## Molecular cloning of HEK, the gene encoding a receptor tyrosine kinase expressed by human lymphoid tumor cell lines

(human EPH/ELK-like receptor tyrosine kinase)

I. P. WICKS\*, D. WILKINSON, E. SALVARIS, AND A. W. BOYD

Lions Clinical Cancer Laboratory, Walter and Eliza Hall Institute of Medical Research, P.O. Royal Melbourne Hospital, Parkville 3050, Australia

Communicated by G. J. V. Nossal, December 2, 1991

ABSTRACT We describe the molecular cloning of <sup>a</sup> receptor tyrosine kinase from a cell line (LK63) derived from a case of human pre-B-cell leukemia. We have previously shown that a monoclonal antibody (IIA4) raised against LK63 recognized a glycosylated, cell-surface 135-kDa molecule (HEK), which displayed tyrosine kinase activity in vitro. The HEK protein was purified by using a IIIA4 antibody column and both N-terminal and internal amino acid sequences were obtained. A 51-mer degenerate oligonucleotide based on the internal amino acid sequence was used to screen an LK63 derived AgtlO cDNA library under low-stringency hybridization conditions. One clone of 2.5 kilobases (kb) was isolated and characterized and used to rescreen the library under morestringent hybridization conditions. A 4.5-kb clone containing the entire HEK coding region was isolated and its complete DNA sequence was determined. The 4.5-kb insert was subcloned into the expression vector CDM8 and transfected into COS cells. COS cells transfected with the sense HEK/CDM8 construct stained specifically with the IIIA4 antibody, thereby confirming that the antigen recognized by the IIIA4 antibody and the expressed protein product of the HEK cDNA clone were identical. DNA sequence analysis revealed that HEK is <sup>a</sup> newly discovered member of the EPH/ELK family of receptor tyrosine kinases. Northern blot analysis of a number of cell lines demonstrated the expression of 5.5- to 6.0-kb HEK transcripts in LK63 and the T-cell lines JM and HSB-2. Southern blot analysis of DNA from LK63 suggested that the HEK gene was neither amplified nor rearranged in the LK63 tumor.

Posttranslational phosphorylation of proteins constitutes an important regulatory mechanism for eukaryotic cells. For example, tyrosine phosphorylation of cdc2 is central to regulation of the cell cycle (1). Kinases that phosphorylate proteins on tyrosine residues (protein tyrosine kinases; PTKs) form a structurally related group of molecules that exhibit functional diversity (2). Genetic alterations that lead to the inappropriate activation or expression of PTKs may be oncogenic and have been associated with some human and murine tumors (3). Some PTKs were first characterized as components of transforming retroviruses, with the subsequent identification of normal cellular counterparts (4). The receptors for many of the known growth factors are PTKs, such as the receptors for epidermal growth factor (5), platelet-derived growth factor (6), colony-stimulating factor 1 (7), and stem cell factor (8). The recent isolation of lineagespecific (e.g., flk-2) (9) and developmentally regulated (e.g., CEK 5) (10) PTKs highlights an emerging role for this group of molecules in differentiation and development.

As part of our study of human B-cell growth and differentiation, we have raised a number of monoclonal antibodies (mAbs) against early B-cell populations. One of these, des-

ignated IIIA4, has been shown to recognize a 135-kDa molecule (HEK) on the surface of the pre-B-cell leukemic cell line LK63 (11). Glycosylation studies demonstrated an estimated molecular mass of <sup>95</sup> kDa for the HEK core protein. By indirect immunofluorescence, the human T-cell line JM also expressed this antigen, but a wide range of other lymphoid and myeloid cell lines were negative. Fluorescence-activated cell sorter analysis of normal human lymphoid and bone marrow cells showed no detectable HEK expression. However, HEK expression was observed on <sup>a</sup> small number of fresh hemopoietic tumor specimens, raising the possibility that IIIA4 recognized the product of a sporadically overexpressed gene. Immunoprecipitated HEK protein was shown to have tyrosine kinase activity in vitro. These data suggested that IIIA4 recognized a membraneassociated PTK. We used the IIIA4 mAb to affinity purify HEK protein from LK63 cells. Amino acid sequencing demonstrated <sup>a</sup> unique N terminus and PTK-related internal sequences.

In this report, we describe the isolation and sequence determination of <sup>a</sup> 4.5-kilobase (kb) cDNA encoding the HEK receptor tyrosine kinase  $(RTK)$ .<sup>†</sup> Identity between the antigen recognized by IIIA4 and the product of the 4.5-kb HEK clone was demonstrated by staining COS cells transfected with HEK subcloned into the expression vector CDM8. Sequence comparison of HEK with other PTKs revealed a high degree of homology with members of the EPH/ELK family of RTKs. Northern blot analysis of RNA from <sup>a</sup> panel of cell lines showed restriction of HEK expression to one pre-B- and to two T-cell lines. This pattern is quite unlike that described for any other member of the EPH/ELK family. The apparent restriction of HEK expression to lymphoid tumor cell lines raises the possibilities that HEK may play a role in some human lymphoid malignancies and also in normal lymphoid function and differentiation.

