Proc. Natl. Acad. Sci. USA
Vol. 90, pp. 2197–2201, March 1993
Biochemistry

A widely expressed human protein-tyrosine phosphatase containing *src* homology 2 domains

(dephosphorylation/cloning/expression/purification)

SULTAN AHMAD*, DENIS BANVILLE*, ZHIZHUANG ZHAO†, EDMOND H. FISCHER†, AND SHI-HSIANG SHEN*‡

*Section of Molecular Biology, Biotechnology Research Institute, National Research Council of Canada, 6100 Royalmount Avenue, Montreal, PQ, Canada H3A 2R2; and †Department of Biochemistry, University of Washington, Seattle, WA 98195

Contributed by Edmond H. Fischer, December 9, 1992

ABSTRACT **A cDNA encoding a nontransmembrane protein-tyrosine phosphatase (PTP; EC 3.1.3.48), termed PTP2C, was isolated from a human umbilical cord cDNA library. The enzyme contains a single phosphatase domain and two adjacent copies of the *src* homology 2 (SH2) domain at its amino terminus. A variant of PTP2C (PTP2Ci) which has four extra amino acid residues within the catalytic domain has been identified also. PTP2C is widely expressed in human tissues and is particularly abundant in heart, brain, and skeletal muscle. The catalytic domain of PTP2C was expressed as a recombinant enzyme in *Escherichia coli* and purified to near homogeneity by two chromatographic steps. The recombinant enzyme was totally specific toward phosphotyrosine residues. The structural similarity between PTP2C and the previously described PTP1C suggests the existence of a subfamily of SH2-containing PTPs; these may play an important role in signal transduction through interaction of their SH2 domains with phosphotyrosine-containing proteins.**

Phosphorylation of several intracellular proteins on tyrosine residues has been implicated in a number of critical physiological and pathological processes, including cell growth, differentiation, and neoplastic transformation (1–5). The state of such phosphorylation depends upon the relative activities of two families of enzymes with opposing actions, namely, protein-tyrosine kinases (PTKs; EC 2.7.1.112) and protein-tyrosine phosphatases (PTPs; EC 3.1.3.48). As with PTKs, there are two classes of PTPs: transmembrane and nontransmembrane enzymes (for review, see refs. 6 and 7). All the nontransmembrane PTPs reported to date have a single highly conserved catalytic domain of ca. 230 amino acid residues, while the transmembrane PTPs have two, with the exception of HPTP-β (8) and DPTP10D (9), which have only one. Segments outside of the catalytic domain display considerable structural diversity.

Recently, we and others (10–13) have identified a nontransmembrane PTP (PTP1C) which contains two copies of the *src* homology 2 (SH2) domain at its N terminus. Likewise, a *Drosophila* gene, designated corkscrew, also encodes a PTP with two SH2 domains (14); it differs from PTP1C in having an insert interrupting the catalytic segment. SH2 domains were originally found in various nonreceptor PTKs and other cytoplasmic signaling proteins (for review, see refs. 15 and 16). They bind strongly to peptide segments containing tyrosine phosphate, but the selectivity of the interaction is determined by the nature of the residues surrounding the phosphotyrosine residue, particularly those at positions +1 and +3. These domains play a major role in signal transduction by allowing the interaction of intracellular tyrosine kinases or other SH2-containing proteins with growth factor receptors, once these have undergone autophosphorylation. The specificity of the interaction will determine which particular signal pathway may become activated. In the present report, we describe another soluble phosphatase, designated as PTP2C; it also contains two SH2 domains and is widely expressed in human tissues.§ We further identified a variant of this enzyme having an in-frame insertion of 12 base pairs within the catalytic domain.

