTTAGGTGCCAGGCTGTAAGCATTCTGAGCTGGCTTGTTGTTTTTAAGTCCTG   2928
TATATGTATGTAGTAGTTTGGGTGTGTATATATAGTAGCATTTCAAAATGGACGTACTGGTTTAACCTCCTATCCTTGGAGAGCAGCTGGCTCTCCACCTTGTTACACATTATGTTAGAG   3048
AGGTAGCGAGCTGCTCTGCTATGTCCTTAAGCCAATATTTACTCATCAGGTCATTATTTTTTACAATGGCCATGGAATAAACCATTTTTACAAAA                              3143
```

FIG. 1. Nucleotide and predicted amino acid sequence of PTP1B cDNA. The cDNA sequence of the PTP1B gene is shown along with the deduced 435-amino acid translation product (the termination codon is marked by an X). Nucleotide numbering is on the right; the sequence is numbered from the first ATG codon observed in the single open reading frame. The underlined amino acid sequences correspond to the oligopeptides chosen for construction of the degenerate oligonucleotides used to screen the cDNA library. The Asn^{321} residue marking the carboxyl terminus of the purified 37,000-Da PTPase 1B is underlined. A putative poly(A) addition site is observed at nucleotides 3123–3128.

acids with a molecular mass 49,966 Da. The amino-terminal 321 amino acids encoded by the cDNA are identical to the empirically determined amino acid sequence of the 37,000-Da PTPase 1B (12).

Significant similarity between the amino-terminal 276 amino acids of PTPase 1B and the translation product of the T cell-specific PTPase cDNA has been noted (12, 13). Amino acid sequence identity in the remainder of the amino acid sequences deduced from the cDNAs is predominantly localized to the carboxyl termini in which there are 8 out of 19 identical amino acid residues. In addition, the carboxyl-terminal regions exhibit overall structural similarity based on Dayhoff alignments and in terms of localized regions of hydrophilicity or hydrophobicity (25). Fig. 2 shows a schematic of the overall similarity of PTPase 1B to the T-cell PTPase and an alignment of the deduced carboxyl termini of the two proteins.

The purified PTPase 1B is 37,000 Da (12), whereas the cDNA sequence predicts a protein of 49,966 Da. To determine whether an mRNA produced from the cDNA would be translated to give a 49,966-Da protein, the cDNA clone was used as a template for *in vitro* transcription and translation. By using transcripts from the 3215-nucleotide cDNA, the major translation product was a protein with an apparent molecular mass of 50,000 Da as measured by SDS/PAGE, whereas no 37,000-Da product was observed (data not shown).

Genomic Structure of PTP1B. To obtain suitable DNA fragments for chromosome mapping, a genomic clone was isolated by screening a human genomic library with the PTP1B cDNA. Two genomic fragments of ≈13 kilobases each were isolated. Restriction analysis indicates that the two clones exhibit nearly 90% overlap. In addition to using the genomic DNA for chromosome mapping (see below), we undertook sequence and restriction analysis to further characterize the PTP1B gene. DNA sequencing shows that the genomic clones encompass the 3′ 2700 nucleotides of the cDNA. Neither clone contains the 5′ portion of the PTP1B gene, however. The carboxyl-terminal 272 amino acids of PTPase 1B plus the stop codon are encoded by five exons. A single exon encodes amino acids Asp^{289}–Glu^{362}, which includes Asn^{321}, the carboxyl-terminal amino acid of the 37,000-Da form of PTPase 1B. Southern analysis of human genomic DNA identified single *Eco*RI, *Hin*dIII, and *Bam*HI restriction fragments when the 1305-nucleotide open reading frame portion of the cDNA was used as a probe (data not shown). The exon/intron structure of PTP1B is shown in Fig. 3.

