Proc. Natl. Acad. Sci. USA
Vol. 88, pp. 4897–4901, June 1991
Cell Biology

Characterization of a member of the immunoglobulin gene superfamily that possibly represents an additional class of growth factor receptor

(protein-tyrosine kinase)

YAH-HUEI CHOU AND MICHAEL J. HAYMAN

Department of Microbiology, Health Science Center, State University of New York at Stony Brook, Stony Brook, NY 11794

Communicated by Peter K. Vogt, March 12, 1991 (received for review January 11, 1991)

ABSTRACT **We have screened cDNA libraries prepared from embryonic chicken tissues to isolate additional genes encoding growth factor receptors. Nucleotide sequencing of a cDNA encoding a gene, which we have termed *klg*, revealed it to represent an additional member of the immunoglobulin gene superfamily, which also possesses extensive sequence similarity to the protein-tyrosine kinase growth factor receptor genes. The *klg* gene was shown to encode a 140-kDa glycoprotein. However, the sequence of the tyrosine kinase domain is unusual in that the aspartate residue located within the highly conserved Asp-Phe-Gly triplet is replaced by an alanine residue. The presence of this aspartate has previously been found to be essential for tyrosine kinase activity. Consistent with the replacement of this aspartate, we were unable to detect any evidence of an associated kinase activity with the *klg*-encoded protein. These observations raise the possibility that the *klg* gene product represents a newly discovered class of receptor that plays a role in signal attenuation rather than signal propagation.**

Cell growth and differentiation is controlled by a complex network of cell–cell and cell–growth factor interactions. Polypeptide growth factors activate target cells by binding to cell-surface receptors, after which the signal is transduced across the cell membrane. Although the biochemical details of these receptor-regulated pathways vary among different signaling systems, the cell-surface receptors themselves can be divided into families of structurally related molecules. The receptor protein-tyrosine kinases (PTKs) represent the largest and best characterized of these cell-surface receptor gene families (1). This family of growth factor receptors comprises an extracellular ligand-binding domain that is linked through a transmembrane domain to a cytoplasmic catalytic domain, which functions as a ligand-stimulated tyrosine kinase.

The PTK catalytic domains of these enzymes are very highly conserved and can be divided into 11 structurally conserved subdomains (2). Mutagenesis studies have shown that within these domains certain invariant amino acid residues are essential for the kinase activity of these enzymes. For example, the lysine residue that is involved in binding of ATP is essential for kinase activity, as are the two aspartate residues found within either the His-Arg-Asp (HRD) or Asp-Phe-Gly (DFG) triplets (3). Biochemical studies have addressed the mechanisms by which the tyrosine kinase activity of these receptors is activated. The formation of dimers on ligand binding has been implicated as playing a critical role (1, 4). Thus, experimentally, when dimers are formed between an inactive and an active kinase, the subsequent signal to the cell is impaired, implying that intermolecular interactions are important for signal transmission. Recently, receptors have been described that could function as the inactive partner in such interactions: for example, the forms of the trkB protein that contain only the extracellular domain (5) or the kinase-inactive versions of the c-kit receptor (6). These molecules represent candidates for proteins with the potential for playing roles in signal attentuation.

We have isolated a gene with a structure reminiscent of that of a growth factor receptor but with an unusual structural change in the kinase domain. The DFG triplet is replaced by the sequence Ala-Leu-Ser (ALS). The other 10 kinase subdomains are conserved. In addition to this unusual change within the "kinase" domain, this protein is a member of the immunoglobulin gene superfamily having seven immunoglobulin-like loops in its extracellular domain. We have given this gene the designation *klg* (*k*inase-*l*ike *g*ene) until a more functional designation can be made. In this report, we describe the isolation and characterization of this gene.*

MATERIALS AND METHODS

Isolation of *klg* cDNA Clones. A 10-day-old chicken embryonic brain λgt10 cDNA library was kindly provided by J. Levy (7) and a λgt11 cDNA library prepared from 10-day-old chicken embryos was kindly provided by B. Vennstrom (8). The 1.0-kilobase (kb) *Pst* I/*Pvu* II DNA fragment of the v-*sea* oncogene (9) was used as a probe.

Nucleotide Sequencing. DNA sequence was determined by dideoxynucleotide chain termination methods with modified T7 DNA polymerase (10). Most cDNAs were sequenced from both single-stranded M13 phage DNA and double-stranded plasmid DNA.

RNA Analysis. Total cellular RNA was isolated from embryonic or 14-day-old chicken tissues and cell lines by the guanidine thiocyanate and $CsCl_2$ procedure (11). Poly(A)-containing mRNA was electrophoresed on agarose/formaldehyde gels, transferred to GeneScreen nylon membranes (DuPont), and hybridized with a *klg*-specific probe.

Antisera. klg restriction enzyme fragments were subcloned into trpE expression vectors and overexpressed as trpE fusion peptides in the C600 strain of bacteria as described (12). Polyclonal rabbit antisera were then raised against these trpE fusion peptides.

Metabolic Labeling, Immune Precipitation, and Western Blotting. Metabolic labeling and immune precipitation were as described (13). For determination of the glycosylation state of the klg proteins, cells were labeled in the presence of tunicamycin (5 μg/ml; Sigma). Western blotting of immunoprecipitates was as described (14).

***In Vitro* Kinase Assays.** Cells were extracted and immunoprecipitated as described above and the immune complexes

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.