MATERIALS AND METHODS

Isolation of PTP2C Clones. For the polymerase chain reaction (PCR), a set of degenerate primers was designed from two highly conserved regions within the catalytic domain of PTPs. One primer corresponded to the amino acid sequence KC(A/D)QYWP and the other, to VHCSAGV. PCR fragments of about 250 bp were amplified from a human umbilical cord cDNA library. PCRs were performed for 30 cycles with a 94°C denaturation and a 50°C annealing temperature under standard conditions. The products were purified by electrophoresis on a 1.6% agarose gel and cloned in the pBluescript II KS (+) vector (Stratagene). One of the cloned PCR fragments was used as a probe to screen the cDNA library at 50°C in 4× SSC/2× Denhardt's solution/0.5% SDS/0.1 M sodium phosphate buffer, pH 7.0, containing sodium dextran sulfate at 100 mg/ml, and sonicated and denatured salmon sperm DNA at 100 μg/ml (1× SSC = 0.15 M NaCl/0.015 M sodium citrate; 1× Denhardt's solution = 0.02% bovine serum albumin/0.02% Ficoll/0.02% polyvinylpyrrolidone). Positive clones were subcloned in the pBluescript KS vector and sequenced with oligonucleotides or internal restriction fragment primers (17), using a T7 DNA polymerase sequencing kit (Pharmacia).

Northern Blot Analysis. A human multiple-tissue Northern blot system (Clontech) was used to determine the level of expression of PTP2C in various tissues. The blot was prehybridized for 3 hr and then hybridized overnight at 50°C in 50% (vol/vol) formamide/2.5× Denhardt's solution/25 mM potassium phosphate buffer, pH 7.4/0.1% SDS/10% sodium dextran sulfate/5× SSPE containing sonicated and denatured salmon sperm DNA at 250 μg/ml and yeast tRNA at 500 μg/ml (1× SSPE = 0.15 M NaCl/0.01 M sodium phosphate, pH 7.4/1 mM EDTA). The probes (PTP2C and β-actin) were labeled with the T7 Quick Prime kit (Pharmacia). The blot was washed twice with 2× SSPE/0.1% SDS, then with 1× SSPE/0.1% SDS, and finally with 0.2× SSPE/0.5% SDS at room temperature for 15 min each before exposure to Kodak X-Omat AR film at −80°C with an intensifying screen.

Expression, Purification, and Assay of Recombinant Enzyme. The PTP2C cDNA clone was cleaved with *Bgl* II (at

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: PTK, protein-tyrosine kinase; PTP, protein-tyrosine phosphatase; SH2, *src* homology 2; MAP kinase, mitogen-activated protein kinase.

‡To whom reprint requests should be addressed.

§The sequence reported in this paper has been deposited in the GenBank data base (accession no. L08807).

