Partial Catalytic Domains of New Protein-tyrosine Kinases Cloned from cDNA Amplified by Polymerase Chain Reaction

Hiroko SASAKI, Kazuko NAGURA, Kiyoshi KOTANI and Terukatsu SASAKI

Department of Biochemistry, Cancer Research Institute, Sapporo Medical University School of Medicine, Sapporo 060, Japan

ABSTRACT

A feature common to all members of the protein-tyrosine kinase (PTK) family is a highly conserved catalytic domain which is characteristic of this group. Degenerate oligonucleotide primers corresponding to two of the most highly conserved regions of the PTK catalytic domain were designed to amplify cDNA sequences of restricted subfamilies of PTKs from rat brain mRNA in the polymerase chain reaction (PCR). A third degenerate oligonucleotide primer corresponding to a highly conserved, PTK subfamily-specific sequence located between the two sequences mentioned above was also used to amplify cDNA sequences of PTKs of novel subfamilies from rat brain mRNA. pBluescript PCR libraries were constructed from the PCR-amplified cDNA. The PCR libraries were then screened by DNA sequencing for PTK-related sequences. Several sequences were identified that, on the basis of sequence comparison with known PTKs in GenBank, may encode new PTKs.

Key words: Protein-tyrosine kinase, Polymerase chain reaction, Catalytic domain, PCR library.

INTRODUCTION

The polymerase chain reaction (PCR) has been applied to the cloning of cDNA sequences of the PTK gene family and has greatly facilitated the isolation of previously unknown PTKs (1-7). Recent advances have indicated that there are nonreceptor PTKs of previously unknown subfamilies such as ZAP70/Syk PTKs and activated Cdc42Hs-associated PTK (8-10). An ever-expanding array

Abbreviations:
PTK, protein-tyrosine kinase; PCR, polymerase chain reaction.

of PTKs of Eph-and other receptor PTK-subfamilies, JAK-subfamily, and other subfamilies are found by the PCR-based cloning strategy. In this study, partial catalytic domains of several new PTKs were cloned from cDNA libraries constructed from PCR-amplified cDNAs derived from rat brain mRNA; degenerate oligonucleotide primers were designed and used in reverse transcription and PCR to amplify catalytic domains of novel PTK subfamilies.

MATERIALS AND METHODS

Template RNA

RNA was prepared by the method of MacDonald *et al.* (11) from two brains of adult rats of the Wistar strain. Frozen brains were powdered in liquid nitrogen by grinding in a mortar. RNA was extracted from the powdered brain in 20 m*l* of 7.5 M guanidine-HCl/0.1 M Tris-HCl, pH 7.0/10 mM dithiothreitol by mixing for 60 sec in a Waring blender at top speed. About 90 μg of RNA was obtained by the procedure. The RNA preparation was used without further purification.

Oligonucleotides

Oligonucleotides were synthesized on an Applied Biosystems 380A oligonucleotide synthesizer using standard chemistry. The crude oligonucleotide preparation was purified by the use of Oligonucleotide Purification Cartridge columns (Applied Biosystems, Foster USA).

Reverse-transcription, DNA amplification, and construction of pBluescript PCR libraries

The first strand cDNA was synthesized complementary to rat brain mRNA. 2.26 μg of total RNA from rat brain in 11 μl water was kept at 94℃ for 10 min. To this solution, 2 μl 10×PCR Reaction Buffer (Perkin Elmer PJ5100), 1 μl 70 mM $MgCl_2$, 1 μl each of dATP, dCTP, dGTP, and dTTP, 0.5 μl RNase inhibitor (Perkin Elmer 2310), 1 μl 50 μM reverse primer (either PTK2-primer or PTK3-primer), and 0.5 μl reverse transcriptase (Perkin Elmer 2610) were added and the mixture was incubated at 42℃ for 1 h. After the incubation, 3 μl 10×PCR Reaction Buffer, 25.7 μl water, 1 μl 50 μM PTK1-primer, and 0.3 μl AmpliTaq DNA Polymerase (Perkin Elmer 2531) were added to the mixture. PCR was performed in a Model PJ2000 DNA Thermal Cycler (Perkin Elmer). The PCR cycle was 1 min at 94℃ (denaturation), 2 min at 55℃ (annealing), and 2 min at 72℃ (elongation). A total of 30 cycles was performed. After the reaction, mineral oil was removed by addition of 100 μl chloroform. The reaction mixture in the upper phase was washed with phenol/chloroform (1 : 1). The amplified DNA

was ethanol-precipitated, air-dried, and dissolved in 3.5 μl water. To the solution, 1 μl 10×BamHI buffer (1 M NaCl, 100 mM Tris-HCl, pH 8.0, 70 mM $MgCl_2$, 20 mM 2-mercaptoethanol, 1 mg/ml bovine serum albumin), 0.25 μl (5 units) BamHI, and 0.25 μl (5 units) EcoRI were added and the mixture was incubated at 30℃ overnight.

