Proc. Natl. Acad. Sci. USA
Vol. 88, pp. 11266–11270, December 1991
Genetics

Cloning and characterization of a receptor-class phosphotyrosine phosphatase gene expressed on central nervous system axons in *Drosophila melanogaster*

(tyrosine phosphorylation/oncogenes/eye development)

ISWAR K. HARIHARAN, PAO-TIEN CHUANG, AND GERALD M. RUBIN

Howard Hughes Medical Institute, Department of Molecular and Cell Biology, University of California, Berkeley, CA 94720

Contributed by Gerald M. Rubin, September 24, 1991

ABSTRACT **We have cloned and characterized cDNAs coding for a receptor-class phosphotyrosine phosphatase gene from *Drosophila melanogaster*. The gene maps to the polytene chromosome bands 99A7–8. The cDNA clones code for a polypeptide of 1301 amino acids with a predicted molecular mass of 145 kDa. The extracellular domain includes two fibronectin-type III-like domains. The cytoplasmic region contains two tandemly repeated phosphotyrosine phosphatase-like domains. Residues shown crucial for catalytic activity are absent in the second domain. This *Drosophila* receptor-class phosphotyrosine phosphatase polypeptide is expressed on axons of the embryonic central nervous system.**

A number of cellular processes including cell proliferation and differentiation appear regulated by the phosphorylation of proteins on tyrosine residues. Biochemical studies have shown that phosphorylation of a variety of cellular proteins is a dynamic process involving competing phosphorylation and dephosphorylation reactions. The phosphorylation reactions are catalyzed by the protein-tyrosine kinases (PTKs). More recently, the enzymes that catalyze the dephosphorylation reactions, the phosphotyrosine phosphatases (PTPs), have been identified (1, 2). Like the PTKs, the PTPs include receptor-class PTPs that are probably activated by extracellular ligands. Other forms are entirely intracellular. The number of different PTPs that have been identified is increasing steadily, leading to speculations that this family may be as large as the PTK family (3).

PTKs are involved in a number of developmental processes in *Drosophila*. The product of the *sev* gene is a receptor-class PTK that receives a signal required for the determination of the R7 photoreceptor cell in the adult compound eye (4). The product of the *flb* locus, which is the Drosophila epidermal growth factor receptor homolog, appears to have roles in both embryonic development (5, 6) and in the development of the adult eye (7). The product of the *tor* gene is required for the development of the terminal regions of the embryo (8). The *abl* homolog has been implicated in neuronal pathfinding because embryos mutant for both *abl* and *fasI* display defects in growth cone guidance (9).

It seems likely that PTPs may also be involved in regulating each of these processes. However, to date, no existing *Drosophila* mutations have been mapped to PTP-encoding genes. Two *Drosophila* receptor phosphatase genes have been cloned by their sequence similarity to mammalian PTPs (10). The identification of more PTPs may facilitate the correlation of existing mutations with lesions in a PTP gene.

In this study, we describe the cloning and characterization of a *Drosophila* receptor PTP*. This protein, PTP protein encoded at *Drosophila* band 99A (DPTP 99A) has fibronectin-type III repeats in its extracellular domain as are found in several molecules implicated in neural adhesion. The DPTP 99A polypeptide is expressed on the surface of axons in the embryonic central nervous system (CNS).

MATERIALS AND METHODS

Amplification of the Phosphatase Domains. Primers for the PCR were derived from the peptide sequences SDYINA and FWRMVWE. The oligonucleotides used were CGGATCCGA(T/C)TA(T/C)AT(T/C)AA(T/C)GC and GGAATTCCA(A/G/C/T)ACCAT(A/G/C/T)A(G/T)CCA(A/G)AA. cDNA derived by reverse transcription of 0.2 μg of poly(A)$^+$ RNA from eye-antennal imaginal discs or 1 μg of genomic DNA was used as a substrate. PCR reactions were done as described by Wilks (11). The product was cloned into M13 vectors, and individual clones were characterized by sequencing.

cDNA Cloning and Sequencing. The sequence of the amplified fragment was used to synthesize a 20-mer oligonucleotide that was used to screen two cDNA libraries—an eye-antennal disc library in λgt10 (constructed by Alan Cowman, University of California, Berkeley) and a 9- to 12-hr embryo cDNA library in λgt11 (12). Of 5×10^5 plaques screened, three positive plaques were identified in the eye-antennal disc library, and two were identified from the embryo library. The longest cDNA clones were obtained from the eye-disc library. Each clone consisted of three *Eco*RI fragments that were subcloned separately. Deletions were made using exonuclease III (Promega Erase-a-Base kit) and sequenced using the Sequenase kit (United States Biochemical). Sequencing across the *Eco*RI sites was accomplished by amplifying, by PCR, short segments from the phage that spanned the *Eco*RI sites. Sequence similarities were determined by using the BLAST network service at the National Center for Biotechnology Information (13).

