MOLECULAR AND CELLULAR BIOLOGY, Dec. 1990, p. 6316–6324
0270-7306/90/126316-09$02.00/0

Vol. 10, No. 12

cDNA Cloning and Characterization of *eck*, an Epithelial Cell Receptor Protein-Tyrosine Kinase in the *eph/elk* Family of Protein Kinases

RICHARD A. LINDBERG[1,2]* AND TONY HUNTER[1]

Molecular Biology and Virology Laboratory, The Salk Institute for Biological Studies, P.O. Box 85800, San Diego, California 92186-5800,[1] and Department of Biology, University of California, San Diego, La Jolla, California 92093[2]

Received 25 June 1990/Accepted 6 September 1990

A human epithelial (HeLa) cDNA library was screened with degenerate oligonucleotides designed to hybridize to highly conserved regions of protein-tyrosine kinases. One cDNA from this screen was shown to contain a putative protein-tyrosine kinase catalytic domain and subsequently used to isolate another cDNA from a human keratinocyte library that encompasses the entire coding region of a 976-amino-acid polypeptide. The predicted protein has an external domain of 534 amino acids with a presumptive N-terminal signal peptide, a transmembrane domain, and a cytoplasmic domain of 418 amino acids that includes a canonical protein-tyrosine kinase catalytic domain. Molecular phylogeny indicates that this protein kinase is closely related to *eph* and *elk* and that this receptor family is more closely related to the non-receptor protein-tyrosine kinase families than to other receptor protein-tyrosine kinases. Antibodies raised against a TrpE fusion protein immunoprecipitated a 130-kDa protein that became phosphorylated on tyrosine in immune complex kinase assays, indicating that this protein is a bona fide protein-tyrosine kinase. Analysis of RNA from 13 adult rat organs showed that the *eck* gene is expressed most highly in tissues that contain a high proportion of epithelial cells, e.g., skin, intestine, lung, and ovary. Several cell lines of epithelial origin were found to express the *eck* protein kinase at the protein and RNA levels. Immunohistochemical analysis of several rat organs also showed staining in epithelial cells. These observations prompted us to name this protein kinase *eck*, for epithelial cell kinase.

Protein-tyrosine kinases (PTKs) play a key role in the regulation of cellular activity. They are involved in transducing signals that act to modify or change these activities. It seems clear that most or all PTKs function in pathways in which information is being passed across the cell membrane from the cell's external environment. The pathways are ill defined thus far, but it is clear that PTKs can modulate the proliferation or differentiation status of the cell. The growth factor receptors, such as the epidermal growth factor (EGF) receptor, exemplify this (see references 28 and 34 for recent reviews). These molecules span the cell membrane, which divides them into extracellular and cytoplasmic domains. Binding of a ligand to the extracellular domain results in activation of a catalytic domain inside the cell, thus effecting transmembrane signaling. This activation elicits an array of changes in cellular activities.

PTKs are also important because of their role in malignant transformation. Many were first identified as the products of oncogenes and still constitute the largest family of known oncogenes (2). A number of PTKs have been transduced by retroviruses, but more importantly, several have been shown to cause cell transformation or are implicated in human malignancies without the intervention of retroviruses. HER 2/*neu* is an example of this; it was first identified as an oncogene in rat tumor cells, in which it was activated by a point mutation in its transmembrane domain (1), and subsequently it has been shown that HER 2/*neu* overexpression is correlated with progression of human breast and ovarian cancers (25).

The number of known PTKs is approaching 50 and will undoubtedly continue to grow. PTKs can be divided into several families on the basis of sequence and overall structure. These are the *src*, *fps*, and *abl* families (the cytoplasmic PTKs) and four families of receptor PTKs. The receptor family prototypes are the EGF receptor (27), insulin receptor (26), platelet-derived growth factor (PDGF) receptor (this family includes the colony-stimulating factor 1 and basic fibroblast growth factor receptors) (33), and the *eph* protein (11). Included among the receptors are a large number of presumptive receptor PTKs with overall structures closely related to the bona fide receptor protein-tyrosine kinases. The number and composition of PTK families are continuing to evolve.

All PTKs have a conserved catalytic domain of about 270 amino acids that is responsible for their enzymatic activity (9). This fact has been used in developing strategies to isolate new PTK sequences, such as low-stringency hybridization screening of cDNA libraries. Within this domain there are short stretches of highly conserved residues whose coding regions are ideally suited for the design of degenerate oligonucleotide probes to be used to identify new PTKs. This strategy was first used for protein-serine/threonine kinases (8) and subsequently for PTKs (12). Other successful strategies include screening cDNA libraries with oligonucleotides deduced from protein sequences (e.g., reference 26), transfection-transformation assays (e.g., reference 1), polymerase chain reaction (e.g., reference 30), and screening of expression libraries with antibodies to phosphotyrosine (14, 17, 18, 23).

Given the prominent role of these enzymes in controlling cell proliferation and their frequent oncogenic activation, it

* Corresponding author.

is important to identify new members of the family. Each cell type probably expresses its own specific repertoire of PTKs, and to understand their physiological function fully, one must identify all of the components in the regulatory network. Epithelial cells are a good system for the study of such pathways because many exist as stem cells that are in a constant equilibrium between a resting and a proliferating-differentiating state. This sensitive equilibrium may contribute to the propensity of this type of cell to become cancerous (~80% of human neoplasias), and the growth factor receptors involved are candidate oncogene products. For these reasons, we decided to identify new receptorlike PTKs in epithelial cells and have found one which we have named *eck*, for epithelial cell kinase.

MATERIALS AND METHODS

Isolation and characterization of cDNA clones. Degenerate oligonucleotide probes were designed to hybridize to regions that were known to be highly conserved in receptorlike PTKs by analysis of protein kinase catalytic domain sequence alignments. The criteria for design of the probes were as follows: the regions chosen were selected to identify receptorlike PTK cDNAs and discriminate against protein-serine/threonine kinases, and the degeneracy was kept as low as possible by use of codon usage tables (7) to eliminate infrequently used codons and was then reduced below 200 by use of known PTK codon usage. The HRDLAAR probe was 5'-C(GT)(TAG)GC(TAGC)GCCA(GA)GTC(TAGC)CG GTG-3'. This probe will not hybridize to *src* family PTK sequences, which have the sequence HRDLRAA in this region. The P(I/V)KW(T/M)APE probe was 5'-TC(GA)GG (TAG)GC(CG)(AG)TCCACTT(CG)A(TC)(AG)GG-3'. These probes were used to hybridize to 60,000 colonies from a HeLa cell Okayama/Berg cDNA library exactly as previously described by Hanks (8). Wash temperatures were increased from 42 to 56°C stepwise, with autoradiography performed on the filters at each step. Clones that were positive with either probe were picked, and plasmid DNA was prepared (see Results for numbers of clones). The plasmids were cleaved with restriction enzymes, followed by Southern analysis with the two probes described above. Fragments that hybridized to the probes were subcloned into the M13 sequencing vector (32) and sequenced by the dideoxy-chain termination method (24). Clones that contained canonical PTK catalytic domain sequences were studied further. A single-strand probe encompassing nucleotides 1681 to 2007 (Fig. 1) of one of these clones was synthesized from the M13 sequencing vector, using the sequencing primer and [γ-^{32}P]dCTP, and used to screen 500,000 clones from the same HeLa cell library. A probe from the clone with the longest insert containing nucleotides 731 to 1031 (Fig. 1) was used to screen a human keratinocyte λgt11 cDNA library (Clontech Laboratories, Inc., Palo Alto, Calif.). The λ clone that extended the farthest 5', which was 3.7 kb in length, contributed the first 1,681 nucleotides to the final *eck* sequence, and the original OB18 insert contributed the rest. Computer analyses were performed on the sequence as described previously (9).