The DNA was then purified by electrophoresis in 3% NuSieve GTG agarose (FMC Bioproducts, Rockland USA). An ethidium bromide-stained band at 180-220 base pairs was cut out and the agarose gel containing the band was dissolved at 65℃. Warmed Tris/EDTA buffer was added to reduce the gel concentration to less than 0.5%. The DNA was extracted from the gel suspension by mixing with phenol, and the extract was washed with phenol/chloroform and then with chloroform. The washed DNA solution was concentrated by mixing with butanol. The DNA was then ethanol precipitated, air-dried, and dissolved in 5 μl Tris/EDTA buffer. The DNA (1 μl) was ligated into 50 ng of EcoRI- and BamHI-cleaved pBluescript II SK (+) (Pharmacia). E. coli XL1-blue competent cells were transformed with the PCR library by electroporation. Plasmid DNA was extracted by alkali-SDS method from E. coli of β-galactosidase-negative white colonies. Only those plasmid DNAs which contain expected EcoRI/BamHI fragment were subjected to sequencing. Sequencing was carried out by the dideoxy chain-termination method (12), using a Sequenase kit (United States Biochemical). In all cases [α-^{32}P]dCTP (Amersham, catalogue no. PB10205) was used. Sequencing was done from the 3' side of the inserts in pBluescript by the use of T7 primer, and was not always completed up to the sequence of the forward primer.

RESULTS AND DISCUSION

Selection of oligonucleotide primer sequences

Three regions of the catalytic domains of the PTKs were used to derive oligonucleotide primers. PTK I and PTK II, shown in Fig. 1, have been identified as the most highly conserved amino acid sequences of the PTK catalytic domains (13), and have been selected for the derivation of oligonucleotide primers. The sequence -V/IHRDLA/RAR/AN- (PTK I) defined the 5' border of the region of the cDNA to be amplified, and the sequence -SDVWSFG- (PTK II) defined the 3' border in the first PCR amplification. Between these two sequences is a highly conserved, subfamily specific sequence, -W (M/T/Y/C/S/A) (A/P/S/) PE- (PTK III), which also defined the 3' border in the second PCR amplification. The amino acid sequences of PTKs in PTK III are conserved across the PTK subfamilies to permit an assignment of any candidate PTK clones to the cytoplasmic or the growth factor receptor subfamilies, or, more recently, to new sub-

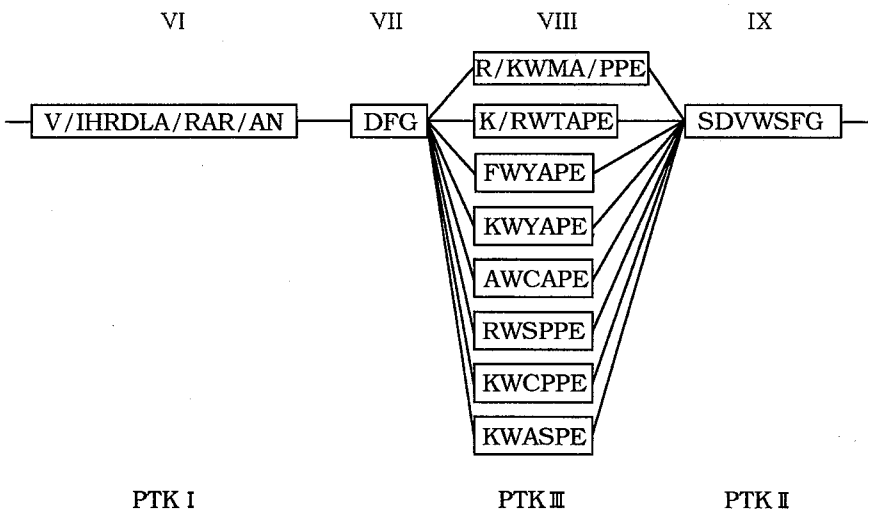

Fig. 1 Conserved regions of PTK catalytic domains used in the PCR amplification of PTK-related sequences. Blocks of identity held in common between the catalytic damains of PTKs are shown as boxes, with the conserved motif written in one-letter code within the boxes; "/" in the amino acid sequence indicates "or". Each conserved motif is labeled above the boxes by the nomenclature of Hanks *et al.* (15). Degenerate oligonucleotide primers were derived from the -FI/VHRDLA- motif (PTK I), -SDVWSFG- motif (PTK II), and -WYAPECIN- motif (PTK III). The motif VIII appears to be subfamily-specific: -R/KWMA/PPE- in receptor PTKs, -K/RWTAPE- in nonreceptor PTKs of the Src subfamily and receptor PTKs of the Eph-subfamily, -FWYAPE- in PTKs of the JAK subfamily, -KWYAPE- in PTKs of the ZAP70/Syk subfamily, -AWCAPE- in activated Cdc42Hs-associated PTK (10), -RWSPPE- in agammaglobulinaemia PTK (16), -KWCPPE- in Tec (17), and -KWASPE- in Itk (18).

families of PTKs.

The oligonucleotide primers derived (Fig. 2 and 3) were based upon the nucleotide sequences corresponding to the three consensus amino acid sequences. All three primers, PTK1-, PTK2-, and PTK3-primers, were designed to avoid amplification of the Src subfamily of PTKs; these primers have mismatches at the 3' terminal base sequences with PTKs of the Src-subfamily. The PTK2-primer targets to amplify PTKs of the ZAP70/Syk subfamily in addition to other PTKs. The PTK1-, PTK2-, and PTK3-primers are mixtures of 128, 64, and 256 different oligonucleotide sequences, respectively. Each oligonucleotide primer was additionally disigned to include a restriction enzyme-specific linker flanking the 5' end of the consensus PTK suquences facilitating the subsequent directional cloning of the amplified fragments into pBluescript.