## MATERIALS AND METHODS

CeUl Lines, mAb IIIA4, HEK Protein Structure and Function. The LK63 and LK63/CD20<sup>+</sup> cell lines were derived from a child with acute lymphoblastic leukemia. LK63/ CD20+ is a tetraploid variant of LK63, which arose spontaneously in vitro and has enhanced HEK expression. In contrast to the parental cell line,  $LK63/CD20^+$  expresses CD20 (12). These lines have cytogenetic features of pre-Bcell leukemia and have not been transformed with Epstein-Barr virus. JM and HSB-2 are human T-cell leukemic cell lines. The IIIA4 mAb was generated against the LK63 cell line and recognized a 135-kDa cell-surface molecule (HEK)

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Abbreviations: HEK, human EPH/ELK-like kinase; PTK, protein tyrosine kinase; RTK, receptor tyrosine kinase; mAb, monoclonal antibody.

<sup>\*</sup>To whom reprint requests should be addressed.

tThe sequence reported in this paper has been deposited in the GenBank data base (accession no. M83941).

with in vitro kinase activity, expressed by LK63, LK63/  $CD20<sup>+</sup>$ , and JM (11).

The IIIA4 mAb was used to purify HEK antigen for amino acid sequencing (11). The amino acid sequences obtained were as follows: N terminus, ELIPQPSNEVNLXD- (S)KXIQ; internal, GYRLPPPMDCPAALYQLMLDC (X, unidentified residue; parentheses, doubtful residue).

LK63 cDNA Library Construction and Screening. A random-primed cDNA library was constructed in λgt10 (Amersham cDNA synthesis and  $\lambda$ gt10 library kits), using 5  $\mu$ g of  $poly(A)^+$ -selected mRNA from LK63/CD20<sup>+</sup> cells. A degenerate oligonucleotide was designed on the basis of the internal (3') HEK protein sequence. The neutral base inosine was included at positions of high codon degeneracy (13). The 51-mer TA(C/T)CGICTICCICCICCIATGGA(C/T)TG(C/ T)CCIGCIGCICTITA(C/T)CA(A/G)CTIATG was endlabeled using  $[\gamma^{-3}P]$ dATP and polynucleotide kinase, followed by separation on a G25 Sephadex column as described (14). Approximately 250,000 plaques were screened in  $2 \times$ SSC  $(1 \times SSC = 0.15 \text{ M NaCl}/0.015 \text{ M sodium citrate})$ hybridization buffer at 37°C as described (14). Washes were

performed in  $2 \times$  SSC/0.1% SDS at 42°C-55°C. The signal from one duplicating plaque persisted after 55°C washes. The DNA from this plaque contained an insert of 2.5 kb (HEK 2.5). HEK 2.5 was labeled with  $[\alpha^{-32}P]ATP$  (Amersham random primer kit) for Northern blot analysis of LK63 cells. The polymerase chain reaction (PCR) was performed using HEK 2.5 and oligonucleotide primers based on conserved motifs within the catalytic domain of PTKs and the <sup>3</sup>' amino acid HEK sequence as described (15). HEK 2.5 was labeled with  $[\alpha^{-32}P]ATP$  (as described above) and used to rescreen the random-primed LK63 cDNA library in  $2 \times$  SSC hybridization buffer at  $65^{\circ}$ C. Thirty-two duplicating positives were isolated and screened by hybridization with a degenerate oligonucleotide based on the N-terminal HEK protein sequence. A 4.5-kb HEK clone (HEK 4.5), which hybridized with the N-terminal oligonucleotide, was chosen for complete characterization.

DNA Sequencing and Analysis of HEK cDNA. HEK 4.5 was subcloned into pGEM7, which had been digested with EcoRI and treated with calf intestinal alkaline phosphatase. Doublestranded DNA was purified on <sup>a</sup> cesium chloride gradient and