Northern (RNA) and Southern blot analyses. For Northern analysis, the probe used was the 267-bp *Sau*3A fragment (nucleotides 2180 to 2447; Fig. 1). RNA was prepared by the guanidinium isothiocyanate lysis method described by Chomczynski and Sacchi (4). Equivalent amounts of whole-cell RNA as judged by A_{260} were loaded onto gels. The 28S and 18S rRNAs were stained with methylene blue after transfer to Nytran (Schleicher & Schuell, Keene, N.H.) to ensure that the RNAs were not degraded and were present in each sample in approximately equal quantities. The RNAs were resolved in formaldehyde gels, transferred, and hybridized by standard procedures (19). Final washes of Northern blots were at 65°C for human RNAs and 42°C for rodent RNAs in 0.2× SSPE. DNA preparation and Southern blot analysis were done as described previously (19).

Antibodies. The rabbit antiserum to *eck* was raised against a TrpE fusion protein made as follows. The 299-bp *Alu*I-*Bam*HI fragment comprising nucleotides 2733 to 3032 was ligated into the pATH1 expression vector (3). The resulting fusion protein, which contained amino acids 874 to 974 of *eck*, was isolated by polyacrylamide gel electrophoresis and used as an antigen to immunize New Zealand White rabbits as described previously (10). Several bleeds from two rabbits gave identical results in immunoprecipitation experiments. The control antiserum used for immunostaining was raised against a TrpE-OB19 fusion protein in a similar manner. OB19 is another PTK identified in the same screen of the HeLa cell library (12).

Labeling and immunoprecipitation of *eck*. A431 human epidermoid carcinoma cells (obtained from Jill Meisenhelder, Salk Institute) were grown in Dulbecco modified Eagle medium with 10% fetal bovine serum. To label proteins with ^{35}S, medium was changed to methionine- and cysteine-free medium containing 5% dialyzed fetal bovine serum and 100 μCi of Tran^{35}Slabel per ml (~1,000 Ci/mmol; ICN Radiochemicals, Irvine, Calif.) for 3 h at 37°C. Cells from a 75% confluent 10-cm-diameter dish were lysed in 2 ml of cold RIPA buffer (6). Then 500 μl of lysate was clarified by centrifugation at 29,000 × *g* for 45 min at 4°C after the addition of 250 μl of Pansorbin (Calbiochem, San Diego, Calif.). *eck* protein was immunoprecipitated by the addition of 2 μl of rabbit antiserum, followed 30 min later by 30 μl of a 50% slurry of protein A-Sepharose (Pharmacia, Piscataway, N.J.). The immunoprecipitates were washed four times in RIPA buffer and boiled for 2 min in 30 μl of 2× sample buffer, and 10 μl was fractionated on 7.5% sodium dodecyl sulfate (SDS)-polyacrylamide gels.

In vitro kinase reactions were carried out with unlabeled immunoprecipitates in a mixture containing 25 mM *N*-2-hydroxyethylpiperazine-N'-2-ethanesulfonic acid (HEPES) (pH 7.4), 10 mM $MnCl_2$, 1 mM EDTA, 0.1% Triton X-100, and 10 μCi of [γ-^{32}P]ATP (4,500 Ci/mmol; ICN) at 30°C for 10 min. The reaction was stopped by boiling in sample buffer and analyzed on 7.5% SDS-polyacrylamide gels. Phosphoamino acid analyses were done as described by Hunter and Sefton (13).

Immunohistochemistry. Tissue localization of the *eck* protein was done on 4-μm unfixed frozen sections with affinity-purified antibodies. Anti-*eck* antibodies were purified by binding to and elution from the TrpE-*eck* fusion protein that had been blotted to nitrocellulose as described previously (10). Controls were anti-TrpE-OB19 and preimmune immunoglobulin G isolated by incubation with and elution from the TrpE-*eck* fusion protein in the same manner. The staining procedure followed was essentially as described previously (10). Briefly, the sections were air dried, hydrated, incubated with antibodies for 1 h, washed for 10 min, incubated with horseradish peroxidase-conjugated goat anti-rabbit immunoglobulin G (Cappel Laboratories, Cochranville, Pa.), washed for 10 min, stained with diaminobenzidine (Polysciences Inc., Warrington, Pa.) for 3 min, washed for 10 min, and counterstained with hematoxylin.