DNA Amplification

Fig. 4 shows the PCR-amplified product of cDNA derived from reverse-transcription of rat brain mRNA. After 30 cycles of amplification, an obvious eth-

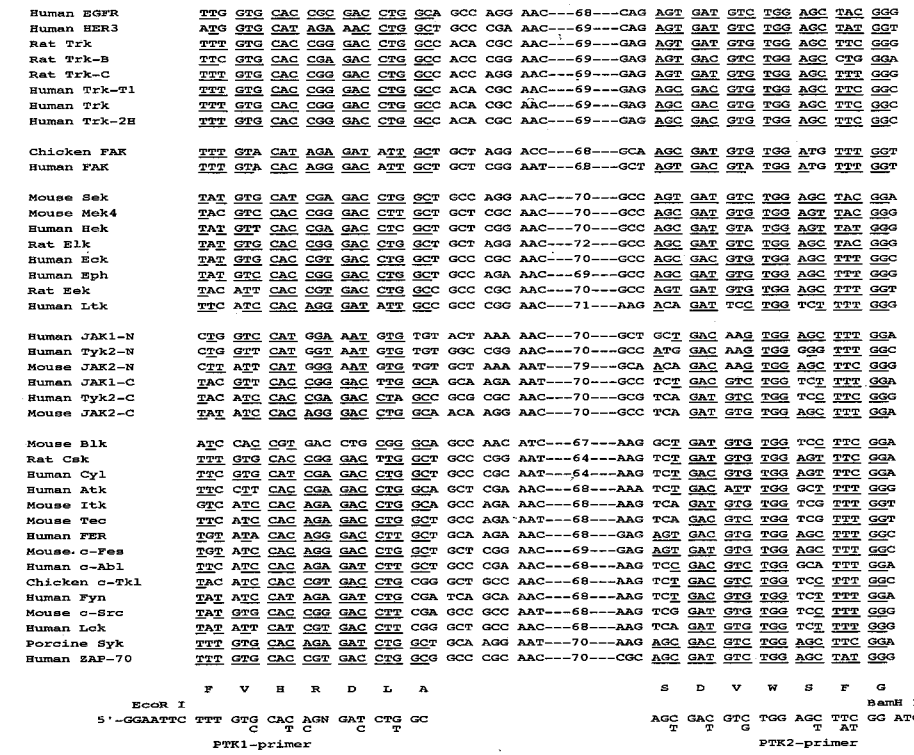

Fig. 2 Derivation of consensus nucleotide sequences for PCR primers. The nncleotide sequences of various PTKs are aligned in the regions of the PTK I (subdomain VI) and PTK II (subdomain IX) motifs. Sequences were taken from the folowing sources: human EGFR (Nature 1984, 309: 418-), human HER3 (Proc Natl Acad Sci USA 1989, 86: 9193-, ibd. 1990, 87: 4905-) , rat Trk (Proc Natl Acad Sci USA 1992, 89: 2374-), rat Trk-B (Mol Cell Biol 1991, 11: 143-), human Trk-T1 (Oncogene 1992, 7: 237-), human Trk (Nature 1986, 319: 743-), human Trk-2H (EMBO J 1988, 7: 147-), chicken FAK (Proc Natl Acad Sci USA 1992, 89: 5192-), human FAK (Biochem Biophys Res Commun 1993, 190: 140-), mouse Sek (Oncogene 1992, 7: 2499-), mouse Mek4 (New Biol 1991, 3: 769-, and EMBL M68513), human Hek (Proc Natl Acad Sci USA 1992, 89: 1611-), rat Elk (Mol Cell Biol 1991, 11: 2496-), human Eck (Mol Cell Biol 1990, 10: 6316-), human Eph (Science 1987, 238: 1717-), rat Eek (Oncogene 1991, 6: 1057-), human Ltk (Oncogene Res, 1990, 5: 199-), human JAKI (Mol Cell Biol 1991, 11: 2057- and Oncogene 1992, 7: 895-), human Tyk2 (Oncogene 1990, 5: 1329- and Cell 1992, 70: 313-), mouse JAK2 (Oncogene 1992, 7: 1347-), mouse Blk (Science 1990, 247: 332-), rat Csk (Nature 1991, 351: 69-), human Cyl (Oncogene 1991, 6: 2013-), human Atk (Nature 1993, 361: 226-), mouse Itk (Proc Natl Acad Sci USA 1992, 89: 11194-), mouse Tec (Oncogene 1990, 5: 1781-), human FER (Mol Cell Biol 1989, 9: 1587-), mouse Fes (Oncogene 1988, 3: 289-), human Abl (Cell 1986, 47: 277-), chicken Tkl (Proc Natl Acad Sci USA 1987, 84: 8778-), human Fyn (Proc Natl Acad Sci USA 1986, 83: 5459-), mouse Src (Science 1987, 237: 411-), human Lck (Eur J Immunol 1986, 16: 1643- and GenBank HUMLCKAA), porcine Syk (J Biol Chem 1991, 266: 15790-), human Zap-70 (Cell 1992, 71: 649-). Degenerate oligonucleotide sequences employed in the PCR amplification of PTK sequences are shown at the bottom and are designated PTK1- and PTK2-primers for the formard and the reverse primers, respectively; alternative nucleotides are shown below the sequence. The symbol N represents positions where all four nucleotides are present in an oligonucleotide primer. The PTK2-primer sequence was reversed and complemented before synthesis of the corresponding oligonucleotide. Restriction sites for EcoRI and BamHI have been built into the PTK1 and PTK2 oligonucleotides, respectively. The nucleotides underlined are those that are identical to sequences in the PTK1- and PTK2-primers. Amino acid Sequences are shown in one-letter code above the sequences of the oligonucleotide primers. The numbers of amino acids between the residues encoded by the first nucleotide at the left and the last nucleotide at the right are indicated at the middle of each nucleotide sequence. -N and -C attached to human JAK1, human Tyk 2 and mouse JAK2 indicate the sequences of the kinase domains at the N-terminal and the C-terminal sides.

```
Human EGFR        GGC AAA GTG CCT ATC AAG TGG ATG GCA TTG GAA TCA ATT TTA
Human HER3        GCC AAG ACT CCA ATT AAG TGG ATG GCC CTT GAG AGT ATC CAC
Human FAK         GGA AAA TTG CCT ATT AAA TGG ATG GCT CCA GAG TCA ATC AAT