Abbreviation: PTK, protein-tyrosine kinase.
*The sequence reported in this paper has been deposited in the GenBank data base (accession no. M63437).

```
AATTCCGCGGCGGTGGCGGTAGCGGCATGGCTCGGGCGCGGGGCGCGCTGTGAGCGGCGCGGGGCGAGGATGGCGGCGCTGCGGGCGCTGCTGCTGCTGCTCGCGGTGGGGGCTCAGGCA   120
                                                                           M  A  A  L  R  A  L  L  L  L  L  A  V  G  A  Q  A    17

GCGATCCGTTTCGCCAAGGAGCCGTACTCGCAGGACGCCCTGCATGGCCGCAGTGCCATCCTCCGCTGCGAGGTGGAGGAGCCGGCCCACGTGGAGTTTGAGTGGCTGCAGAACGGGCTC   240
 A  I  R  F  A  K  E  P  Y  S  Q  D  A  L  H  G  R  S  A  I  L  R  C  E  V  E  E  P  A  H  V  E  F  E  W  L  Q  N  G  L    57

CCCATCCAGGACACGGAGCAGCGCTTCAAGGAAGGCAGCAACCTCCAGTTTGCTGCTGTCGACCGGCATCGGGATGCTGGGAGCTTCCAGTGCGTGGCAAGGAACGTGCAGACCGGGGAG   360
 P  I  Q  D  T  E  Q  R  F  K  E  G  S  N  L  Q  F  A  A  V  D  R  H  R  D  A  G  S  F  Q  C  V  A  R  N  V  Q  T  G  E    97

GAGGCTCGCACGGCCAACGCGTCCTTCAATATCAAGTGGATGGAGACGGGCAGCGTGGTGCTGAAGCAGCCGGCCAGCGCAGCTGAGATCCAGCCCTCCTCCACGGTGGTGCTGCGGTGT   480
 E  A  R  T  A  N  A  S  F  N  I  K  W  M  E  T  G  S  V  V  L  K  Q  P  A  S  A  A  E  I  Q  P  S  S  T  V  V  L  R  C   137

CACATCGACGGCCACCCGCGCCCCACGTGGCAGTGGTTTCGGGACGGCGCCCCGCTCCCCGACGGCCGCGGCACCTATTCTGTCAGCAGCAAGGAGCGGACCCTCACCCTGCGCGGCGCC   600
 H  I  D  G  H  P  R  P  T  W  Q  W  F  R  D  G  A  P  L  P  D  G  R  G  T  Y  S  V  S  S  K  E  R  T  L  T  L  R  G  A   177

GGCCCCGATGACAACGGTCTCTACTACTGCTCGGCCCGGCCAAGGGCCGTTGGCTCCGTGTGCAGCCAGGACAACTTCACGCTGAACATCATCGATGAGAGCTTCCCCCAGGCGGTGGTC   720
 G  P  D  D  N  G  L  Y  Y  C  S  A  R  P  R  A  V  G  S  V  C  S  Q  D  N  F  T  L  N  I  I  D  E  S  F  P  Q  A  V  V   217

GTCCCGGAGGACCTCATCGTGACCAAGAACGAAGAGGCCATGTTTGATTGCCAGTTTGCAGCCGTGCCCCCTCCCACGCAGGAGTGGCTCTTCGAAGACAGCCCCATCACCAACAGATCC   840
 V  P  E  D  L  I  V  T  K  N  E  E  A  M  F  D  C  Q  F  A  A  V  P  P  P  T  Q  E  W  L  F  E  D  S  P  I  T  N  R  S   257

AAGACAACTGTGTTTGCCAATGGCTCCCTGCTGATCACCCAGGTGAAGGCCCGCAGCACCGGTGTCTACAAGTGCATTGGCCATGGGCAGAAGGGGAAAGCTCTCGTGCTGAAGGCGACT   960
 K  T  T  V  F  A  N  G  S  L  L  I  T  Q  V  K  A  R  S  T  G  V  Y  K  C  I  G  H  G  Q  K  G  K  A  L  V  L  K  A  T   297

CTGCGGCTGGCAGAGATCGAAGAGATGGCACCCTTCTCCCCGAAGGTGTTGACAGCGAACCAGGGACACCGTGTGTCCTGTGCCTGCCCGCGGGGTGTACCCACGCCCCAGGTCTGGTGG  1080
 L  R  L  A  E  I  E  E  M  A  P  F  S  P  K  V  L  T  A  N  Q  G  H  R  V  S  C  A  C  P  R  G  V  P  T  P  Q  V  W  W   337

GAGAGAAACCAGGAGCGGGTCCCCACTGCTGGCAGGGTGTACCAAGAGGCAGAGCAGCTCGTTTTCACTAGCATCACCGAGGCAGATGCGGGGATCTATACATGCCATGCTGCCAACAAG  1200
 E  R  N  Q  E  R  V  P  T  A  G  R  V  Y  Q  E  A  E  Q  L  V  F  T  S  I  T  E  A  D  A  G  I  Y  T  C  H  A  A  N  K   377

GCTGGGGAGAAGAAACAGGAGCTGAGCATCACCGTGGCTACGGTACCCAAGTGGGTGGAGATGCCCAAGGACAGCCAGCTGGAGGAGAGCAAGCCAGGATACCTGCATTGCCTCAGCAAG  1320
 A  G  E  K  K  Q  E  L  S  I  T  V  A  T  V  P  K  W  V  E  M  P  K  D  S  Q  L  E  E  S  K  P  G  Y  L  H  C  L  S  K   417

GCCTCCCTGAAACCTACAGTCACCTGGTACCGCAACGGCGTCTCCATCTCCGAGGACTCACGCTTTGAAATCTCAGAGAATGGGACACTGCGCATCAACAACGTGGAGGTGTACGATGGG  1440
 A  S  L  K  P  T  V  T  W  Y  R  N  G  V  S  I  S  E  D  S  R  F  E  I  S  E  N  G  T  L  R  I  N  N  V  E  V  Y  D  G   457

ACGATGTACAAGTGTGTGAGCAGCACGCCGGCAGGCAGCATTGAGGGTTATGCACGGGTACATGTGCTAGAGAAGTTAAAGTTCACTCCTCCACCCCAGCCACTGCAGTGCATGGAGTTC  1560
 T  M  Y  K  C  V  S  S  T  P  A  G  S  I  E  G  Y  A  R  V  H  V  L  E  K  L  K  F  T  P  P  P  Q  P  L  Q  C  M  E  F   497

AACAAGGAGGTTACGGTGTCCTGCTCAGCTACTGGCCGGGAAAAACCCACCATCCAGTGGACTAAAACAGATGGGAGCAGCCTGCCATCCCATGTCAGCCACAGAGCTGGCATCTTGTCC  1680
 N  K  E  V  T  V  S  C  S  A  T  G  R  E  K  P  T  I  Q  W  T  K  T  D  G  S  S  L  P  S  H  V  S  H  R  A  G  I  L  S   537

TTCCACAAGGTGAGCAGGAGCGACTCAGGGAACTACACGTGCATTGCCTCCAACAGCCCGCAGGGTGAGATCCGTGCCACCGTGCAGCTGGTGGTGGCAGTTTATGTCACCTTCAAGCTG  1800
