Molecular cloning and expression of a unique rabbit osteoclastic phosphotyrosyl phosphatase

Li-Wha WU, David J. BAYLINK and K.-H. William LAU*

Departments of Medicine and Biochemistry, Loma Linda University and Mineral Metabolism, Jerry L. Pettis Memorial V.A. Medical Center, 11201 Benton Street, Loma Linda, CA 92357, U.S.A.

Tyrosyl phosphorylation plays an important regulatory role in osteoclast formation and activity. Phosphotyrosyl phosphatases (PTPs), in addition to tyrosyl kinases, are key determinants of intracellular tyrosyl phosphorylation levels. To identify the PTP that might play an important regulatory role in osteoclasts, we sought to clone an osteoclast-specific PTP. A putative full-length clone encoding a unique PTP (referred to as PTP-oc) was isolated from a 10-day-old rabbit osteoclastic cDNA library and sequenced. A single open reading frame predicts a protein with 405 amino acid residues containing a putative extracellular domain, a single transmembrane region, and an intracellular portion. PTP-oc is structurally unique in that, unlike most known transmembrane PTPs, it has a short extracellular region (eight residues), lacks a signal peptide proximal to the N-terminus, and contains only a single 'PTP catalytic domain'. The PTP catalytic domain shows 45–50 % sequence identity with the catalytic domain of human HPTPβ and with the first catalytic domain of LCA. The PTP-oc gene exists as a single copy in the rabbit genome. The corresponding mRNA (3.8 kb) is expressed in osteoclasts but not in other bone-derived cells (e.g. osteoblasts and stromal cells). The 3.8 kb PTP-oc mRNA transcript was also expressed in the rabbit brain, kidney and spleen. However, the brain and kidney, but not osteoclasts or spleen, also expressed a larger transcript (6.5 kb). The PTP catalytic domain of PTP-oc was expressed as a GST–cPTP-oc fusion protein. *In vitro* phosphatase assays indicated that the purified fusion protein exhibited phosphatase activities at neutral pH values toward *p*-nitrophenyl phosphate, phosphotyrosyl Raytide, and phosphotyrosyl histone, whereas it had no appreciable activity toward phosphoseryl casein. In summary, we have: (a) cloned and sequenced the putative full-length cDNA of a unique PTP (PTP-oc) from rabbit osteoclasts; (b) shown that the mature 3.8 kb PTP-oc mRNA was expressed primarily in osteoclasts and the spleen; and (c) shown that the PTP-oc fusion protein exhibited a phosphotyrosine-specific phosphatase activity. In conclusion, PTP-oc represents a structurally unique subfamily of transmembrane PTPs.

INTRODUCTION

Bone resorption is required for bone growth (modelling), remodelling and repair processes. The bone resorption process is mediated by osteoclasts [1]. Recent observations indicated that at least two members of the protein tyrosyl kinase (PTK) family may be involved in normal regulation of osteoclast formation and activity. Studies with the *op/op* variant of murine osteopetrosis have shown that production of macrophage-colony stimulating factor (M-CSF) (colony stimulating factor-1) and activation of its receptor are required for normal osteoclast formation [2–4]. The M-CSF receptor, which is encoded by the proto-oncogene c-*fms*, contains an intrinsic PTK activity [5]. In addition, recent studies with the c-*src* gene knock-out mice have indicated that the c-*src* PTK activity is essential for the manifestation of osteoclastic bone resorption and ruffled border formation, but not for osteoclast formation [6,7]. The ruffled border is the highly specialized area of the osteoclast membrane, where bone resorption occurs. The pp60$^{c\text{-}src}$ PTK activity, which is regulated by its tyrosyl phosphorylation status, has been shown to correlate with the number of bone-resorbing osteoclasts [8]. Accordingly, these findings strongly suggest that the intracellular protein tyrosyl phosphorylation status in osteoclasts could play an important regulatory role with respect to osteoclast formation and activity.

The steady-state cellular level of protein tyrosyl phosphorylation is controlled by the balance of the PTK and phosphotyrosyl phosphatase (PTP) activities. At one time, tyrosyl phosphorylation was thought to be regulated largely by PTK activities. However, a large body of evidence has accumulated to indicate that PTPs do not play a mere counteracting role to the action of PTKs; but instead are important determinants of various cellular functions [9–12]. Like PTKs, PTPs are widely distributed in various mammalian cells and tissues, and belong to a superfamily of distinct but structurally related enzymes [13,14]. There are at least two distinct PTP families: one family consists of cytosolic enzymes with a single catalytic domain (e.g. human PTP-1B, T-cell PTP, PTP-2C, etc.) [15–17]; and the other family are transmembrane PTPs which structurally are composed of extracellular, transmembrane and cytoplasmic domains. The cytoplasmic portion of the known transmembrane PTPs contains two tandem catalytic domains, with the exception of HPTPβ and DPTP10D, both of which contain only a single catalytic domain [18–20]. Recent studies suggest that the membrane-proximal catalytic domain (PDI) is catalytically active, whereas the membrane-distal catalytic domain (PDII) by itself has no measurable enzymic activity but may have regulatory functions [21,22]. While the cytoplasmic core phosphatase domains of PTPs are highly conserved, the extracellular domains of the receptor-like PTPs are unrelated to each other [23,24]. Although

Abbreviations used: DTT, dithiothreitol; FN-III, fibronectin type-III; GAPDH, glyceraldehyde-3′-phosphate dehydrogenase; GST, glutathione S-transferase; IPTG, isopropyl β-D-thiogalactopyranoside; M-CSF, macrophage-colony stimulating factor; ORF, open reading frame; pNPP, *p*-nitrophenyl phosphate; PTK, protein tyrosyl kinase; PTP, phosphotyrosyl phosphatase; RT-PCR, reverse-transcription PCR; UTR, untranslated region.

* To whom correspondence should be addressed at: Mineral Metabolism Unit (151), Jerry L. Pettis Memorial V.A. Medical Center, 11201 Benton Street, Loma Linda, CA 92357, U.S.A.

The complete cDNA sequence has been deposited in GeneBank, and received the accession number U32587.

numerous PTPs have been cloned from various cell types, the identity of the PTPs in osteoclasts have not been investigated.

While the importance of PTKs, such as pp60$^{c\text{-}src}$ and *c-fms*, in osteoclasts has been well documented [2–8], the potential involvement of PTPs in osteoclast formation and activity has been less clear. Accordingly, as an initial approach to examining the identity and potential functions of PTPs in osteoclasts, we sought to identify and clone a unique PTP in osteoclasts. We reasoned that an osteoclast-specific PTP could have specialized functions in osteoclasts. In this paper, we report the cloning and complete sequence of the full-length cDNA of a novel rabbit osteoclastic PTP, termed PTP-oc.

EXPERIMENTAL

Materials

All oligonucleotides were synthesized by Integrated DNA Technologies, Inc. (Coralville, IA, U.S.A.). Superscript II reverse transcriptase and dithiothreitol (DTT) were from Gibco/BRL Life Technologies (Grand Island, NY, U.S.A.). Restriction enzymes, large fragment of DNA polymerase I (Klenow fragment) and RNasin were purchased from Promega (Los Angeles, CA, U.S.A.). QAIGEN plasmid kits were obtained from QAIGEN, Inc. (Chatsworth, CA, U.S.A.). SDS was ordered from National Diagnostics, Inc. (Atlanta, GA, U.S.A.). Sequence kits and formamide were from United States Biochemical (Cleveland, OH, U.S.A.). Vent (exo$^-$) DNA polymerase and *Eco*RI linker were products of New England Biolabs (Beverly, MA, U.S.A.). Nitrocellulose and MagnaGraph nylon filters were from Micro Separations Inc. (Westborough, MA, U.S.A.). PMSF, BSA (fraction V), isopropyl β-D-thiogalactopyranoside (IPTG), *p*-nitrophenyl phosphate (pNPP), glutathione (reduced form), antipain, leupeptin, pepstatin, histone type IIA, casein, the catalytic subunit of protein kinase A, and Norit A were products of Sigma Chemical Co. (St. Louis, MO, U.S.A.). Glyceraldehyde-3′-phosphate dehydrogenase (GAPDH) cDNA was from the American Type Culture Collection (Rockville, MD, U.S.A.). Raytide and pp60$^{c\text{-}src}$ were obtained from Oncogene Science, Inc. (Uniondale, NY, U.S.A.). [γ-^{32}P]ATP (7000 Ci/mmol) and [α-^{32}P]dCTP (3000 Ci/mmol) were purchased from DuPont NEN (Boston, MA, U.S.A.). dNTP, pGEX-5X-3 vector and glutathione–Sepharose 4B were from Pharmacia Biotech Inc. (Piscataway, NJ, U.S.A.). Bradford protein reagent was from Bio-Rad Laboratories (Hercules, CA, U.S.A.). All other reagents were of either reagent or molecular biology grade and were obtained from Fisher Scientific (Tustin, CA, U.S.A.) or Sigma Chemical Co.

Cell cultures

Rabbit osteoclasts were isolated from bones of 10-day-old rabbits as described previously [25]. Osteoblasts were obtained by outgrowth of cells from explants of bone chips of 10-day-old rabbits. Stromal cells were prepared from bone marrows of 10-day-old rabbits.