Rat Trk           ACC ATG CTG CCC ATC CGC TGG ATG CCT CCA GAG AGC ATC CTC
Rat Trk-B         ACA ATG TTG CCC ATC CGA TGG ATG CCT CCA GAG AGC ATC ATG
Rat Trk-C         ACC ATG CTC CCC ATC CGC TGG ATG CCA CCT GAA AGC ATC ATG
Human Trk-T1      ACC ATG CTG CCC ATT CGC TGG ATG CCG CCC GAG AGC ATC CTG
Human Trk         ACC ATG CTG CCC ATT CGC TGG ATG CCG CCC GAG AGC ATC CTG
Human Trk-2H      ACC ATG CTG CCC ATT CGC TGG ATG CCG CCC GAG AGC ATC CTG
Human Ltk         GCC TTG CTC CCA GTC AAG TGG ATG CCC CCA GAG GCC TTC CTG

Mouse Sek         GGC AAG ATT CCT ATC CGG TGG ACT GCG CCA GAA GCA ATT GCG
Mouse Mek4        GGA AAG ATA CCA ATA AGG TGG ACA TCA CCA GAA GCA ATG TCC
Human Hek         GGG AAG ATC CCA ATC AGG TGG ACA TCA CCA GAA GCT ATA GCC
Rat Elk           GGG AAG ATA CCT GTT AGA TGG ACA GCT CCA GAG GCC ATC GCC
Human Eck         GGC AAG ATC CCC ATC CGC TGG ACC GCC CCG GAG GCC ATT TCC
Human Eph         GGA AAG ATC CCT ATC CGT TGG ACA GCC CCT GAA GCC ATT GCC
Rat Eek           GGG AAG ATC CCT ATC CGA TGG ACA GCA CCA GAG GCC ATT GCC

Human JAK1-N      TGC ATT GAA CGA ATC CCA TGG ATT GCT CCT GAG TGT GTT GAG
Human Tyk2-N      CGG GTG GAG AGG ATC CCC TGG CTG GCC CCC GAA TGC CTA CCA
Mouse JAK2-N      CTT CAG GAG AGA ATA CCA TGG GTA CCA CCT GAG TGC ATT GAG
Human JAK1-C      CGG GAC AGC CCT GTG TTT TGG TAT GCT CCA GAA TGT TTA ATG
Human Tyk2-C      GGG GAC AGC CCC GTG TTC TGG TAT GCC CCA GAG TGC CTG AAG
Mouse JAK2-C      GGG GAA AGC CCC ATA TTC TGG TAC GCA CCT GAA TCC TTG ACG

Chicken Tk1       GCT AAA TTC CCC ATC AAG TGG ACA GCA CCG GAG GCT ATC AAT
Mouse Blk         GCC AAG TTC CCC ATC AAG TGG ACC GCC CCG GAG GCC ATC CAC
Mouse Src         GCC AAA TTC CCC ATC AAG TGG ACC GCC CCT GAA GCT GCT CTG
Human Src         GCC AAA TTC CCC ATC AAG TGG ACG GCT CCA GAA GCT GCC CTC
Mouse Lck         GCC AAA TTT CCC ATT AAG TGG ACA GCA CCA GAA GCC ATT AAC
Human Arg         GCC AAA TTT CCT ATT AAG TGG ACA GCA CCA GAG AGT CTT GCC
Human Hck         GCC AAG TTC CCC ATC AAG TGG ACA GCT CCT GAA GCC ATC AAC
Human Lyn         GCT AAG TTC CCT ATT AAG TGG ACG GCT CCA GAA GCA ATC AAC
Human Yes         GCA AAA TTT CCA ATC AAA TGG ACA GCT CCT GAA GCT GCA CTG
Human Fyn         GCA AAG TTC CCC ATC AAG TGG ACG GCC CCC GAG GCA GCC CTG
Human Fgr         TCC AAG TTC CCC ATC AAG TGG ACA GCC CCA GAA GCT GCC CTC
Human FER         AAG CAG ATT CCC ATT AAA TGG ACA GCA CCG GAA GCT CTT AAT
Mouse Fes         AGA CAA GTC CCT GTT AAG TGG ACT GCC CCT GAG GCC CTT AAC
Human Abl         GCC AAG TTC CCC ATC AAA TGG ACT GCA CCC GAG AGC CTG GCC
Human Cyl         GGC AAG CTG CCA GTC AAG TGG ACA GCC CCT GAG GCC CTG AGA
Rat CSK           GGC AAA CTG CCA GTC AAG TGG ACA GCT CCT GAA GCC TTG AGA

Mouse Tec         GCC AAG TTC CCT GTG AAG TGG TGT CCC CCA GAA GTG TTT AAT
Mouse Itk         ACC AAA TTC CCA GTG AAG TGG GCA TCC CCA GAG GTG TTC TCC
Human Atk         TCC AAA TTT CCA GTC CGG TGG TCC CCA CCG GAA GTC CTG ATG
Porcine Syk       GGG AAG TGG CCC GTG AAA TGG TAC GCT CCG GAA TGC ATC AAC
Human ZAP-70      GGG AAG TGG CCG CTC AAG TGG TAC GCA CCC GAA TGC ATC AAC

                                          W  Y  A  P  E  C  I  N
                                                                BamH I
PTK3-primer                   5'-TGG TAC GCN CCN GAA TGT ATC AA GGATCC TG-3'
                                     T           G  C  T
```