 F  H  K  V  S  R  S  D  S  G  N  Y  T  C  I  A  S  N  S  P  Q  G  E  I  R  A  T  V  Q  L  V  V  A  V  Y  V  T  F  K  L   577

GAGCCGGAGCCCACAACCGTGTACCAGGGGCACACGGCCATGTTCCAGTGCCAGGCAGAGGGGGACCCCGTGCCACATATCCAATGGAAAGGGAAGGACAAGATTTTGGATCCCAGCAAG  1920
 E  P  E  P  T  T  V  Y  Q  G  H  T  A  M  F  Q  C  Q  A  E  G  D  P  V  P  H  I  Q  W  K  G  K  D  K  I  L  D  P  S  K   617

CTTCTGCCCAGGATTCAGATCATGCCCAATGGCTCCCTGGTCATTTACGACGTCACCACCGAGGACTCTGGGAAGTACACCTGTATTGCTGGCAACAGCTGCAACATCAAGCACCGCGAG  2040
 L  L  P  R  I  Q  I  M  P  N  G  S  L  V  I  Y  D  V  T  T  E  D  S  G  K  Y  T  C  I  A  G  N  S  C  N  I  K  H  R  E   657

GCCTTCCTCTATGTTGTAGACAAGCCGGCAGCAGAAGAAGATGAGGGACCCAGCAGCCACACGCCCTACAAGATGATCCAGACCATTGGGCTCTCAGTGGGAGCAGCTGTCGCCTACATC  2160
 A  F  L  Y  V  V  D  K  P  A  A  E  E  D  E  G  P  S  S  H  T  P  Y  K  M  I  Q  T  I  G  L  S  V  G  A  A  V  A  Y  I   697

ATTATCGTGCTGGGGCTGATGTTCTACTGCAAGAAGCGGAGGAAAGCCAAGAGGCTGAAGAAGCACCCGGAGGGCGAGGAGCCCGAAATGGAGTGTCTGAATGGTGGTACTTTGCTGCAG  2280
 I  I  V  L  G  L  M  F  Y  C  K  K  R  R  K  A  K  R  L  K  K  H  P  E  G  E  E  P  E  M  E  C  L  N  G  G  T  L  L  Q   737

AATGGGCAGACGACGGCGGAGATCCAGGAGGAGGTGGCCTTGACCAACCTGGGCAGCAGCTCTGGGGCCAGCAAGCGGCACAGCGCCCGTGACAAGATGCACTTCCCCCGTTCCAACCTG  2400
 N  G  Q  T  T  A  E  I  Q  E  E  V  A  L  T  N  L  G  S  S  S  G  A  S  K  R  H  S  A  R  D  K  M  H  F  P  R  S  N  L   777

CAAACAATCACTACCTTGGGTAGGGGGGAGTTTGGTGAAGTCTTCCTGGCCAAGGCAAAAGGTGCAGAGGATGCCGAGGGCGAAGCGCTGGTGCTGGTGAAGAGCCTGCAGACAAGGGAT  2520
 Q  T  I  T  T  L  G  R  G  E  F  G  E  V  F  L  A  K  A  K  G  A  E  D  A  E  G  E  A  L  V  L  V  K  S  L  Q  T  R  D   817

GAGCAGCTGCAGCTGGACTTCAGGCGCGAGGCAGAGATGTTTGGCAAGCTGAACCACGTCAACGTGGTGCGGCTGCTGGGGCTGTGCCGCGAGGCTGAGCCACACTACATGGTCCTGGAG  2640
 E  Q  L  Q  L  D  F  R  R  E  A  E  M  F  G  K  L  N  H  V  N  V  V  R  L  L  G  L  C  R  E  A  E  P  H  Y  M  V  L  E   857

TATGTGGACTTGGGGGACCTGAAGCAGTTCCTGAGGATCTCCAAGAGCAAGGATGAGTCCCTGAAGCCACAGCCCCTCAGCACCAAGCACAAGGTGTCTCTCTGCACGCAGGTGGCCCTG  2760
 Y  V  D  L  G  D  L  K  Q  F  L  R  I  S  K  S  K  D  E  S  L  K  P  Q  P  L  S  T  K  H  K  V  S  L  C  T  Q  V  A  L   897

GGGATGGAGCACCTCTCCAACGGCAGGTTCGTGCACCGGGACTTGGCGGCCAGGAACTGCCTGGTGAGCGCCCAGAGGCAGGTCAAGGTCTCAGCCCTGAGCCTCAGCAAGGACGTGTAC  2880
 G  M  E  H  L  S  N  G  R  F  V  H  R  D  L  A  A  R  N  C  L  V  S  A  Q  R  Q  V  K  V  S  A  L  S  L  S  K  D  V  Y   937

AACAGCGAGTACTACCACTTCCGCCAGGCCTGGATCCCGCTGCGCTGGATGCCGCCCGAAGCCGTGCTGGAGGACGAGTTCTCCACCAAGTCAGACGTGTGGTCCTTCGGGGTGCTCATG  3000
 N  S  E  Y  Y  H  F  R  Q  A  W  I  P  L  R  W  M  P  P  E  A  V  L  E  D  E  F  S  T  K  S  D  V  W  S  F  G  V  L  M   977

TGGGAGGTGTTCACGCAGGGGGAGATGCCGTACGCCCCGCTGGCGGATGACGAGGTCCTAGCAGGTCTGAAGTCGGGAAAGACGAAGCTCCCGCAGCCCGAGGGCTGCCCTTCACGCCTC  3120
 W  E  V  F  T  Q  G  E  M  P  Y  A  P  L  A  D  D  E  V  L  A  G  L  K  S  G  K  T  K  L  P  Q  P  E  G  C  P  S  R  L  1017

ACCAAGCTGATGCAGCGCTGCTGGGCGCCCAGCCCCAAGGACCGACCCTCCTTCAGCGAGCTCGCTGCCGCCCTCGGGGACAGCCCGGCCGACAGCAAAGCCTGAGCACCTTGCTGGGGA  3240
 T  K  L  M  Q  R  C  W  A  P  S  P  K  D  R  P  S  F  S  E  L  A  A  A  L  G  D  S  P  A  D  S  K  A  .              1051

AGGATGCGGCAGGCTGGGTGGGGGCTCTGCGGCTGCCCCCGGTGCTCCTGCGGCCTTCGCTGAGCCCGACCCAACGCACGGCACCAGCGGCTGCTGACTCAAAGCACAAGCCAGTGTACT  3360
TCCCCCACACACCCCCTCTCCTCCGCGGGTCCAAAAAGCCGCGGGCTGCTGGCACCGGGCTCTCCCTCCCCGCTCGCCACGTGCATATCTGTCCGTCCGTCTGTGCTAGGCATGCCTGCG  3480
GGCACGGGCCCGCCATCAGGAACACGAAGTGCCTGGCATATGAAGGAGACGCTCCACCACCTTTCCCCACATCCCTTCCCCGTGCCTTTTTTTGTACGACTCCTGGGGGCAGCGCTCTCC  3600
TCCTCCTCCCCCTCCCCGGCGCGCTGCCTCCTCCCCACCATAGTATTTCTGTTACTGGAACGCAGCCTAGAGTGGGTTAGTTTGTTTGCTTGAGCGGAGGGTACCGCAGCCAAGAATGGT  3720
TGGGTTTTTGTTGGGTTTATATGTGCTGCTCTCAATAAAGAAGGCTTTTTTTTTTTAAAAAAAAAAA
```