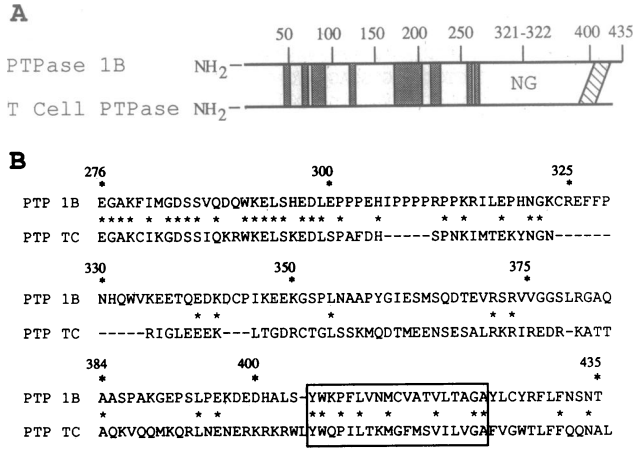

FIG. 2. Amino acid alignments of PTPase 1B and T-cell PTPase. (*A*) The primary translation product of each cDNA is represented by the individual lines. The crosshatched boxes connecting the two lines indicate oligopeptides of seven or more amino acids that are identical. The numbers above the lines refer to PTPase 1B amino acid numbering. The PTPase 1B Asn^{321}-Gly^{322} peptide pair, also present in the T-cell PTPase, is indicated by NG. The hatched box represents the hydrophobic region at the carboxyl terminus of the two proteins. (*B*) Alignment of the deduced carboxyl termini of PTPase 1B (labeled PTP 1B) and the T-cell PTPase (labeled PTP TC). Amino acids are indicated by the one-letter code, and the amino acid identities between the two regions are indicated by stars. The hydrophobic region containing 8 of 19 identical amino acids is boxed.

Chromosomal Mapping of the PTP1B Gene. The genomic clone, labeled with biotinylated dUTP, was used to probe human metaphase chromosomes. Localization to a single chromosomal locus was determined by nonisotopic, fluorescence detection of *in situ* hybridization. Greater than 90% of metaphase figures showed hybridization, which was consistently on a chromosomal pair having the small metacentric morphology of F group chromosomes. Chromosome banding analysis was performed primarily by diaminophenylindole staining of 5-bromodeoxyuridine-incorporated chromosomes and was confirmed with Giemsa/trypsin banding. As illustrated in Fig. 4 *A* and *B*, hybridization could be unambiguously visualized in one cell by the identical labeling of both sister chromatids on each of the two homologous chromosomes. Banding analysis demonstrated this to be chromosome 20 and allowed regional localization of the signal to q13.1–q13.2. Fig. 4*C* illustrates the position of the chromosome signal on several enlarged chromosomes; Fig. 4*D* summarizes the data from 10 banded chromosomes, scored independently by two observers with essentially identical results. Placement of the signal was also analyzed by measurement along the length of the chromosome; based on the average of 27 determinations, the PTP1B gene would map in the middle of q13.1. However, because banding analysis is more precise (45), we conclude that PTP1B is localized to 20q13.1–13.2.

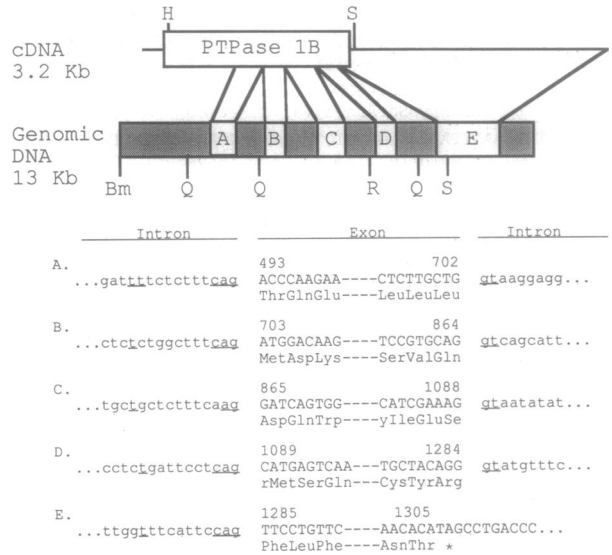

FIG. 3. PTP1B cDNA and genomic DNA structures. (*Upper*) Schematic representations of the cDNA clone, with the open box representing the 1305-nucleotide open reading frame, and the 13-kilobase genomic fragment. The genomic figure is not drawn to scale. Exons, identified by determining exon/intron junctions through comparison of cDNA and genomic DNA sequences, are shown as open boxes labeled A–E and are mapped onto the cDNA. Various restriction endonuclease sites are shown for both clones (Bm, *Bam*HI; R, *Eco*RI; H, *Hin*dIII; Q, *Sac* I; S, *Sph* I). (*Lower*) The DNA sequences of the intron/exon and exon/intron junctions for the five exons (A–E) identified are shown. Intron sequences are in lowercase letters; exon sequences are in uppercase letters. Numbers above the exon sequences correspond to the cDNA numbering in Fig. 1. Consensus nucleotides for intron sequences (26) are underlined. The reading frame and translation products are shown beneath the exon sequences. Exon E has only an intron/exon junction because it contains the stop codon and the 3′ untranslated region.