Nucleotide sequence accession number. The sequence re-

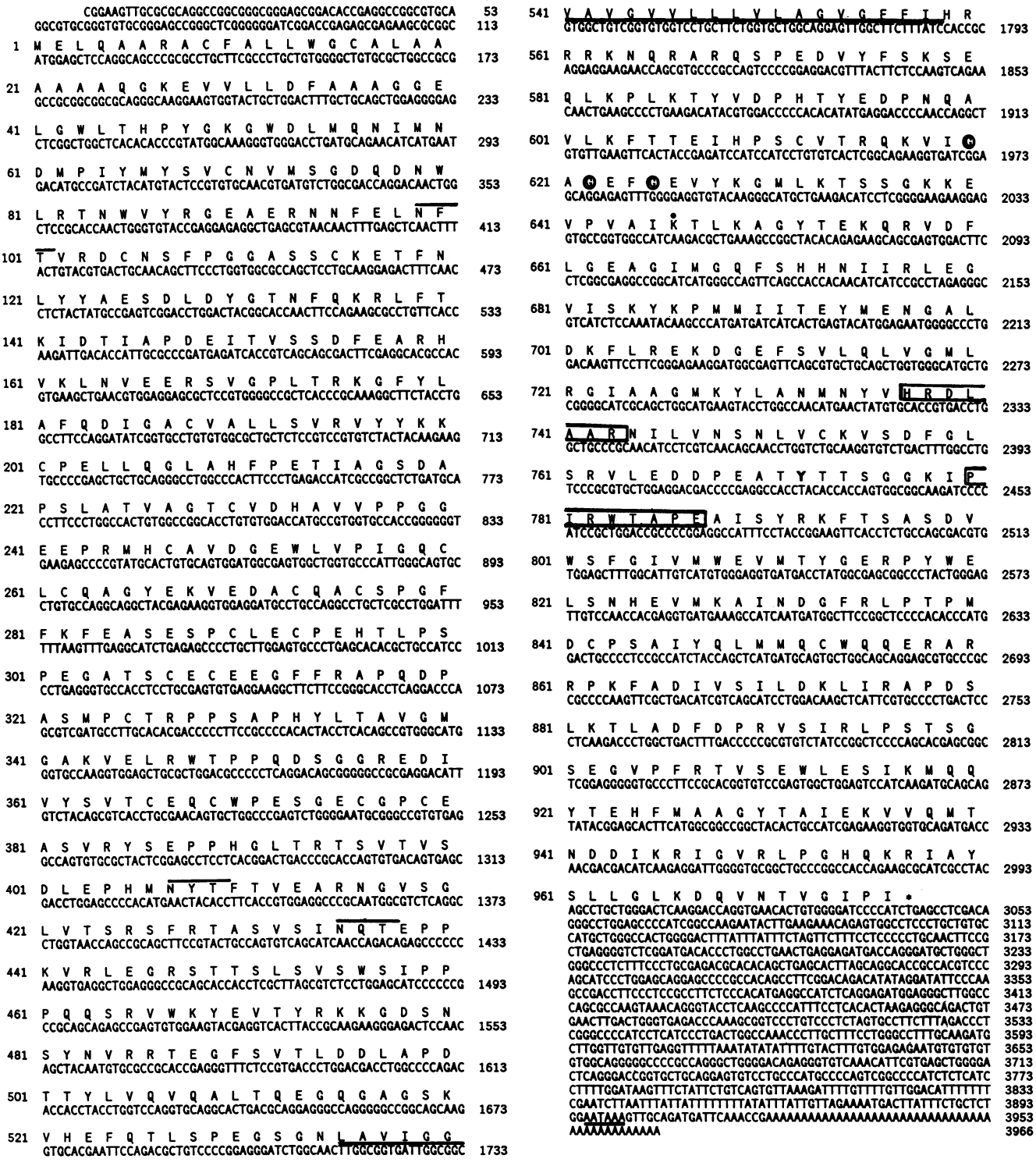

FIG. 1. Complete nucleotide and amino acid sequences of the *eck* cDNA. Amino acids are numbered to the left, starting with the initiation codon; nucleotides are numbered to the right. Potential N-linked glycosylation sites are overlined. The 24 amino acids comprising the transmembrane domain are underlined. The circled glycines and the dotted lysine represent conserved residues of the ATP-binding site. The boxed residues are the conserved regions flanking the putative autophosphorylation site (the bold tyrosine) to which the oligonucleotide probes were designed. An AATAAA polyadenylation signal in the 3′ untranslated region is underlined.

ported here has been deposited in the DDBJ, EMBL, and GenBank data banks under accession number M36395.

RESULTS

Isolation and characterization of cDNA clones. A HeLa cell Okayama/Berg cDNA library was screened with degenerate oligonucleotides that were designed to hybridize to regions that code for highly conserved motifs in the catalytic domains of PTKs. Amino acids chosen to be encoded by the probes were HRDLAAR and P(I/V)KW(T/M)APE. The former was designed to identify receptor PTKs rather than *src* family cDNAs, whereas the latter could potentially recognize all PTK families. These regions were chosen because they are conserved in PTKs yet are divergent enough from protein-serine/threonine kinases to avoid cross-hybridization. Forty clones were picked that were positive with either or both probes. Sequence analysis was performed on inserts of the plasmids from 12 clones. Two of these coded for putative PTKs on the basis of having canonical sequences: OB18, which encodes a portion of *eck*, and another clone, which we have called OB19 (12). Both clones were isolated by hybridization with the HRDLAAR probe (in our experience and that of others, the HRDLAAR probe seems to be particularly good at selecting PTK sequences). OB18 had a perfect match with one of the HRD LAAR oligonucleotide isomers and was negative with the other probe, which had a minimum of four mismatches. This clone was used to screen 500,000 clones of the HeLa cell library. The clone with the longest insert was isolated, sequenced, and found to contain 3,555 bp of *eck* cDNA, including a poly(A) tail. This cDNA had a long open reading frame but did not contain the entire coding region of *eck*; to find cDNAs that extended further 5′, a probe from the 5′ end of this clone was used to screen a human λgt11 keratinocyte cDNA library. The λ clone whose insert extended the farthest 5′ had a 3.7-kb insert and proved to be identical to OB18 in overlapping regions by restriction mapping and partial sequence data. Sequence data from the λgt11 and OB18 inserts were combined to yield a sequence that encompasses the entire coding region (Fig. 1).

***eck* encodes a new receptor PTK in the *eph/elk* family.** The composite *eck* cDNA has an open reading frame encoding 976 amino acids flanked by a 5′ untranslated region of 113 nucleotides and 925 nucleotides in the 3′ untranslated region (Fig. 1). The first ATG is in a context that follows Kozak's rules (15) and is followed by what appears to be a cleavable signal sequence that includes 17 hydrophobic amino acids (29). There is a classical transmembrane domain consisting of 24 hydrophobic amino acids followed by a basic stop-transfer motif. This domain divides the molecule into (i) an external domain that consists of 534 amino acids and has three potential signals for N-linked glycosylation and (ii) a cytoplasmic domain that contains a canonical PTK catalytic domain. All consensus sequences (9) in PTKs are present in *eck*, including the ATP-binding site (G-620–X–G-622–X–X–G-625 and K-646), the probe regions, and a putative autophosphorylation site at Y-772. A C-terminal tail of 98 amino acids follows the catalytic domain. The 3′ untranslated region has 925 nucleotides including the poly(A) tail, which is preceded by a polyadenylation signal 20 nucleotides upstream.

The *eck*, *eph* (11), and *elk* (17) catalytic domain sequences were entered into a PTK data base, and a phylogenetic tree was constructed; the tips of the branches correspond to individual PTKs and are for simplicity denoted only by family groupings (Fig. 2). These three PTKs form a family branch separate from the other families and interestingly appear to be more closely related to the nonreceptor PTKs than to the receptor PTKs. The identities of the *eck* catalytic domain with other PTKs are as follows: *elk*, 70%; *eph*, 60%; c-*src*, 43%; c-*abl*, 39%; c-*fes*, 39%; EGF receptor, 36%; insulin receptor, 36%; and PDGF receptor, 34%. *elk* is a rat sequence, PDGF receptor is from mouse cells, and the remaining sequences are from human cells. These data indicate that *elk*, *eph*, and *eck* represent a new family of receptor PTKs.

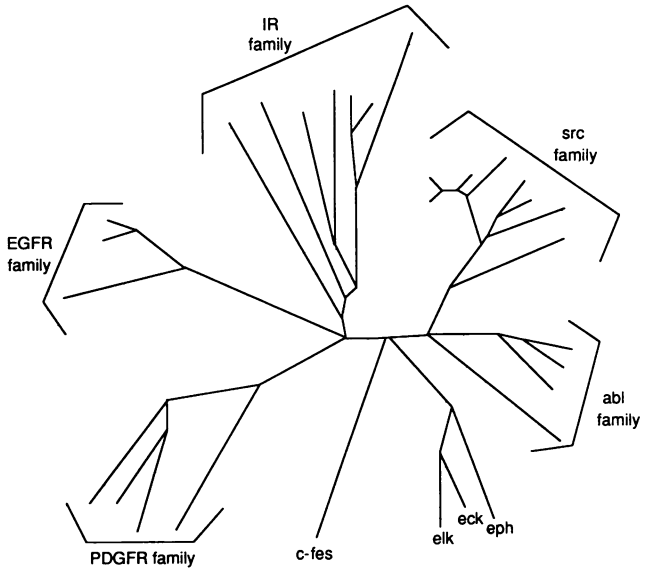

FIG. 2. Molecular phylogeny based on the sequence of PTK catalytic domains. The phylogenetic tree was constructed exactly as described by Hanks et al. (9). The tips of the branches represent individual PTKs and correspond to those depicted previously (9). Three new sequences were added to the original data base: *eph*, *elk*, and *eck*. IR, Insulin receptor; EGFR, EGF receptor; PDGFR, PDGF receptor.