FIG. 1. Nucleotide sequence and deduced amino acid sequence of the *klg* cDNA clone 13-10. The nucleotides are numbered on the right-hand side beginning with the 5′ region of the clone 13-10 and ending at base pair 3788. The amino acid sequence shown begins at the first ATG in the nucleotide sequence and the amino acid numbers are shown below the nucleotide numbers. The hydrophobic stretches, which include putative signal peptide and transmembrane domain, are double underlined. Cysteine residues involved in immunoglobulin domain formation are marked by circles above the appropriate residues and the consensus N-glycosylation sites are underlined. The position of the ALS sequence, which replaces the highly conserved amino acid residue sequence DFG, is indicated by asterisks above the amino acid residues.

were resuspended in 30 μl of kinase reaction buffer {10 μCi of [γ-^{32}P]ATP (3000 Ci/mmol; 1 Ci = 37 GBq; Amersham), 10 mM Tris·HCl (pH 7.4), and either 5 mM $MgCl_2$ or 5 mM $MnCl_2$} supplemented with various concentrations of unlabeled ATP from 0 to 300 μM.

***In Vitro* Transcription and Translation.** The 13-10 cDNA clone was subcloned into the *Eco*RI site of the pBluescript vector, linearized by *Sma* I digestion, capped, and transcribed from the T3 promoter. Translation of this capped RNA was in a micrococcal nuclease-pretreated rabbit reticulocyte according to the manufacturer's instructions (Promega).

DNA Transfection. The *Eco*RI fragment from *klg* cDNA 13-10 was subcloned into the pMT2 mammalian expression vector kindly provided by Genetics Institute (15) and then transfected into 10^6 COS M6 monkey cells by electroporation. The cell lysates were harvested 48 hr after transfection and *klg* expression was assessed by immunoprecipitation as described above.