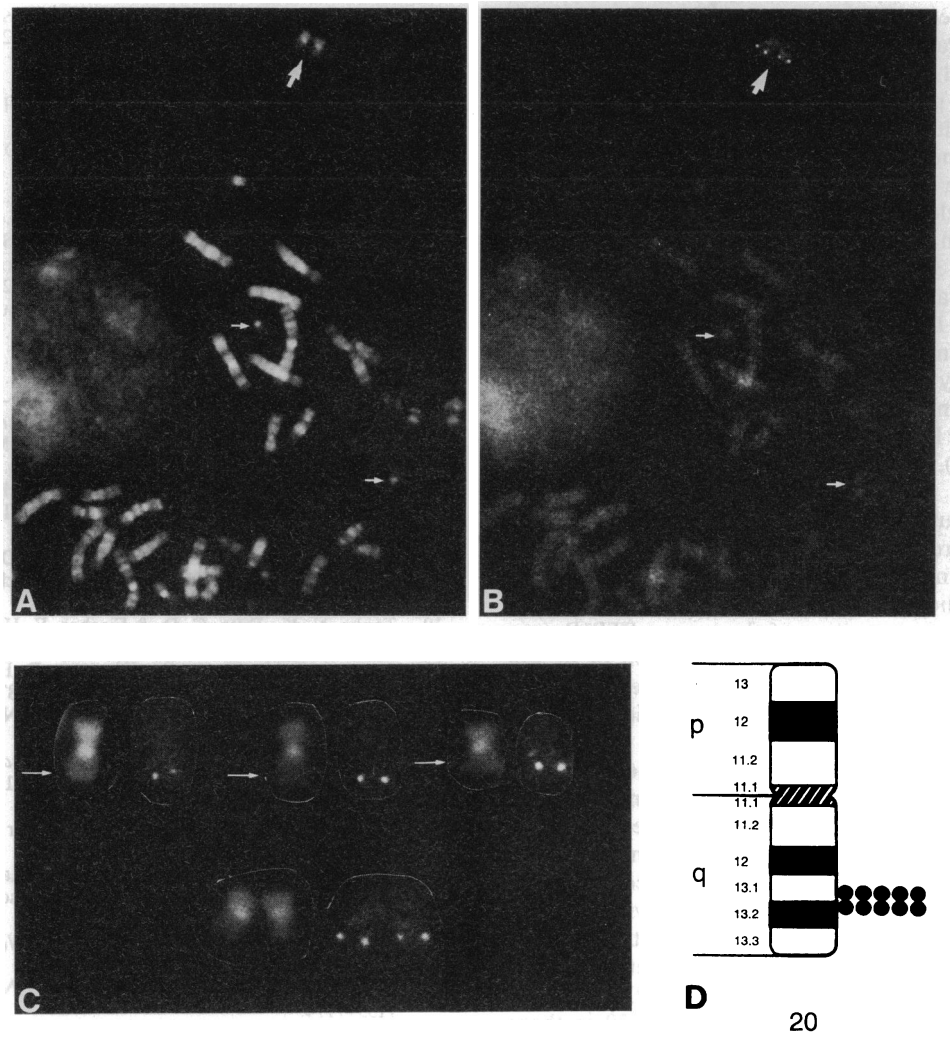

FIG. 4. Localization of the PTP1B gene by fluorescence detection of *in situ* hybridization. (*A*) Diaminophenylindole staining of metaphase chromosomes from normal human lymphocytes. Bands are sufficient to allow identification of individual chromosomes. The large arrow at the top of the panel points to a pair of no. 20 chromosomes. The small arrows indicate chromosome 19, which is distinguished by a more prominent centromere and paler arms. (*B*) Fluorescein/avidin detection of hybridization with the biotin-labeled PTP1B genomic DNA probe. Both homologs of chromosome 20 label on both sister chromatids, and no other chromosomal sites are labeled. (*C*) Higher magnification of several no. 20 chromosomes showing alignment of the total chromosome diaminophenylindole fluorescence on the left of each pair with the fluorescein avidin hybridization signals on the right. (*D*) Schematic representation of chromosome 20 with a summary of the PTP1B gene placement based on analysis of 10 banded metaphase chromosomes.