\* M D C CATGGATGGTAACTTCTCCAGCAATCAGAGCGCTCCCCC TCACATCAGTGGCATGCTTCATGGAGATATGCTCCTCTCACTGCCCTC TGCACCAGCAACATGGATTGT 108 10 20 Q L <sup>S</sup> <sup>I</sup> L <sup>L</sup> L L S C S V L D S F G E L <sup>I</sup> <sup>P</sup> Q <sup>P</sup> S N E V N L L D <sup>S</sup> K T <sup>I</sup> Q  $216$ CAGCTCTCCATCCTCCTCCTTCTC-AGCTGCTCTGTTCTCGACAGCTTCGGGAACTGATTCCGCAGCCTTCCAATGAAGTCAATCTACTGGATTCAAAAACAATTCAA 40 50 60 70 <sup>G</sup> <sup>E</sup> <sup>L</sup> <sup>G</sup> <sup>W</sup> <sup>I</sup> <sup>3</sup> <sup>Y</sup> <sup>P</sup> S <sup>H</sup> <sup>G</sup> <sup>W</sup> <sup>E</sup> <sup>E</sup> <sup>I</sup> <sup>S</sup> <sup>G</sup> V <sup>D</sup> <sup>E</sup> <sup>H</sup> <sup>Y</sup> <sup>T</sup> <sup>P</sup> <sup>I</sup> <sup>R</sup> <sup>T</sup> Y <sup>Q</sup> V <sup>C</sup> <sup>N</sup> <sup>V</sup> <sup>M</sup> <sup>D</sup>  $324$ 80 90 110<br>H S Q N N W L R T N W V P R N S A Q K I Y V E L K F T L R D C N S I P L CACAGTCAAAACAATTGGCTGAGAACAAACTGGGTCCCCAGGAACTCAGCTCAGAAGATTTATGTGGAGCTCAAGTTCACTCTACGAGACTGCAATAGCATTCCATTG 43 2 120 130 140 V L G T C K E T F N L Y Y M E S D D D H G V K F R E H Q F T K <sup>I</sup> D T <sup>I</sup> A GTTTTAGGAACTTGCAAGGAGACATTCAACCTGTACTACATGGAGTCTGATGATGATCATGGGGTGAAATTTCGAGAGCATCAGTTTACAAAGATTGACACCATTGCA 54 0 150 160 160<br>A D E S F T Q M D L G D R I L K L N T E I R E V G P V N K K G F Y L A F CCTOATGARAGTTCCCCARATEROGATEROADSCONE CONTRACT TRANSPACT CONTRACT TRANSPACT TRANSPACT TRANSPACT TRANSPACT TRANSPACT TRANSPACT TO A CONTRACT TRANSPACT TO A CONTRACT TRANSPACT TO A CONTRACT TRANSPACT TO A CONTRACT TRANSPACT 864 972 CCTCACAGTTCTACTCAGGAAGATGGTTCAATGAACTGCAGGTGTGAGAATMATTACTTCCGGGCAGACAAAGACCCTCCATCCATGGCTTGTACCCGACCTCCATCT 1080 <sup>330</sup> v v v 350 <sup>360</sup> <sup>S</sup> <sup>P</sup> R N V <sup>I</sup> <sup>S</sup> N <sup>I</sup> N <sup>E</sup> T <sup>S</sup> V <sup>I</sup> L <sup>D</sup> W <sup>S</sup> W <sup>P</sup> <sup>L</sup> <sup>D</sup> T G G R K D V <sup>T</sup> F N <sup>I</sup> <sup>I</sup> <sup>C</sup> TCACCAAGAAATGTTATCTCTAATATAAACGAGACCTCAGTTATCCTGGACTGGAGTTGGCCCCTGGACACAGGAGGCCGGAAAGATGTTACCTTCAACATCATATGT 10188 <sup>370</sup> 380 v v v <sup>K</sup> <sup>K</sup> <sup>C</sup> <sup>G</sup> W <sup>N</sup> <sup>I</sup> K Q C <sup>E</sup> <sup>P</sup> C <sup>S</sup> <sup>P</sup> N V <sup>R</sup> <sup>F</sup> <sup>L</sup> P R <sup>Q</sup> <sup>F</sup> <sup>G</sup> <sup>L</sup> <sup>T</sup> <sup>N</sup> <sup>T</sup> T V T V <sup>T</sup> D <sup>L</sup> AAAAAATGTGGGTGGAATATAAAACAGTGTGAGCCATGCAGCCCAAATGTCCGCTTCCTCCCTCGACAGTTTGGACTCACCAACACCACGGTGACAGTGACAGACCTT 1296 <sup>400</sup> v v v <sup>410</sup> 420 <sup>430</sup> L A H T N Y T F <sup>E</sup> <sup>I</sup> <sup>D</sup> A V N G V <sup>S</sup> <sup>E</sup> L <sup>S</sup> <sup>S</sup> <sup>P</sup> <sup>P</sup> R Q F A A V <sup>S</sup> <sup>I</sup> <sup>T</sup> T N Q A CTGGCACATACTAACTACACCTTTGAGATTGATGCCGTTAATGGGGTGTCAGAGCTGAGCTCCCCACCAAGACAGTTTGCTGCGGTCAGCATCACAACTAATCAGGC T 1404 440 450 460 470 A P <sup>S</sup> P V L T <sup>I</sup> K K <sup>D</sup> R T S R N <sup>S</sup> <sup>I</sup> <sup>S</sup> <sup>L</sup> S W Q E <sup>P</sup> <sup>E</sup> H <sup>P</sup> N G <sup>I</sup> <sup>I</sup> L D Y <sup>E</sup> GCTCCATCACCTGTCCTGACGATTAAGAAAGATCGGACCTCCAGAAATAGCATC TCTTTGTCCTGGCAAGAACCTGAACATCCTAATGGGATCATATTGGACTACGAG 1512 480 490 v v v 500 V K <sup>Y</sup> <sup>Y</sup> <sup>E</sup> K Q E Q <sup>E</sup> T <sup>S</sup> <sup>Y</sup> <sup>T</sup> <sup>I</sup> L R A R G <sup>T</sup> N V <sup>T</sup> <sup>I</sup> <sup>S</sup> <sup>S</sup> L K <sup>P</sup> <sup>D</sup> T <sup>I</sup> <sup>Y</sup> V <sup>F</sup> GTCAAATACTATGAAAAGCAGGAACAAGAAACAAGTTATACCATTCTGAGGGCAAGAGGCACAAATGTTACCATCAGTAGCCTCAAGCCTGACACTATATACGTATTA 162 0 510 540 520 520 520 520 530<br>Q I R A R T A A G Y G T N S R K F E F E T S P D S F S I S G E S S Q V V<br>CAAATCCGAACAGCCCCCCCCCCGAACAGACGAACAGCGCAAGTTTGAACTTTGAAACTAGTCCAGACTCTTTCTCCATCTCTGGTGAAAGTAGCCAA<u>GTCGTC</u> 1728 550 570<br>M I A I S A V A I I L L T V V L V L I G R F C G Y K S K H G A D E