Fig. 3 Derivation of a consensus sequence for the PCR reverse primer used to amplify PTKs of the ZAP70/Syk subfamily. The nucleotide sepuences of various PTKs are aligned in the region of the PTK III (subdomain VIII) motif. Sequences were taken from the following sources: human Src (Mol Cell Bio 1985, 5: 1122-), mouse Lck (Cell 1985, 43: 393-), human Arg (Proc Natl Acad Sci USA 1990, 87: 5802-), human Hck (Mol Cell Biol 1987, 7: 2276-), human Lyn (Mol Cell Biol 1987, 7: 237-), human Yes (Mol Cell Biol 1987, 7: 41-), and human Fgr (Mol Cell Biol 1988, 8: 259-). Other sources of the sequences are described in the legend to Fig. 2. Degenerate oligonucleotide sequences used as a PCR reverse primer, PTK3-primer, are shown at the bottom; alternative nucleotides are shown below the sequence. The symbol N represents positions where all for nucleotides are present in an oligonucleotide primer. The amino acid sequence is shown in one-letter code above the sequence of oligonucleotide primer. The nucleotides underlined are those that are identical to sequences in the PTK3-primer. The PTK3-primer sequence was reversed and complemented before synthesis of the corresponding oligonucleotide. Restriction site for BamHI has been built into the PTK3 oligonucleotide. -N and -C attached to human JAK1, human Tyk 2, and mouse JAK2 indicate the sequences of the kinase domains at the N-terminal and C-terminal sides.

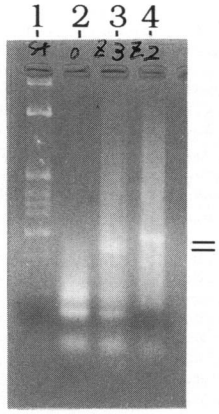

Fig. 4 Gel electrophoresis of amplified PTK sequences. After 30 cycles of amplification, one-tenth aliquot of the PCR product was electrophoresed in a 3% agarose gel and was detected by ethidium bromide staining. Reverse transcription and PCR were performed as described in Materials and Methods. Rat brain total RNA was used as a template. The PCR primers were employed in pairs: PTK1- and PTK3-primers in lanes 2 and 3, or PTK1- and PTK2-primers in lane 4. The DNA size standards in lane 1 were HinfI digestion products of pBR 322. Markers at right point to approx. 220-base-pair and approx. 180-base-pair DNA fragments amplified in the PCR reaction.

idium bromide-stained DNA band of the expected size (about 220 and 180 base pairs with the PTK2-primer, and the PTK3-primer, respectively) was detected. After BamHI/EcoRI digestion of the PCR products, the DNA was cloned into BamHI/EcoRI-cleaved, phosphatase-treated pBluescript II SK (+) DNA. The pBluescript clones were sequenced by the dideoxy method (12), and the nucleotide and encoded protein sequences were compared with known PTK sequences by e-mail service employing blastx and blastn search programs against the GenBank nucleotide sequence database. The deduced protein sequences of representative clones are shown in Fig. 5, along with related known PTK sequences.