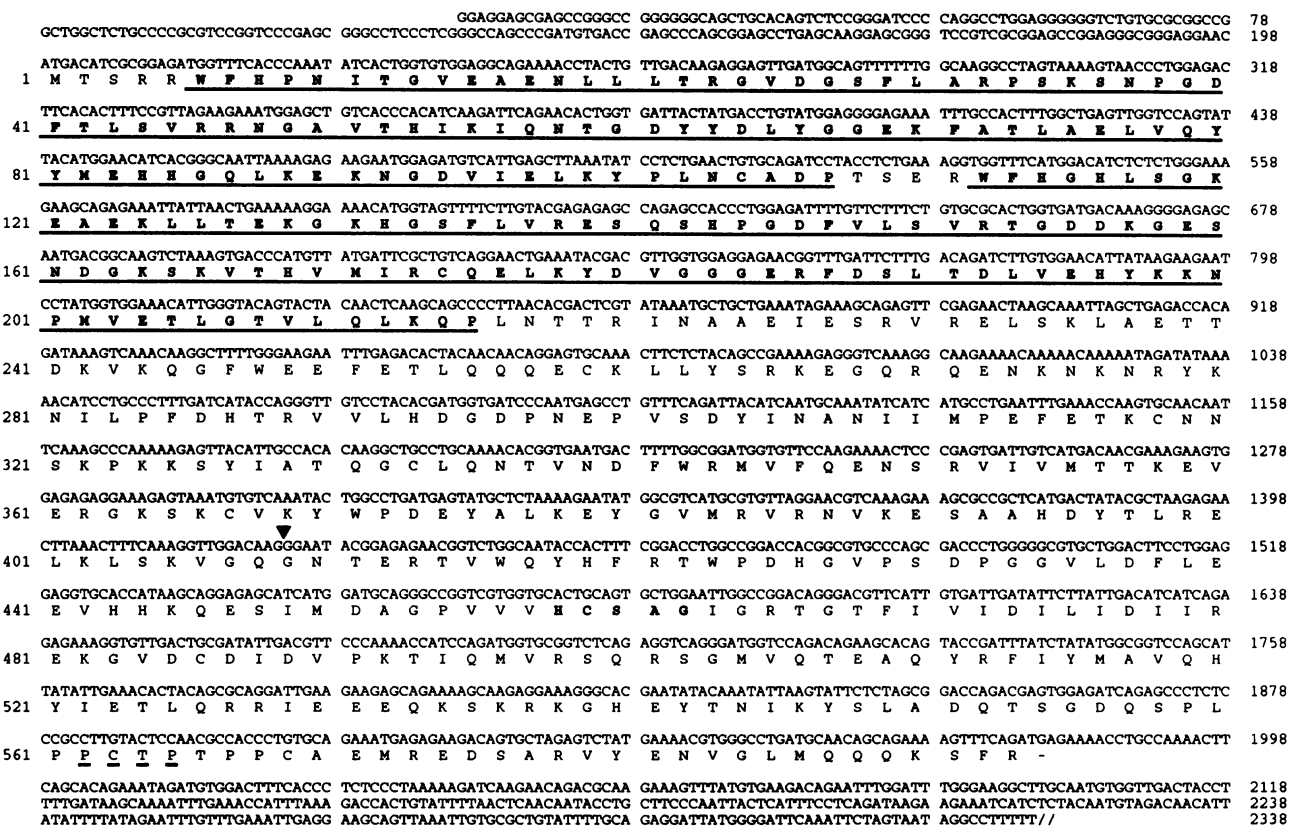

FIG. 1. Nucleotide sequence of the entire coding region plus part of the noncoding regions of the PTP2C cDNA clone and the amino acid sequence deduced. Amino acid residues are numbered on the left, nucleotide positions on the right. The two adjacent copies of the SH2 domain are underlined and boldfaced. The active site of the PTP is boldfaced (HCSAG, residues 458–462). The position of the 12-bp insert in PTP2Ci is indicated by an inverted triangle. A putative phosphorylation site for mitogen-activated protein (MAP) kinase is indicated by a broken underline.

nucleotide 771) and the ends were filled with the Klenow fragment of DNA polymerase. The DNA was redigested with *Hin*dIII. The expression vector pET-3c (18) was cleaved with *Eco*RI, filled in, and redigested with *Hin*dIII. The truncated PTP2C fragment was inserted in frame into the treated vector pET-3c. The resultant plasmid, designated pΔSH2-PTP2C, was used to transform *Escherichia coli* BL21 (DE3) cells. Expression of the recombinant enzyme (ΔSH2-PTP2C) was carried out as described (10). The recombinant protein consists of 406 amino acids in which the first four residues (MARI) were derived from the vector. For purification, *E. coli* cells containing the recombinant enzyme were sonicated in a buffer consisting of 25 mM Tris·HCl at pH 7.5, 10 mM 2-mercaptoethanol, 2.0 mM EDTA, 1.0 mM benzamidine, 0.1 mM phenylmethylsulfonyl fluoride, leupeptin at 20 μg/ml, pepstatin A at 1.0 μM, and aprotinin at 0.027 trypsin inhibitor unit/ml. After centrifugation at 100,000 × *g* for 30 min, the supernatant was loaded onto a fast-flow Q-Sepharose column (Pharmacia) equilibrated with 25 mM Tris·HCl, pH 7.5/2.0 mM 2-mercaptoethanol/1.0 mM EDTA. The flow-through was collected, adjusted to pH 6.0, and then applied to a Mono-S column (Pharmacia) equilibrated with 20 mM 2-(*N*-morpholino)ethanesulfonic acid (Mes)–NaOH, pH 6.0/1.0 mM EDTA/2.0 mM 2-mercaptoethanol. Elution was carried out with a gradient of 0–1.0 M NaCl. Fractions containing phosphatase activity were collected. Control cells containing the vector without PTP2C sequence were treated in a similar way.