Reverse-transcription PCR (RT-PCR)

For RT-PCR, degenerate oligonucleotide primers (oligos I and II) were designed from two conserved regions within the catalytic domains of the previously characterized PTPs. The sense oligo I (5′-AARUGYSMNCARUAYUGGCC) corresponded to the amino acid sequence KC(A/D/H)QYWP, whereas the antisense oligo II (5′-CCNAYRCCBGCRCTRCAGT) corresponded to (I/V)HCSAG(V/A)GGSA. Total RNA from the isolated rabbit osteoclasts was used as the template for the reverse transcriptase. Reverse transcription was initiated by adding 200 units of Superscript II reverse transcriptase to a reaction mixture of 20 μl containing 50 mM Tris/HCl (pH 8.3), 75 mM KCl, 3 mM $MgCl_2$, 0.5 μg of total RNA, 4 μM oligo II, 20 units of RNasin, 10 mM DTT and 1 mM dNTP. The reaction was carried out at 42 °C for 20 min. The first-strand cDNA was purified by phenol–chloroform extraction, and recovered with ethanol–ammonium acetate precipitation.

The PCR was carried out in a 50 μl reaction mixture containing 10 mM KCl, 20 mM Tris/HCl (pH 8.8), 10 mM $(NH_4)_2SO_4$, 2 mM $MgSO_4$, 0.1 % Triton X-100, 2 units of Vent (exo$^-$) DNA polymerase, 0.8 mM dNTP, 0.1 mg/ml BSA, 4 μM each of oligos I and II and with the first-strand cDNA derived from rabbit osteoclast mRNA as the template. Before cycling, the reaction was subjected to a hot start for 5 min at 100 °C, then 2 min at 49 °C, before the addition of Vent (Exo$^-$) polymerase. This procedure was followed by 30 cycles of extension for 1 min at 74 °C, denaturing for 1 min at 94 °C, and annealing for 1 min at 49 °C. The resulting PCR products were purified by phenol–chloroform extraction, blunt-ended by Klenow then subcloned into the *Sma*I-digested pUC119 vector.

Nucleotide sequence determination

Plasmids were prepared by either the modified alkaline lysis method [26] or QAIGEN plasmid isolation kits. DNA sequences were determined by the dideoxynucleotide chain-termination method [27] with a series of specific synthetic oligonucleotides as primers (17–20-mers).

Analyses of nucleotide and amino acid sequences

All DNA and deduced protein sequence searches were done with the release 6.85 of PC/GENE sequence analysis software package developed by IntelliGenetics (Mountain View, CA, U.S.A.). The EMBL 41 and SWISS-PROT 41 databases were searched, respectively, with the Lipman–Pearson FASTA program [28]. Hydropathy index computation for PTP-oc was based on the method developed by Klein et al. [29]. The alignment of multiple protein sequences were achieved with the CLUSTAL program using the Dayoff's matrix for scoring and a K-tuple value of 2 and a gap penalty of 10 [30].

Screening of the rabbit osteoclastic cDNA library and isolation of cDNA clones

One of the PCR subclones (i.e. PCR1), which encoded a novel PTP fragment, was used as the probe to screen a λZAPII rabbit osteoclastic cDNA library (a generous gift from Dr. M. Kumegawa; [25]). Nitrocellulose filter lifts of the phage plates, each with 20000 plaques/100 mm^2, were washed with 5 × SSC, 0.5 % SDS and 1 mM EDTA (pH 8) at 42 °C for 2 h prior to hybridization (SSC: 0.15 M NaCl, 0.015 M sodium citrate). The nitrocellulose membranes were hybridized overnight at 42 °C in a hybridization buffer consisting of 5 × SSPE, 5 × Denhardt's solution, 0.1 % SDS, 100 μg/ml salmon sperm DNA and 10^6 c.p.m./ml of probe [SSPE: 0.15 M NaCl/10 mM sodium phosphate (pH 7.4)/1 mM EDTA]. The filters were then washed in 2 × SSC/0.1 % SDS at room temperature four times, each for 10 min, followed by washing once in 1 × SSC/0.1 % SDS at 65 °C for 80 min. The final washes were with 0.2 × SSC/0.1 % SDS at 65 °C twice, each for 90 min. Putative positive phage clones were plaque-purified. The pBluescript-derived plasmids containing the osteoclastic cDNA inserts were recovered from λZAPII phage by *in vivo* excision according to the manufacturer's

(Strategene) instruction. Among 500000 plaques, 16 clones were picked, selected, and classified into five groups by restriction enzyme mapping. To obtain full-length clones, the cDNA clone containing the longest cDNA insert (clone 4-21) was selected for sequencing. Both strands of the cDNA insert in the plasmid were sequenced by the dideoxy chain-termination method using a series of oligonucleotide primers.

Southern analysis

Rabbit genomic DNA was cleaved with the restriction enzymes *Dra*II, *Sac*I, *Puv*II, *Kpn*I, *Bam*HI and the combination of *Bam*HI and *Hin*dIII. Following electrophoresis, the digested genomic DNA was denatured and transferred to a MagnaGraph nylon filter as described elsewhere [31]. The filter was first hybridized to the ^{32}P-labelled 1.6 kb fragment derived from the 5′-end of the cDNA insert (5′ cDNA probe in Figure 3). The same filter was stripped and rehybridized to the labelled 1.7 kb segment derived from the 3′-end of the cDNA insert (3′ cDNA probe in Figure 3). Both cDNA fragments were labelled by the random-priming method [32]. Hybridization of DNA blots were carried out overnight at 42 °C in a buffer containing 50 % formamide, 7 % SDS, 0.25 M sodium phosphate (pH 7.2), 0.25 M NaCl, 1 mM EDTA (pH 8), 200 μg/ml denatured salmon sperm DNA and ^{32}P-labelled probe (10^6–10^7 c.p.m./ml). The filters were washed with 2 × SSPE/0.1 % SDS twice at room temperature, and twice at 50 °C, each time for 20 min. The final washes were carried out at 50 °C in 0.2 × SSPE/0.1 % SDS twice, each for 30 min. Autoradiography was performed with a DuPont Cronex III intensifying screen at −80 °C.

Northern blot analysis

Total RNA from various rabbit tissues and isolated cells was prepared and electrophoresed on a 1.1 % agarose/formaldehyde gel and transblotted on to a MagnaGraph nylon filter as described previously [31]. The blot was hybridized with ^{32}P-labelled cDNA fragments in 50 % formaldehyde, 5 × SSPE, 5 × Denhardt's solution, 0.1 % SDS, 10 % dextran sulphate and 200 μg/ml fragmented salmon sperm DNA overnight at 42 °C. The blot was washed once with 6 × SSPE/0.1 % SDS at room temperature, twice with 2 × SSPE/0.1 % SDS at room temperature, and twice with the same buffer at 42 °C. The filter was finally washed twice with 0.2 × SSPE/0.1 % SDS at 42 °C for 20 min each, and then subjected to autoradiography. To show the integrity of the loaded RNA samples, the same filter was stripped and rehybridized with a ^{32}P-labelled human GAPDH cDNA probe.

Production and purification of PTP-oc fusion proteins

For production of bacterial glutathione S-transferase (GST) fusion proteins containing the catalytic region of PTP-oc (i.e. cPTP-oc), the DNA fragment encoding the putative catalytic domain (nucleotides 665–1538) was cloned in-frame to the gene coding for GST in the pGEX-5X-3 vector. *Eco*RI linkers were used such that a novel *Eco*RI site was introduced at the 5′ end of the catalytic domain in the PTP-oc cDNA clone. The 0.9 kb DNA fragment containing the putative catalytic domain was then released from digestion of the *Eco*RI-mutated PTP-oc cDNA with *Eco*RI and *Pvu*II. The pGEX–cPTP-oc construct was generated by ligation of the *Eco*RI/*Sma*I-linearized pGEX-5X-3 to the 0.9 kb fragment.

The reading frame of the pGEX–cPTP construct was confirmed by DNA sequencing. The pGEX-5X-3 (producing GST protein as a negative control) and the pGEX-cPTP-oc plasmids were each transformed into HB101 *Escherichia coli* cells. The GST and its fusion proteins were generated and purified as described [33]. Briefly, tenfold diluted overnight cultures of transformed HB101 cells were grown to mid-logarithmic phase at 30 °C and expression of fusion proteins was then induced with 0.1 mM IPTG for 5–6 h. Following induction, the cultures were harvested by centrifugation at 3000 ***g*** for 10 min. Cell pellets were re-suspended in a 1/50 culture volume of the lysis buffer consisting of 20 mM Hepes (pH 7.6), 100 mM KCl, 0.2 mM EDTA, 20 % (v/v) glycerol, 1 mM DTT, 1 mM PMSF, 10 μg/ml leupeptin, 10 μg/ml pepstatin and 10 μg/ml antipain. The GST and fusion proteins were batch-purified using glutathione–Sepharose 4B as described in the manufacturer's instruction manual. The purified proteins were kept at 4 °C until assayed for enzyme activity.

Phosphatase assays

The phosphatase activity of the GST–cPTP-oc fusion protein was assayed with (a) pNPP, (b) [^{32}P]tyrosyl-phosphorylated Raytide, and (c) [^{32}P]tyrosyl-phosphorylated histone as the *in vitro* substrate at neutral pH as previously described [20,34–36]. To determine the specificity of the phosphatase activity, [^{32}P]seryl-phosphorylated casein was included as a substrate.

(a) pNPP phosphatase assay

The reaction mixture consisted of 20 mM Tris/HCl (pH 7.2), 0.1 % 2-mercaptoethanol, 0.05 % Triton X-100, 0.1 mg/ml BSA, an aliquot of fusion protein and pNPP at the indicated concentration, in a final volume of 100 μl. The reaction was carried out at 30 °C for 0–20 min and terminated by addition of 1 ml of 0.2 M NaOH. Absorbance at 410 nm was determined and a molar absorption coefficient of 1.78×10^4 $M^{-1} \cdot cm^{-1}$ was used to calculate the concentration of the *p*-nitrophenolate ion produced in the reaction.