K<br><u>ATGATCGCCATTTCAGTGGGAGTAATTATTCTCCTCACTGTTGTCATCTATTGTTTTGATTGG</u>GAGGTTCTGTGGCTATAAACATGGGGCAGATGAAAAA 1836 580 590 600 610 <sup>R</sup> <sup>L</sup> <sup>H</sup> <sup>F</sup> <sup>G</sup> <sup>N</sup> G <sup>H</sup> <sup>L</sup> <sup>K</sup> <sup>L</sup> <sup>P</sup> <sup>G</sup> <sup>L</sup> <sup>R</sup> <sup>T</sup> <sup>Y</sup> V <sup>D</sup> <sup>P</sup> <sup>H</sup> <sup>T</sup> <sup>Y</sup> <sup>E</sup> <sup>D</sup> <sup>P</sup> <sup>T</sup> Q A V <sup>H</sup> <sup>E</sup> <sup>F</sup> A K <sup>E</sup> AGACTTCATTTTGGCAATGGGCATTTAAAACTTCCAGGTCTCAGGACTTATGTTGACCCACATACATATGAAGACCCTACCCAAGCTGTTCATGAGTTTGCCAAGGAA 19414 620 620 ● ● ● ● ● 650 650 650 8<br>L D A T N I S I D K V V G A G E F G E V C S G R K L P S K K E I S V A<br>TTGGATGCCACCAACATATCATTAAAGTTGTTGGAGCAGGTGAAGGTGTGCAGTGGTCGCTTAAAAATTTCCTTCAAAAAAAGAGATTTCAGTGCCC 2052 660 670 680 <sup>I</sup> K T L K V G <sup>Y</sup> <sup>T</sup> <sup>E</sup> K Q R R D <sup>F</sup> <sup>L</sup> G E A <sup>S</sup> <sup>I</sup> M G Q <sup>F</sup> D H P N <sup>I</sup> <sup>I</sup> R L E G ATTAAAACCCTGAAAGTTGGCTACACAGAAAAGCAGAGGAGAGACTTCCTGGGAGAAGCAAGCATTATGGGACAGTTTGACCACCCCMATATCATTCGACTGGAAGGA 22160 690 700 700<br>V V T K S K P V M I V T E Y M E N G S L D S F L R K H D A Q F T V I Q L<br>GTTGTTACCAAAAGTAAGCCAGTTTATGATTGTCACAGAATAGTATGGAATAGGTTCCTTGGATAGTTTCCTAGGTAAACACGATGCCCAGTTTACTGTCATTCAGCTA 730 730<br>V G M L R S H K Y L S D M G Y V H R D L A A R N I L I N S N L V<br>GTGGGGATGCTTCGAGGGATAAGCTLCTGCCATGAAGTACCTGTCAGACATGTCACCGAGACCTCGCTGCTCGGAACATCTTGATCAACAGTAACTTGGTG 2376 760 760 770 770<br>C K V S D F G L S R V L E D D P E A A Y T T R G G K I P I R W T S P E A<br>TGTAAGGTTTCTGATTTCGGACTTTCGCGTGTGCAGGATGACCCAGAAGCTGCTTATACAACAAGGAGGGAAGATCCCAATCAGGTGGACATCACCAGAAGCT 2484  $\begin{array}{ccccccccc} & R & V & L & E & D & D & P & E \\ \texttt{SCGTGTCCTGGAGGATACCCAGAA} & & & & & & & & \\ \texttt{S} & D & V & W & S & Y & G & I \\ \texttt{CAGCGATATGGGAGTATGGGATTATGCTGTTT} & & & & & & \\ \texttt{S} & & & & & & & & \\ \texttt{T} & & & & & & & & \\ \texttt{T} & & & & & & & & \\ \texttt{T} & & & & & & & & \\ \texttt{T} & & & & & & & & \\ \end{array}$ 820 800<br>I A Y R K F T S A S D V W S Y G I V L W E V M S V G E R P Y W E M S N Q<br>ATAGCCTACCGCAAGTTCACCCAGCGATGTATGGAGTTATGGGATTGTTCTCTGGGAGGTGATGTCTTATGGAGAGACCATACTGGGAGATGTCCAATCAC 2592 861 <sup>D</sup> V <sup>I</sup> K A V D E G Y R L P <sup>P</sup> P M D C P A A L <sup>Y</sup> Q L M L D C W Q K <sup>D</sup> R N N GATGTAATTAAAGCTGTAGATGAGGGCTATCGACTGCCACCCCCCATGGACTGCCCAGCTGCCTTGTATCAGCTGATGCTGGACTGCTGGCAGAAAGACAGGAACAAC 22700 870 880 890 900 <sup>R</sup> <sup>P</sup> K <sup>F</sup> <sup>E</sup> Q <sup>I</sup> V <sup>S</sup> <sup>I</sup> <sup>L</sup> <sup>D</sup> K <sup>L</sup> <sup>I</sup> <sup>R</sup> <sup>N</sup> <sup>P</sup> G <sup>S</sup> <sup>L</sup> <sup>K</sup> <sup>I</sup> <sup>I</sup> <sup>T</sup> <sup>S</sup> A A A R <sup>P</sup> <sup>S</sup> <sup>N</sup> <sup>L</sup> <sup>L</sup> <sup>L</sup> AGACCCAAGTTTGAGCAGATTGTTAGTATTCTGGACAAGCTTATCCGGAATCCCGGCAGCCTGAAGATCATCACCAGTGCAGCCGCAAGGCCATCAAACCTTCTTCTG 2808 910 920 930 <sup>D</sup> <sup>Q</sup> <sup>S</sup> <sup>N</sup> V <sup>D</sup> <sup>I</sup> <sup>S</sup> <sup>T</sup> <sup>F</sup> <sup>R</sup> <sup>T</sup> <sup>T</sup> G <sup>D</sup> W <sup>L</sup> <sup>N</sup> <sup>G</sup> V R <sup>T</sup> A <sup>H</sup> C <sup>K</sup> <sup>E</sup> <sup>I</sup> <sup>F</sup> <sup>T</sup> <sup>G</sup> V <sup>E</sup> <sup>Y</sup> <sup>S</sup> <sup>S</sup> GACCAAAGCAATGTGGATATCTCTACCTTCCGCACAACAGGTGACTGGCTTAATGGTGTCCGGACAGCACACTGCAAGGAAATCTTCACGGGCGTGGAGTACAGTTC T 2 916 940 950 960 970 C <sup>D</sup> T <sup>I</sup> A K <sup>I</sup> <sup>S</sup> <sup>T</sup> <sup>D</sup> <sup>D</sup> M K K V G V <sup>T</sup> V V G <sup>P</sup> Q K K <sup>I</sup> <sup>I</sup> S <sup>S</sup> <sup>I</sup> K A L E T Q TGTGACACAATAGCCAAGATTTCCACAGATGACATGAAAAAGGTTGGTGTCACCGTGGTTGGGCCACAGAAGAAGATCATCAGTAGCATTAAAGC TCTAGAAACGCAA 3024 980 980<br>S K N G P V P V <sup>★</sup><br>CAAAGAATGCCCCAGTTCCCGTGTAAAGCACGACGGAAGTGCTTCTGGACGGAAGTGGTCGCTGTGGAAGCCGTCAAGTCATCCTGCAGACAATAATTCTGGA 3132