In addition to having sequence homology, *eck* and *eph* (the only two family members whose complete sequences are known) show similarities in domain topology (Fig. 3B). Seventeen cysteine residues in their external domains are conserved and clustered in a region (C-187–C-379; Fig. 3A) that could be considered a cysteine-rich box (11), but the cysteine spacing is distinct from that in the EGF and insulin receptor boxes. This separates them from the other receptor families; the EGF receptor family has two cysteine-rich boxes, the insulin receptor family has a heterotetrameric disulfide-linked structure with one cysteine-rich box, and the PDGF receptor family, which includes the fibroblast growth factor receptor (16), has immunoglobulinlike repeats (see reference 28 for a review). The *eph* and *eck* proteins also have similar positionings of their initiating methionines and transmembrane domains, and their C-terminal tails are of similar lengths (Fig. 3A). The identities between the two proteins in the external domain and the C-terminal tail are 44 and 54%, respectively. Comparison of overall structure with that of other family members will have to await reporting of full-length sequences. Genomic Southern analysis suggests that *eck* is a single-copy gene but that there are many related genes. *Pst*I-digested human DNA, when probed with the 267-bp *Sau*3A catalytic domain fragment described in Materials and Methods, revealed 13 bands when washed at 50°C

A

```
eck MELQAARACFALLWGCALAAAAAAQGKEVVLLDFAAAGGELGWLTHPYGKGWDLMQNIMNDMPIYMYS   68
eph --RRWPLGLGLV-LL--PLPPGALRA---T-M-TSK-Q------LE-PKE--SEQEQ-L-GT-L---Q   67

eck VCNVMSG.DQDNWLRTNWVYRGEAE.RNNFELNFTVRDCNSFPGGASS..CKETFNLYYAESDLDYGT  132
eph D-PMQGRR-TEH---S--I----EAS-VHV--QF-----K------GPLG----P--L-M---Q-V-I  135

eck NFQKRLFTKIDTIAPDEITVSSDFEARHVKLNVEERSVGPLTRKGFYLAFQDIGACVALLSVRVYYKK  200
eph QLRRP--Q-VT-V-A-QSFTIR-LASGS------RC-L-R---R-L----HNP------V----F-QR  203

eck CPELLQGLAHFPETIAGSDAPSLATVAGTCVDHAVV.PPGGEEPRMHCAVDGEWLVPIGQCLCQAGYE  267
eph ---T-N---Q--D-LP-PAG..-VE-----LP--RAS-RPSGA-----SP-------V-R-H-EP---  269

eck K..VEDACQACSPGFFKFEASESPCLECPEHTLPSPEGATSCECEEGFFRAPQDPASMPCTRPPSAPH  333
eph EGGSGE--V--PS-SYRMDMDTPH--T--QQSTAES----I-T--S-HY---GEGPQVA--G-----R  337

eck YLTAVGMGAKVELRWTPPQDSGGREDIVYSVTCEQCWP...ESGECGPCEASVRYSEPPHGLTRTSVT  398
eph N-SFSAS-TQLS---E--A-T---QDVR---R-S--QGTAQDG-P-Q--GVG-HF-PGAR---TPA-H  405

eck VSDLEPHMNYTFTVEARNGVSGLVTSRSFRTASVSINQTEPPKVR.LEGRSTTSLSVSWSIPPPQQS.  464
eph -NG---YA----N---Q------GS-GHAS- ----SMGHAESLSG-SL-LVKKEPRQLELTWAGSRP  472

eck RVWKYEVTYRKKGDSNSYNVRRTEGFS.VTLDDLAPDTTYLVQVQALTQEGQGAGSKVHEFQTLSPEG  531
eph -SPGANL--ELHVLNQDEERYQMVLEPR-L-TE-Q-----I-R-RM--PL-P-PF-PD---R-SP-VS  540

eck SGNLAVIGGVAVGVVLLLVLAGVGFFIHRRRKNQRARQ.......SPEDVYFSKSEQLKPLKTYVDPH  592
eph R-LTGGEIVAVIFGL--GAALLL-ILVF-S-RA--Q--QRHVTAPPMWIPRT-CA-A-CGTSRHTRTL  608

eck TYED...PNQAVLKFTTEIHPSCVTRQKVIGAGEFGEVYKGMLKTSSGKKEVPVAIKTLKAGYTEKQR  657
eph HR-PWTL-GGWSNFPSR-LD-AWLMVDT---E-------R-TLRLP-.QDCKT-------DTSPGG-W  675

eck VDFLGEAGIMGQFSHHNIIRLEGVISKYKPMMIITEYMENGALDKFLREKDGEFSVLQLVGMLRGIAA  725
eph WN--R--T-------PH-LH----VT-R--I-----F-------A----REDQLVPG---A--Q---S  743

eck GMKYLANMNYVHRDLAARNILVNSNLVCKVSDFGLSRVLEDDPEATYTTSGGKIPIRWTAPEAISYRK  793
eph --N--S-H---------------Q--C--------T-L-D-FD.G--E-Q--------------AH-I  810

eck FTSASDVWSFGIVMWEVMTYGERPYWELSNHEVMKAINDGFRLPTPMDCPSAIYQLMMQCWQQERARR  861
eph --T--------------LSF-DK--G-M--Q----S-E--Y---P-V---APL-E--KN--AYD----  878

eck PKFADIVSILDKLIRAPDSLKTLADFDPRVSIRLPSTSGSEGVPFRTVSEWLESIKMQQYTEHFMAAG  929
eph -H-QKLQAH-EQ-LAN-H--R-I-N--P--TL----L---D-I-Y----------R-KR-IL--HS--  946

eck YTAIEKVVQMTNDDIKRIGVRLPGHQKRIAYSLLGLKDQVNTVGIPI                       976
eph LDTMGC-LEL-AE-LTQM-ITL-------LC-IQ-F--                               984
```

B

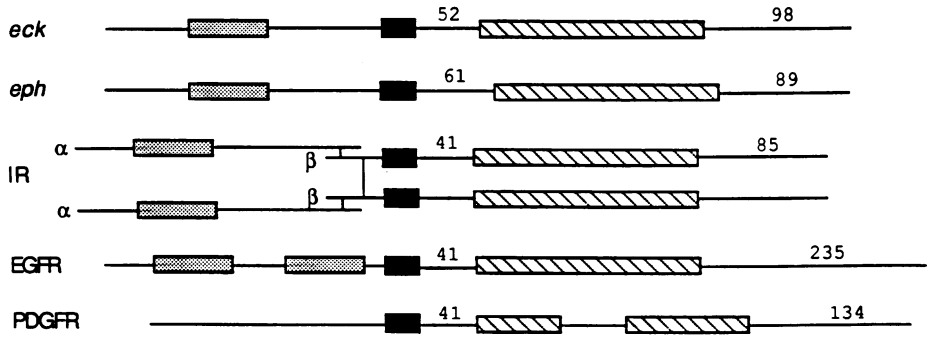

FIG. 3. Comparison of the protein sequences and overall structures of *eck* and *eph* (11). (A) Amino acid sequences were aligned by eye; gaps are marked by periods, and amino acid identities are marked by dashes. Putative signal sequences are boxed, extracellular cysteines are underlined and bolded, the transmembrane domains are underlined, and the ends of the catalytic domains are delineated by the curved arrows. (B) Structural features of *eck* and *eph* compared with those of representatives of other receptor PTK families. Symbols: ▩, cysteine-rich region; ■, transmembrane domain; ▨, catalytic domain. The numbers indicate how many amino acids lie outside the catalytic domains in the cytoplasmic part of the molecule. IR, Insulin receptor; EGFR, EGF receptor; PDGFR, PDGF receptor.