RNA Blot Analysis. RNA [1 μg of poly(A)$^+$ or 10 μg of total RNA] was electrophoresed on a formaldehyde gel and transferred to nitrocellulose as described (14). The probe used was a 2.8-kilobase (kb) *Eco*RI fragment from the middle of the cDNA.

Antibodies. A 970-base-pair (bp) *Hin*cII–*Eco*RI fragment corresponding to amino acids 707–1032 was fused in frame to the *Schistosoma japonicum* glutathione-*S*-transferase gene in the pGEX-1 vector, and the fusion protein was purified by affinity chromatography on a glutathione column (15). Purified fusion protein (5 μg per mouse) was injected s.c. with Ribi adjuvant (Immunochem Research) into five female

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, phosphotyrosine phosphatase; CNS, central nervous system; LCA, leukocyte common antigen; LAR, leukocyte antigen-related tyrosine phosphatase; DPTP 99A, PTP encoded at *Drosophila melanogaster* band 99A.

*The sequence reported in this paper has been deposited in the GenBank data base (accession no. M81795).

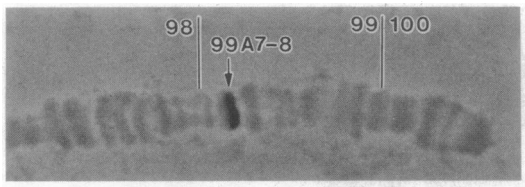


FIG. 1. *In situ* hybridization of a G6 cDNA probe to polytene chromosomes localizes the gene to band 99A7-8. (×800.)

Swiss–Webster mice. Booster immunizations were given at 2-week intervals after the primary dose. Antisera after the third boost (at 1/200 dilution) were used to stain whole-mount embryos (16).

RESULTS

Isolation of DPTP 99A cDNA Clones. To identify additional *Drosophila* PTP genes, we used the PCR to amplify regions of the phosphatase domain from either cDNA prepared from poly(A)$^+$ RNA prepared from eye-antennal imaginal discs or from genomic DNA. A clone, designated G6, was obtained from both sources, and its sequence did not correspond to any previously described PTP. Using the amplified region as a probe, we screened the cDNA libraries. The two longest cDNA clones were obtained from the eye disc library. These clones were characterized further.

To determine the cytological location of the G6-containing gene, a cDNA probe was hybridized to polytene chromosomes. The gene was mapped to bands 99A 7-8 (Fig. 1). We will, henceforth, refer to this gene as DPTP 99A.

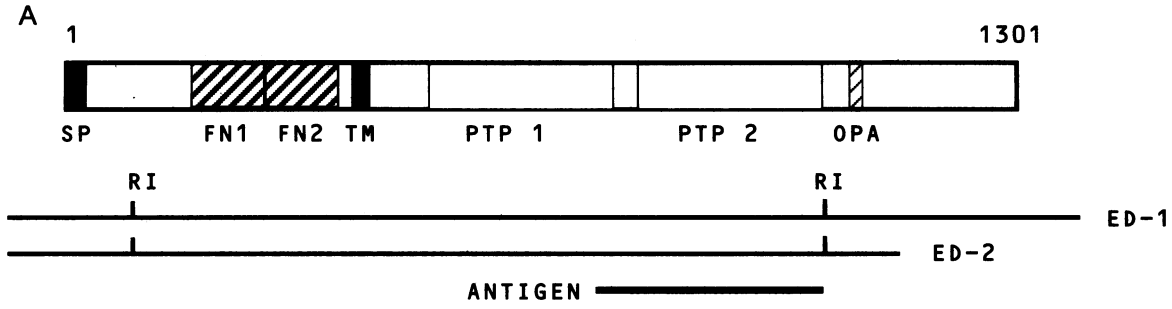
Structures of DPTP 99A Gene and Protein. Sequence analysis of the cDNA clones revealed a long open reading frame of 3903 nucleotides. The likely initiation codon is preceded by three in-frame termination codons, indicating that the entire coding region is included in the cDNA. The sequence predicts a 1301-amino acid polypeptide with a predicted molecular size of 145 kDa (Fig. 2 *A* and *B*).