Activity assays were performed at pH 5.0 with low molecular weight and peptide substrates and at pH 7.0 with protein substrates as described (19). One unit of activity is defined as the amount of enzyme causing the release of 1.0 nmol of phosphate per min.

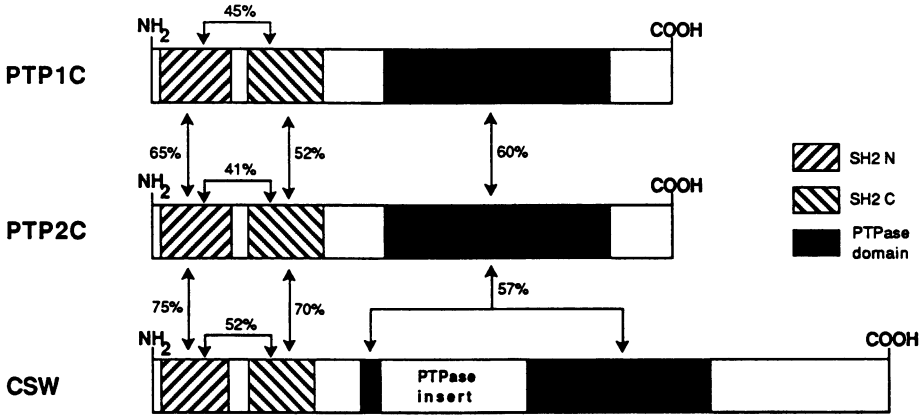

FIG. 2. Comparison of PTP2C with PTP1C and csw-PTP (CSW). The SH2 domains and the catalytic domains are schematically aligned and the percent identity between the three proteins is indicated in the middle. The insert in the catalytic domain of csw-PTP was not taken into account.

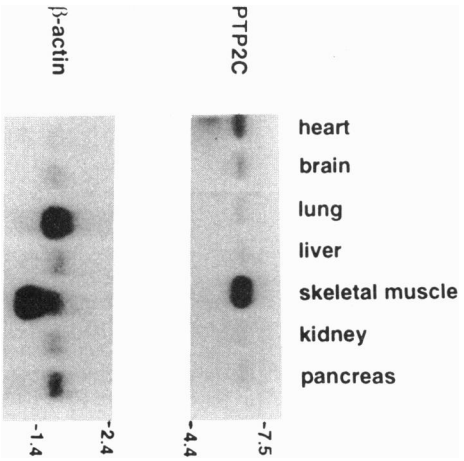

FIG. 3. Expression of PTP2C in various human tissues. Each lane contains 2 μg of poly(A)$^+$ RNA. PTP2C expression was examined by hybridization with the 1.83-kb PTP2C-1 cDNA as a probe (*Upper*). A β-actin probe was used to detect the β-actin mRNA on the same blot (*Lower*). The numbers on the right refer to the sizes of RNA, given in kb.

RESULTS

Identification of the PTP2C cDNA Clone. PCR amplification of the human umbilical cord cDNA library yielded a fragment of about 250 bp as expected. When the same cDNA library was screened with the PCR fragment, one positive clone, PTP2C-1, was obtained. This clone contained a 1.8-kb insert and had an open reading frame with a putative initiation codon starting at nucleotide 198. Sequence analysis revealed that PTP2C-1 contained two tandem SH2 domains plus the entire catalytic domain characteristic of PTPs but lacked the C terminus of the protein. A convenient *Pst* I restriction site was found near the 3′ end. Thus, a 254-bp fragment was excised and used as a probe to screen the aforementioned library. Two additional overlapping clones, PTP2C-2 and PTP2C-3, were obtained, which together with PTP2C-1 covered the entire coding region of the PTP2C cDNA.