(0.2× SSPE) but only 1 band when washed at 68°C (data not shown).

The *eck* gene is expressed primarily in epithelial cells. Northern blot analysis of RNAs from both human and rodent cells with the 267-bp *Sau*3A probe detected a single RNA, which migrated slightly faster than the 28S rRNA, making its size ~4.7 kb (data not shown). The probe used in these analyses was from a conserved portion of the catalytic domain, yet only a 4.7-kb band was observed; this signal remained on blots of mouse RNA at reasonably high strin-

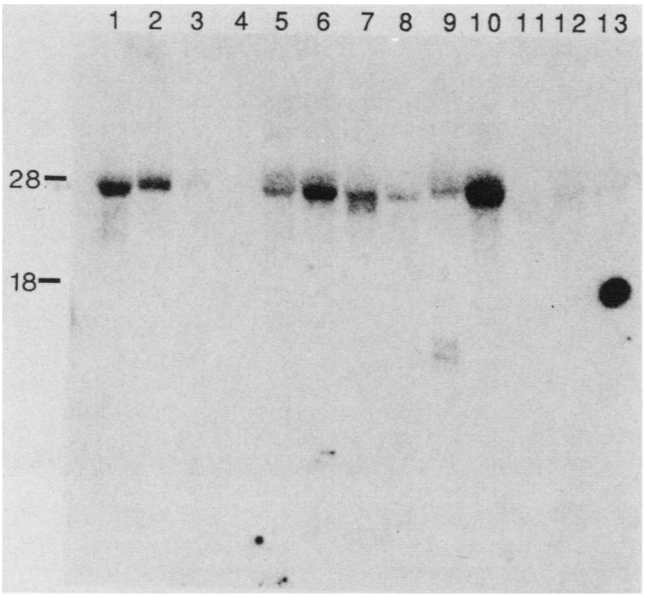

FIG. 4. Northern blot of RNA isolated from rat tissues. Whole-cell RNA (20 μg) was fractionated, transferred, and probed as described in Materials and Methods. Positions of the 28S and 18S rRNAs are marked at the left. Lanes: 1, skin; 2, ovary; 3, heart; 4, skeletal muscle; 5, brain; 6, small intestine; 7, kidney; 8, spleen; 9, submaxillary gland; 10, lung; 11, liver; 12, testes; 13, thymus. The blot was exposed at −70°C for 4 days to presensitized film with an intensifying screen. The slight variability in the mobility of the *eck* mRNA appears to be due to its closeness to 28S rRNA.

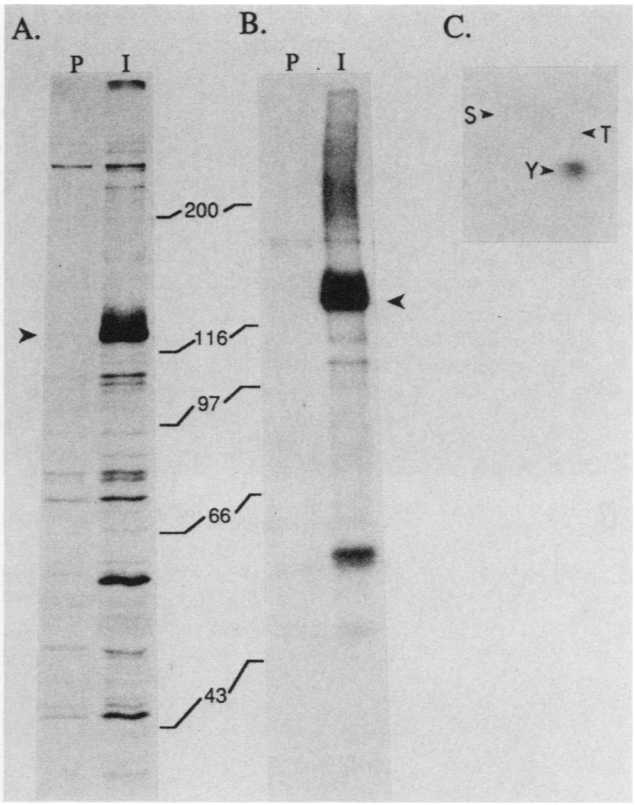

FIG. 5. Immunoprecipitation of the *eck* protein. Anti-*eck* antiserum was used to immunoprecipitate *eck* from A431 cell lysates as described in Materials and Methods. Arrowheads identify the *eck* protein; immunoprecipitates were separated on 7.5% SDS-polyacrylamide gels. Lanes: P, preimmune serum; I, immune serum. Markers are indicated in kilodaltons. S, T, and Y mark the positions of phosphoserine, phosphothreonine, and phosphotyrosine, respectively. The nature of the specifically precipitated ^{35}S-labeled ~60-kDa band, which appears to be phosphorylated in vitro, has not been investigated. (A) Cells were labeled with [^{35}S]methionine-cysteine prior to lysis. The gel was treated with diphenyloxazole and exposed to presensitized XAR film for 2 days at −70°C. (B) Unlabeled immunoprecipitates were used for in vitro kinase reactions. The gel was exposed for 45 min. (C) The band marked by the arrowhead in panel B was subjected to phosphoamino acid analyses.

gencies (65°C, 0.2× SSPE). These results suggest that *eck* is highly conserved across species and that this 4.7-kb band is the major *eck* mRNA. However, we cannot rule out that there are other *eck* mRNAs, which lack catalytic domain sequences similar to those found recently for other receptor protein-tyrosine kinases.

The tissue distribution of *eck* expression was investigated by Northern analysis of RNAs isolated from several rat organs (Fig. 4). *eck* was expressed most highly in lung, skin, small intestine, and ovary. It was detectable in kidney, brain, spleen, and submaxillary gland but was present at very low levels in heart, skeletal muscle, liver, testes, and thymus. Cell lines tested that expressed *eck* were A431 human epidermoid carcinoma, HeLa human epithelial carcinoma, 132N human astrocytoma, mouse BALB MK keratinocytes (kindly provided by Bernard Weissman, University of North Carolina), and primary rat Schwann cells. Cell lines with little or undetectable expression were MG63 human osteosarcoma, HepG2 human hepatoma, CEM human B lymphocytes, K562 human pre-erythroid/myeloid, and Rat 2 fibroblasts (data not shown). Thus, *eck* is expressed predominantly in cell lines of epithelial origin and in tissues that contain significant proportions of epithelial cells.