Conceptual translation of the sequence reveals a polypeptide with the organization of a receptor-class PTP (Fig. 2*A*). The likely N-terminal methionine residue is followed by a hydrophobic region that may serve as a signal peptide. The putative extracellular domain is separated from the remainder of the protein by a stretch of 25 hydrophobic residues that are likely to form a transmembrane domain. Charged residues are found at both ends of this segment. Two tandemly repeated phosphatase domains (≈250 amino acids each) are separated by a stretch of 35 amino acids. The C-terminal region includes a glutamine-rich *opa* repeat.

The Extracellular Domain. The most remarkable feature of the extracellular domain is the presence of two fibronectin type III repeats of ≈100 amino acids each (Fig. 3*A*). Fibronectin type III repeats are typically 90–100 amino acids long with the location of aromatic amino acids and nearby hydrophobic residues being conserved (17, 18). The repeats found in DPTP 99A are compared in Fig. 3*A* to the second fibronectin type III repeat from the receptor phosphatase leukocyte antigen-related tyrosine phosphatase (LAR) (19) and the ninth repeat of chicken tenascin (20), which is a major extracellular glycoprotein. The two repeats found in DPTP 99A are identical to each other at 26 of 100 positions, and conservative substitutions are found at a further 7 positions.

The Intracellular Domain. Two domains, each ≈250 amino acids long with sequence similarity to catalytic PTP domains are found in the predicted intracellular portion of the polypeptide. The first PTP domain is markedly similar to the first PTP domain of human PTP γ (21), being identical at 161 of 251 residues (64%) with conservative substitutions at 28 of the remaining 90 residues (Fig. 3*B*). The second domain is not as closely related to any other PTP domain.

The most notable feature of the second domain is the absence of sequence motifs that are highly conserved in a number of PTP domains and that appear essential for catalytic activity (22). The motif HCSAGVGRT is found in the first PTP domain of all the receptor PTPs. Amino acid substitutions at several of these residues (cysteine, alanine,



B

```
MPRPQHHALLRAMLKLLLFASIAEHCATALPTNSSNSPSSPSPFTVASLPPTTASSSSSPAVISTSSFDRNLADLVNPEAETSGSGWESLETEFNLATTV  100
DSSTQKTAKEPVLGTAATSIEQQDQPPDVPATTLAFANAFPVPVAGEMGNGNGNYNDATPPYAAVDDNYVPSKPQNLTILDVSANSITMSWHPPKNQNGA  200
IAGYHVFHIHDNQTGVEIVKNSRNSVETLIHFELQNLRPYTDYRVIVKAFTTKNEGEPSDQIAQRTDVGGPSAPAIVNLTCHSQESITIRWRRPYEFYNT  300
IDFYIIKTRLAGQDTHRDIRINASAKELETAMILQNLTTNSYYEVKVAAATFSVINPKKIVLGKFSESRIIQLQPNCEKLQPLLRQSHNDYNLAVLVGII  400
FSCFGIILIIMAFFLWSRKCFHAAYYYLDDPPHHPNAPQVDWEVPVKIGDEIRAAVPVNEFAKHVASLHADGDIGFSREYEAIQNECISDDLPCEHSQHP  500
ENKRKNRYLNITAYDHSRVHLHPTPGQKKNLDYINANFIDGYQKGHAFIGTQGPLPDTFDCFWRMIWEQRVAIIVMITNLVERGRPKCDMYWPKDGVETY  600
GVIQVKLIEEEVMSTYTVRTLQIKHLKLKKKKQCNTEKLVYQYHYTNWPDHGTPDHPLPVLNFVKKSSAANPAEAGPIVVHCSAGVGRTGTYIVLDAMLK  700
QIQQKNIVNVFGFLRHIRAQRNFLVQTEEQYIFLHDALVEAIASGETNLMAEQVEELKNCTPYLEQQYKNIIQFQPKDIHIASAMKQVNSIKNRGAIFPI  800
EGSRVHLTPKPGEDGSDYINASWLHGFRRLRDFIVTQHPMAHTIKDFWQMVWDHNAQTVVLLSSLDDINFAQFWPDEATPIESDHYRVKFLNKTNKSDYV  900
SRDFVIQSIQDDYELTVKMLHCPSWPEMSNPNSIYDFIVDVHERCNDYRNGPIVIVDRYGGAQACTFCAISSLAIEMEYCSTANVYQYAKLYHNKRPGVW 1000
TSSEDIRVIYNILSFLPGNLNLLKRTALRTEFEDVTTATPDLYSKICSNGNVPQHVILQQQQLHMLQLQQQHLETQQQQQQQQQQQQQQQQTALNETVST 1100
PSTDTNPSLLPILSLLPPTVAPLSSSSSTTPPTPSTPTPQPPQTIQLSSHSPSDLSHQISSTVANAASPVTPATASASAGATPTTPMTPTVPPTIPTIPS 1200
LASQNSLTLTNANFHTVTNNAADLMEHQQQQMLALMQQQTQLQQQYNTHPQQHHNNVGDLLMNNADNSPTASPTITNNNHITNNNVTSAAATDAQNLDIV 1300
G                                                                                                        1301
```