DNA Sequence Analysis. The nucleotide sequence of PTP2C (including 198 bp of the 5′ noncoding region and 350 bp of the 3′ noncoding region) and the amino acid sequence deduced from it are presented in Fig. 1. The open reading frame consists of 1779 nucleotides potentially encoding a protein of 593 amino acids with a predicted molecular mass of 68 kDa. The deduced amino acid sequence contains no membrane-spanning segment or signal peptide at the N terminus, indicating that PTP2C, like PTP1C, is a nontransmembrane enzyme.

Interestingly, the PTP2C-2 clone, which spans nucleotides 951–2162, possesses 12 extra base pairs. This results in the insertion of an additional four amino acid residues (ALLQ) within the catalytic domain between Gln-408 and Gly-409 (Fig. 1). The existence of this PTP2C variant (designated as PTP2Ci) at the mRNA level was confirmed by sequencing the product of reverse-transcript PCR, using two oligonucleotides flanking the 12-bp sequence as primers.

Sequence analysis revealed striking similarities among the overall structures of PTP1C, PTP2C, and the *Drosophila* corkscrew PTP (csw-PTP) (Fig. 2). The catalytic domain of PTP2C shares an overall identity of 60% and 57% with the catalytic domains of PTP1C and csw-PTP, respectively. As with two other members of this family, PTP2C has two SH2 domains of approximately 100 residues located near the N terminus, the first starting at Trp-6 and the second at Trp-112. Its SH2 domains are more similar to those of csw-PTP than to those of PTP1C. The whole sequence of PTP2C contains several potential sites of phosphorylation by tyrosine kinases as well as by serine/threonine kinases. A putative phosphorylation site for MAP kinase (20) located at the C terminus of the molecule is indicated in Fig. 1.

Expression of PTP2C Transcripts in Tissues. Northern analysis showed that PTP2C is expressed in all the human tissues investigated (Fig. 3). After normalization of the signal obtained with a β-actin probe, heart, brain, and skeletal muscle appear to express the highest levels. The expression of PTP2C was confirmed also in several human cell lines, including breast carcinoma (ZR-75-1), acute myelogenous leukemia (KG-1), and epidermoid carcinoma A431 cells (data not shown). The size of the transcript from the PTP2C gene is approximately 7.0 kb, with the 3′ noncoding region spanning about 5 kb.

Purification and Activity of Recombinant Enzyme Expressed in *E. coli*. Purification is summarized in Table 1. As indicated in the footnote of Table 1, *E. coli* cells transformed with recombinant vector displayed much higher phosphatase activity than those transformed with control vector, especially when the flow-through fraction of Q-Sepharose column was compared. In addition to increasing the specific activity of the recombinant enzyme by approximately 50-fold, the anion-exchange chromatography step resulted in a ca. 10-fold increase in total activity. This indicated that the phosphatase in the crude extract was inhibited. Further purification of ΔSH2-PTP2C on a Mono-S column produced a single peak and resulted in a nearly homogeneous protein pattern on SDS/polyacrylamide gel electrophoresis (Fig. 4). A molecular mass of 46 kDa calculated from its mobility on SDS gel was consistent with the size predicted from the amino acid sequence.

Table 2 illustrates the phosphatase activity of ΔSH2-PTP2C toward a variety of substrates. As expected, the enzyme was entirely specific for the dephosphorylation of tyrosine residues. While its activity toward two low-molecular-weight substrates, *p*-nitrophenyl phosphate and phosphotyrosine, was comparable to that of PTP1C, its activity toward a tyrosine-phosphorylated nonapeptide [ENDY(P)INASL] and two protein substrates (reduced, carboxamidomethylated, and maleylated lysozyme and myelin basic protein) was more than 10-fold higher (19). Nevertheless, the values with the two protein substrates were considerably lower than those reported for other soluble PTPs, such as PTP1B and human T-cell PTP, which reached values around 10,000 units/mg (21, 22). Three tyrosine-phosphorylated substrates of potentially physiological relevance—namely, the intracellular domains of epidermal growth factor receptor and insulin receptor and lipocortin I—were also