Clones amplified by PCR with PTK1- and PTK2-primers

All clones were amplified from rat brain mRNA. TrK-B cDNA (GenBank M55291) was amplified most often by the PCR (data not shown). Trk-C cDNA (GenBank L03813) was also amplified by the PCR (data not shown). Clone Mar9-1 is identical in amino acid sequence to MMMPK6 (GenBank X57240), a 172 base-pair fragment of a mouse cDNA for a putative receptor PTK. The complete nucleotide sequence of the human equivalent of the cDNA was recently described (14). The sequence indicates that the PTK has a discoidin-like extracellular domains. Clones Mar18-4 and Mar19-31 were also very similar to MMMPK6. Clone Brain15 is identical in both nucleotide sequence and the translated amino acid sequence to rat Eek (GenBank X59290), a receptor PTK of the Eph-subfamily, except for the sequence derived from the primers. Clone 3Mar18-1 is identical in amino acid sequence to mouse Sek (GenBank X65138), a brain receptor PTK of the Eph-subfamily. Clone 3Mar10-4 is identical in amino acid sequence to MMM310 TKR (GenBank X69674), a 171 base-pair fragment of a mouse cDNA for a putative receptor PTK of the Eph

```
Mar9-1          FVHRDLATRNCLVGENFTIKIADFG.MSRNLYAG...DYYRVQGRAV.LPIRWMAWECILMGKFTTASDVWSF
Discoidin-PTK   FVHRDLATRNCLVGENFTIKIADFG.MSRNLYAG...DYYRVQGRAV.LPIRWMAWECILMGKFTTASDVWAF
MMMPK6               ATRNCLVGENFTIKIADFG.MSRNLYAG...DYYRVQGRAV.LPIRWMAWECILMGKFTTAS
Mar18-4                           FG.MSRNLYAG...NYYRVQGRAV.LSIRWMAWECILMGKFTTASDVWSF
Mar19-31                    GENFTIKIADFG.MSRNLYAG...DYYRVQVRAV.LPIRWMAWECILMGKFTTASDVWSF
Human Trk-2H    FVHRDLATRNCLVGQGLVVKIGDFG.MSRDIYST...DYYRVGGRTM.LPIRWMPPESILYRKFTTESDVWSF
Human Trk       FVHRDLATRNCLVGQGLVVKIGDFG.MSRDIYST...DYYRVGGRTM.LPIRWMPPESILYRKFTTESDVWSF
Human Trk-T1    FVHRDLATRNCLVGQGLVVKIGDFG.MSRDIYST...DYYRVGGRTM.LPIRWMPPESILYRKFTTESDVWSF
Rat Trk         FVHRDLATRNCLVGQGLVVKIGDFG.MSRDIYST...DYYRVGGRTM.LPIRWMPPESILYRKFSTESDVWSF
Rat Trk-C       FVHRDLATRNCLVGANLLVKIGDFG.MSRDVYST...DYYRVGGHTM.LPIRWMPPESIMYRKFTTESDVWSF
Rat Trk-B       FVHRDLATRNCLVGENLLVKIGDFG.MSRDVYST...DYYRVGGHTM.LPIRWMPPESIMYRKFTTESDVWSL

Mar19-26                                    RVLEDD.LEAGYTTR..GGKIPIRWTAREAIEYRKFTSVSDVWSY
3Mar10-4        FVHXDLAARNILINSNLVCKXSXFX.LSRVLEDD.PEAAYTTR..GGKIPIRWTAPEAIAFRKFTSASDVWSY
MMM310TKR            ARNILINSNLVCKVSDFG.LSRVLEDD.PEAAYTTR..GGKIPIRWTAPEAIAFRKFTSAS
3Mar18-1                          135G.MSRVLEDD.PEAAYTTR..GGKIPIRWTAPEAIAYRKFTSASDVWSY
Mouse Sek       YVHRDLAARNILVNSNLVCKVSDFG.MSRVLEDD.PEAAYTTR..GGKIPIRWTAPEAIAYRKFTSASDVWSY
Human Hek       YVHRDLAARNILINSNLVCKVSDFG.LSRVLEDD.PEAAYTTR..GGKIPIRWTSPEAIAYRKFTSASDVWSY
Human Eck       YVHRDLAARNILVNSNLVCKVSDFG.LSRVLEDD.PEATYTTS..GGKIPIRWTAPEAISYRKFTSASDVWSF
Brain15         FVHRDLAARNILXDGRXVCKVSDFG.LSRALEDD.PEAAYTTA..GGKIPIRWTAPEAIAFRTFSSASDVWDL
Rat Eek         YIHRDLAARNILVDGRLVCKVSDFG.LSRALEDD.PEAAYTTA..GGKIPIRWTAPEAIAFRTFSSASDVWSF
Rat Elk         YVHRDLAARNILVNSNLVCKVSDFG.LSRYLQDDTSDPTYTSSL.GGKIPVRWTAPEAIAYRKFTSASDVWSY
Human Eph       YVHRDLAARNILVNQNLCCKVSDFG.LTR LLDD.FDGTYETQ..GGKIPIRWTAPEAIAHRIFTTASDVWSF

3Mar9-3         FVHSDLAVRNILVASPECVKLGDFG.LSRYIEDE...DYYKAS..VTRLPIKWMSPESINFRRFTTASDVSSY
3Mar18-5                                    RYIEDE...DYYKAS..VTRLPIKWMSPESINFRSFTTASEVWSF
Chicken FAK     FVHRDIAARNVLVSATDCVKLGDFG.LSRYMEDS...TYYKAS.KGK.LPIKWMAPESINFRRFTSASDVWMF
Human FAK       FVHRDIAARNVLVSSNDCVKLGDFG.LSRYMEDS...TYYKAS.KGK.LPIKWMAPESINFRRFTSASDVWMF

22Apr6          FVHSDLAARNILVESEAHVKIADFG.LAKLLPLGK..DYYVVRDPGQ.SPIFWYVPECI
22Apr10         FVHRDLAARNLLLATRDLVKIGDFG.LMRALPQND..DHYVMQEHRK.VPFAWYAPECI
Cdc42Hs-PTK     FIHRDLAARNLLLATRDLVKIGDFG.LMRALPQND..DHYVMQEHRK.VPFAWCAPESL

Porcine Syk     FVHRDLAARNVLLVTQHYAKISDFG.LSKALRADE..NYYKAQTHGK.WPVKWYAPECI
3May49          FVHSDLAARNVLLVNRHYAKISDFG.LSKALGADD..SYYTARSAGK.LPLKWYXPECI
Human ZAP-70    FVHRDLAARNVLLVNRHYAKISDFG.LSKALGADD..SYYTARSAGK.WPLKWYAPECI

Human JAK1      YVHRDLAARNVLVESEHQVKIGDFG.LTKAIETDK..EYYTVKDDRD.SPVFWYAPECL
Mouse JAK2      YIHRDLATRNILVENENRVKIGDFG.LTKVLPQDK..EYYKVKEPGE.SPIFWYAPESL
Human Tyk2      YIHRDLAARNVLLDNDRLVKIGDFG.LAKAVPEGH..EYYRVREDGD.SPVFWYAPECL

24Apr35         FVHRDLAARNFLVWENHVVKVADFG.LSRLITG....DTYTXHS.GAKFPIN
23Apr28         FLHRDLSARNFLEGENHLVKAPDF...SRTITG....DTY
Human c-Abl     FIHRDLAARNCLVGENHLVKVADFG.LSRLMTG....DTYTAHA.GAKFPIKWTAPESL

2May41          FVHSDLALRNCLLTADLTVKVGDYG.LAHCKYRE...D.YLVTADQLWVQLAWY PECI
21Apr15         FVHRDLALRNCYLTSDLNVKVGDYG.IGFSRYKE...D.YIDTDDKKIXXPEWYXPECI
Human IGF1R     FVHRDLAARNCMVAEDFTVKIGDFG.MTRDIYET...DYYRKGGKG.LIPVRWMSPESL
Human INS.R     FVHRDLAARNCMVAHDFTVKIGDFG.MTRDIYET...DYYRKGGKG.LIPVRWMAPESL

22Apr17         FVHRDLASRNVLVKSPNHVKITDFR.LGRLLESRX...KKIQCXPG.KMPIKWYGXXCI
21Apr13         FVHRDLAARNVLVKSPNHVKITDFGTLARLLEGD..EKEYHAD..GGKMPIKWYAPECI
Human EGFR      LVHRDLAARNVLVKTPQHVKITDFG.LAKLLGAEE..KEYHAE..GGKVPIKWMALESI
Rat Neu         LVHRDLAARNVLVKSPNHVKITDFG.LARLLDIDE..TEYHAD..GGKVPIKWMALESI
```