(b) Raytide phosphatase assay

Raytide was phosphorylated at the tyrosine residue with [γ-^{32}P]ATP and pp60$^{c\text{-}src}$ PTK as described by the supplier. Following the phosphorylation, the ^{32}P-labelled Raytide was purified to remove the unincorporated [γ-^{32}P]ATP as described elsewhere [37]. The dephosphorylation reaction of [^{32}P]tyrosyl-phosphorylated Raytide was initiated by adding an aliquot of purified fusion protein in a final volume of 50 μl of reaction mixture including 25 mM Hepes (pH 7.6), 5 mM EDTA (pH 8.0), 10 mM DTT and ^{32}P-labelled Raytide [(1–2) × 10^4 c.p.m.]. The reaction was carried out at 30 °C for the indicated length of time, and was terminated by addition of 0.75 ml of a charcoal mixture [0.9 M HCl/90 mM sodium pyrophosphate/2 mM NaH_2PO_4 and 1.4 % (w/v) Norit A]. Following centrifugation, 0.1 ml of supernatant was spotted on a Whatman glass filter circle and [^{32}P]P_i was measured by Čerenkov counting. The relative enzyme activity was reported as the percentage of [^{32}P]P_i released.

(c) Phosphoprotein phosphatase assay

Phosphotyrosyl protein phosphatase assay was measured with tyrosyl phosphorylated histone. Histone type IIA was phosphorylated with [^{32}P]ATP by pp60$^{c\text{-}src}$ as previously described [37]. The phosphoseryl protein phosphatase activity was assayed with ^{32}P-labelled casein, which was phosphorylated at the serine/threonine residues by the catalytic subunit of protein kinase A as described in [37]. Dephosphorylation of phosphorylated proteins was performed under the same conditions as that for Raytide with the exception that 30000 c.p.m. of ^{32}P-labelled histone or ^{32}P-labelled casein, instead of Raytide,

was used as the substrate. The protein phosphatase assay was terminated by adding 100 μl of ice-cold 20 % trichloroacetic acid and 100 μg of BSA into the reaction mixture. Following centrifugation, 50 μl of the supernatant was spotted on a Whatman glass filter circle and [^{32}P]P_i was measured by Čerenkov counting. The relative enzyme activity was reported as the percentage of [^{32}P]P_i released.

Other methods

Protein concentrations were determined by the method of Bradford [38] using BSA as the standard. SDS/PAGE was performed as described elsewhere [31].

RESULTS

Molecular cloning of PTP-oc cDNA

Degenerate oligonucleotides which corresponded to the two conserved motifs [i.e. KC(A/D/H)QYWP and (I/V)HCSAG (V/A)G] within the 'PTP catalytic domain' were synthesized and used in a RT-PCR with rabbit osteoclast mRNA as the template. The PCR product was approx. 300 bp in length, which agreed with the anticipated nucleotide size between the two conserved motifs. The resulting PCR product was subcloned into pUC119 vector and sequenced. Of the 37 PCR subclones sequenced two (i.e. PCR1 and PCR12) were found to encode potentially novel members of the PTP family. To test whether the mRNA recognized by these putative novel PTP probes would display tissue-specific expression, Northern blot analysis of total RNA isolated from various rabbit tissues was performed. The cDNA fragments from PCR1 and PCR12 were each used as the probe for Northern hybridization. Both PCR1 and PCR12 probes hybridized with similar mRNA species, and showed identical tissue-specific expression patterns. In this regard, the probes hybridized with either a single or two mRNA transcripts in the spleen, brain and kidney, but not in thymus, heart, lung, liver, and haematopoietic blood cells (results not shown). Thus, the two PCR probes could be related and recognized the same PTP mRNA species. Accordingly, only the insert of PCR1 subclone was used as the probe to screen the rabbit osteoclast cDNA library.

One half of a million plaques were screened, and five positive cDNA clones were selected for further analysis. Restriction enzyme mapping and partial sequence analysis indicated that these clones could be related (results not shown). Because we are interested in cloning the full-length cDNA of the putative novel PTP, we isolated and purified the cDNA insert of clone 4-21, which contained the longest cDNA insert (i.e. approx. 4 kb), for sequence determination. This cDNA is referred to as PTP-oc.

Sequence analysis of PTP-oc

The complete nucleotide sequence of the PTP-oc cDNA was determined and confirmed by sequencing both strands. The nucleotide and the corresponding predicted amino acid sequences are shown in Figure 1. A single open reading frame (ORF) of 1215 bp was identified in this cDNA insert of 3623 bp in size. As observed in other mammalian PTPs [20,39,40], the PTP-oc cDNA sequence had two in-frame methionine codons (ATG), separated by nine nucleotides at the beginning of the ORF. The nucleotide sequence around the first ATG conforms to the canonical translation initiation signal, (A/G)CC*ATG*G [41]. The putative initiation methionine codon is preceded by several in-frame termination codons. The ORF was preceded by a 5′-untranslated region (UTR) of 391 bp in length and followed by a relatively

```
   1                                                    attaaaccttatggtgctcttaattacctac
  32 atttggatttttatgtttactccaataactaatgatatgaaaaaaataaaacatctaat
  92 taaatttttaaaacattctttctttactcagaaccagctcctccaaaatcactctttgca
 152 gtgaacaaaactcagacttcagtgactctgctgtgggtggaagagggcgtagctgatttc
 212 ttcgaagtctttctgtcaacaagttggctctggtctggaaaccaaactccaggagccagtt
 272 gctgtttcttctcatgttgtgaccatctctagtctccttccagccactgcttacaactgt
 332 agtgtcaccagctttagccacgacagccccagtgttcctacatttatagctgtctcaaca
       M  V  T  E  M  N  P  N  V  V  V  I  S  V  L  A  I  L  S  T    20
 392 ATGGTTACAGAGATGAACCCTAACGTGGTAGTAATCTCAGTGCTGGCCATCCTTAGCACA
       L  L  I  G  L  L  L  V  T  L  I  I  L  R  K  K  H  L  Q  M    40
 452 CTTCTAATTGGTCTGCTGCTTGTTACTCTCATCATTCTTAGGAAGAAACATCTGCAGATG
       A  R  E  C  G  A  G  T  F  V  N  F  A  S  L  E  R  D  G  K    60
 512 GCTAGGGAGTGTGGAGCAGGAACATTTGTCAATTTTGCATCTTTAGAGAGGGATGGAAAG
       L  P  Y  N  W  R  R  S  I  F  A  F  L  T  L  L  P  S  C  L    80
 572 CTTCCCTACAACTGGCGTAGAAGCATCTTTGCTTTCTTAACTCTGCTGCCTTCATGTCTT
       W  T  D  Y  L  L  A  F  Y  I  N  P  W  S  K  N  G  L  K  K   100
 632 TGGACTGATTATCTTTTGGCATTTTATATTAACCCTTGGAGTAAAAATGGCTTAAAGAAG
       R  K  L  T  N  P  V  Q  L  D  D  F  D  A  Y  I  K  D  M  A   120
 692 AGGAAACTAACTAACCCAGTTCAACTGGATGACTTTGATGCTTACATCAAGGATATGGCC
       K  D  S  D  Y  K  F  S  L  Q  F  E  E  L  K  L  I  G  L  D   140
 752 AAAGACTCTGACTATAAATTTTCCCTTCAATTTGAGGAGTTGAAATTGATCGGACTGGAT
       I  P  H  F  A  A  D  L  P  L  N  R  C  K  N  R  Y  T  N  I   160
 812 ATTCCACATTTTGCTGCCGATCTTCCACTGAACCGATGTAAAAACCGTTACACAAACATC
       L  P  Y  D  F  S  R  V  R  L  L  S  M  N  E  E  E  G  A  D   180
 872 TTGCCATATGACTTTAGCCGAGTGAGATTACTCTCCATGAATGAAGAGGAAGGTGCGGAC
       Y  I  N  A  N  Y  I  P  G  Y  N  S  P  Q  E  Y  I  A  T  Q   200
 932 TATATTAATGCCAACTATATTCCTGGGTACAACTCACCCCAGGAGTACATTGCCACCCAG
       G  P  L  P  E  T  R  N  D  F  W  K  M  V  L  Q  Q  K  S  Q   220
 992 GGGCCGCTGCCTGAAACCAGAAATGACTTCTGGAAGATGGTCCTACAACAGAAGTCGCAG
       M  I  V  M  L  T  Q  C  N  E  K  R  R  V  K  C  D  H  Y  W   240
1052 ATGATTGTCATGCTCACTCAATGCAATGAGAAAAGGAGGGTGAAATGCGACCATTATTGG
       P  F  T  E  E  P  I  A  Y  G  D  I  T  V  E  M  I  S  E  E   260
1112 CCATTCACGGAAGAACCCATCGCCTACGGGGACATCACTGTGGAGATGATCTCGGAGGAA
       E  Q  D  D  W  A  H  R  H  F  R  I  N  Y  A  D  E  M  Q  D   280
1172 GAACAGGATGACTGGGCCCATAGACACTTCCGGATCAACTACGCTGATGAGATGCAGGAT
       V  M  H  F  N  Y  T  A  W  P  D  H  G  V  P  T  A  N  A  A   300
1232 GTGATGCATTTTAATTACACTGCATGGCCTGATCACGGTGTACCCACGGCCAATGCCGCT
       E  S  I  L  Q  F  V  H  M  V  R  Q  Q  A  T  K  S  K  G  P   320
1292 GAAAGTATCCTGCAGTTTGTACACATGGTCCGACAGCAAGCCACCAAGAGCAAAGGCCCC
       M  I  I  H  C  S  A  G  V  G  R  T  G  T  F  I  A  L  D  R   340
1352 ATGATCATTCACTGCAGCGCTGGAGTGGGGCGGACAGGAACCTTCATTGCCCTGGACAGG
       L  L  Q  H  I  R  D  H  E  F  V  D  I  L  G  L  V  S  E  M   360
1412 CTCCTGCAGCACATTCGGGACCATGAGTTTGTCGACATCTTAGGGCTGGTGTCGGAAATG
       R  S  Y  R  M  S  M  V  Q  T  E  E  Q  Y  I  F  I  H  Q  C   380
1472 AGGTCATACAGGATGTCTATGGTACAGACAGAGGAGCAGTACATTTTTATCCATCAGTGT
       V  Q  L  M  W  M  K  K  K  Q  Q  F  C  I  S  D  V  I  Y  E   400
1532 GTGCAGCTGATGTGGATGAAGAAGAAGCAGCAATTCTGCATCAGTGACGTCATATACGAA
       N  V  S  K  S  *
1592 AATGTTAGCAAGTCCTAGttcaggatccagagcagagaggacgtgatctgcacccatcct
1652 cccttgcttccagacattttggggagccctgctagtcattttgctaacaggagcccctgc
1712 tttgtagtatgtggccaaggagatatttttatctcatagaagcactgagaagacttagcc
1772 ttaaagagcctacagtgtcttttggactctttcacttctggaaatttaataatggacaaa
1832 cccaacagaacacctgaaaggtcaagatgcactctgcaaagggcagcgagtatggcactt
1892 ctcaagactttcaggccctttgctagttgggctgagtttttttgtgttcatttttttaaag
1952 tgcaataatttctgtatatgattttatcagacagttgaattgttttctgcccacaccatt
2012 gaccgaccccatagcccaggaaggaacaggcatggttagcattgaattatacctcattgt
2072 tcaaaaaaagcatggtcacacatcaagaaatagcaattctacttcaagtaaatggaccca
2132 gcaatgtctgtactgcaaacctaagctaccagatcagaagcggggagggtggatgggtaa
2192 gagagagggcttctacccacagatcaaccacaaatcttttcctattcaaaatatgaaaag
2252 ctgtaatttaacttcaaagtagaatagaaaaatatttgtattaagtggtctagttcttga
2312 tggttttcttctttattaacagttcgatgttttttccttggccccttggaataatgtga
2372 ctgtccaggttctttttcaagaaaccagatctggttcagaagagcgtcaagtctcattct
2432 tcaaactctgttgctgtttgagcaatcttgggtgcctacactttgtattccttttctgt
2492 cacttgaacacctgtgaaagctatttctttgtgaactataggctgtgaaaatgcactttc
2552 tttcccccccccaaagagctgggaatttatgaaattatgacaaagcatgctaggacaatt
2612 ctttggctacatttcctgtaatattgtcaagtatctctatggattaccaaggagatttct
2672 tttttgttatttctgtttgactgtcagtttcattttaagcaatgtaactagtaacatttc
2732 attctttagattttgtataattacagtacatgattgttattgtgacatgaatgctgtca
2792 aaatggacattgatggcattgtgaagcctgttcctttgtgtcacttcctgagaaataggt
2852 ggtggtatagatataccaaggaaaacatttggagtggaaagtaaacacaagctggccgct
2912 tcccggtggctgtcataggaagccttgcttctctgtgtctgatcactgcccattccttgg
2972 cttcctgccattcccgatcacctttatgaaaccgggtatacaggataagcttaatgcgct
3032 atcagagacgttagtgatggtgaacgatgccaagtgaacagctctgcaccccactgtctc
3092 ttgtcccgagaaaccctgtccacaacagaccttgacaacattcctggcaagacggtcagt
3152 gagtcccagtgtcacctcgcaatgtgatttgcagaggatgtgcacaccctgcaggttaac
3212 agcttgggctcctcggggattgtcacttctcactcacccttcctctacttcatgcttccc
3272 tgctgctgcttcctcctcttctcacccctcttttatttcccttcttggacattttaat
3332 catctatcacaccatactcaacaacaaacatgattacagaaaggcggcaaatgtatgtgt
3392 gacattcttctgttatcacacaaggatagcccaagagtggagcctgtgacttcggaagag
3452 ctggtgctgtgtgtggtgtggggtgggagagcagaaatgcagttccaacactgc
3512 cacacaggagcttacagaggttgtgtaacctggaactttacttcttgtgaataaagcac
3572 ttcataaataaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa
```