Table 1. Summary of ΔSH2-PTP2C purification

Step	Volume, ml	Protein, mg/ml	Activity,* units/ml	Specific activity, units/mg	Purification, fold	Yield, %
Crude	10	1.8	380	210	1	—
Q-Sepharose	30	0.10	1,200	12,000	57	100
Mono-S	2	0.17	10,900	64,000	300	61

*Activity was assayed with 10 mM *p*-nitrophenyl phosphate; the specific activities of crude extract and Q-Sepharose break-through fraction of *E. coli* cells transformed with control vector were 20 and 110 units/mg, respectively.

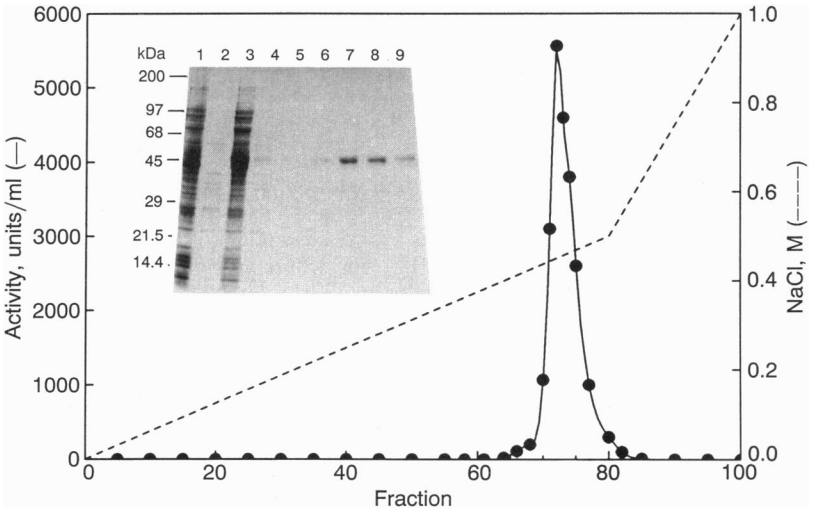

FIG. 4. Purification of ΔSH2-PTP2C on a Mono-S column. The solid line indicates PTP activity toward 10 mM *p*-nitrophenyl phosphate and the broken line indicates the NaCl salt gradient. (*Inset*) SDS/9–19% polyacrylamide gel electrophoresis of samples from *E. coli* cells transformed with control vector (lanes 1 and 2) and recombinant vector containing ΔSH2-PTP gene (lanes 3–9). Lanes 1 and 3, crude extract; lanes 2 and 4, fast-flow Q-Sepharose flow-through fraction; lanes 5–9, Mono-S fractions 68, 70, 72, 74, and 76, respectively.

susceptible. In contrast, those peptides or proteins phosphorylated on serine or threonine were totally resistant.

DISCUSSION

PTP2C belongs to the subfamily of PTPs having two SH2 domains. Unlike PTP1C, which is restricted to hematopoietic and epithelial cells (11–13), PTP2C is rather widely distributed. The ubiquitous nature of this enzyme may be indicative of a fundamental role in signal transduction. The identity between the SH2 domains of PTP1C and 2C is 50–60%, higher than the identity between the two SH2 domains within the same molecule (41% and 45%, respectively). Nevertheless, the differences in structure between the two enzymes suggest that they might be directed toward different targets. Therefore, they could be involved in different signaling pathways.

