The Drosophila kekkon Genes: Novel Members of both the Leucine-Rich Repeat and Immunoglobulin Superfamilies Expressed in the CNS

Michele Musacchio and Norbert Perrimon
Department of Genetics, Howard Hughes Medical Institute, 200 Longwood Avenue, Boston, Massachusetts 02115

We have identified two members of a novel class of genes in Drosophila that encode putative transmembrane proteins with six leucine-rich repeats and a single immunoglobulin loop. These two molecules, Kek1 and Kek2, show striking conservation in their extracellular domains and have large and more divergent intracellular regions. Both genes are expressed in neurons as they differentiate in the embryonic central nervous system (CNS). kek1 is also expressed in other patterned epithelia, such as the follicle cells of the developing egg chamber, where it is found in a dorsal-ventral gradient around the oocyte. The homology of the kek genes to other known adhesion and signaling molecules, together with their expression patterns, suggests that both genes are involved in interactions at the cell surface. Genetic analysis reveals that deletion of the kek1 gene causes no obvious developmental defects. The coexpression of kek2 in the CNS leads us to suggest that Kek1 is part of a family of cell surface proteins with redundant function.

INTRODUCTION

Morphogenesis in the central nervous system (CNS) results in extremely specific connections between neurons, achieved through specific cell recognition during stages of neuronal segregation and differentiation. One consequence of this precise connectivity is the highly ordered pattern of CNS axon tracts. Drosophila has been used to dissect the molecular mechanisms that generate this pattern (for reviews see Goodman and Shatz, 1993; Tear et al., 1993; Keynes and Cook, 1995). Studies have shown that mesectodermal derivatives are a source of cues in the initiation and organization of axon tracts (Klambt et al., 1991, 1992; Seeger et al., 1993), and individual genes in both glia and neurons have been shown to contribute to functions of fasciculation of identified axon bundles and tracts (Lin et al., 1994; Menene and Klambt, 1994; Klambt et al., 1992). However, few molecules involved in all neurons mediating the general response to guidance cues in the CNS have been found by genetic analysis alone. How is this response organized in terms of the function of individual molecules? The semaphorin, connectin, and fasll genes represent functions in guidance and connectivity in the CNS and PNS as defined by in vitro assays and misexpression in vivo (Nose et al., 1992, 1994; Mathies et al., 1995; Chiba et al., 1995). A better understanding of the organization of molecular interactions underlying cell recognition in the CNS will require the characterization of additional neuronal surface proteins and of how they function in different contexts.

In this study we have taken an enhancer trap approach to look specifically at genes expressed in the CNS. By identifying genes in a broad sampling of expression patterns, enhancer trap technology has contributed to a genetic and molecular understanding of proteins involved in tissue differentiation, particularly in the CNS (Bellen et al., 1989; Wilson et al., 1989). By including genes that are reused in different developmental contexts, this analysis can identify genes that might otherwise be overlooked by classical genetic analysis.

We describe two members of a new class of putative transmembrane molecules, first expressed in overlapping patterns in the CNS and containing motifs from both the leucine-rich repeat (LRR) and immunoglobulin (Ig) superfamilies. Because these genes each represent the union of two superfamilies within one molecule, we have named them kekkon (keh·kon), the Japanese word for marriage. The predicted proteins Kek1 and Kek2 are strikingly similar in their extracellular domains, each containing six LRRs and one C2-type Ig domain. The combination of LRR and Ig motifs in a transmembrane protein is shared only by the vertebrate trk family of receptor tyrosine kinases (Schneider and Schweiger, 1991). Kek1 and Kek2 each have large and more...
dissolved intracellular domains with no other similarities to known proteins. We characterize the expression of kek1 in detail, showing that it is regulated in developing epithelia in addition to the CNS, including the imaginal discs and the follicular epithelium of the developing ovaries. We demonstrate that the absence of the kek1 gene product causes no overt phenotype. This, together with the presence in the CNS of the structurally similar gene kek2, suggests that kek genes, or additional genes, may overlap in function in the assembly of the CNS.