Figure 1 Nucleotide sequence and predicted amino acid sequence of the PTP-oc

The coding sequence is indicated with uppercase letters, whereas the non-coding sequence is indicated with lowercase letters. The DNA sequence is numbered on the left, and amino acids are numbered on the right; amino acid sequence is showed as single-letter codes above the nucleotides. The termination codon is marked with an asterisk and the polyadenylation signal (aataaa) is underlined in the nucleotide sequence. The region of 28 amino acid residues that is specific for PTP-oc and is not shared by the kidney isoenzyme GLEPP1 is shown in boldface and italic font. The transmembrane segment of the protein is also underlined. The active site of phosphatase catalytic domain is in boldface font.

```
PDPTPOC.RAB      ADLPLNRCKNRYTNILPYDFSRVRLLSMN-EEEGADYINANYI-------
PDHPTPβ.HUM      ALLPENRGKNRYNNILPYDATRVKLSNVD-DDPCSDYINASYI-------
PDILAR.HUM       SNLEVNKPKNRYANVIAYDHSRVILTSID-GVPGSDYINANYI-------
PDICD45.HUM      ARKPFNQNKNRYVDILPYDYNRVELSEIN-GDAGSNYINASYI-------
PDPTP1B.HUM      AKLPKNKNRNRYRDVSPFDHSRIKLHQED-----NDYINASLIKME----
PDPTP2C.HUM      GQRQENKNKNRYKNILPFDHTRVVLHDGDPNEPVSDYINANIIMPEFETK
                      *   ***       *  *  *            ****  *

PDPTPOC.RAB      -PGYNSPQEYIATQGPLPETRNDFWKMVLQQKSQMIVMLTQCNEKRRVKC
PDHPTPβ.HUM      -PGNNFRREYIVTQGPLPGTKDDFWKMVWEQNVHNIVMVTQCVEKGRVKC
PDILAR.HUM       -DGYRKQNAYIATQGPLPETMGDFWRMVWEQRTATVVMMTRLEEKSRVKC
PDICD45.HUM      -DGFKEPRKYIAAQGPRDETVDDFWRMIWEQKATVIVMVTRCEEGNRNKC
PDPTP1B.HUM      ----EAQRSYILTQGPLPNTCGHFWEMVWEQKSRGVVMLNRVMEKGSLKC
PDPTP2C.HUM      CNNSKPKKSYIATQGCLQNTVNDFWRMVFQENSRVIVMTTKEVERGKSKC
                           **  **    *   ** *         **     *    **

PDPTPOC.RAB      DHYWPFTEE---PI-AYGDITVEMISEEEQDDWAHRHFRINYADEMQD--
PDHPTPβ.HUM      DHYWPADQD---SL-YYGDLILQMLSESVLPEWTIREFKICGEEQLDAHR
PDILAR.HUM       DQYWPAR-----GTETCGLIQVTLLDTVELATYTVRTFAL-HKSGSSEKR
PDICD45.HUM      AEYWPSMEE---GTRAFGDVVVKINQHKRCPDYIIQKLNIVNKKEKATGR
PDPTP1B.HUM      AQYWPQKEEKEMIFED-TNLKLTLISEDIKSYYTVRQLELENLTTQET-R
PDPTP2C.HUM      VKYWP----DEYALKEYGVMRVRNVKESAAHDYTLRELKLSKVGQGNTER
                   ***

PDPTPOC.RAB      -VMHFNYTAWPDHGVPTANAAESILQFVHMVRQQATKSKG--PMIIHCSA
PDHPTPβ.HUM      LIRHFHYTVWPDHGVP--ETTQSLIQFVRTVRDYINRSPGAGPTVVHCSA
PDILAR.HUM       ELRQFQFMAWPDHGVP--EYPTPILAFLRRVKACNPLD--AGPMVVHCSA
PDICD45.HUM      EVTHIQFTSWPDHGVP--EDPHLLLKLRRRVNAFSNFF--SGPIVVHCSA
PDPTP1B.HUM      EILHFHYTTWPDFGVP--ESPASFLNFLFKVRESGSLSPEHGPVVVHCSA
PDPTP2C.HUM      TVWQYHFRTWPDHGVPSD--PGGVLDFLEEVHHKQESIMDAGPVVVHCSA
                       *** ***              *           *   ****

PDPTPOC.RAB      GVGRTGTFIALDRLLQHIRDHEF---VDILGLVSEMRSYRMSMVQTEEQY
PDHPTPβ.HUM      GVGRTGTFIALDRILQQLDSKDS---VDIYGAVHDLRLHRVHMVQTECQY
PDILAR.HUM       GVGRTGCFIVIDAMLERMKHEKT---VDIYGHVTCMRSQRNYMVQTEDQY
PDICD45.HUM      GVGRTGTYIGIDAMLEGLEAENK---VDVYGYVVKLRRQRCLMVQVEAQY
PDPTP1B.HUM      GEGRSGTFCLADTCLLLMDKRKDPSSVDIKKVLLEMRKFRMGLIQTADQL
PDPTP2C.HUM      GIGRTGTFIVIDILIDIIREKGVDCDIDVPKTIQMVRSQRSGMVQTEAQY
                 * ** *     *              *       *  *    *   *

PDPTPOC.RAB      IFIHQCV-QLMWMKK------KQQFCI-----------------
PDHPTPβ.HUM      VYLHQCVRDVLRARKLRSEQENPLFPIYENVNPEYHRDPVYSRH
PDILAR.HUM       VFIHEALLE--------------AATCGHTEV----------P
PDICD45.HUM      ILIHQALVE--------------YNQFGETEV----------N
PDPTP1B.HUM      RFSYLAVI-------------------------------------
PDPTP2C.HUM      RFIY-----------------------------------MA
```