Table 2. Phosphatase activity of ΔSH2-PTP2C

Substrate	Protein kinase producing substrate		Activity,* units/mg
	Tyr	Ser/Thr	
p-NPP, 10 mM			64,000
Phosphotyrosine (2 mM)			15,000
ENDY(P)INASL† (5 μM)	sEGFR		7,800
RCM-lysozyme (1 μM)	BIRK		460
MBP (1 μM)	BIRK		100
sEGFR	sEGFR		+
BIRK	BIRK		+
Lipocortin I	sEGFR		+
LRRAS(P)LG (2 μM)		PKA	0
RTKRSGS(P)VYEPLKI (1 μM)		PKC	0
RRRDDDS(P)DDD (2 μM)		CKII	0
RRREEET(P)EEE (1 μM)		CKII	0
Lipocortin I		PKC	−
Phosphorylase		Phosph kin	−
Casein		CKII	−

Abbreviations: *p*-NPP, *p*-nitrophenyl phosphate; RCM-lysozyme, reduced, carboxamidomethylated, and maleylated lysozyme; MBP, myelin basic protein; sEGFR, intracellular domain of epidermal growth factor receptor; BIRK, intracellular domain of insulin receptor; PKA, protein kinase A; PKC, protein kinase C; CKII, casein kinase II; Phosph kin, phosphorylase kinase.

*When numerical values are not given, dephosphorylation was detected by SDS/polyacrylamide gel electrophoresis. + and − represent the relative ability or inability, respectively, of ΔSH2-PTP2C to dephosphorylate these substrates.

†G. Daum, F. Solca, C. D. Diltz, Z.Z., D. E. Cool, and E.H.F., unpublished results.

The role of the SH2 domains in this family of PTPs is still unknown. First, they could serve to maintain the phosphatases in an inactive state, if the enzyme were phosphorylated on tyrosine—interaction of the phosphotyrosine residue with one of the SH2 domains could shield the catalytic site. Such an autoinhibitory mechanism was postulated for the repression of c-src activity (15, 23). However, ΔSH2-PTP2C, which lacks the SH2 domains, is still not fully active. While the activity of PTP1C could be increased 10- to 20-fold by limited proteolysis of its C-terminal segment, there is no evidence that this activation resulted from the removal of a phosphotyrosine residue (19). Second, the SH2 domains could allow the enzyme to interact with another protein serving the role of a regulatory or modulatory subunit. As of now, no such binding protein has been identified. Finally, the SH2 domains could help to target the enzyme toward specific substrates. It is interesting to note that when the SH2 domain of phospholipase C_γ was allowed to interact with the site of autophosphorylation of the epidermal growth factor receptor, the resulting complex was resistant to the action of PTPs (24). This suggests that the SH2-containing PTPs would interact preferentially with substrates that are phosphorylated on more than one tyrosine; interaction of their SH2 domains with one of these phosphorylated residues might position the enzyme to attack the other. The situation would be analogous to the interaction of phospholipase C_γ or GTPase-activating protein with growth factor receptors, which allows these proteins to be phosphorylated on adjacent tyrosine residues (15, 16).

The C-terminal noncatalytic segment of PTP2C displays almost no homology with that of PTP1C but shows significant sequence identity (ca. 30%) to that of csw-PTP. Whether this segment of PTP2C has any regulatory function, as found for PTP1C (19), remains to be determined. Nevertheless, the existence of putative sites of phosphorylation by MAP kinase in the C termini of both PTP2C and csw-PTP suggests that these enzymes might be regulated by MAP kinase.

In the present studies, a variant of PTP2C, termed PTP2Ci, was identified also. It is identical to PTP2C except for an insert of four amino acid residues (ALLQ) within its catalytic domain. The position of this insertion corresponds to an exon–intron junction in the analogous PTP1C gene (D.B., unpublished results), making it likely that the insertion results from an alternative splicing event. Whether this results in a change in catalytic activity or substrate specificity remains to be determined.

During the preparation of this manuscript, the amino acid sequence of several peptide fragments derived from a rat liver enzyme, PTP-SH2β, was reported (25). Comparison of the published peptide sequences, totalling 272 residues, with the

sequence of human PTP2C reveals that they are identical except for the substitution of Lys-389 in PTP2C for an arginine in PTP-SH2β. It can be assumed, therefore, that PTP-SH2β is the rat homologue of human PTP2C.