RESULTS

Isolation of *klg* cDNAs and Nucleotide Sequence Analysis. The transforming gene of S13 avian erythroblastosis virus, v-*sea*, is a tyrosine kinase (9, 16). We decided to isolate the c-*sea* gene and other *sea*-related genes by screening cDNA libraries with a v-*sea*-specific probe. Initially, three overlapping clones containing *Eco*RI insert sizes of 1.0 (clone 20), 1.2 (clone 18), and 1.8 (clone 21) kb were isolated from an initial screen of 6 × 10^5 recombinant λgt10 plaques derived from an embryonic brain cDNA library. Northern blot analysis indicated that these clones hybridize to a 4.0-kb mRNA (see below and Fig. 4*A*). To isolate full-length clones, a chicken embryonic fibroblast cDNA library was screened. Two additional clones were isolated, clone 19 and clone 13-10. Clone 13-10 contained an ≈4.0-kb *Eco*RI insert. These positive cDNA clones showed overlapping restriction enzyme digestion patterns that were significantly different from that of the v-*sea* oncogene. These data indicated that these clones most likely represented *sea*-related genes rather than the c-*sea* gene. To confirm this, their nucleotide sequences were determined.

The cDNAs were found to be derived from the same gene. The complete nucleotide sequence of the clone 13-10 cDNA extends for 3788 nucleotides and contains an open reading frame of 1051 amino acid residues flanked by 69 nucleotides of 5′ untranslated sequences and 563 nucleotides of 3′ untranslated sequences (Fig. 1). A polyadenylation signal (AATAAA) and a poly(A) tail are found at the 3′ end of this cDNA. An ATG codon that could potentially serve as an initiator of protein synthesis is located at position 70 and is followed by the open reading frame. This open reading frame is predicted to encode a protein with a molecular mass of 115,600 Da that has several features in common with cell-surface growth factor receptors. It contains an NH_2-terminal hydrophobic leader sequence (residues 1–22), a single membrane-spanning region (residues 682–707), and an intracellular PTK-like domain (residues 775–1051). Seven extracellular immunoglobulin-like loops are found in the extracellular domain that conform to the consensus of a cysteine residue followed in 11 or 12 residues by a tryptophan and then another cysteine residue ≈50 residues after the first cysteine located within the consensus sequence DXGXYXC (residues 40–88, 137–187, 234–282, 324–372, 414–462, 505–551, and 594–645). Seven potential N-linked glycosylation sites, NXS/T, were also found (residues 103, 202, 255, 264, 444, 548, and 627). Within the predicted cytoplasmic domain, the deduced amino acid sequence contains sequences motifs considered to be hallmarks of a PTK (2). Fig. 2 shows a comparison of this region of *klg* with that of the tyrosine

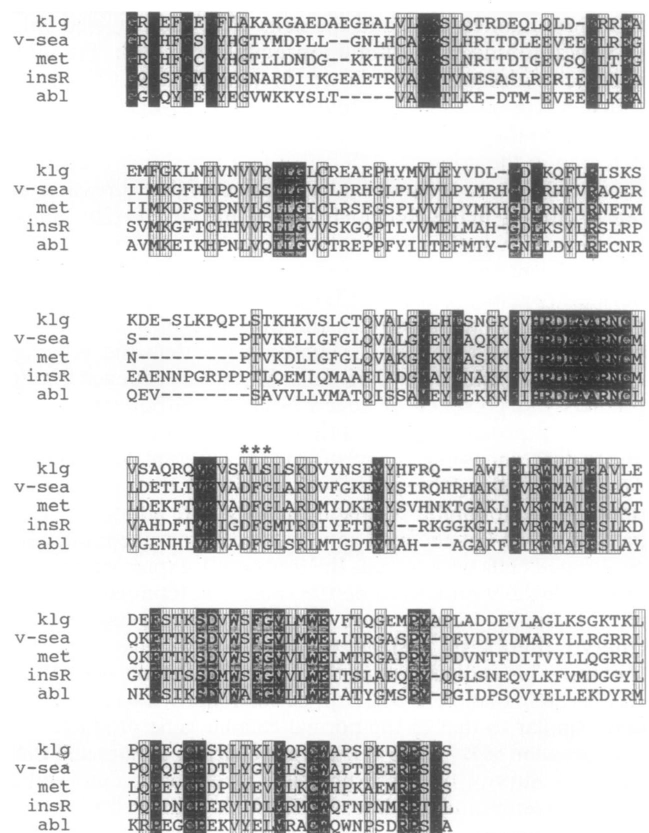

FIG. 2. Comparison of the amino acid sequence of the klg protein with the corresponding kinase domains of v-*sea*, c-*met*, insulin receptor, and c-*abl*. A minimal number of gaps were introduced to optimize alignment. The dark shaded box represents identical residues in all of the proteins shown here. The stippled box represents conservative residues shared by these proteins. Asterisks indicate the positions of the ALS residues found in the *klg* gene.

kinase domains of the avian v-*sea* (9), and human *MET* (17), insulin receptor (18), and *ABL* (19) genes. The best sequence similarity was with v-*sea* but even this only amounted to 39%. These similarities include the ATP binding domain (residues 784–792), with the consensus lysine at residue 811, the HRDLAARN sequence unique to tyrosine kinases (residues 909–916), and two tyrosine residues in equivalent positions to those found in PTK receptors (residues 941 and 942). In addition, it appears to contain a small 8-amino acid insert within the kinase domain (residues 875–882). However, despite this strong similarity to PTKs, it lacks the highly conserved amino acid triplet DFG in its catalytic domain and this is replaced by the sequence ALS (residues 929–931). Altogether five different cDNA clones for this gene were isolated from two different cDNA libraries. Sequencing confirmed that they all contained the ALS sequence, ruling out the possibility that this difference represents a cloning and/or sequencing artifact.