MATERIALS AND METHODS

Isolation of Genomic DNA and cDNAs

Nucleic acid blotting and hybridizations were performed as described in Sambrook et al. (1989). Genomic DNA was prepared from homozygous 15A6 flies for plasmid rescue as described in Wilson et al. (1989). A 1.2 kb ClaI–HindIII subclone (15C3) from rescued plasmids was used to isolate four unique, overlapping clones from the λEMBL3 genomic library (Blackman et al., 1987), spanning ~38 kb. Two of them, phages 2-W and 6-2, are depicted in Fig. 3. The 15C3 fragment was used to screen a 12- to 24-hr embryonic cDNA library under standard conditions for high stringency, and an 8- to 12-hr embryonic cDNA library (both libraries from Brown and Kafatos, 1988) at lowered stringency, with only the final washes modified from standard conditions (55°C in 0.2× SSC, 0.5% SDS). Of six cDNAs isolated at high stringency, four are identical (3.5 kb) and are represented by NB1; two are shorter at the 3’ end as determined by restriction analysis. All hybridized to phase 6-2 and correspond to kek1. A 4.2-kb clone, NB7, isolated at lowered stringency, did not hybridize back to phase 6-2 under standard conditions. NB7 was subsequently used to isolate nine additional cDNAs at high stringency, corresponding to kek2, of which six are identical to NB7; the three shorter forms are missing 5’ sequences as determined by restriction analysis. Their hybridization to 15C3 is presumably due to three regions of contiguous homology spanning 5, 8, and 21 bases, separated by single base mismatches at the 3’ end of the 15C3 probe, as revealed by DNA sequencing. Northern analysis was done as described in Perkins et al. (1992).

DNA Sequencing

DNA sequencing was carried out using Sequenase (U. S. Biochemical Corp.) and a dideoxy chain termination protocol (Del Sal et al., 1989) with the following modifications: double-stranded plasmid DNA template was denatured at 70°C for 15 min, and template and primer were annealed at 37°C for 15 min. All sequencing was done with primer walking on both strands of DNA. Phage 2-W and 6-2 were subcloned in SalI fragments into pBluescript SK(+) (Stratagene) for sequencing. The insertion site of the transposon was determined from the sequence of the ClaI plasmid rescued from 15A6 genomic DNA and lies 684 bases 5’ to a consensus TATA box and 718 bases 5’ to a consensus cap site (Arkhipova, 1995).

DNA sequence analysis was performed using the Wisconsin Genetics Computer Group (WigCG) sequence analysis package (Devereux et al., 1984) and amino acid sequence alignments were determined using the Lasergene program (DNAStar). Homology searches were done using the BLAST Network Service (Altschul et al., 1990). [Note: the complete sequence of the N81 (kek1) and N87 (kek2) cDNAs have been deposited with GenBank Data Library under Accession Nos. U42767 and U42768, respectively.]

Stocks and Genetic Protocols

To revert the ry+ marker associated with the 15A6 P[lArB] insertion by excision, 15A6/15A6; ry506/ry506 males were crossed en masse with Sp/CyO; Δ2-3[ry+]Sb/TM2, U bx virginal females. Male progeny of genotype CyO/15A6; Δ2-3[ry+], Sb/ry506 were crossed to 2B65/CyO; ry506/ry506 virginal females. 2B65 is a lethal ry+ P[lArB] insertion on the second chromosome. Single [ry+, Cy] revertant males were isolated to establish stocks. Of 500 revertant independent lines, 464 were either partially or fully viable and 36 were associated with lethality, which was determined by the absence of Cy− progeny. Among them are ra5, rj22, rk9, and rm2, which are described in Table 1.