FIG. 1. Nucleotide sequence and deduced amino acid sequence of the HEK coding region with partial <sup>3</sup>' and <sup>5</sup>' untranslated sequence. Numbers at right indicate positions of nucleotides and numbers above amino acids refer to amino acid sequence. Single underline indicates the presumed signal peptide. Double underline indicates the presumed transmembrane region. Dashed overline indicates identity between the predicted amino acid sequence and the sequence obtained from purified HEK protein. Triangles indicate potential sites for N-linked glycosylation within the extracellular domain. Dots indicate the GXGXXG element of the putative ATPbinding site. Diamond indicates a putative autophosphorylation site. Asterisks indicate stop codons.

used as the template in dideoxynucleotide chain-termination sequence reactions (16). Sense and antisense oligonucleotide primers were used to complete sequencing with T7 DNA polymerase (Promega). Protein sequence alignment was performed using the GAP program (University of Wisconsin Genetics Computer Group).

Expression of HEK in COS Cells. The HEK 4.5 EcoRI insert was blunt ended with Klenow DNA polymerase <sup>I</sup> and dATP plus TTP, followed by ligation to BstXI adaptors. The adapted insert was ligated to BstXI-digested CDM8 (17). Sense and antisense constructs were prepared and transfected into COS cells using DEAE-dextran/chloroquine with dimethyl sulfoxide (17). Two days posttransfection, COS cells were stained with IIIA4 followed by fluorescein isothiocyanate-conjugated sheep anti-mouse immunoglobulin (Silenus, Australia) and examined under a fluorescence microscope.

Northern and Southern Blot Analyses of Cell Lines.  $Poly(A)^+$ -selected mRNA was isolated as described (18) and fractionated on a 1% formaldehyde agarose gel before transfer onto Hybond-C extra membrane (Amersham). Filters were probed with HEK 4.5 and subsequently with a glyceraldehyde-3-phosphate dehydrogenase insert as a control. DNA was prepared by lysis with guanidine hydrochloride (19), transferred to Zetaprobe membrane, and hybridized under conditions suggested by the manufacturer (Bio-Rad). To minimize cross-hybridization with other tyrosine kinases in Southern analysis of genomic DNA, PCR was used to generate <sup>a</sup> 1.1-kb HEK probe that spans <sup>a</sup> less highly conserved region of the molecule (nucleotides 1109-2241; see Fig. 1). The autoradiogram of the Southern blot was digitized using the MACSCAN program on a Macintosh IIx computer.

Scatchard Analysis of I1A4 Binding to Cell Lines. Binding of 125I-labeled IIIA4 to cell lines was performed in competition with unlabeled IIIA4 as described (20).

## RESULTS

Isolation and Characterization of cDNA Clones for HEK. One duplicating signal was obtained from screening  $\approx$  250,000 plaques of an LK63-derived  $\lambda$ gt10 cDNA library under relaxed conditions with a degenerate 51-mer oligonucleotide. This plaque contained a 2.5-kb insert (HEK 2.5), which hybridized with a single 5.5- to 6.0-kb mRNA species in Northern blot analysis of cell lines expressing HEK-i.e., LK63 and JM (data not shown). PCR using HEK 2.5 and oligonucleotide primers based on conserved motifs within the catalytic domains of tyrosine kinases (15) gave DNA products of the appropriate size (data not shown). These results

indicated HEK 2.5 was truncated at the <sup>5</sup>' end. HEK 2.5 was used to rescreen the library under more stringent conditions and <sup>a</sup> 4.5-kb HEK (HEK 4.5) clone was isolated. This clone hybridized with a degenerate oligonucleotide based on the N-terminal protein sequence and produced DNA bands of the predicted sizes in PCRs using the primers referred to above (data not shown). These data indicated that the 4.5-kb clone probably contained the complete HEK coding region.