Characterization of the *klg* Gene Product. To characterize the product of the *klg* cDNA clone 13-10, a subclone containing the entire 3.8-kb insert in the pBluescript vector was used to make *in vitro* transcripts that were then translated in a rabbit reticulocyte lysate. An ≈120-kDa protein that could be immunoprecipitated by a *klg*-specific antiserum (SB98) was detected (Fig. 3, lane 1). The size of this *in vitro*-synthesized molecule was as predicted for the nonglycosylated polypeptide, assuming translation was initiated at the ATG at position 70. We also determined that the 13-10 cDNA clone could direct the synthesis of a protein after transient expression in COS cells. The cDNA was transfected into

COS M6 monkey cells and the cells were radiolabeled with [^{35}S]methionine and immunoprecipitated with the *klg*-specific antiserum. As shown in lane 9, the COS cells transiently expressed high levels of an ≈135-kDa protein in addition to the 120-kDa protein. These proteins could not be detected in the control COS cells (data not shown). Expression studies (see below) had identified high expression of mRNA for *klg* in the chicken liver cell line DU249 (20). Thus, to determine whether the products seen by *in vitro* translation and transient expression were the same molecular size as the normal *klg* gene product, the DU249 cells were also analyzed. Fig. 3 (lane 7) demonstrates the presence of a *klg*-specific protein of 140 kDa that is slightly larger than the proteins expressed transiently in COS cells. Since the sequence for *klg* predicts the presence of seven N-linked carbohydrate side chains, it is possible that the differences in sizes observed were due to differences in glycosylation in the different cell systems. To examine this, the cells were grown in the presence of tunicamycin to inhibit N-linked carbohydrate addition. Comparison of the *klg*-encoded gene products from the tunicamycin-treated cells now reveals the transiently expressed protein and the DU249 protein to be the same size (compare lanes 5 and 3). Similar results were obtained with normal chicken embryo fibroblasts instead of the DU249 cells (data not shown). This indicates that the *klg* cDNA clone encodes a glycoprotein that has a polypeptide sequence with a molecular mass similar to that of the normal cellular gene product.

Expression of *klg* Gene in Different Chicken Tissues and Cell Lines. A survey of *klg* expression in different embryonic chicken tissues and cell lines by RNA slot blot hybridization demonstrated expression at a relatively low level in a variety of tissues and cell lines, including bone marrow, spleen, bursa, thymus, and brain (data not shown). The *klg* gene was expressed at slightly higher amounts in fibroblasts and at significantly higher amounts in the MC29-transformed liver cell line DU249 (data not shown). To confirm and extend this analysis, Northern blots were prepared, which identified a 4.0-kb mRNA present in both chicken embryo fibroblasts and DU249 (Fig. 4*A*, lanes 1 and 2). Analysis of several experiments indicated the level of expression of the mRNA in DU249 cells to be 3–5 times higher than that found in the chicken embryo fibroblasts. However, no mRNA could be detected in normal liver cells from adult animals (lane 3).

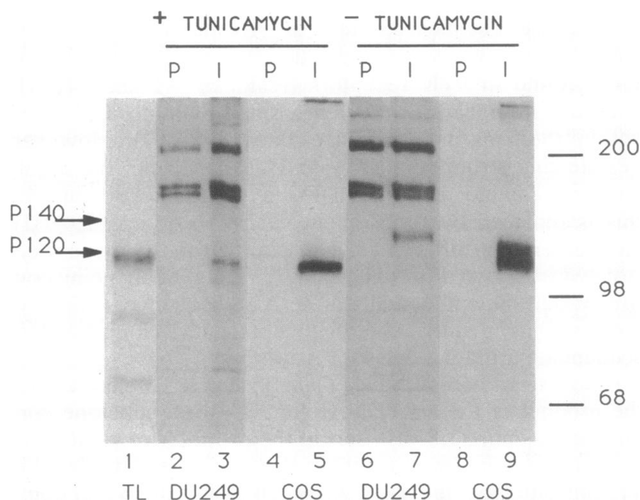

FIG. 3. Comparison of immunoprecipitation of klg protein from the DU249 cells and COS cells. Immunoprecipitations were done with preimmune serum (lanes P) or *klg*-specific antiserum (lanes I). Lanes: 1, *in vitro* translated products from reticulocyte lysates primed with *in vitro* transcribed RNA from clone 13-10; 2–5, DU249 or COS cells treated with tunicamycin (5 μg/ml); 6–9, untreated DU249 and COS cells.