To revert the ry+ marker by X rays, 0- to 5-day-old 15A6/15A6; ry506/ry506 males were irradiated with 3000 rad (Torrex, 120 D X-ray source) and then crossed to CyO/2B65; ry506/ry506 virginal females. About 80,000 CyO/15A6; ry506/ry506 progeny were screened for the loss of the ry− eye-color marker. Twenty-five independent revertant lines were established successfully; among them are the zygotic lethal lines 3K2X and 3K2S, which are described in Table 1.

In the complementation tests between two zygotic lethal mutations, over 400 progeny were scored under uncrowded culture conditions for the presence of straight-winged (Cy+−) flies.

In Situ Hybridization and Immunohistochemistry

In situ hybridization to polytene chromosomes was carried out as described in Perrimon et al. (1991) using P[lArB] as a probe. In situ hybridization to RNA in embryos was performed as described by Tautz and Pfleifl (1989).

Overnight embryo collections were dechorionated, fixed for 5 min with 4% paraformaldehyde in PBS + 0.1% Triton X-100, and devitellinized using standard methods. Mousenomiconal antibodies BP102 (a gift from A. Bieher and C. S. Goodman) and 22C10 (Goodman et al., 1987) and antibodies to β-galactosidase (β-gal) (Promega) were detected with a horse anti-mouse antibody (Vector) and stained using the Vector Elite ABC kit. Third-instar larvae from both excision- and X-ray-derived lines and ovaries from the 15A6 line were stained for β-gal activity by dissection in Ringer’s solution, fixed for 5 min with 4% paraformaldehyde in PBS, rinsed, and then incubated in X-gal solution at 37°C for 4–12 hr.

RESULTS

Isolation of the 15A6 Enhancer Trap Line

To identify genes that play a role during axonal outgrowth in the CNS, we screened a collection of insertion lines generated with the P[lArB] enhancer trap transposon (Perrimon et al., 1991). We were particularly interested in lines that express β-galactosidase specifically in subsets of CNS cells after the first neuroblasts have segregated and before complete formation of the axon scaffold (stages 9–12 of embry-
New Class of LRR/Ig Genes in Drosophila

TABLE 1
Characterization of 15A6 Revertants

<table>
<thead>
<tr>
<th></th>
<th>/gal stain</th>
<th>nubbin⁺</th>
<th>RA5</th>
<th>3K2X</th>
<th>RJ22</th>
<th>RM2</th>
</tr>
</thead>
<tbody>
<tr>
<td>RA5</td>
<td>+</td>
<td>−</td>
<td>−</td>
<td>−</td>
<td>−</td>
<td>−</td>
</tr>
<tr>
<td>3K2X</td>
<td>−</td>
<td>+</td>
<td>−</td>
<td>−</td>
<td>−</td>
<td>−</td>
</tr>
<tr>
<td>RJ22</td>
<td>−</td>
<td>+</td>
<td>−</td>
<td>−</td>
<td>+/−</td>
<td>+</td>
</tr>
<tr>
<td>3K2S</td>
<td>−</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>−</td>
<td>−</td>
</tr>
<tr>
<td>RM2</td>
<td>−</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>−</td>
</tr>
<tr>
<td>RK9</td>
<td>−</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>−</td>
</tr>
<tr>
<td>Df(2L)Pr</td>
<td>na</td>
<td>nd</td>
<td>nd</td>
<td>nd</td>
<td>+</td>
<td></td>
</tr>
</tbody>
</table>

Note. The RA5, 3K2X, RJ22, 3K2S, and RM2 chromosomes were recovered as ry⁺, lethal revertants of 15A6 (see Materials and Methods). Lines were characterized for retention of lacZ expression and interse complementation pattern. The complementation pattern with Df(2L)Pr, which deletes region 32F1-3 to 33F1-2 (Lindsley and Zimm, 1992), and the viable visible mutation nubbin, which maps to the region are indicated. In two combinations partial complementation was detected. The survival rate of RJ22/3K2S transheterozygous animals is 0.5 and 0.1% for the RJ22/RA5. Among the revertant lines, nine alleles of the RM2 complementation group were isolated. na, not applicable; nd, not determined.