The sequence of the coding region for HEK, together with partial <sup>3</sup>' and <sup>5</sup>' untranslated sequence, is shown in Fig. 1. An open reading frame of 2952 nucleotides extends from the initiation methionine at position 100 to the first termination codon at position 3051. Translation of the cDNA results in <sup>a</sup> predicted protein of 983 amino acids. There is identity between the amino acids obtained by sequencing the purified HEK protein and the predicted amino acid product of the  $cDNA$  clone (see Fig. 1). The predicted molecular mass of the translated protein (minus the putative signal peptide) is 92.8 kDa. This is in good agreement with previous results demonstrating a core protein of  $\approx$ 95 kDa in both tunicamycin- and endoglycosidase-treated LK63 cells (11).

The predicted protein product of the HEK cDNA clone has the features of a type la integral membrane protein (21). Two predominantly hydrophobic regions indicate a putative signal peptide (amino acids 1-20) and a transmembrane segment (amino acids 542-565). The extracellular domain of 521 amino acids contains five possible sites for N-linked glycosylation. The N-terminal region (amino acids 21-376) of the extracellular domain is rich in cysteine residues. The C-terminal region (amino acids 326-511) of the extracellular domain contains two repeats homologous to those found in fibronectin type III (10). The cytoplasmic domain (amino acids 566-983) of HEK contains the characteristic features of PTK catalytic domains, including a typical element of the ATPbinding site (GXGXXG), the DFG motif, and <sup>a</sup> putative site for tyrosine autophosphorylation at position 779 (2, 22).

HEK has the following overall protein sequence identity with each of the four fully sequenced members of the EPH/ELK family: chicken CEK <sup>5</sup> (10), 56.4%; rat ELK (23), 56.1%; rat ECK (24), 50.6%; human EPH (25), 42.3%. Protein sequence alignment between HEK and <sup>a</sup> close relative, ELK, is shown in Fig. 2. As expected, homology between these molecules is greatest within the catalytic domains. Outside the catalytic domains, there is conservation of a number of short motifs, which may have structural or functional significance. There is conservation of the number and spatial arrangement of cysteine residues within the extracellular domains of HEK, ELK, and other members of the EPH family (10). These cysteine residues cluster within the N-ter-



FIG. 2. Protein sequence alignment of HEK with ELK, <sup>a</sup> closely related gene within the EPH/ELK family. Alignment was performed with the GAP program. Amino acid positions are numbered on the right. Dots in the sequence indicate gaps introduced to optimize alignment. Dashes indicate identity between amino acids. Asterisks indicate stop codons. Dots above the line of amino acids indicate residues contributing to the two repeats of homology with fibronectin type 111 within the C-terminal regions of the extracellular domains. Triangles above the line of amino acids highlight conserved cysteine residues within the N-terminal region.



FIG. 3. Expression of HEK in COS cells. The HEK 4.5-kb cDNA clone was subcloned into the expression vector CDM8. COS cells were transfected with this construct using DEAE-dextran/chloroquine and dimethyl sulfoxide. Two days after transfection, cells were stained in situ with the HIMA4 mAb followed by fluorescein isothiocyanate-conjugated sheep anti-mouse immunoglobulin and photographed under light microscopy (A) or fluorescence microscopy (B).  $(\times 320.)$ 

minal portion of the extracellular domains (10). The C-terminal regions of the extracellular domains contain repeats homologous to those found in fibronectin type <sup>111</sup> (10). HEK has a cysteine in the C-terminal tail (amino acid 928), rather than the tyrosine, which is conserved in this position between other members of the EPH/ELK family. This may be significant in that phosphorylation of C-terminal tyrosine residues outside the catalytic domain can regulate tyrosine kinase activity (26). However, HEK has <sup>a</sup> C-terminal tyrosine at position 937 and a number of other cytoplasmic tyrosine residues, which may be in context for autophosphorylation (27).

Transfection and Expression of HEK in COS Cells. It was important to demonstrate that the cDNA clone isolated did indeed encode the molecule recognized by the IIIA4 mAb. HEK 4.5 was subcloned into the expression vector CDM8 and transfected into COS cells in both sense and antisense orientations. As shown in Fig. 3, COS cells transfected with HEK in the sense orientation stained specifically with IIIA4, confirming that our cDNA clone contains the full coding sequence and is identical to the molecule recognized by IIIA4. COS cells transfected with HEK in the antisense orientation did not stain with IIIA4 (data not shown).

Expression of HEK in Human Lymphoid Cell Lines. Cellsurface staining with IIIA4 revealed a highly restricted pattern of HEK expression on LK63 (a pre-B-cell line) and JM (a T-cell line). To further explore the expression of HEK, we





FIG. 4. Northern blot analysis of HEK expression in cell lines. Poly(A)+ RNA from huextra membrane. The filter was hybridized with the HEK 4.5-kb  $cDNA$  (*Upper*). The same filter was hybridized with glyceraldehyde-3-phosphate-dehydroge-GAPDH nase (GAPDH) as a quantitative control (Lower). REH, NALM-1, and FAKEM are pre-B-cell leukemic cell lines. BALL-1 is an early B-cell leukemic cell line. RAMOS is <sup>a</sup> mature B-cell leukemic cell line. HSB-2, HPB-ALL, and JM are T-cell leukemic cell lines. MW, molecular size markers (kb).

performed Northern blot analysis with HEK 4.5 (Fig. 4). A single 5.5- to 6.0-kb band was seen in both LK63 and JM cells. These results indicate that HEK 4.5 lacks 1-1.5 kb of untranslated sequence. However, there was a less intense band of the same size in another T-cell line-HSB-2-which did not stain with IIIA4. There were no HEK transcripts detected in <sup>a</sup> range of other lymphoid cell lines. We determined the number of HEK molecules on HSB-2, LK63/CD20<sup>+</sup>, and other cells by Scatchard analysis of HIIA4 mAb binding. The LK63/CD20<sup>+</sup> cells had  $\approx$ 15,000 sites per cell and JM cells had 9500 sites per cell. In contrast, HSB-2 had  $\approx$ 1070 sites per cell, which is too low for detection by immunofluorescence against the autofluorescence background of this cell line. The affinity constants for antibody binding were in the range of  $2.5-4.0 \times 10^9$ . Raji and K562 cells showed no detectable antibody binding above background.