FIG. 2. (*A*) Structure of DPTP 99A polypeptide. ED-1 and ED-2 are the cDNA clones obtained from the eye-antennal disc cDNA library; RI indicates the *Eco*RI sites. The 1301-amino acid polypeptide predicted by the sequence is schematically represented. SP, signal peptide—amino acids 2–29; fibronectin (FN)1 and FN2, two fibronectin type III repeats (amino acids 171–270 and 271–370, respectively); TM, transmembrane domain (amino acids 393–417); PTP 1 and PTP 2, two PTP domains (amino acids 497–747 and 783–1032, respectively); and OPA, *opa* repeat (amino acids 1076–1091). (*B*) Amino acid sequence of DPTP 99A. The putative signal peptide and transmembrane region are underlined.

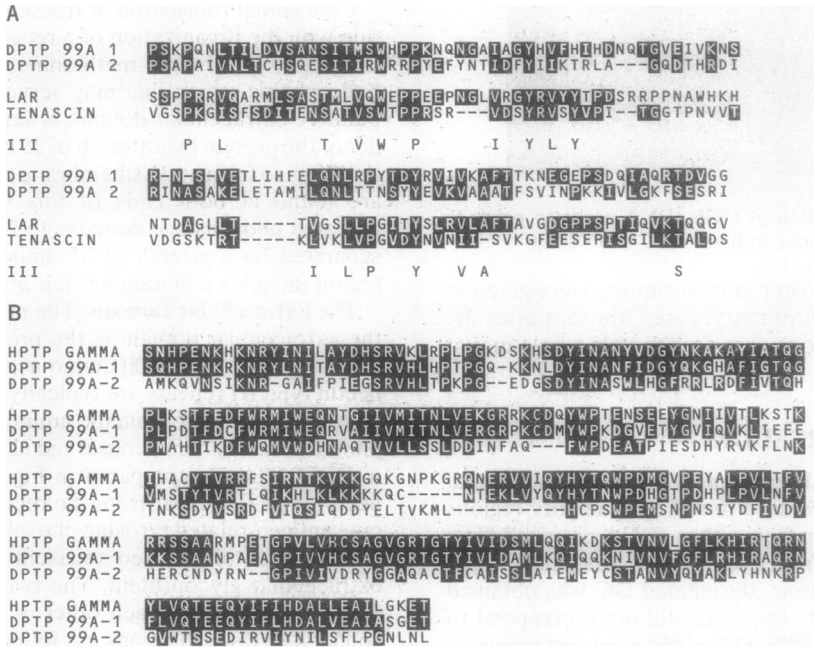


FIG. 3. (*A*) Sequence comparisons between the two fibronectin type III repeats of DPTP 99A (amino acids 171–270 and 271–370), the ninth repeat of chicken tenascin, and the second repeat of human LAR. The consensus sequence derived from examination of a number of fibronectin type III repeats is also shown (III). (*B*) Sequence of the PTP domains of DPTP 99A (amino acids 497–747 and 783–1032) are compared with the first PTP domain of human PTPγ (HPTP GAMMA). Black shading indicates residues that are identical; gray shading indicates conservative substitutions (G/A; V/L/I/M; D/E; R/K; Q/N; S/T; F/Y/W).

arginine, and the first glycine residues) abolish catalytic activity. No corresponding motif is obvious in the appropriate position in the second domain of DPTP 99A (Fig. 3*B*), suggesting that this domain may not have any catalytic activity. In this respect, it resembles the second PTP-like domain of human PTP-γ, where the cysteine that is essential for catalytic activity is replaced by asparagine.