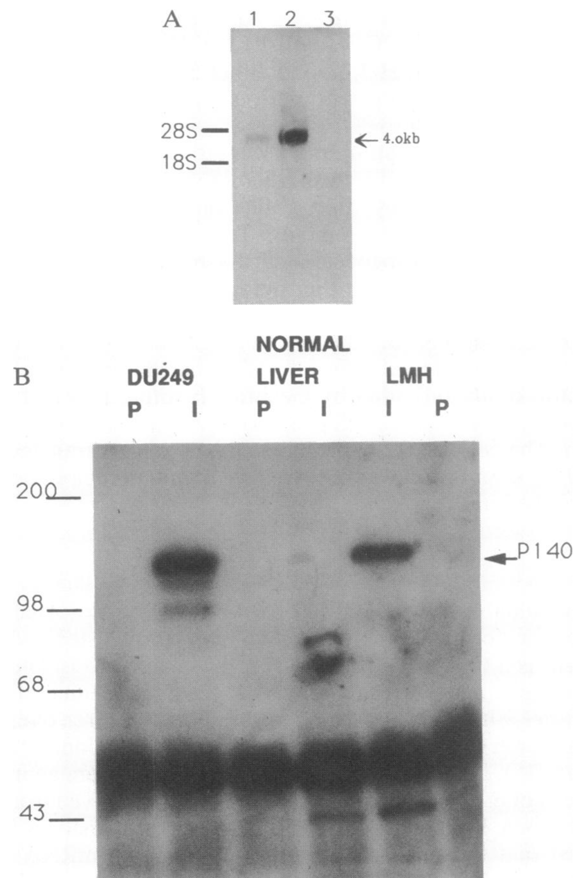

FIG. 4. (*A*) Northern blot analysis. Poly(A)$^+$ RNA was prepared from chicken embryo fibroblasts, the MC29-transformed chicken liver cell line (DU249), and normal chicken liver cells. The RNA was electrophoresed and then hybridized with a *klg*-specific probe. Lanes: 1, 10 μg of chicken embryo fibroblast RNA; 2, 10 μg of DU249 RNA; 3, 30 μg of normal liver RNA. (*B*) Detection of the *klg* gene product in transformed and normal liver cells. Total protein lysate (820 μg) was used for immunoprecipitation and the immune complexes were then analyzed by Western blotting with *klg*-specific antiserum. Bound immunoglobulin was detected with ^{125}I-labeled protein A. P, preimmune serum; I, immune serum.

Steady-state levels of klg protein in the DU249 cell line, normal adult liver, and the chemically transformed liver cell line LMH (21) were also determined. Cell lysates, normalized for equal amounts of protein, were first immunoprecipitated with either normal or *klg*-specific antiserum. The immune complexes were then run on SDS/polyacrylamide gels and Western blotted with the *klg*-specific antiserum (this assay was found to be more sensitive than analysis of whole cell lysates). Fig. 4*B* demonstrates that the DU249 and LMH liver cell lines express easily detectable amounts of klg protein but, even on the long exposure shown, no klg protein could be detected in the normal adult liver sample.

Lack of Detectable Enzymatic Activities Associated with *klg*. Although the *klg* gene shows extensive sequence similarity to PTK growth factor receptors, we have been unable to detect any kinase activity *in vitro* or phosphorylation on tyrosine *in vivo* associated with this protein isolated either from normal chicken embryo fibroblasts or from the DU249 cell line. The pp60$^{v\text{-}src}$ protein was used as a positive control in these experiments and kinase activity could be detected. Attempts to label the protein *in vivo* in DU249 cells with either $^{32}P_i$ or inorganic [^{35}S]sulfate demonstrated that the klg protein could be labeled but at levels so low that it precluded further analysis of the phosphorylated amino acid or determination of whether the sulfate was covalently attached to tyrosine or

carbohydate. Because of the proposed sequence similarities between kinases and sulfur transferases (22), we tested the ability of the klg protein to function as a tyrosylprotein sulfotransferase using ^{35}S-labeled 3′-phosphoadenosine 5′-phosphosulfate as cosubstrate, but again the results were negative. Thus, although it is possible that the klg protein has unusual reaction condition requirements that we have been unable to provide, we consider it more likely that the klg protein does not function as a tyrosine kinase.

DISCUSSION

The nucleotide sequence of the *klg* cDNA clones predicts the gene product to be a member of the immunoglobulin gene superfamily (23), since the extracellular domain contains seven C2-type immunoglobulin-like domains, the sequences of which are unique. This family contains members that are either growth factor receptors, such as the platelet-derived growth factor receptor, or molecules involved in cellular adhesion, such as the neural cell adhesion molecule (23). All of these molecules are cell-surface proteins and preliminary live cell proteolysis experiments indicate that klg is also exposed at the cell surface (data not shown). The sequence of the presumed cytoplasmic domain of the *klg* gene product has many features in common with that of protein tyrosine kinases and, thus, although it could play roles in adhesion, its structure is most reminiscent of a growth factor receptor.

Expression studies indicated that the *klg* gene is expressed in a wide variety of embryonic tissues. The highest expression was found in two independently isolated transformed liver cell lines; however, increased expression was not found in other transformed cells (data not shown). No expression could be detected in normal adult liver, whereas low level expression could be detected in embryonic liver (data not shown). Both of the liver cell lines were isolated from tumors that were induced in adult animals (20, 21). This indicates that this expression of the *klg* gene is unusual and may represent the reexpression of a gene that is normally preferentially expressed embryonically. Predominantly embryonic expression patterns have recently been described for other tyrosine kinases that have been implicated in controlling development (24).

While the overall structure of the klg protein is similar to that of a growth factor receptor, the lack of the aspartate residue within the DFG triplet argues against this protein being a standard ligand-activated tyrosine kinase. Mutagenesis studies with the fps tyrosine kinase demonstrated the essential role this aspartate plays in the kinase activity (3). Recently, other kinases have been isolated in which the DFG sequence is replaced by the amino acid sequence DNA. These proteins, however, have retained the aspartate residue and possess kinase activity (C. Hovens and A. F. Wilks, personal communication). The cytoplasmic domain of the *klg* gene product contains several tyrosine residues that could serve as sites for phosphorylation. Thus, it is possible that this protein could interact with other tyrosine kinases as a substrate. This interaction could then function to either suppress or potentiate a signal response.