Southern Blot Analysis. To investigate the basis for overexpression of HEK in the lymphoid tumor cell line LK63, we performed Southern analysis of genomic DNA (Fig. 5). A 1.1-kb fragment covering a less conserved region of  $HEK$  was used as a probe to minimize background arising from conserved regions of the catalytic domains of related tyrosine kinase molecules. Compared with normal peripheral blood mononuclear cell DNA, there is no apparent amplification or rearrangement of the HEK gene in the LK63 or LK63/CD20<sup>+</sup> tumor cell lines.

## DISCUSSION

We report the molecular characterization of <sup>a</sup> RTK designated HEK, which is expressed by a number of human lymphoid tumor cell lines. The HIIA4 mAb was used to characterize the biochemical structure and function of HEK on the human pre-B-cell tumor cell line LK63. The amino acid sequence was obtained from HEK protein, which had



FIG. 5. Southern blot analy-<br>sis of HEK in cell lines and normal human peripheral blood cell DNA. Samples were digested with  $Hind\overline{11}$  (lanes 1-3) or BamHI (lanes 4-6), run on a 1% agarose gel, and transferred to Zetaprobe membrane. The membrane was hybridized with a 1.1-kb fragment of HEK, extending from nucleotides 1109 to 2241 (see Fig. 1). Lanes: 1 and 4, normal peripheral blood; 2 and 5, LK63 cells; <sup>3</sup> and 6, LK63/ CD<sub>20</sub>+ cells.

been affinity purified using the IIIA4 mAb. A degenerate 51-mer oligonucleotide based on the HEK amino acid sequence was used to screen <sup>a</sup> cDNA library derived from LK63. A 4.5-kb cDNA clone containing the complete HEK coding region was isolated, sequenced, and expressed in COS cells. The predicted translation product of the HEK cDNA clone has the structural features of a membrane-associated, receptor-type tyrosine kinase that is homologous to the EPH/ELK family of PTKs (23, 25). Northern blot analysis showed restriction of HEK expression to the human lymphoid tumor cell lines LK63, JM, and HSB-2.

The intrinsic tyrosine kinase function and conservation of structural features of the tyrosine kinase domain have encouraged molecular strategies for the identification of additional PTKs. cDNAs encoding EPH (25) and related PTKs ELK (23), ECK (24), ERK and EEK (28), and CEK <sup>5</sup> (10) were identified through molecular approaches and form a subfamily of RTKs. The members of this family that have been fully characterized to date (EPH, ELK, ECK, CEK 5, and HEK) are type la integral membrane proteins with highly conserved tyrosine catalytic domains. The extracellular domains of these molecules characteristically contain an N-terminal region that is rich in strictly conserved cysteine residues and a C-terminal region containing two fibronectin type 111-like repeats (10). The functional significance of these structural features is not yet known.

In spite of highly conserved structural similarities, the pattern of expression for members of this family is quite distinct. A human EPH probe detected EPH mRNA transcripts in rat liver, lung, kidney, and testis and in human epithelial tumors (29). ELK expression was limited to rat brain and testis (23). ECK was expressed more widely, but it was most abundant in rat lung, skin, small intestine, and ovary as well as in epithelial cell lines (24). Human ERK and EEK probes identified transcripts in rat lung and brain tissue, respectively (28). CEK <sup>5</sup> protein was detected as early as day 2 in chicken embryos and at day 10 it was broadly expressed in embryonic tissues (10). Expression of CEK <sup>5</sup> in embryonic brain gradually declined with development, whereas a rapid fall in CEK 5 expression in developing muscle was closely correlated with muscle differentiation. In adult chickens, CEK <sup>5</sup> was most abundant in brain, with lower levels in liver, intestine, and gizzard. A notable feature is the lack of reported expression of members of this family in hemopoietic or lymphoid tissues. This fact, together with the sequence differences apparent from alignment of HEK with ELK, ECK, and CEK 5, suggests that HEK is probably not the human equivalent of these molecules but is a bona fide addition to the EPH/ELK RTK family. The high degree of sequence conservation seen across members of the EPH/ ELK RTK family may suggest shared functional activities. However, the varied patterns of expression may also imply tissue-specific functions and ligands.

Overexpression of EPH has been documented in colonic, lung, and breast carcinomas (25). There was no evidence for gene amplification or rearrangement of the EPH gene in these tumors. In addition, the sequence of EPH cloned from normal liver tissue was identical to that originally derived from a hepatoma cell line (29). However, the in vitro overexpression of EPH in NIH 3T3 cells produced tumors when these cells were injected into nude mice (30). These data suggest that the dysregulated expression of an otherwise normal member of the EPH/ELK RTK family may be oncogenic. Likewise, we found no evidence for HEK gene amplification or rearrangement in lymphoid tumor cell lines.

By analogy, aberrant HEK expression may contribute to the genesis of some human lymphoid tumors.