Expression of DPTP 99A Gene. To study expression from the DPTP 99A gene, we analyzed poly(A)$^+$ RNA by RNA blot analysis (Fig. 4). Two transcripts (6 kb and 7.5 kb) were detected in RNA obtained from imaginal discs. Of these, only the 6-kb transcript is detected in RNA prepared from 0- to 20-hr-old embryos. In addition, two transcripts of 5 kb and 3.7 kb are found in the embryo that are not prominent in RNA from imaginal discs. These differences suggest that tissue-specific processing of the DPTP 99A RNA occurs.

To analyze the distribution of the DPTP 99A polypeptide, we raised antibodies to this polypeptide. A segment of the protein derived mainly from the second phosphatase domain, which appears quite dissimilar to other known phosphatase domains, was expressed in *Escherichia coli* as a fusion to the *Schistosoma japonicum* glutathione-*S*-transferase gene (Fig. 2*A*). The purified fusion protein was used to immunize mice. Staining of embryos was detected by using sera raised from two of the five mice. The preimmune sera showed no reactivity.

The antisera stain axons in the embryonic CNS (Fig. 5). The staining appears during the later stages of germ-band shortening and increases in intensity. Staining of the anterior and posterior commissures and the longitudinal connectives is seen. There is also staining of the axons in the commissures in the neuropile of the supraoesophageal ganglia—the supraoesophageal and frontal commissures. No staining is observed in either the nerve roots or the axons of the peripheral nervous system.

Although it is clear that the antibodies recognizing CNS axons were obtained only after immunization with the fusion protein, it is impossible to exclude, in the absence of a mutant that fails to express any DPTP 99A protein, the possibility that some cross-reactive epitope may also be detected. No expression above background was evident in the eye-antennal disc. This finding is surprising because the transcript was detected in RNA from imaginal discs, and the cDNA clones were obtained from a library prepared from eye-antennal discs. One possibility is that the fixation conditions used to stain discs destroys the epitopes recognized by the antiserum. An alternate possibility is that DPTP 99A is expressed at a low level in many cells, which may not be distinguishable from background, but is expressed at higher levels on embryonic CNS axons.

DISCUSSION

We have cloned and characterized a gene coding for a receptor-class PTP that maps to bands 99A7–8. The presumed extracellular domain has two copies of a motif similar to the type III repeats found in fibronectin and several neural adhesion molecules. The intracellular domain includes two

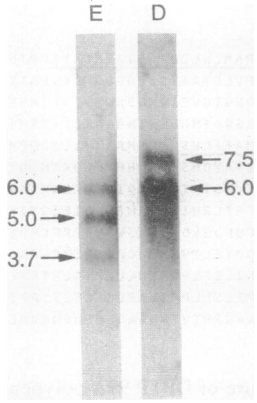


FIG. 4. Expression of the DPTP 99A gene. Blots of RNA from 0- to 20-hr embryos (E) or imaginal discs (D) were hybridized with a DPTP 99A probe.