Figure 2 Amino acid sequence alignment of PTP-oc with other mammalian PTPs

The catalytic domain (PD) sequences of the following PTPs are aligned: rabbit PTP-oc (146–394 amino acids), human HPTPβ (1722–1997 amino acids), human LAR (1359–1606 amino acids), human CD45 (669–919 amino acids), human PTP-1B (35–275 amino acids) and human PTP-2C (267–517 amino acids). For those PTPs with two catalytic domains, only the first catalytic domain (PDI) is used. Gaps (hyphens) are added for optimal alignment. The amino acid residues that are identical in at least five of six PTPs are in boldface font. Residues shared by all six sequences are indicated by asterisks below the sequences.

long (i.e. 1917 bp) 3′-UTR with stop codons in all three ORFs and a putative polyadenylation signal (AATAAA).

Molecular structure of PTP-oc

The ORF predicted a protein of 405 amino acid residues with a calculated molecular mass of 47276 Da. The hydropathy index computation analysis identified a single stretch of strongly hydrophobic amino acid residues (indicated in Figure 1 by underlining), a characteristic of a transmembrane domain. These 25 hydrophobic residues are followed by several basic residues (i.e. RKK), that are consistent with the stop-transfer signal associated with the membrane-spanning domains [42]. In contrast to the previously described generic transmembrane PTPs, the PTP-oc lacks at the N-terminus a signal peptide that targets the protein to the secretory pathway. The putative extracellular domain is very short (i.e. eight residues) and the cytoplasmic region is comprised of 372 residues. Several consensus motifs, including the two regions that were used to design degenerate oligonucleotides for RT-PCR, were identified in the cytoplasmic region. No SH2 domain was found. The 'signature sequence' of

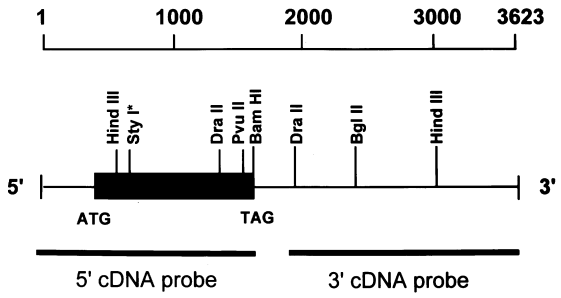

Figure 3 Restriction map of the rabbit PTP-oc cDNA

The restriction map of the rabbit cDNA which resides on the 3623 bp fragment is shown. The first nucleotide reported corresponds to the 5′ end of the PTP-oc cDNA insert. The ORF of the PTP-oc protein is denoted with a black box, while the 5′ and 3′ UTRs are shown by thin lines. The thick bars indicate the length and position of 5′ and 3′ probes used in Northern and Southern blot hybridizations.

the active site (i.e. HCSAGVGRTG), which is highly conserved among the mammalian PTPs [13], was found within the putative 'PTP catalytic domain' of PTP-oc. These findings, taken together, are consistent with the interpretation that PTP-oc is a transmembrane PTP.

Sequence comparison between PTP-oc and other PTPs

Comparison and alignment of the predicted amino acid sequence of PTP-oc with the existing protein sequences in the up-to-date SWISS-PROT databank using the FASTA program indicated that the PTP-oc was most related to human transmembrane HPTPβ and *Drosophila* receptor-linked PTP10D. The sequence similarity (less than 50% identity) was found primarily in the regions containing the 'PTP catalytic domain'. Unlike most known transmembrane PTPs, which have two tandem catalytic domains in their cytoplasmic portions, the cytoplasmic domain of PTP-oc contains only a single 'PTP catalytic domain'. Thus far, the human transmembrane HPTPβ and *Drosophila* PTP10D are the only other known examples of a transmembrane PTP containing only a single catalytic domain in the cytoplasmic regions. Because the catalytic domains of cytosolic and transmembrane PTPs are known to be highly conserved, we compared the core phosphatase domains of the rabbit PTP-oc (shown in Figure 2) with those of three human receptor-like PTPs (i.e. HPTPβ, LAR:PDI, LCA/CD45:PDI) and two human cytosolic PTPs (i.e. PTP-1B, PTP-2C) [16,23,39,43,44]. The predicted catalytic domain of PTP-oc shares higher sequence identity with that of transmembrane PTPs (e.g. 48% with HPTPβ and 45% with LCA/CD45) than that of cytosolic PTPs (e.g. 38% with PTP-1B). While most transmembrane PTPs (except HPTPϵ and HPTPα) have a large extracellular domain of several hundred residues which contains numerous potential glycosylation sites (i.e. N-X-S or N-X-T), Ig-like motifs, or fibronectin type-III (FN-III) motifs [23,36,44], the extracellular domain of PTP-oc is very short (i.e. eight residues), which shares no significant sequence similarity with the known transmembrane PTPs.

Tissue specificity of PTP-oc mRNA

To examine whether expression of PTP-oc mRNA is specific to osteoclasts but not other bone-derived cells, a 1.6 kb 5′ cDNA probe (as shown in Figure 3) was used as a probe in the Northern blot analysis of total RNA isolated from osteoclasts, osteoblasts

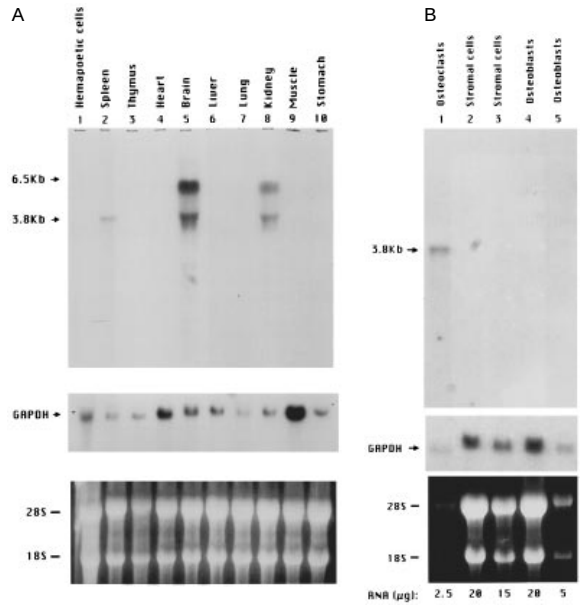

Figure 4 Expression of PTP-oc mRNA in rabbit tissues and bone cells

In (**A**), 20 μg of total RNA from various rabbit tissues were subjected to electrophoresis in a 1.1.% formaldehyde–agarose gel and transferred to a MagnaGraph filter. The RNA blot was hybridized with a ^{32}P-labelled 5′-PTP-oc cDNA probe (top panel). The same blot was stripped and rehybridized with GAPDH cDNA (middle panel). The ethidium bromide-stained 18 S and 28 S ribosomal RNAs used as an index for RNA loading are indicated at the left (bottom panel). In (**B**), 2–20 μg of total RNA from osteoclasts, osteoblasts and bone-marrow stromal cells were subjected to Northern blotting. The blot was hybridized with the 5′-PTP-oc cDNA probe (top panel), then with GAPDH cDNA probe (middle panel). The ethidium bromide-stained 18 S and 28 S rRNAs are indicated (bottom panel).

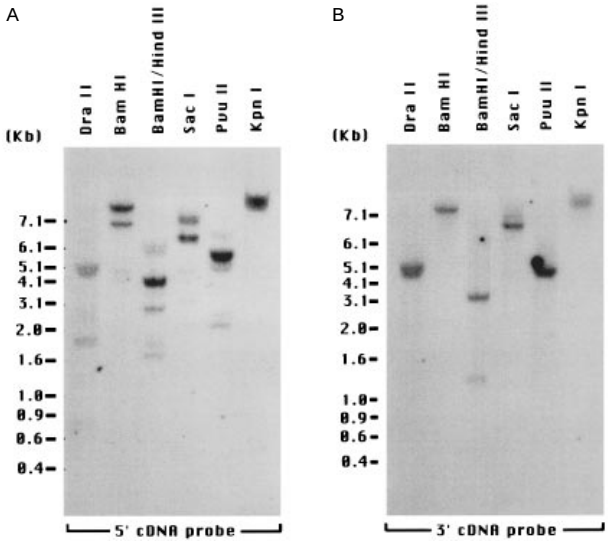

Figure 5 Southern analysis of rabbit osteoblast genomic DNA

Genomic DNA (10 μg) was digested with the indicated restriction enzymes. Digested DNA was transferred to a nylon filter and probed with a ^{32}P-labelled 5′- (**A**) or 3′-PTP-oc cDNA probe (**B**). The numbers indicate the size of DNA markers in kb.

and bone-marrow stromal cells of 10-day-old rabbits. Figure 4 shows that a major transcript of approx. 3.8 kb was found in osteoclasts, but not in osteoblasts or bone-marrow stromal cells, indicating that the expression of the PTP-oc mRNA is specific for osteoclasts and not for other bone-derived cells. We should point out that much less osteoclast RNA (lane 1 in the bottom panel of Figure 4B) than RNAs from osteoblasts and stromal cells (lanes 2–5 in the bottom panel of Figure 4B) has been loaded on to this gel. Accordingly, the lack of the 3.8 kb transcript in osteoblasts and bone-marrow stromal cells was probably not due to insufficient RNA loading. These findings also indicated that this PTP-oc mRNA was expressed in high abundance in osteoclasts. Considering the size of the poly(A) tail in most mammalian mRNA (which is approx. 200 bp in length), the message size of PTP-oc (i.e. 3.8 kb) corresponded well with the determined length of the PTP-oc cDNA clone (3623 bp). Thus, we tentatively conclude that the PTP-oc clone is a full-length clone.