***eck* has in vitro kinase activity and autophosphorylates on tyrosine residues.** Polyclonal rabbit antisera were raised against a bacterially expressed TrpE fusion protein, containing 101 amino acids from the C-terminal tail of *eck* (residues 874 to 974; Fig. 1). This region was chosen because it does not include any amino acids required for activity by analogy with other PTKs and is therefore from a region unique to *eck*. The leucine at position 871 corresponds to the last residue believed to be required for activity (31). *eck* was immunoprecipitated from ^{35}S-labeled A431 cells and was found to migrate as a doublet at approximately 125 to 130 kDa on SDS-polyacrylamide gels (Fig. 5A). This doublet was precipitated by two different antisera but not by the corresponding preimmune sera. Proteins of similar size were also immunoprecipitated from lysates from HeLa cells, mouse C127 epithelial cells, and mouse BALB MK keratinocytes but not from NIH 3T3 mouse fibroblasts. The ability of the antibodies to recognize *eck* across species implies that the noncatalytic C-terminal tail is conserved and is perhaps important in *eck* function. When an unlabeled lysate of A431 cells was immunoprecipitated and in vitro kinase reactions with [γ-^{32}P]ATP were performed, the same two bands became labeled (Fig. 5B). Phosphotryptic peptide maps of these bands were identical, indicating that the bands probably represent different forms of the same protein (data not shown). The peptide maps were performed at two pHs, and at both pHs there was one major spot and five minor ones, which were identical when the patterns generated from the two bands were compared. Phosphoamino acid analysis showed the bands were phosphorylated primarily on ty-

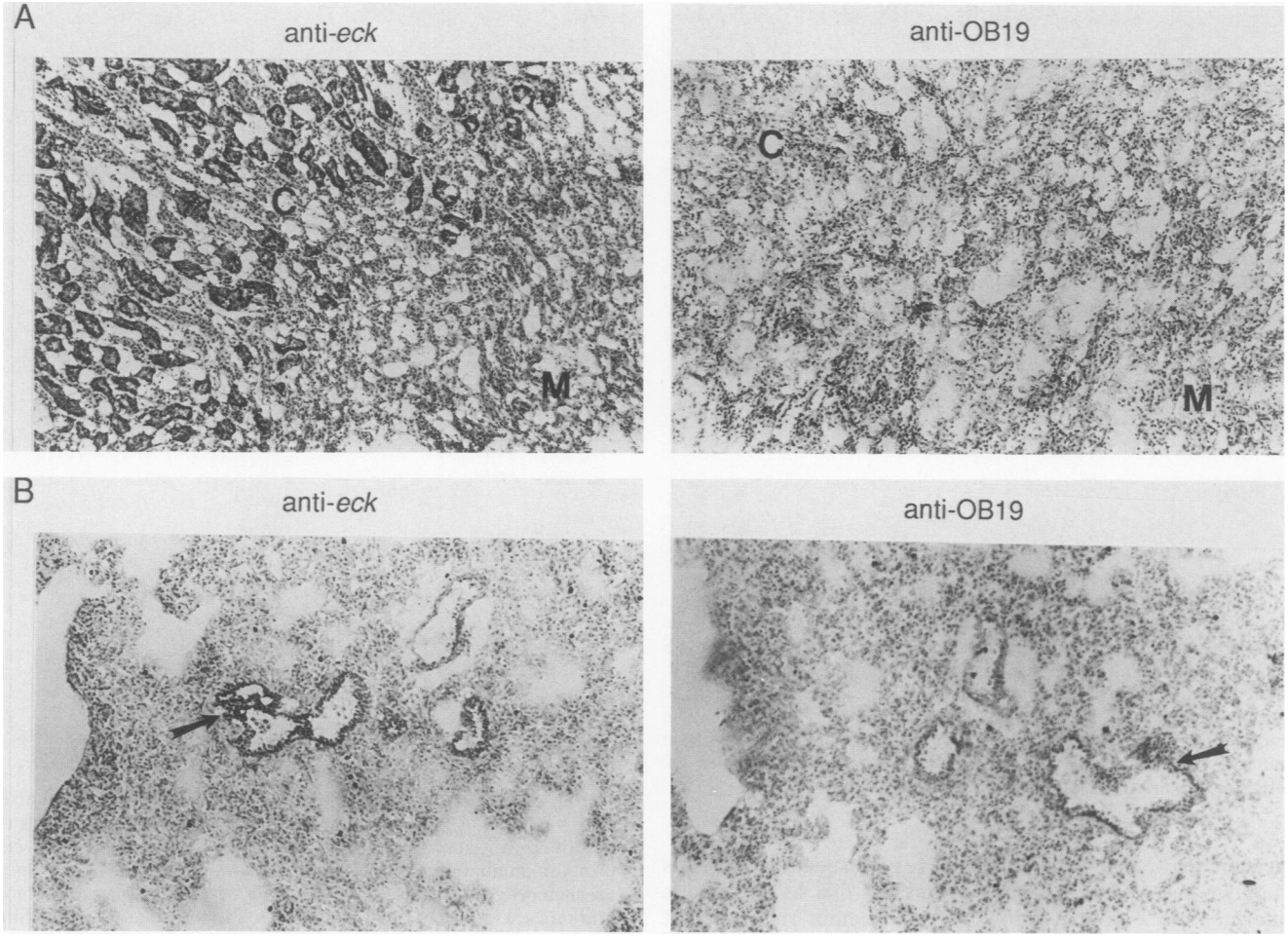

FIG. 6. Localization of *eck* to specific cells in the kidney and the lung. Immunohistochemistry was performed on adjacent tissue sections from rat with affinity-purified anti-*eck* and control (anti-OB19) antibodies as described in Materials and Methods. Magnification, ~55×. (A) Kidney sections treated with affinity-purified rabbit sera. The cortex (C) and medulla (M) are marked. (B) Lung sections. The arrow indicates a bronchiole; the upper tubular structure is a pulmonary arteriole.

rosine residues (Fig. 5C). This finding suggests that *eck* has intrinsic PTK activity and autophosphorylates in vitro.

Localization of the *eck* protein by immunohistochemistry. To determine which cells expressed the *eck* protein in tissues, we probed sections cut from several rat organs with antibodies affinity purified from the anti-TrpE-*eck* serum on the TrpE-*eck* fusion protein. The strongest staining occurred in the epithelial cells of the convoluted tubules of the kidney (Fig. 6A). Other cells in the cortex and the cells in the medulla did not stain with the antibodies. As a control, we used antibodies affinity purified from anti-TrpE-OB19 serum on the TrpE-*eck* fusion protein. This antiserum was raised against a TrpE fusion protein containing the catalytic domain of a PTK whose cDNA was isolated during the screen described in Materials and Methods and is a more appropriate control than preimmune serum in that it has antibodies to TrpE determinants. Neither anti-TrpE-OB19 nor preimmune serum showed any cell-specific staining in any tissues. The lung also showed specific staining in what appears to be the epithelium of terminal bronchioles of proximal respiratory bronchioles (Fig. 6B). Other organs tested that did not show strong staining in specific cells were the liver, which had light staining in all cells; the small intestine, which had a light staining of mucosa compared with submucosal structures such as Peyer's patches; and skin, which stains darker than controls but without cell specificity (data not shown).

DISCUSSION

In this communication, we describe the isolation of a cDNA that encodes a new receptorlike PTK, its primary structure and expression pattern, its protein kinase activity and specificity, and its localization to specific cells in tissues. Because expression of both mRNA and protein seems to be primarily in epithelial cells, we have named it *eck*, for epithelial cell kinase.