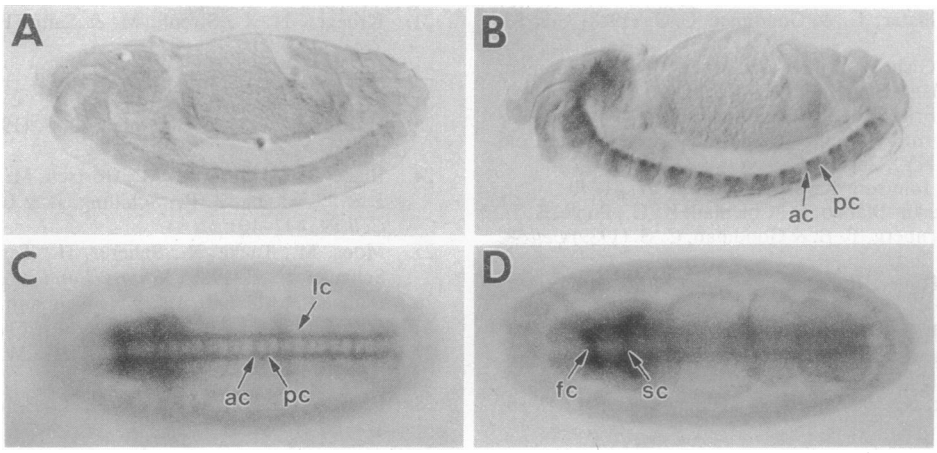


FIG. 5. Expression of DPTP 99A polypeptide. (*A*) Staining with preimmune serum. (*B–D*) Staining with serum after immunization. Stainings of the anterior commissures (ac), posterior commissures (pc), longitudinal connectives (lc), frontal commissure (fc), and supraoesophageal commissure (sc) are indicated. *A* and *B* are lateral views; *C* and *D* are ventral views; anterior is at left in *A–D*. (×100.)

regions of ≈250 amino acids, each that have sequence similarity to the "core region" of PTPs. However, the second of these domains lacks crucial amino acid residues that have been shown essential for catalytic activity, suggesting that this domain may lack PTP activity. An antibody raised to the DPTP 99A polypeptide stains CNS axons.

A significant portion of the extracellular domain of DPTP 99A is made up of the two fibronectin-type III repeats. Remarkably, many of the receptor PTP isolated so far have fibronectin-type III repeats in their extracellular domains (21). No ligands have been identified so far for any of the receptor PTPs. Most receptor PTKs do not have this motif in their extracellular domains, raising the possibility that the PTPs and the PTKs may interact with different classes of ligands.

Motifs similar to the type III repeats of fibronectin have also been found in several putative neural adhesion molecules including fasciclin II (23), neuroglian (24) L1, and N-cellular adhesion molecule (N-CAM) (25). Because N-CAM has been shown to mediate homophilic interactions, it is conceivable that these repeats may be involved in promoting cell adhesion, perhaps by promoting homophilic interactions. However, the family of receptors, which includes the growth hormone receptor and the prolactin receptor, that bind small soluble polypeptide hormones all appear to possess a fibronectin type III domain (18). Moreover, fibronectin type III repeats are found in twitchin, an intracellular protein of *Caenorhabditis elegans* that regulates myosin function and muscle assembly, suggesting that this motif may occur in a variety of contexts to facilitate particular forms of protein–protein interactions (26).

There is a high degree of similarity between the first phosphatase domain of DPTP 99A and human PTPγ. However, their second phosphatase domains seem far less related. Because the extracellular domain of human PTPγ has not yet been characterized, whether similarity is confined to the first phosphatase domain is unclear.

The tandem arrangement of two phosphatase domains in DPTP 99A and several other receptor PTPs is intriguing. In the receptor PTPs LAR and leukocyte common antigen (LCA), only the first domain appears to possess catalytic activity (22). This situation is also probably so for DPTP 99A because several motifs that are highly conserved among catalytic domains are absent in the second PTP domain of DPTP 99A. However, within related groups of PTPs, the second PTP domains are often more similar than the first domains (21). For instance, the first PTP domain of *Drosophila* LAR is 72% identical to human PTP δ and 71% identical to human LAR. The corresponding second domains display identities of 85% and 83%, respectively. This high degree of conservation hints at some physiological role for this second domain—possibly a regulatory one.

The DPTP 99A polypeptide appears to be expressed exclusively on CNS axons. This pattern of expression is very similar to that of the *abl* tyrosine kinase (27). Although homozygous mutants at the *abl* locus do not have any gross defects in the morphogenesis of their CNS, embryos mutant for both *abl* and *fas1* have grossly abnormal axon scaffolds characterized by a loss of most axons in the commissures and many axons in the longitudinal tracts between segments (9). The similar patterns of expression of DPTP 99A and *abl* raises the possibility that they may interact by regulating the phosphorylation states of shared substrates or that DPTP 99A may regulate the tyrosine kinase activity of *abl* by regulating its tyrosine phosphorylation state. Indeed, there is evidence for interaction between the receptor PTP LCA and *src*-class PTKs in T lymphocytes; LCA may activate the *lck* and *fyn* tyrosine kinases by dephosphorylating tyrosine residues (28).