DISCUSSION

Malignant transformation is believed to result from the cumulative effects of multiple genetic lesions. Two classes of genetic lesions have been clearly defined. The best understood class is the dominant, activating mutations that generate oncogenes (27). The second class is the loss of function mutations that lead to neoplastic growth; these mutations identify growth suppressor genes (28, 29), also termed tumor suppressor genes or antioncogenes. One mechanism for generating loss of function mutations is deletion of a putative growth suppressor gene, which may be accompanied by observable chromosomal abnormalities. The putative growth suppressor genes *RB* (retinoblastoma) and *DCC* (deleted in colorectal carcinomas) have been isolated based on this criterion (30–33). To determine if PTPases function as growth suppressors, we have isolated PTP1B cDNA and genomic clones for functional studies and chromosome mapping studies, respectively.

Analysis of the cDNA and genomic DNA indicates that the primary translation product of the human PTP1B gene is a 435-amino acid protein with a predicted molecular mass of 49,966 Da. This result is in contrast to the 321-amino acid PTPase 1B protein previously purified and sequenced (12). The nature of the discrepancy in protein sizes is not clear. The 37,000-Da form could result from a posttranslational processing of the primary translation product, be encoded by an alternatively spliced RNA, or be a consequence of purification. Posttranslational processing would require proteolytic cleavage specifically at the Asn^{321}—Gly^{322} peptide bond; at present such proteolytic specificity is unknown. It is unlikely that the truncated protein comes from an alternatively spliced messenger RNA since the codon for Asn^{321} is found in the middle of an exon without any surrounding splice consensus sequences. Asn—Gly peptide bonds are susceptible to alkali-catalyzed cleavage generating peptides with a carboxyl-terminal Asn (34), which suggests that the 321-amino acid form may be generated during purification.

Recently an identical cDNA encoding a 435-amino acid form of PTPase 1B was reported by Chernoff *et al.* (35), and the sequence of a rat brain cDNA encoding a PTPase has also been described (36). The 432-amino acid rat brain PTPase has >90% amino acid similarity to the 435-amino acid PTPase 1B reported herein, and the carboxyl-terminal 114 amino acids of the two proteins have nearly 60% amino acid similarity. The

evolutionary conservation between human PTPase 1B and the rat brain PTPase plus the overall structural similarity between the T-cell PTPase and PTPase 1B suggest that the carboxyl-terminal portion of these molecules is critical for their function. Although it is unclear whether both the 49,966-Da form and the 37,000-Da form of PTPase 1B are present *in vivo*, we have observed that PTPase 1B expressed from the cDNA clones in NIH 3T3 cells appears to be a 50,000-Da protein (S.B.-S., unpublished observations). While the purified 37,000-Da species is enzymatically active *in vitro* (9), the carboxyl-terminal 114 amino acids of the 49,966-Da PTPase 1B may perform a necessary role, perhaps in regulating enzymatic activity, controlling access to substrates, or possibly targeting the protein to the cytoplasmic surface of the membrane.