The combination of two overlapping cDNA clones provides the entire coding region of *eck* and accounts for ~4 kb of the 4.7-kb mRNA. Presumably, the missing sequences are in an extended 5′ untranslated region. Although there is no in-frame termination codon upstream of the proposed initiation codon, we believe all the coding sequences to be present from the following lines of evidence: (i) the first methionine follows Kozak's rules (15); (ii) there is a classical signal peptide sequence following this methionine; (iii) it has a similarly positioned initiating codon and migrates in SDS-polyacrylamide gels to give the same apparent molecular weight as its closest reported homolog, *eph* (20); (iv) and

finally, immunoprecipitation of in vitro translation products from RNA synthesized from a reconstructed full-length *eck* cDNA reveals a protein that migrates only slightly faster in SDS-polyacrylamide gels, a difference that could be accounted for by glycosylation at the three sites predicted in Fig. 1 (data not shown).

The primary structure of the predicted protein products puts *eck* in a new receptor PTK family by the criteria of catalytic domain homology and overall structure. The only full-length sequence of other family members reported thus far is that of *eph* (11), which was isolated from an erythropoietin-producing human hepatoma, but other cDNAs representing additional predicted members have been isolated. *elk* is a related cDNA that was isolated from a rat brain cDNA library (17), and the predicted product of this gene also has a receptorlike structure with an external domain related to those of the *eph* and *eck* products (Vladimir Lhotak and Tony Pawson, personal communication). Polymerase chain reaction products from rat RNA have identified eight different *eck* family protein kinases (Steve Hanks and Cary Lai, personal communications). In addition, *eek*, an *eph*- and *elk*-related kinase, has been isolated from a rat brain cDNA library (Joanne Chan and Valerie Watt, personal communication).

Receptor PTKs can be regulated by phosphorylation on amino acids in the juxtamembrane, catalytic, and C-terminal tail domains (for a review, see reference 28). *eck* has a tyrosine residue in the same position as the autophosphorylation site phosphorylated by many PTKs; the prototype is *src* Y-416 and corresponds to Y-772 in *eck*. Some PTK receptors are transmodulated by phosphorylation in the juxtamembrane domain, e.g., the EGF receptor by protein kinase C. Inspection of this region in *eck* does not reveal any obvious consensus protein kinase C phosphorylation sites, and the almost total lack of homology with *eph* in this region suggests this family may not share regulatory phosphorylation in this domain. Phosphorylation of the C-terminal tail is another mode of PTK regulation. In this region, *eck* and *eph* are 54% identical, which may indicate that they have similar regulatory mechanisms or interact with similar cytoplasmic proteins, or both. *eck* has three tyrosines in its C-terminal tail, one of which is in common with *eph*. These could be autophosphorylation sites, but they do not appear to be in a good context in comparison with known PTK phosphorylation sites. There are many serine and threonine residues conserved between *eck* and *eph* in this domain, which could be sites of phosphorylation by protein-serine/threonine kinases.

Presumably, a family of ligands exists that activates these receptors, but we have not identified any candidates. The similar structures and the conservation of cysteine residues in the putative ligand-binding domains of *eck* and *eph* indicates that they may bind the same or similar ligands. These molecules have a higher percent identity (44%) in this region than do the PDGF α and β receptors (31%), which are capable of binding the same ligand (5, 22). We are currently attempting to determine the location of the *eck* protein in polarized epithelial cells. This information should be relevant to the nature of the ligand. Where the receptor is expressed in stratified epithelium will also indicate whether the ligand is soluble or another cell surface protein.

The RNA expression data from rat organs shows that the *eck* gene is expressed in many places in the body but to different extents. Presumably, it is expressed in certain cell types in each of these tissues. *eck* was expressed in all cell lines tested that were of epithelial origin. However, it was expressed at low levels in some lines not of epithelial origin, e.g., an osteosarcoma and a hepatoma cell line. The immunohistochemistry data also indicate that *eck* is expressed in epithelial cells in vivo. The only cells in kidney and lung sections that reacted with anti-*eck* antibodies were specific types of epithelial cell. However, staining of other tissues that expressed *eck* mRNA was less convincing and not cell type specific. There is not a very good concordance between the level of *eck* mRNA detected by Northern blotting and the amount of immunoreactive protein detected by immunohistochemistry in tissues. For example, skin and small intestine had higher mRNA levels than did kidney but did not show as intense staining. To attempt to reconcile the RNA and staining data, we performed immune complex kinase assays on tissue extracts. The amount of *eck* autophosphorylating activity correlated quite well with the amount of *eck* mRNA by Northern analysis and furthermore confirmed the presence of *eck* protein in tissues that were negative by immunohistochemistry (R. A. Lindberg, unpublished observation). This leads us to believe that the immunohistochemistry data do not necessarily reflect total *eck* expression in tissues. There could be several explanations for the failure to detect *eck* protein expression by immunostaining. For example, the proportion of cells expressing *eck* in some tissues, such as skin, could be much higher than that in kidney, but the amount of *eck* protein per cell might be less in skin than in the convoluted tubule epithelial cells, or the accessibility or preservation of the antigen may differ in different tissue types. In sum, our data imply that *eck* may not be expressed in all epithelial cells and is expressed in some cells not of epithelial origin but is expressed predominantly in epithelial cells.

The expression of the members in this new PTK family may be largely exclusive from one another. Tissue Northern analyses show different patterns of expression. *eph* is expressed in lung, liver, kidney, and testes (21), *elk* is expressed mainly in brain and testes (17), and *eck* is expressed at highest levels in lung, skin, ovary, and small intestine. Exactly which cells in these tissues express these receptors and what their functions are will need to be known to determine whether they perform similar functions in different cells. Finding the ligand for these receptors will be the most critical step in discovering the role that they play in the cell and proving that they really are receptors. By analogy with genes for other receptor PTKs, *eck* is a putative proto-oncogene, and therefore it will also be important to determine whether it has a role in carcinogenesis. It has recently been reported that NIH 3T3 cells that overexpress the *eph* gene acquire tumorigenic ability in nude mice (20), suggesting that this family of PTK genes can function as oncogenes. Given the frequency human malignancies in which epithelial cells are the culprits, it will be extremely important to determine whether overexpression of *eck* is a contributing factor.

ACKNOWLEDGMENTS

We thank Steve Hanks for the HeLa cell library, the computer analyses, and advice and support throughout the work described here. We thank Jill Meisenhelder, Detlev Jähner, Kathy Gould, Bill Boyle, and numerous others in the Hunter-Eckhart laboratory for materials and help with techniques. We also thank Marianne Bowman for tutelage with the immunohistochemistry.

This work was supported by a Public Health Service predoctoral training grant and by Public Health Service grant CA-39780, both from the National Institutes of Health, by a grant from the American Business Foundation for Cancer Research to T.H., and by Cancer Center core grant CA14195 from the National Cancer Institute.