Dimerization has recently been demonstrated to play an important role in ligand-activated signal transduction and experimental data indicate that both members of the dimer complex have to be active enzymes for the signal to be actively transmitted (1). An *in vivo* equivalent of these experiments has been recently described. The white spotted locus in mice, *w*, is encoded by the c-*kit* growth factor receptor tyrosine kinase (25). In the w^{42} mouse, the w^{42} allele is the result of a single amino acid change of the normally invariant aspartate within the HRD triplet of the c-*kit* gene (6). This mutation eliminates the kinase activity of the c-*kit* gene and, most interestingly, heterozygous w^{+}/w^{42} mice are severely compromised. This suggests that the w^{42} allele represents a dominant loss of function mutation that functions by participating in the formation of inactive dimers with the wild-type receptor (6). Other dominant loss of function mutations in this kinase have been described and may function in a similar fashion (26). These observations raise the possibility that signaling by growth factor PTKs can be regulated by the formation of inactive dimers. Analysis of the sequence of the *erbB-3* receptor (27, 28) reveals that it is lacking the same aspartate residue in the HRD triplet that is mutated in the w^{42} locus and thus possibly represents another example of an inactive kinase. Taken together, these data imply that the *klg* gene and possibly the *erbB-3* gene represent a class of ligand-activated receptors that could interact with other receptors and play roles in signal attenuation.

We would like to thank Alison Crowe, Steven Anderson, and Paula Enrietto for their helpful comments and critical reading of the manuscript. We are also indebted to Andrew Wilks for his insightful comments and for the communication of results prior to publication. This work was supported by Grants CA28146 and CA42573 from the National Institutes of Health.

1. Ullrich, A. & Schlessinger, J. (1990) *Cell* **61**, 203–212.
2. Hanks, S. K., Quinn, A. M. & Hunter, T. (1988) *Science* **241**, 42–52.
3. Moran, M. F., Koch, C. A., Sadowski, I. & Pawson, T. (1988) *Oncogene* **3**, 665–672.
4. Schlessinger, J. (1988) *Trends Biochem. Sci.* **13**, 443–447.
5. Klein, R., Conway, D., Parada, L. F. & Barbacid, M. (1990) *Cell* **61**, 647–656.
6. Tan, J. C., Nocka, K., Ray, P., Traktman, P. & Besmer, P. (1990) *Science* **247**, 209–212.
7. Levy, J. B., Dorai, T., Wang, L.-H. & Brugge, J. S. (1987) *Mol. Cell. Biol.* **7**, 4142–4145.
8. Sap, J., Munoz, A., Damm, K., Goldberg, Y., Ghysdael, J., Leutz, A., Beug, H. & Vennstrom, B. (1986) *Nature (London)* **324**, 635–640.
9. Smith, D., Vogt, P. K. & Hayman, M. J. (1989) *Proc. Natl. Acad. Sci. USA* **86**, 5291–5295.
10. Sanger, F., Nicklen, S. & Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. USA* **74**, 5463–5467.
11. Maniatis, T., Fritsch, E. F. & Sambrook, J. (1982) *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Lab., Cold Spring Harbor, NY).
12. Spindler, K. R., Rosser, D. S. E. & Berk, A. J. (1984) *J. Virol.* **49**, 132–141.
13. Hayman, M. J. (1978) *Virology* **85**, 475–486.
14. Towbin, H., Staehelin, T. & Gordon, J. (1979) *Proc. Natl. Acad. Sci. USA* **76**, 4350–4353.
15. Kaufman, R. J., Murtha, P. & Davies, M. V. (1987) *EMBO J.* **6**, 187–193.
16. Hayman, M. J., Kitchener, G., Vogt, P. K. & Beug, H. (1985) *Proc. Natl. Acad. Sci. USA* **82**, 8237–8241.
17. Park, M., Dean, M., Kaul, K., Braun, M. J., Gonda, M. A. & Vande Woude, G. (1987) *Proc. Natl. Acad. Sci. USA* **84**, 6379–6383.
18. Ullrich, A., Bell, J. R., Chen, E. Y., Herrera, R., Petruzzelli, L. M., Dull, T. J., Gray, A., Coussens, L., Liao, Y. C., Tsubokawa, M., Mason, A., Seeburg, P. H., Grunfeld, C., Rosen, O. M. & Ramachandran, J. (1985) *Nature (London)* **313**, 756–761.
19. Shtivelman, E., Lifshitz, B., Gale, R. P., Roe, B. A. & Canaani, E. (1986) *Cell* **47**, 277–284.
20. Langlois, A. J., Lapis, K., Ishizaki, R., Beard, J. W. & Bolognesi, D. P. (1974) *Cancer Res.* **34**, 1457–1464.
21. Kawaguchi, T., Nomura, K., Hirayama, Y. & Kitagawa, T. (1987) *Cancer Res.* **47**, 4460–4464.
22. Niehrs, C. & Huttner, W. B. (1990) *EMBO J.* **9**, 35–42.
23. Williams, A. F. & Barclay, A. N. (1988) *Annu. Rev. Immunol.* **6**, 381–405.
24. Pasquale, E. B. & Singer, S. J. (1989) *Proc. Natl. Acad. Sci. USA* **86**, 5449–5453.
25. Besmer, P., Murphy, J. E., George, P. C., Qiu, F., Bergold, P. J., Lederman, L., Snyder, H. W., Brodeur, D., Zuckerman, E. E. & Hardy, W. D. (1986) *Nature (London)* **320**, 415–421.
26. Nocka, K., Tan, J. C., Chiu, E., Chu, T. Y., Ray, P., Traktman, P. & Besmer, P. (1990) *EMBO J.* **9**, 1805–1813.
27. Plowman, G. D., Whitney, G. S., Neubauer, M. G., Green, J. M., McDonald, V. L., Todaro, G. J. & Shoyab, M. (1990) *Proc. Natl. Acad. Sci. USA* **87**, 4905–4909.
28. Kraus, M. H., Issing, W., Miki, T., Popescu, N. C. & Aaronson, S. A. (1989) *Proc. Natl. Acad. Sci. USA* **86**, 9193–9197.