LacZ is expressed during larval development; see Campos-Ortega and Hartenstein, 1985). Among these, the 15A6 line was selected because it expresses lacZ initially in a segmentally repeated set of cells at the CNS midline at early stages. At later stages, lacZ eventually becomes expressed in a large set of CNS cells and some peripheral nervous system (PNS) cells.

During embryogenesis, lacZ is first strongly detected at the beginning of germ band retraction (stage 10) in a few cells in each segment at the dorsal midline of the CNS (Figs. 1A and 1C). These cells do not double-stain with the monoclonal antibody 22C10 (mAb 22C10), which at this stage recognizes the neurons aCC and pCC (data not shown, Goodman and Doe, 1993). However, by early stage 12, the dorsal midline staining evolves to a ventral position, and some midline-staining cells begin to double-stain with mAb 22C10 (data not shown). This, together with their apparent migration from dorsal to ventral, suggests that they include the ventral unpaired medial cells (VUMs), which are among the first CNS neurons to differentiate (Goodman and Doe, 1993). LacZ expression evolves by stage 12 to include many neurons of the CNS (Fig. 1E), and is maintained throughout larval development. None of the midline or longitudinal glial cells express lacZ, as determined by their size and position (Figs. 1K and 1L).

In the PNS, a small subset of the ventral and lateral cell clusters express lacZ by stage 14. In abdominal segments, lacZ is also expressed in the cap cells, which are nonneuronal support cells that contact the chordotonal neurons and are identifiable by their distinct morphology and position (Fig. 1G). In the head region, both neuronal and nonneuronal cells of the antennomaxillary complex begin expressing lacZ by stage 14 (Fig. 1I).

In addition to these expression patterns, the 15A6 lacZ insertion is expressed during larval development in all imaginal discs while those tissues undergo patterning (Fig. 2). In the eye–antennal disc, the 15A6 enhancer trap is expressed throughout the differentiating retinal epithelium, from the posterior border up until several cells behind the morphogenetic furrow (Fig. 2C). LacZ is also expressed in a large patch that includes the presumptive ocelli and ocellar bristles (Fig. 2C).

The 15A6 line shows dynamic lacZ expression during the patterning of the follicle cell epithelium in the ovary. During oogenesis, lacZ is weakly expressed in the nurse cells and more strongly in the columnar follicle cells surrounding the developing oocyte (Figs. 2A and 2B). By stage 7 of oogenesis, lacZ expression is apparent in follicle cells at the posterior pole surrounding the oocyte. However, by stage 9, this posterior expression disappears and lacZ is found heavily expressed in follicle cells located above the anteriorly placed oocyte nucleus. Interestingly, this pattern of lacZ expression reflects the onset of dorsoventral polarity in the egg chamber (see review by Lehmann, 1995), and at stage 9 lacZ is distributed in a gradient (Fig. 2A). The follicle cells at this position later differentiate into the dorsal filaments, and lacZ is expressed during their formation in the 15A6 line. Other follicle cells that express lacZ include the border cells (data not shown), which migrate extensively throughout oogenesis.

These results show that the 15A6 insertion is specifically regulated in the embryonic nervous system and in several developing epithelia. To explore the possibility that 15A6 is inserted near a gene that is expressed during patterning and differentiation of these tissues, we cloned the genomic region surrounding this P[IarB] insertion.