LITERATURE CITED

1. **Bargmann, C. I., M.-C. Hung, and R. A. Weinberg.** 1986. Multiple independent activations of the *neu* oncogene by a point mutation altering the transmembrane domain of p185. Cell **45**:649–657.
2. **Bishop, J. M.** 1987. The molecular genetics of cancer. Science **235**:305–311.
3. **Boyle, W. J., J. S. Lipsick, and M. A. Baluda.** 1986. Antibodies to the evolutionarily conserved amino-terminal region of the v-*myb*-encoded protein detect the c-*myb* protein in widely divergent metazoan species. Proc. Natl. Acad. Sci. USA **83**: 4685–4689.
4. **Chomczynski, P., and N. Sacchi.** 1987. Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction. Anal. Biochem. **162**:156–159.
5. **Claesson-Welsh, L., A. Ericksson, A. Moren, L. Severinsson, B. Ek, A. Ostman, C. Betsholtz, and C. H. Heldin.** 1988. cDNA cloning and expression of a human platelet-derived growth factor (PDGF) receptor specific for B-chain-containing PDGF molecules. Mol. Cell. Biol. **8**:3476–3486.
6. **Gould, K. L., and T. Hunter.** 1988. Platelet-derived growth factor induces multisite phosphorylation of $pp60^{c\text{-}src}$ and increases its protein-tyrosine kinase activity. Mol. Cell. Biol. **8**:3345–3356.
7. **Grantham, R., C. Gautier, M. Gouy, M. Jacobzone, and R. Mercier.** 1981. Codon catalog usage is a genome strategy modulated for gene expressivity. Nucleic Acids Res. **1**:r43–r74.
8. **Hanks, S. K.** 1987. Homology probing: identification of cDNA clones encoding members of the protein-serine kinase family. Proc. Natl. Acad. Sci. USA **84**:388–392.
9. **Hanks, S. K., A. M. Quinn, and T. Hunter.** 1988. The protein kinase family: conserved features and deduced phylogeny of the catalytic domains. Science **241**:42–52.
10. **Harlow, E., and D. Lane.** 1988. Antibodies. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
11. **Hirai, H., Y. Maru, K. Hagiwara, J. Nishida, and F. Takaku.** 1987. A novel putative tyrosine kinase receptor encoded by the *eph* gene. Science **238**:1717–1720.
12. **Hunter, T., P. Angel, W. J. Boyle, R. Chiu, E. Freed, K. L. Gould, C. M. Isacke, M. Karin, R. A. Lindberg, and P. van der Geer.** 1988. Targets for signal-transducing protein kinases. Cold Spring Harbor Symp. Quant. Biol. **53**:131–142.
13. **Hunter, T., and B. M. Sefton.** 1980. The transforming gene product of Rous sarcoma virus phosphorylates tyrosine. Proc. Natl. Acad. Sci. USA **77**:1311–1315.
14. **Kornbluth, S., K. E. Paulson, and H. Hanafusa.** 1988. Novel tyrosine kinase identified by phosphotyrosine antibody screening of cDNA libraries. Mol. Cell. Biol. **8**:5541–5544.
15. **Kozak, M.** 1986. Point mutations define a sequence flanking the AUG initiator codon that modulate translation by eukaryotic ribosomes. Cell **44**:283–292.
16. **Lee, P. L., D. E. Johnson, L. S. Cousens, V. A. Fried, and L. T. Williams.** 1989. Purification and complementary DNA cloning of a receptor for basic fibroblast growth factor. Science **245**:57–60.
17. **Letwin, K., S. P. Yee, and T. Pawson.** 1988. Novel protein-tyrosine kinase cDNAs related to *fps/fes* and *eph* cloned using anti-phosphotyrosine antibody. Oncogene **3**:621–627.
18. **Lindberg, R. A., D. P. Thompson, and T. Hunter.** 1988. Identification of cDNA clones that code for protein-tyrosine kinases by screening expression libraries with antibodies against phosphotyrosine. Oncogene **3**:629–633.
19. **Maniatis, T., E. F. Fritsch, and J. Sambrook.** 1982. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
20. **Maru, Y., H. Hirai, and F. Takaku.** 1990. Overexpression confers an oncogenic potential upon the *eph* gene. Oncogene **5**:445–447.
21. **Maru, Y., H. Hirai, M. Yoshida, and F. Takaku.** 1988. Evolution, expression, and chromosomal location of a novel receptor tyrosine kinase gene, *eph*. Mol. Cell. Biol. **8**:3770–3776.
22. **Matsui, T., M. Heideran, T. Miki, N. Popescu, W. La Rochelle, M. Kraus, J. Pierce, and S. A. Aaronson.** 1989. Isolation of a novel receptor cDNA establishes the existence of two PDGF receptor genes. Science **243**:800–804.
23. **Pasquale, E. B., and S. J. Singer.** 1989. Identification of a developmentally regulated protein-tyrosine kinase by using anti-phosphotyrosine antibodies to screen a cDNA expression library. Proc. Natl. Acad. Sci. USA **86**:5449–5453.
24. **Sanger, F., S. Nicklen, and A. R. Coulson.** 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA **74**:5463–5467.
25. **Slamon, D. J., W. Godolphin, L. A. Jones, J. A. Holt, S. G. Wong, D. E. Keith, W. J. Levin, S. G. Stuart, J. Udove, A. Ullrich, and M. F. Press.** 1989. Studies of the HER-2/*neu* proto-oncogene in human breast and ovarian cancer. Science **244**:707–712.
26. **Ullrich, A., J. R. Bell, E. Y. Chen, R. Herrera, L. M. Petruzzelli, T. J. Dull, A. Gray, L. Coussens, Y. C. Liao, M. Tsubokawa, A. Mason, P. H. Seeburg, C. Grunfeld, O. M. Rosen, and J. Ramachandran.** 1985. Human insulin receptor and its relationship to the tyrosine kinase family of oncogenes. Nature (London) **313**:756–761.
27. **Ullrich, A., L. Coussens, J. S. Hayflick, T. J. Dull, A. Gray, A. W. Tam, J. Lee, Y. Yarden, T. A. Libermann, J. Schlessinger, J. Downward, E. L. V. Mayes, N. Whittle, M. D. Waterfield, and P. H. Seeburg.** 1984. Human epidermal growth factor receptor cDNA sequence and aberrant expression of the amplified gene in A431 epidermoid carcinoma cells. Nature (London) **309**:418–425.
28. **Ullrich, A., and J. Schlessinger.** 1990. Signal transduction by receptors with tyrosine kinase activity. Cell **61**:203–212.
29. **von Heijne, G.** 1984. A new method for predicting signal sequence cleavage sites. Nucleic Acids Res. **14**:4683–4690.
30. **Wilks, A. F.** 1989. Isolation of two putative protein tyrosine kinases by application of the polymerase chain reaction. Proc. Natl. Acad. Sci. USA **86**:1603–1607.
31. **Yaciuk, P., and D. Shalloway.** 1986. Features of the $pp60^{v\text{-}src}$ carboxyl terminus that are required for transformation. Mol. Cell. Biol. **6**:2807–2819.
32. **Yanisch-Perron, C., J. Vieira, and J. Messing.** 1985. Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC19 vectors. Gene **33**:103–119.
33. **Yarden, Y., J. A. Escobedo, W.-J. Kuang, T. L. Yang-Feng, T. O. Daniel, P. M. Tremble, E. Y. Chen, M. E. Ando, R. N. Harkins, U. Francke, V. A. Fried, A. Ullrich, and L. T. Williams.** 1986. Structure of the receptor for platelet-derived growth factor helps define a family of closely related growth factor receptors. Nature (London) **323**:226–232.
34. **Yarden, Y., and A. Ullrich.** 1988. Growth factor receptor tyrosine kinases. Annu. Rev. Biochem. **57**:443–478.