We are grateful to Dr. Morag Park for providing the human umbilical cord cDNA library. We thank Su-Fan Pen, Denis L'Abbe, Robert Larocque, and Rino Stocco for their skillful technical assistance.

1. Aaronson, S. A. (1991) *Science* **254,** 1146–1153.
2. Cantley, L. C., Auger, K. R., Carpenter, C., Duckworth, B., Graziani, A., Kapeller, R. & Soltoff, S. (1991) *Cell* **64,** 281–302.
3. Nishibe, S., Wahl, M. I., Hernandez-Sotomayor, S. M. T., Tonks, N. K., Rhee, S. G. & Carpenter, G. (1990) *Science* **250,** 1253–1256.
4. Huganir, R. L., Delcour, A. H., Greengard, P. & Hess, G. P. (1986) *Nature (London)* **321,** 775–776.
5. Hopfield, J. F., Tank, D. W., Greengard, P. & Huganir, R. L. (1988) *Nature (London)* **336,** 677–680.
6. Fischer, E. H., Charbonneau, H. & Tonks, N. (1991) *Science* **253,** 401–406.
7. Pot, D. A. & Dixon, J. E. (1992) *Biochim. Biophys. Acta* **1136,** 35–43.
8. Krueger, N. X., Streuli, M. & Saito, H. (1990) *EMBO J.* **9,** 3241–3252.
9. Yang, X., Seow, K. T., Bahri, S. M., Oon, S. H. & Chia, W. (1992) *Cell* **67,** 661–673.
10. Shen, S.-H., Bastien, L., Posner, B. & Chretien, P. (1991) *Nature (London)* **352,** 736–739.
11. Mathews, R. J., Bowne, D. B., Flores, E. & Thomas, M. L. (1992) *Mol. Cell. Biol.* **12,** 2396–2405.
12. Yi, T., Cleveland, J. L. & Ihle, J. N. (1992) *Mol. Cell. Biol.* **12,** 836–846.
13. Plutzki, J., Neel, B. G. & Rosenberg, R. D. (1992) *Proc. Natl. Acad. Sci. USA* **89,** 1123–1127.
14. Perkins, L. A., Larsen, I. & Perrimon, N. (1992) *Cell* **70,** 225–236.
15. Koch, C. A., Anderson, D., Moran, M. F., Ellis, C. & Pawson, T. (1991) *Science* **252,** 668–674.
16. Pawson, T. & Gish, G. D. (1992) *Cell* **71,** 359–362.
17. Chretien, P. & Shen, S.-H. (1992) *DNA Cell Biol.* **11,** 337–343.
18. Stidies, F. W., Rosenber, A. H., Dunn, J. J. & Dubendorff, J. W. (1990) *Methods Enzymol.* **185,** 60–89.
19. Zhao, Z., Bouchard, P., Diltz, C. D., Shen, S.-H. & Fischer, E. H. (1993) *J. Biol. Chem.*, in press.
20. Clark, L. I., Sanghera, J. S. & Pelech, S. L. (1991) *J. Biol. Chem.* **266,** 1580–1584.
21. Tonks, N. K., Diltz, C. D. & Fischer, E. H. (1988) *J. Biol. Chem.* **263,** 6731–6737.
22. Zander, N. K., Lorenzen, J. A., Cool, D. E., Tonks, N. K., Daum, G., Krebs, E. G. & Fischer, E. H. (1991) *Biochemistry* **30,** 6964–6970.
23. Roussel, R. R., Brodeur, S. R., Shalloway, D. & Laudano, A. P. (1991) *Proc. Natl. Acad. Sci. USA* **88,** 10696–10700.
24. Rotin, D., Margolis, B., Mohammadi, M., Daly, R. J., Daum, G., Li, N., Fischer, E. H., Burgess, W. H., Ullrich, A. & Schlessinger, J. (1992) *EMBO J.* **11,** 559–567.
25. Hiraga, A., Munakata, H., Hata, K., Suzuki, Y. & Tsuiki, S. (1992) *Eur. J. Biochem.* **209,** 195–206.