To investigate the tissue-specific expression pattern of the PTP-oc mRNA, the same 5′ cDNA probe (see Figure 3) was used as a probe in multi-tissue Northern blot analysis. Among the tissues tested so far, a transcript of 3.8 kb was found only in rabbit brain, kidney and spleen, in addition to osteoclasts (Figure 4A). However, the 3.8 kb transcript in the spleen was less abundant than that in the brain and kidney. The brain and kidney also expressed, in addition to this 3.8 kb transcript, a related, but larger (6.5 kb) transcript. The spleen, like osteoclasts, did not express this 6.5 kb transcript (Figure 4A). Neither the 3.8 kb nor the 6.5 kb transcript was found in the other test tissues, i.e. haematopoietic blood cells, thymus, heart, lung, liver, stomach and muscle (Figure 4A). To confirm the findings of the tissue-specific expression of PTP-oc, we performed Northern blot analysis using the 3′ probe (see Figure 3), which corresponded to the 3′ end of the PTP-oc cDNA, and obtained identical results to those obtained with the 5′ probe (results not shown).

Southern analysis of the PTP-oc gene

To determine the copy number of the PTP-oc gene, the rabbit genomic DNA was digested with *Bam*HI, *Dra*II, *Kpn*I, *Pvu*II, *Sac*I, and the combination of *Bam*HI and *Hin*dIII. The digested DNA fragments were subjected to Southern analyses. Both the 5′ and 3′ probes (as shown in Figure 3) were used for the hybridization. Figure 5(A) shows that multiple bands with different intensities were seen in the blot hybridized with the 5′ probe. This may reflect the presence of introns within the 5′ region of the gene or of the related genes. Hybridization of the same blot with the 3′ probe indicated that a single unique band was obtained in the digestions with *Bam*HI, *Dra*II, *Kpn*I, *Pvu*II and *Sac*I, alone; whereas two bands were evident in the digestion with the combination of *Bam*HI and *Hin*dIII (Figure 5B). The simplicity of the DNA restriction patterns while using 3′ probe for Southern analysis indicates that the PTP-oc gene exists as a single copy in the rabbit genome.

Evidence that the PTP-oc gene encodes an active PTP

To confirm that PTP-oc indeed encodes a PTP, a bacterial fusion protein (referred to as GST–cPTP-oc) that contained a putative catalytic domain of the PTP-oc linked to the GST protein was produced for analysis of PTP activity. SDS/PAGE analysis revealed that a protein band corresponding to the predicted size (62 kDa) of the GST–cPTP-oc fusion protein was seen in the lysate of IPTG-induced pGEX-cPTP-containing transformants (lane 5 and 6 in Figure 6). The GST–cPTP-oc fusion protein was purified with glutathione–Sepharose affinity chromatography (lane 6 in Figure 6). In addition to the 62 kDa fusion protein, a

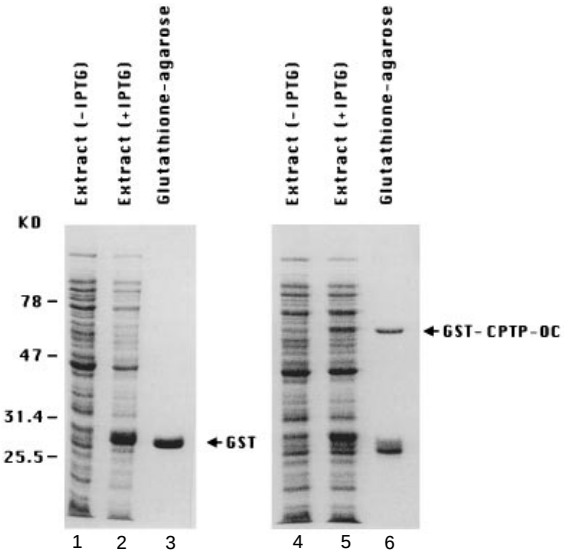

Figure 6 SDS/PAGE analysis of recombinant GST–cPTP-oc

GST–cPTP-oc fusion protein or GST alone was expressed and purified as described in the Experimental section. The purified proteins were then electrophoresed on a 10% SDS/PAGE and stained with Coomassie Blue. Molecular sizes are indicated in kDa. Lanes 1 and 4, 10 μl of bacterial lysate before IPTG induction; lanes 2 and 5, 10 μl of IPTG-induced bacterial lysate; lanes 3 and 6, 5 μl of GST and GST–cPTP-oc purified by glutathione–agarose.

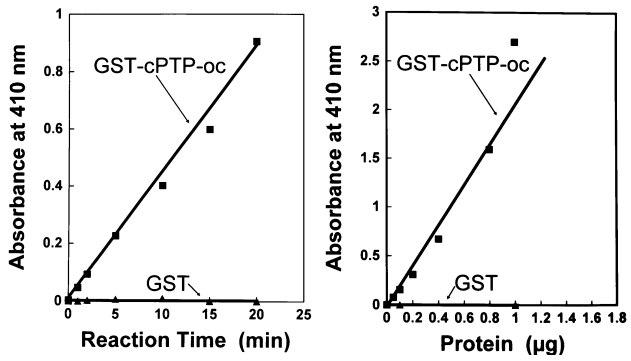

Figure 7 pNPP phosphatase activity of the GST–cPTP-oc fusion protein

GST–cPTP-oc and GST were expressed in *E. coli* and purified in parallel to near homogeneity by affinity chromatography using glutathione–Sepharose (Figure 6). Dephosphorylation of pNPP, expressed as increased absorbance at 410 nm, was measured as a function of time (left-hand panel) and of fusion protein concentration (right-hand panel).

few unexpected smaller proteins were present with the purified fusion protein preparations, presumably resulting from auto-cleavage of the fusion protein. GST protein was also produced and purified as a control following induction of pGEX-5X-3-containing bacteria by IPTG (lanes 1–3 in Figure 6).

Because previous studies showed that PTPs have significant activities on pNPP, tyrosyl-phosphorylated peptides, and phosphotyrosyl proteins *in vitro*, and because the physiological substrate of PTP-oc has not been identified, we measured the phosphatase activity of the fusion protein at neutral pH using three artificial substrates: (a) pNPP, (b) tyrosyl-phosphorylated Raytide, and (c) tyrosyl-phosphorylated histone. Figure 7 shows that the GST–cPTP-oc fusion protein was highly efficient in dephosphorylating pNPP, and that the dephosphorylation was

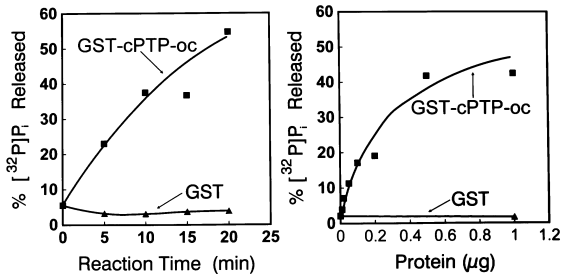

Figure 8 Dephosphorylation of [^{32}P]tyrosyl-phosphorylated Raytide by GST–cPTP-oc fusion protein

The phosphatase activity is expressed as a percentage of [^{32}P]P_i released. The left-hand panel shows the dephosphorylation rate as a function of time; and the right-hand panel shows it as a function of the fusion protein concentration.

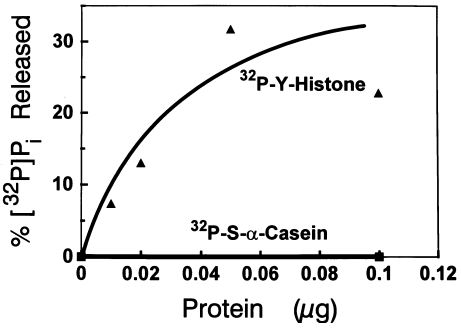

Figure 9 Dephosphorylation of [^{32}P]tyrosyl-phosphorylated histone and [^{32}P]seryl-phosphorylated casein by GST–cPTP-oc

The dephosphorylation rate is shown as percentage of [^{32}P]P_i released. The GST protein did not show any detectable activity with either protein substrates (results not shown).

proportional to the amount of the fusion protein (right-hand panel) and to the reaction time (left-hand panel). GST alone had no activity on pNPP. The apparent V_{max} and K_m of the enzyme activity with pNPP as the substrate were estimated to be 493 units/mg and 1.37 mM pNPP, respectively. Sodium orthovanadate, a known inhibitor for mammalian PTPs [45], at 50 μM completely inhibited the phosphatase activity of the fusion protein (results not shown). The apparent K_i for vanadate was estimated to be 0.51 μM. Similarly, Figure 8 indicates that the GST–cPTP-oc, but not the GST, was effective at dephosphorylating the [^{32}P]tyrosyl-phosphorylated Raytide in a time- (left-hand panel) and dose-dependent (right-hand panel) manner. Figure 9 shows that while PTP fusion protein exhibited strong phosphatase activity toward [^{32}P]tyrosyl-phosphorylated histone, it was ineffective in dephosphorylating [^{32}P]seryl-phosphorylated casein under the same conditions.