The cytoplasmic PTPases have been proposed to be growth suppressors (6, 7). One criterion for a candidate growth suppressor would be to demonstrate that the gene was inactivated by mutation or deletion in specific cancers. Identifying the chromosomal locus of a growth suppressor gene and correlating that locus with specific chromosomal abnormalities implicated in cancer is an essential first step. The chromosomal map position of the human PTP1B gene has been determined by fluorescence detection of *in situ* hybridization to banded metaphase chromosomes. PTP1B is a single-copy gene located on chromosome 20 in the region q13.1–q13.2. This localizes PTP1B to a region on chromosome 20 that also contains several other genes encoding factors implicated in signal transduction, including a *src* gene (37, 38), the hemopoietic cell kinase (*hck*) gene (39), and a phospholipase C (*plc1*) gene (40). While chromosome 20q deletions spanning the q11–q13.3 region have been observed in bone marrow cells from some patients with myeloid disorders (41, 42), precise determination of the putative tumor suppressor gene identified by these deletions awaits further analysis. Clearly, genes other than PTP1B could be altered or otherwise mutated on chromosome 20 in these patients. However, the loss of a PTPase activity could contribute to a neoplastic state through aberrantly increased levels of phosphotyrosine. Phenotypically, this would appear similar to neoplastic transformation induced by an oncogenic tyrosine kinase such as v-*fms* (43, 44) but would instead result from the loss of a growth-suppressing function. Additional analysis of PTPase expression in normal and malignant cells will be necessary to determine what role, if any, PTPases may play in various cancers. With the availability of cDNA clones and knowledge of the chromosome locus, it will be possible to directly test whether PTPase 1B can function as a growth suppressor and to determine if the PTP1B gene is an "anti-oncogene."

We thank Drs. Jonathon Chernoff and Ray Erikson for communicating their results, described in ref. 35, prior to publication. We also thank Drs. Rene Bernards, Deborah Cool, Edmond Fischer, Steve Friend, Edwin Krebs, and Nicholas Tonks for helpful discussions and Drs. Bryan Roberts, Robert Weinberg, and Sam Yin for critical reading of the manuscript. A portion of this work was supported by the National Institutes of Health Grant GM41991 (to J.B.L.).

1. Yarden, Y. & Ullrich, A. (1988) *Annu. Rev. Biochem.* **57**, 443–478.
2. Snyder, M. A., Bishop, J. M., McGrath, J. P. & Levinson, A. D. (1985) *Mol. Cell. Biol.* **5**, 1772–1779.
3. Weinmaster, G., Zoller, M. J. & Pawson, T. (1986) *EMBO J.* **5**, 69–76.
4. Sefton, B. M., Hunter, T., Beemon, K. & Eckhart, W. (1980) *Cell* **20**, 807–816.
5. Cooper, J. A., Nakamura, K. D., Hunter, T. & Weber, M. J. (1983) *J. Virol.* **46**, 15–28.
6. Charbonneau, H., Tonks, N. K., Walsh, K. A. & Fischer, E. H. (1988) *Proc. Natl. Acad. Sci. USA* **85**, 7182–7186.
7. Hunter, T. (1989) *Cell* **58**, 1013–1016.
8. Klarlund, J. K. (1985) *Cell* **41**, 707–717.
9. Tonks, N. K., Diltz, C. D. & Fischer, E. H. (1988) *J. Biol. Chem.* **263**, 6722–6730.
10. Lau, K. H. W., Farley, J. R. & Baylink, D. J. (1989) *Biochem. J.* **257**, 23–36.
11. Jones, S. W., Erikson, R. L., Ingebritsen, V. M. & Ingebritsen, T. S. (1989) *J. Biol. Chem.* **264**, 7747–7753.
12. 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.
13. 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.
14. Tonks, N. K., Charbonneau, H., Diltz, C. D., Fischer, E. H. & Walsh, K. A. (1988) *Biochemistry* **27**, 8695–8701.
15. Kiener, P. A. & Mittler, R. S. (1989) *J. Immunol.* **143**, 23–28.
16. Streuli, M., Krueger, N. X., Tsai, A. Y. M. & Saito, H. (1989) *Proc. Natl. Acad. Sci. USA* **86**, 8698–8702.
17. Ostergaard, H. L., Shackleford, D. A., Hurley, T. R., Johnson, P., Hyman, R., Sefton, B. M. & Trowbridge, I. S. (1989) *Proc. Natl. Acad. Sci. USA* **86**, 8959–8963.
18. Tonks, N. K., Charbonneau, H., Diltz, C. D., Kumar, S., Cicirelli, M. F., Krebs, E. G., Walsh, K. A. & Fischer, E. H. (1989) *Adv. Protein Phosphatases* **5**, 149–180.
19. Tonks, N. K., Diltz, C. D. & Fischer, E. H. (1988) *J. Biol. Chem.* **263**, 6731–6737.
20. Ausubel, F. M., Brent, R., Kingston, R. E., Moore, D. E., Seidman, J. G., Smith, J. A. & Struhl, K. (1987) *Current Protocols in Molecular Biology* (Greene Publ. Assoc., New York; Wiley-Interscience, New York).