Fig. 5 Comparison of deduced amino acid sequences encoded by clones isolated from the PCR libraries with amino acid sequences of known PTKs. Examples of each species of clones isolated from the PCR libraries are aligned to maximize homology. The conventinal one-letter amino acid code is used; X denotes unidentified amino acid due to an inaccuracy in reading base sequence. Conserved residues are underlined. Dots correspond to gaps introduced to optimize sequence alignment, and each dot represents one amino acid. Sequences were taken from the following source: discoidin-PTK (14), MMMPK6 (4 and GenBank X57240), MMM310TKR (GenBank X69674), Cdc42Hs-PTK (10), human IGF1R (EMBO J 1986, 5: 2503-), human INS. R (Nature 1985, 313: 756-), and rat Neu (GenBank X03362 and X03363). Other sources of the sequences are described in the legends to Figs. 2 and 3.

-subfamily. Clone Mar19-26 is similar but not identical in amino acid sequence to Sek PTK and may represent a new recepotr PTK of the Eph-subfamily. Clones 3Mar9-3 and 3Mar18-5 are similar to each other and were most related to focal adhesion kinase (FAK) in the homology search; the homology to FAKs is not high and may represent a new PTK.

Clones amplified by PCR with PTK1- and PTK3-primers

The PTK3-primer used in the PCR was designed to amplify PTKs with a PTK III sequence of -WYAPEC-, which is characteristic of PTKs of ZAP70/Syk and JAK subfamilies. All clones are amplified from rat brain mRNA. Clone 3May49 was identical in the translated amino acid sequence to human ZAP-70. It remains unknown whether the clone was derived from mRNA of neuronal cells or mRNA of brain lymphocytes. Clone 22Apr10 was isolated 7 times, has a unique sequence of -AWXAPE-, and has no closely related PTKs at the time of sequencing; PSKs of JAK and ZAP70/Syk subfamilies gave the highest score in the homology search on the clone. A complete coding sequence of the clone was recently described (10) ; the results indicated that the clone represents cDNA of a novel non-receptor PTK with specific affinity to activated Cdc42Hs, a Ras-related Rho subfamily of GTP/GDP-binding proteins. Clone 22Apr6 was isolated 3 times and had no closely related PTKs in the database; PTKs of the JAK subfamily gave the highest score in the homology search on the clone. No sequence with high homology score was found for clones 2May41, 21Apr15, 24Apr35, 23Apr28, 22Apr17, and 21Apr13. For clones 2May41 and 21Apr15, insulin receptor PTK and insulin like growth factor receptor PTK gave the highest homology score. Abl PTK gave the highest score in the homology search on clones 24Apr35 and 23Apr28. For clones 22Apr 17 and 21Apr13, EGFR PTK gave the highest homology score.