The PCR screen for PTP genes was conducted in collaboration with Mark Seeger and Corey Goodman. We thank Todd Laverty for performing the hybridization to polytene chromosomes. We are grateful to Bill Chia and Kai Zinn and members of their laboratories for the free exchange of information during this work. Alex Kolodkin dissected the nerve cords. We thank Matthew Freeman, Bruce Hay, Mark Seeger, Mike Simon, and Jessica Treisman for comments on the manuscript. I.K.H. was funded by a Medical Foundation Travelling Fellowship of the University of Sydney and currently by the Leukemia Society of America.

1. Tonks, N. K., Diltz, C. D. & Fischer, E. H. (1988) *J. Biol. Chem.* **263**, 6731–6737.
2. Charbonneau, H., Tonks, N. K., Walsh, K. A. & Fischer, E. H. (1988) *Proc. Natl. Acad. Sci. USA* **85**, 7182–7186.
3. Hunter, T. (1989) *Cell* **58**, 1013–1016.
4. Hafen, E., Basler, K., Edstroem, J. E. & Rubin, G. M. (1987) *Science* **236**, 55–63.
5. Schejter, E. D. & Shilo, B. Z. (1989) *Cell* **56**, 1093–1104.
6. Price, J. V., Clifford, R. J. & Schupbach, T. (1989) *Cell* **56**, 1085–1092.
7. Baker, N. E. & Rubin, G. M. (1989) *Nature (London)* **340**, 150–153.
8. Sprenger, F., Stevens, L. M. & Nusslein-Volhard, C. (1989) *Nature (London)* **338**, 478–483.
9. Elkins, T., Zinn, K., McAllister, L., Hoffman, F. M. & Goodman, C. S. (1990) *Cell* **60**, 565–575.
10. Streuli, M., Krueger, N. X., Tsai, A. Y. M. & Saito, H. (1989) *Proc. Natl. Acad. Sci. USA* **86**, 8698–8702.
11. Wilks, A. F. (1989) *Proc. Natl. Acad. Sci. USA* **86**, 1603–1607.

12. Zinn, K., McAllister, L. & Goodman, C. S. (1988) *Cell* **53**, 577–587.
13. Altschul, S. F., Gish, W., Miller, W., Myers, E. W. & Lipman, D. J. (1990) *J. Mol. Biol.* **215**, 403–410.
14. Sambrook, J., Fritsch, E. & Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Lab., Cold Spring Harbor, NY), 2nd Ed.
15. Smith, D. B. & Johnson, K. S. (1988) *Gene* **67**, 31–40.
16. Patel, N. H., Martin-Blanco, E., Coleman, K. G., Poole, S. J., Ellis, M. C., Kornberg, T. B. & Goodman, C. S. (1989) *Cell* **58**, 955–968.
17. Norton, P. A., Hynes, R. O. & Rees, D. J. G. (1990) *Cell* **61**, 15–16.
18. Patthy, L. (1990) *Cell* **61**, 13–14.
19. Streuli, M., Krueger, N. X., Hall, L. R., Schlossman, S. F. & Saito, H. (1988) *J. Exp. Med.* **168**, 1553–1562.
20. Spring, J., Beck, K. & Chiquet-Ehrismann, R. (1989) *Cell* **59**, 325–334.
21. Krueger, N. X., Streuli, M. & Saito, H. (1990) *EMBO J.* **9**, 3241–3252.
22. Streuli, M., Krueger, N. X., Thai, T., Tang, M. & Saito, H. (1990) *EMBO J.* **9**, 2399–2407.
23. Harrelson, A. L. & Goodman, C. S. (1988) *Science* **242**, 700–708.
24. Bieber, A. J., Snow, P. M., Hortsch, M., Patel, N. H., Jacobs, J. R., Traquina, Z. R., Schilling, J. & Goodman, C. S. (1989) *Cell* **59**, 447–460.
25. Moos, M., Tacke, R., Scherer, H., Teplow, D., Früh, K. & Schachner, M. (1988) *Nature (London)* **334**, 701–703.
26. Benian, G. M., Kiff, J. E., Neckelmann, N., Moerman, D. G. & Waterston, R. H. (1989) *Nature (London)* **342**, 45–50.
27. Gertler, F. B., Bennett, R. L., Clark, M. J. & Hoffman, F. M. (1989) *Cell* **58**, 103–113.
28. Klausner, R. D. & Samelson, L. E. (1991) *Cell* **64**, 875–878.