DISCUSSION

In this study, we have cloned and completely sequenced a putative full-length PTP (i.e. PTP-oc) cDNA from a rabbit osteoclast cDNA library. This gene indeed encodes a PTP because the GST–cPTP-oc fusion protein displayed strong activity toward pNPP, tyrosyl-phosphorylated Raytide and histone at neutral pH. It is a phosphotyrosine-specific phosphatase and not a dual-specificity phosphatase since the PTP-oc fusion protein was unable to dephosphorylate phosphoseryl casein under the same

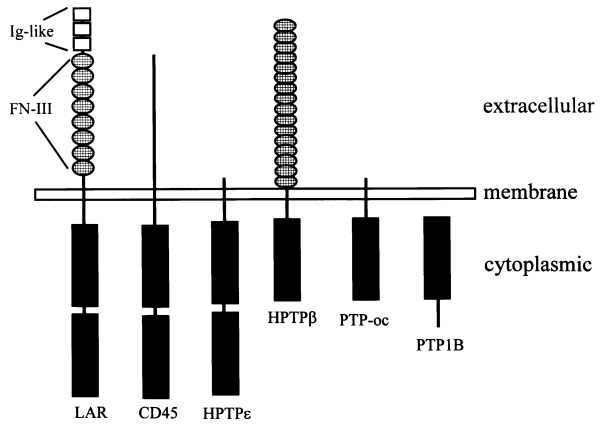

Figure 10 Schematic comparison of the PTP family and PTP-oc

The immunoglobin (Ig)-like, fibronectin (FN) type III-like and catalytic domains are represented by white boxes, hatched ovals, and large black boxes, respectively.

conditions. The PTP catalytic domain of PTP-oc showed a relatively low level of sequence identity with that of the known PTPs. This low level of identity is unlikely to be attributed to the species variance, since the species difference usually accounts for less than 10 % sequence variance among PTP homologues. For instance, human RPTPγ and its murine homologue are 95 % identical in amino acids with 90 % identity in nucleotide sequence [24].

There is circumstantial evidence supporting the possibility that PTP-oc is a member of the transmembrane PTP family. Accordingly, PTP-oc contains a stretch of hydrophobic amino acid residues, followed by several highly charged residues, at a region proximal to the N-terminus. These structural properties are well known characteristics of a transmembrane domain. On the other hand, the lack of a signal peptide and a large extracellular domain are incompatible with the premise that PTP-oc is a transmembrane protein. In addition, the C-termini of some cytosolic PTPs (e.g. PTP-1B and T-cell PTP) have been shown to contain a hydrophobic domain, which serves as an interacting domain with endoplasmic reticulum [46]. (Activation of these cytosolic PTPs requires the removal of the hydrophobic C-termini [39,47]). Thus, we cannot entirely rule out the possibility that PTP-oc is a cytosolic enzyme, which has a regulatory, hydrophobic domain at the N-terminus. However, we favour the possibility that PTP-oc is a 'transmembrane' PTP because the catalytic domain of PTP-oc shares a higher degree of sequence similarity with that of transmembrane PTPs than that with cytosolic PTPs. Future immunocytochemical studies are required to determine its subcellular localization within the osteoclast.

There are three noteworthy observations regarding the molecular structure of PTP-oc when compared with that of other mammalian PTPs (Figure 10): first, PTP-oc is probably one of the smallest 'transmembrane' PTPs, since it is much smaller than most membrane PTPs, whose size is in the range of 80000 to 250000 Da [18,23]. Secondly, unlike most known transmembrane PTPs, the putative intracellular portion of the PTP-oc contains only a single catalytic domain rather than two tandem repeats. Thirdly, in contrast to most transmembrane PTPs whose extracellular domain is relatively large (several hundred residues in length) and contains multiple glycosylation sites, Ig-like and/or FN-III domains, or carbonic anhydrase-like domains [23,24,27,44], the extracellular domain of PTP-oc is very short and lacks these domains. Taken together, these distinct molecular features suggest that PTP-oc may represent a new subfamily of transmembrane PTPs.

The expression of the PTP-oc mRNA appeared to be restricted to the osteoclasts, the spleen, the kidney and the brain. However, it is intriguing to note that the kidney and brain (but not spleen and osteoclasts) expressed, in addition to the 3.8 kb PTP-oc mRNA transcript, also a larger 6.5 kb related mRNA transcript. It has been suggested that mouse osteoclast progenitor cells are derived from the mouse spleen haematopoietic stem cells [48]. Accordingly, that the PTP-oc mRNA is present in the spleen and osteoclasts may suggest that the expression of the PTP-oc (and not the 6.5 kb transcript) is unique to osteoclasts and precursor cells.

The 6.5 kb transcript, referred to as GLEPP1, has recently been cloned from a rabbit kidney cDNA library [49]. Comparison of the nucleotide and predicted amino acid sequences of the kidney GLEPP1 with those of the PTP-oc revealed that the intracellular domains of these two isoenzymes shared approx. 90 % sequence identity, indicating that these two isoenzymes are related. However, the kidney enzyme contains a large extracellular domain, comprising eight repeats of an FN-III-like motif and 15 putative N-glycosylation sites, whereas PTP-oc contains only a very short (i.e. eight residues) putative extracellular domain. Most importantly, PTP-oc contains a unique region of 28 amino acids (residues 66–93 of Figure 1, shown as bold and italic letters) inserted in the intracellular domain, which is absent in the kidney isoenzyme. Thus, PTP-oc appears to be a truncated version of the kidney GLEPP1, presumably resulting from an alternative splicing at the 5′ end of the gene. Accordingly, PTP-oc may be a splicing variant of the kidney isoenzyme. The presence of a splicing variant in osteoclasts and percursor cells (i.e. spleen cells) might play a tissue-specific regulatory role. In this regard, it has been shown that alternative splicing plays a key role in producing alternative variant transcripts for transmembrane PTPs, including human CD45/LCA [50] and rat PTP-P1/PTP-PS [51].

The physiological relevance of a short 'extracellular' domain is unclear. The short extracellular domain does not appear to be unique to PTP-oc since the extracellular domain of HPTPϵ is also very short, i.e. 27 amino acid residues in length [23]. It has been suggested that the extracellular domains, especially the Ig-like and FN-III repeats, of transmembrane PTPs may function as cell-surface receptors or specific binding sites for ligands [23,27,44]. Consistent with this speculation are the findings that Ig-like and FN-III repeats are found at the ligand binding sites of neural cell adhesion receptors [52,53]. Accordingly, it is unlikely that the short extracellular domains of PTP-oc and HPTPϵ function as cell-surface receptors. However, we cannot overlook the possibility that the extracellular domain of PTP-oc (and HPTPϵ) could be part of a multi-subunit receptor complex forming between PTP-oc and one or more molecules with large extracellular regions. Alternatively, because the 33 N-terminal amino acid residues (i.e. the eight residues of the extracellular domain and the 25 residues of the transmembrane domain) of the PTP-oc could function as a weak signal for mitochondria-association according to the prediction by the TRANSPEP program of PC GENE, it may be speculated that PTP-oc could be a mitochondria-associated protein in which the N-terminus may serve as an attaching site with the mitochondrial membrane. Additional work is needed to address these possibilities.

While we are beginning to understand the molecular structure of this unique osteoclastic PTP-oc, its physiological significance and regulation are unknown. However, the expression of PTP-oc mRNA occurred primarily in osteoclasts, suggesting that this

PTP-oc could play a functional role in regulating tyrosyl phosphorylation levels in osteoclasts. Because protein tyrosyl phosphorylation mediated by $pp60^{c\text{-}src}$ PTK has been shown to be absolutely essential for osteoclast activity [6–8], we speculate that this PTP-oc might be involved in the regulation of the protein tyrosyl phosphorylation either by countering the effects of the PTKs (e.g. $pp60^{c\text{-}src}$) through the specific dephosphorylation of the cellular substrates of PTKs, or by altering the PTK activity of $pp60^{c\text{-}src}$ through specific dephosphorylation of this PTK. (The PTK activity of $pp60^{c\text{-}src}$ is inhibited by phosphorylation at Tyr^{527} [54]). Supporting the latter speculation is the recent demonstration that the $pp60^{c\text{-}src}$ dephosphorylation and activation can be mediated by the receptor-like transmembrane PTP [55]. Nevertheless, it seems likely that the physiological role of this PTP-oc will remain undetermined until its physiological substrate(s) is identified.

We thank Drs. Anne Lambline, Todd Starich, Wei Li, Yoko Honda, Edwin Landale and Subburaman Mohan for helpful discussions. Thanks also go to Drs. Donna Strong and Tom Linkhart for suggestions and to the Medical Media staff at the Jerry L. Pettis Memorial VAMC for assistance in the preparation of the manuscript. We also thank Dr. Masayoshi Kumegawa and the Yamanouchi Pharmaceutical Company for providing us with the rabbit osteoclast cDNA library. This work was supported in part by research grants from the National Institute of Health (DE08681) and the Veteran Administration.