ACKNOWLEDGMENTS

We thank Drs. Hiroyuki Tatsumi and Masaho Ishino for making e-mail service available to us from our personal computer. We also thank Prof. Kei Fujinaga for kindly allowing us to use DNA Thermal Cycler and Ms. Sachie Hirai for the synthesis of oligonucleotides.

REFFRENCES

1. WILKS AF. Two putative protein-tyrosine kinases identified by application of the polymerase chain reaction. Proc Natl Acad Sci USA 1989, 86: 1603-1607.
2. WILKS AF, KURBAN RR, HOVENS CM, RALPH SJ. The application of the polymerase chain reaction to cloning members of the protein tyrosine kinase family. Gene 1989, 85:

67-74.

3. HOWARD OMZ, DEAN M, YOUNG H, RAMSBURG M, TURPIN JA, MICHIEL DF, KELVIN DJ, LEE L, FARRAR WL. Characterization of a class 3 tyrosine kinase. Oncogene 1992, 7: 895-900.
4. GILARDI-HEBENSTREIT P, NIETO MA, FRAIN M, MATTÉI M-G, CHESTIER A, WILKINSON DG, CHARNAY P. An Eph-related receptor protein tyrosine kinase gene segmentally expressed in the developing mouse hindbrain. Oncogene 1992, 7: 2499-2506.
5. MASIAKOWSKI P, CARROLL RD. A novel family of cell surface receptors with tyrosine kinase-like domain. J Biol Chem 1992, 267: 26181-26190.
6. HOVENS CH, STACKER SA, ANDRES A-C, HARPUR AG, ZIEMIECKI A, WILKS AF. RYK, a receptor tyrosine kinase-related molecule with unusual kinase domain motifs. Proc Natl Acad Sci USA 1992, 89: 11818-11822.
7. BIESECKER LG, GOTTSCHALK LR, EMERSON SG. Identification of four murine cDNAs encoding putative protein kinases from primitive embryonic stem cells differentiated in vitro. Proc Natl Acad Sci USA 1993, 90: 7044-7048.
8. CHAN AC, IWASHIMA M, TURCK CW, WEISS A. ZAP-70: A 70 kd protein-tyrosine kinase that associates with the TCR ζ chain. Cell 1992, 71: 649-662.
9. TANIGUCHI T, KOBAYASHI T, KONDO J, TAKAHASHI K, NAKAMURA H, SUZUKI J, NAGAI K, YAMADA T, NAKAMURA S, YAMAMURA H. Molecular cloning of a porcine gene *syk* that encodes a 72-KDa protein-tyrosine kinase showing high susceptibility to proteolysis. J Biol Chem 1991, 266: 15790-15796.
10. MANSER E, LEUNG T, SALIHUDDIN H, TAN L, LIM L. A non-receptor tyrosine kinase that inhibits the GTPase activity of $p21^{cdc42}$. Nature 1993, 363: 364-367.
11. MACDONALD RJ, SWIFT GH, PRZYBYLA AE, CHIRGWIN JM. Isolation of RNA using guanidium salts. Methods in Enzymol 1987, 152: 219-227.
12. SANGER F, NICKLEN S, COULSON AR. DNA sequencing with chain-terminating inhibitors. Proc Natl Acad Sci USA 1977, 74: 5463-5467.
13. HANKS SK, QUINN AM. Protein kinase catalytic domain sequence database: Identification of conserved features of primary structure and classification of family members. Methods in Enzymol 1991, 200: 38-62.
14. JOHNSON JD, EDMAN JC, RUTTER WJ. A receptor tyrosine kinase found in breast carcinoma cells has an extracellular discoidin I-like domain. Proc Natl Acad Sci USA 1993, 90: 5677-5681.
15. HANKS SK, QUINN AM, HUNTER T. The protein kinase family: Conserved features and deduced phylogeny of the catalytic domains. Science 1988, 241: 42-
16. VETRIE D, VOŘECHOVSKÝ I, SIDERAS P, HOLLAND J, DAVIES A, FLINTER F, HAMMARSTRÖM L, KINNON C, LEVINSKY R, BOBROW M, SMITH CIE, BENTLEY DR. The gene involved in X-linked agammaglobulinaemia is a member of the *src* family of protein-tyrosine kinases. Nature 1993, 361: 226-233.
17. MANO H, ISHIKAWA F, NISHIDA J, HIRAI H, TAKAKU F. A novel protein-tyrosine kinase, *tec*, is preferentially expressed in liver. Oncogene 1990, 5: 1781-1786.
18. SILICIANO JD, MORROW TA, DESIDERIO SV. *itk*, a T-cell specific tyrosine kinase gene inducible by interleukin 2. Proc Natl Acad Sci USA 1992, 89: 11194-11198.