21. Feinberg, A. P. & Vogelstein, B. (1983) *Anal. Biochem.* **132**, 6–13.
22. Sanger, F., Nicklen, S. & Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. USA* **74**, 5463–5467.
23. Lawrence, J. B., Villnave, C. A. & Singer, R. H. (1988) *Cell* **52**, 51–61.
24. Staunton, D. E., Fisher, R., Le Beau, M., Lawrence, J. B., Burton, D., Franke, U., Dustin, M. & Thorley-Lawson, D. A. (1989) *J. Exp. Med.* **169**, 1087–1099.
25. Hopp, T. P. & Woods, K. R. (1981) *Proc. Natl. Acad. Sci. USA* **78**, 3824–3828.
26. Shapiro, M. B. & Senapathy, P. (1987) *Nucleic Acids Res.* **15**, 7155–7174.
27. Bishop, J. M. (1987) *Science* **235**, 305–311.
28. Hansen, M. F. & Cavanee, W. K. (1988) *Cell* **53**, 172–173.
29. Weinberg, R. A. (1989) *Biochemistry* **28**, 8263–8269.
30. Friend, S. H., Bernards, R., Rogelj, S., Weinberg, R. A., Rapaport, J. M., Albert, D. M. & Dryja, T. P. (1986) *Nature (London)* **328**, 643–646.
31. Fung, Y.-K. T., Murphree, A. L., T'Ang, L.-J., Qian, J., Hinrichs, S. H. & Benedict, W. F. (1987) *Science* **236**, 1657–1661.
32. Lee, W.-H., Bookstein, R., Hong, F., Young, L.-J., Shew, J.-Y. & Lee, E. Y.-H. P. (1987) *Science* **235**, 1394–1399.
33. Fearon, E. R., Cho, K. R., Nigro, J. M., Kern, S. E., Simons, J. W., Ruppert, J. M., Hamilton, S. R., Preisinger, A. C., Thomas, G., Kinzler, K. & Vogelstein, B. (1990) *Science* **247**, 49–56.
34. Bornstein, P. & Balian, G. (1977) *Methods Enzymol.* **47**, 132–145.
35. Chernoff, J., Schievella, A., Jost, C. A. & Erikson, R. L. (1990) *Proc. Natl. Acad. Sci. USA* **87**, 2735–2739.
36. Guan, K., Haun, R. S., Watson, S. J., Geahlen, R. L. & Dixon, J. E. (1990) *Proc. Natl. Acad. Sci. USA* **87**, 1501–1505.
37. Le Beau, M. M., Westbrook, C. A., Diaz, M. O. & Rowley, J. D. (1984) *Nature (London)* **312**, 70–71.
38. Parker, R. C., Mardon, G., Lebo, R. V., Varmus, H. E. & Bishop, J. M. (1985) *Mol. Cell. Biol.* **5**, 831–838.
39. Quintrell, N., Lebo, R., Varmus, H. E., Bishop, J. M., Pettenati, M. J., Le Beau, M. M., Diaz, M. O. & Rowley, J. D. (1987) *Mol. Cell. Biol.* **7**, 2267–2275.
40. Bristol, A., Hall, S. M., Kriz, R. M., Stahl, M. L., Fan, Y. S., Byers, M. G., Eddy, R. L., Shows, T. B. & Knopf, J. L. (1988) *Cold Spring Harbor Symp. Quant. Biol.* **53**, 915–920.
41. Davis, M. P., Dewald, G. W., Pierre, R. V. & Hoagland, H. C. (1984) *Cancer Genet. Cytogenet.* **12**, 63–71.
42. Mitelman, F. (1984) *Cytogenet. Cell Genet.* **36**, 1–516.
43. Wheeler, E. F., Rettenmier, C. W., Look, A. T. & Sherr, C. J. (1986) *Nature (London)* **324**, 377–380.
44. Wheeler, E. F., Askew, D., May, S., Ihle, J. N. & Sherr, C. J. (1987) *Mol. Cell. Biol.* **7**, 1673–1680.
45. Lawrence, J. B., Singer, R. H. & McNeil, J. A. (1990) *Science*, in press.