REFERENCES

1 Rodan, G. A. and Martin, T. J. (1981) Calcif. Tissue Int. **33**, 349–351
2 Yoshida, H., Hayashi, S., Kunisada, T., Ogawa, M., Nishikawa, S., Okamura, H., Suda, T., Shultz, L. D. and Nishikawa, S. (1990) Nature (London) **345**, 442–444
3 Hofstetter, W., Wetterwald, A., Cecchini, M. C., Felix, R., Fleisch, H. and Mueller, C. (1992) Proc. Natl. Acad. Sci. U.S.A. **89**, 9637–9641
4 Tanaka, S., Takahashi, N., Udagawa, N., Tamura, T., Akatsu, T., Stanley, E. R., Kurokawa, T. and Suda, T. (1993) J. Clin. Invest. **91**, 257–263
5 Sherr, C. J., Rettenmier, C. W., Sacca, R., Roussel, M. F., Look, A. T. and Stanley, E. R. (1985) Cell **41**, 665–676
6 Soriano, P., Montgomery, C., Geske, R. and Bradley, A. (1991) Cell **64**, 693–702
7 Boyce, B. F., Toshiyuki, Y., Lowe, C., Soriano, P. and Mundy, G. R. (1992) J. Clin. Invest. **90**, 1622–1627
8 Yoneda, T., Niewolna, M., Lowe, C., Izbicka, E. and Mundy, G. R. (1993) Mol. Endocrinol. **7**, 1313–1318
9 Shultz, L. D., Schweitzer, P. A., Rajan, T. V., Yi, T., Ihle, J. N., Matthews, R. J., Thomas, M. L. and Beier, R. (1993) Cell **73**, 1445–1454
10 Tsui, H. W., Siminovitch, K. A., de Souza, L. and Tsui, F. W. L. (1993) Nature Genet. **4**, 124–129
11 Flint, A. J., Gebbink, M. F., Franza, B. R., Jr., Hill, D. E. and Tonks, N. K. (1993) EMBO J. **12**, 1937–1946
12 Xiao, S., Rose, W., Sasaoko, T., Maegawa, H., Burke, T. R., Jr., Roller, P. P., Shoelson, S. E. and Olefsky, J. M. (1994) J. Biol. Chem. **269**, 21244–21248
13 Fischer, E. H., Charbonneau, H. and Tonks, N. K. (1991) Science **253**, 401–406
14 Lau, K.-H. W. and Baylink, D. J. (1993) Crit. Rev. Oncogenesis **4**, 451–471
15 Charbonneau, H., Tonks, N. K., Kumar, S., Diltz, C. D., Harrylock, M., Cool, D. E., Krebs, E. G., Fischer, E. H. and Walsh, K. A. (1989) Proc. Natl. Acad. Sci. U.S.A. **86**, 5252–5256
16 Cool, D. E., Tonks, N. K., Charbonneau, H., Walsh, K. A., Fischer, E. H. and Krebs, E. G. (1989) Proc. Natl. Acad. Sci. U.S.A. **86**, 5257–5261
17 Ahmad, S., Banville, D., Zhao, Z., Fischer, E. H. and Shen, S.-H. (1993) Proc. Natl. Acad. Sci. U.S.A. **90**, 2197–2201
18 Saito, H., Streuli, M., Krugger, N. X., Itoh, M. and Tsai, A. Y. M. (1992) Biochem. Soc. Trans. **20**, 165–169
19 Wang, Y. and Pallen, C. J. (1992) J. Biol. Chem. **267**, 16696–16702
20 Streuli, M., Krueger, N. X., Tsai, A. Y. M. and Saito, H. (1989) Proc. Natl. Acad. Sci. U.S.A. **86**, 8698–8702
21 Streuli, M., Krueger, N. X., Thai, T., Tang, M. and Saito, H. (1990) EMBO J. **9**, 2399–2407
22 Itoh, M., Streuli, M., Krueger, N. X. and Saito, H. (1992) J. Biol. Chem. **267**, 12356–12363
23 Krueger, N. X., Streuli, M. and Saito, H. (1990) EMBO J. **9**, 3241–3252
24 Barnea, G., Silvennoinen, O., Shaanan, B., Honegger, A. M., Canoll, P. D., D'eustachio, P., Morse, B., Levy, J. B., Laforgia, S., Huebner, K., Musacchio, J. M., Sap, J. and Schlessinger, J. (1993) Mol. Cell. Biol. **13**, 1497–1506
25 Tezuka, K., Sato, T., Kamioka, H., Nijweide, P. J., Tanaka, K., Matsuo, T., Ohta, M., Kurihara, N., Hakeda, Y. and Kumegawa, M. (1992) Biochem. Biophys. Res. Commun. **186**, 911–917
26 Hackett, P. B., Fuchs, J. A. and Messing, J. W. (1988) An Introduction to Recombinant DNA Techniques: Basic Experiments in Gene Manipulation, 2nd edn., The Benjamin/Cummings Publishing Company, Inc., Menlo Park, CA
27 Sanger, F., Nicklen, S. and Coulson, A. R. (1977) Proc. Natl. Acad. Sci. U.S.A. **74**, 5463–5468
28 Lipman, D. J. and Pearson, W. R. (1985) Science **227**, 1435–1441
29 Klein, P., Kanehisa, M. and DeLisi, C. (1985) Biochim. Biophys. Acta **815**, 468–476
30 Higgins, D. G. and Sharp, P. M. (1988) Gene **73**, 237–244
31 Sambrook, J., Fritsch, E. F. and Maniatis, T. (1989) Molecular Cloning: A Laboratory Manual, vols. 1–3, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY
32 Feinberg, A. P. and Vogelstein, B. (1983) Anal. Biochem. **132**, 6–13
33 Smith, D. B. and Johnson, K. S. (1988) Gene **67**, 31–40
34 Mei, L., Doherty, C. A. and Huganir, R. L. (1994) J. Biol. Chem. **269**, 12254–12262
35 Guan, K. L., Broyles, S. S. and Dixon, J. E. (1991) Nature (London) **350**, 359–362
36 Matozaki, T., Suzuki, T., Uchida, T., Inazawa, J., Ariyama, T., Matsuda, K., Horita, K., Noguchi, H., Mizuno, H., Sakamoto, C. and Kasuga, M. (1994) J. Biol. Chem. **269**, 2075–2081
37 Chan, C. P., Gallis, B., Blumenthal, D. K., Pallen, C. J., Wang, J. H. and Krebs, E. G. (1986) J. Biol. Chem. **261**, 9890–9895
38 Bradford, M. M. (1976) Anal. Biochem. **72**, 248–254
39 Chernoff, J. C., Schievella, A. R., Jost, C. A., Erikson, R. L. and Neel, B. G. (1990) Proc. Natl. Acad. Sci. U.S.A. **87**, 2735–2739
40 Keyse, S. M. and Emslie, E. A. (1992) Nature (London) **359**, 644–647
41 Kozak, M. (1987) Nucleic Acids Res. **15**, 8125–8148
42 van der Geer, P., Hunter, T. and Lindberg, R. A. (1994) Annu. Rev. Cell Biol. **10**, 251–337
43 Charbonneau, H., Tonks, N. K., Walsh, K. A. and Fischer, E. H. (1988) Proc. Natl. Acad. Sci. U.S.A. **85**, 7182–7186
44 Streuli, M., Krueger, N. X., Hall, L. R., Schlossman, S. F. and Saito, H. (1988) J. Exp. Med. **168**, 1553–1562
45 Gordon, J. A. (1991) Methods Enzymol. **201**, 477–482
46 Frangioni, J. V., Beahm, P. H., Shifrin, V., Jost, C. H. and Neel, B. G. (1992) Cell **68**, 545–560
47 Cool, D. E., Tonks, N. K., Charbonneau, H., Fischer, E. H. and Krebs, E. G. (1990) Proc. Natl. Acad. Sci. U.S.A. **87**, 7280–7284
48 Udagawa, N., Takahashi, N., Akatsu, T., Tanaka, H., Sasaki, T., Nishihara, T., Koga, T., Martin, T. J. and Suda, T. (1989) Endocrinology **125**, 1805–1813
49 Thomas, P. E., Wharram, B. L., Goyal, M., Wiggins, J. E., Holzman, L. B. and Wiggins, R. C. (1994) J. Biol. Chem. **269**, 19953–19962
50 Streuli, M., Hall, L. R., Saga, Y., Schlossman, S. F. and Saito, H. (1987) J. Exp. Med. **166**, 1548–1566
51 Pan, M.-G., Rim, C., Lu, K. P., Florio, T. and Stork, P. J. S. (1993) J. Biol. Chem. **268**, 19284–19291
52 Cunningham, B. A., Hemperly, J. J., Murray, B. A., Prediger, E. A., Brackenbury, R. and Edelman, G. M. (1987) Science **236**, 799–806
53 Moos, M., Tacke, R., Scherer, H., Teplow, D., Fruh, K. and Schachner, M. (1988) Nature (London) **334**, 701–703
54 Nada, S., Okada, M., MacAuley, A., Cooper, J. A. and Nakagawa, H. (1991) Nature (London) **351**, 69–72
55 Zheng, X. M., Wang, Y. and Pallen, C. J. (1992) Nature (London) **359**, 336–339

Received 13 November 1995/25 January 1996; accepted 30 January 1996