An extensive survey is required to document the pattern of HEK expression in normal and malignant human tissues. Also, whether overexpression of HEK gives rise to tumors in experimental animals remains to be determined. Given the expression of HEK in both pre-B and T-cell lines, HEK and its ligand may play physiological roles within the B- and T-cell lineages.

We thank David Gearing, Tracy Willson, Glenn Begley, Brian Seed, and Andrew Wilks for helpful advice and discussion. The assistance of Vladimir Brusic with computer sequence analysis is gratefully acknowledged. We thank June Evans for typing this manuscript. This work was supported by the National Health and Medical Research Council of Australia, the Anti-Cancer Council of Victoria, and the International Union Against Cancer.

- 1. Draetta, G., Piwnica-Worms, H., Morrison, D., Druker, B., Roberts, J. & Beach, D. (1988) Nature (London) 336, 738-744.
- 2. Hanks, S. K., Quinn, A. M. & Hunter, T. (1988) Science 241, 42-52.
- 3. Bishop, J. M. (1991) Cell 64, 235–248.<br>4. Bishop, J. M. (1983) Anny. Rev. Bioch
- 4. Bishop, J. M. (1983) Annu. Rev. Biochem. 52, 301–354.<br>5. Carnenter. G. & Cohen. S. (1990) J. Biol. Chem. 265.
- 5. Carpenter, G. & Cohen, S. (1990) J. Biol. Chem. 265, 7709- 7712.
- 6. Williams, L. T. (1989) Science 243, 1564-1570.
- 7. Yeung, Y. G., Jubinsky, P. T., Sengupta, A., Yeung, D. C. Y. & Stanley, E. R. (1987) Proc. Natl. Acad. Sci. USA 84, 1268-1271.
- 8. Yarden, Y., Kuang, W. J., Yang-Feng, T., Coussens, L., Munemitsu, S., Dull, T. J., Chen, E., Schlessinger, J., Francke, U. & Ullrich, A. (1987) EMBO J. 6, 3341-3351.
- 9. Mathews, W., Jordan, C. T., Wiegand, G. W., Pardoll, D. & Lemischka, I. R. (1991) Cell 65, 1143-1152.
- 10. Pasquale, E. B. (1991) Cell Regul. 2, 523-534.
- 11. Boyd, A. W., Ward, L. D., Wicks, I. P., Simpson, R. L., Salvaris, E., Wilks, A., Welch, K., Loudovaris, M., Rockman, S. & Busmanis, I. (1991) J. Biol. Chem., in press.
- 12. Salvaris, E., Novotny, J. R., Welch, K., Campbell, L. & Boyd, A. W. (1991) Leuk. Res., in press.
- 13. Martin, F. H., Castro, M. M., Aboul-ela, F. &Tinoco, I. (1985) Nucleic Acids Res. 13, 8927.
- 14. Gearing, D. P., Gough, N. M., King, J. A., Hilton, D. J., Nicola, N. A., Simpson, R. J., Nice, E. C., Kelso, A. & Metcalf, D. (1987) EMBO J. 6, 3995-4002.
- 15. Wilks, A. (1988) Proc. Natl. Acad. Sci. USA 86, 1603-1607.<br>16. Sanger, F., Nicklen, S. & Coulson, A. R. (1977) Proc. Natl.
- Sanger, F., Nicklen, S. & Coulson, A. R. (1977) Proc. Natl. Acad. Sci. USA 74, 5463-5467.
- 17. Seed, B. & Aruffo, A. (1987) Proc. Natl. Acad. Sci. USA 84, 3365-3369.
- 18. van Driel, I., Wilks, A. F., Pietersz, G. A. & Goding, J. W. (1985) Proc. Nail. Acad. Sci. USA 82, 8619-8623.
- 19. Bowtell, D. D. (1987) Anal. Biochem. **162, 463–465.**<br>20. Trucco, M. & de Petris, S. (1981) in *Immunological*
- Trucco, M. & de Petris, S. (1981) in Immunological Methods, eds. Leflovits, I. & Pernis, B. (Academic, New York), Vol. 2, pp. 1-26.
- 
- 21. Singer, S. J. (1990) Annu. Rev. Cell Biol. 6, 247–296.<br>22. Hanks. S. K. & Ouinn. A. M. (1991) Methods Enzyr 22. Hanks, S. K. & Quinn, A. M. (1991) Methods Enzymol. 200, 38-62.
- 23. Lhotak, V., Greer, P., Letwin, K. & Pawson, T. (1991) Mol. Cell. Biol. 11, 2496-2502.
- 24. Lindberg, R. A. & Hunter, T. (1990) Mol. Cell. Biol. 10, 6316-6324.
- 25. Hirai, H., Maru, Y., Hagiwara, K., Nishida, J. & Takaku, F. (1987) Science 238, 1717-1720.
- 26. Cantley, L. C., Auger, K. R., Carpenter, C., Duckworth, B., Graziani, A., Kapeller, R. & Soltoff, S. (1991) Cell 64, 281-302.
- 27. Ullrich, A. & Schlessinger, J. (1990) Cell 61, 203-212.<br>28. Chan. J. & Watt. V. M. (1991) Oncogene 6, 1057-106
- 28. Chan, J. & Watt, V. M. (1991) Oncogene 6, 1057-1061.<br>29. Maru, Y., Hirai, H., Yoshida, M. C. & Takaku, F. (1981)
- 29. Maru, Y., Hirai, H., Yoshida, M. C. & Takaku, F. (1988) Mol. Cell. Biol. 8, 3770-3776.
- 30. Maru, Y., Hirai, H. & Takaku, F. (1990) Oncogene 5,445-447.