Cloning of the Enhancer Trap Insertion and Isolation of the NB1 cDNA

To begin a molecular and genetic characterization of the region, we mapped the 15A6 P[IarB] insertion to a single site at 33F1-2 on the second chromosome by in situ hybridization to salivary gland polytene chromosomes. Genomic DNA from the 15A6 line was cloned by plasmid rescue
FIG. 1. Embryonic expression patterns of the 15A6 insertion line and kek1 RNA. (A, C, E, K, and L) Embryos from the 15A6 enhancer trap insertion line stained for β-gal expression with a monoclonal antibody using HRP immunodetection. (B, D, and F) Wild-type embryos hybridized with the kek1 cDNA (NB1) labeled with digoxigenin and detected with anti-digoxigenin antibodies and alkaline phosphatase immunocytochemistry. (A–D) Cells detected by both methods are at a dorsal position in the CNS midline in these stage 11 embryos. (E and F) Note the similarity of expression in ventral midline cells between 15A6 and NB-1 in stage 14 wild-type embryos. Lateral staining cells visible in F are present but out of the plane of focus in E. (G, I) 15A6 embryos at stage 15 stained with X-gal (blue) and mAb22C10 (brown), showing PNS expression. Inset in G shows deeper plane of focus where cap cells are positioned (arrowhead in H). Arrowhead points to antennomaxillary complex in I. (H, J) RNA in situ showing corresponding lateral PNS staining (H), and antennomaxillary complex (J, arrowheads). (K and L) High-magnification DIC images of dorsal and ventral focal planes of β-gal expression in 15A6 at stage 14, showing the positions of commissures (K, arrowheads) and longitudinal axon tracts (L, arrows). Note the absence of β-gal-staining nuclei at or above these axon tracts, positions expected for longitudinal and midline glia. A, B, G, and H are sagittal views. C–F are ventral views. I, J, K, and L are dorsal views. Anterior is to the left in all panels.

(Wilson et al., 1989). A 1.2-kb Clal genomic insert (called 15C3) was isolated and subsequently used to isolate 38 kb of contiguous genomic DNA from a phage library (Fig. 3A; see Materials and Methods).

The genomic fragment 15C3 (Fig. 3A) identifies a major transcript of 5.5 kb on a Northern blot of poly(A)+ RNA from 9- to 13-hr embryos (data not shown), a time point when abundant lacZ expression is detected in the CNS of 15A6 embryos. Whole-mount in situ RNA hybridization experiments revealed that the same genomic fragment de-
FIG. 3. Genomic organization of the kek1 locus. (A) The P[lArB] transposon is shown inserted into the 15A6 region of genomic DNA. Details on the structure of P[lArB] can be found in Wilson et al. (1989). The map is derived from Southern analyses of genomic DNA from 15A6, 12A8 (a different P[lArB] insertion line from the same screen), and wild-type Ore-R. Southern blots were probed with either P[lArB], 15C3, phages 2-W and 6-2, or the NB-1 cDNA. Direction of transcription of the NB1 transcript is indicated by a thick arrow. The lacZ transcription unit is indicated by a thin arrow; ry*: rosy gene; AmpR: ampicillin resistance gene used for plasmid rescue. (B) The position of the RA5 breakpoint (shaded region) within the ry gene was mapped using the above probes.

FIG. 2. Expression pattern of 15A6 during oogenesis and in imaginal discs. (A, B) 15A6 is expressed in a dorsal-to-ventral gradient of β-gal expression in columnar follicle cells during oogenesis. (A) Early stage 10 egg chamber. (B) Series of egg chambers, showing posterior expression of 15A6 expression in early developmental stages. (C, D) Imaginal disc expression pattern. (C) Eye-antenna disc of third-instar larva. Staining is localized posterior to the morphogenetic furrow (arrowhead) and in the presumptive ocellar region (arrow). (D) Wing disc pattern. Among other places, staining is in the vicinity of the L2 and L3 veins on both dorsal and ventral surfaces on the fate map at this stage. In all panels, β-gal expression is detected by X-gal staining.

FIG. 4. Expression of kek2 during embryogenesis. (A) Dorsal focal plane of an early stage 14 embryo, showing prominent cells at dorsal surface of the CNS (arrowheads). (B) Ventral focal plane of the same embryo, showing small set of ventral midline cells (arrow) and kek2 expression in ventral muscle group (arrowhead). (C) Stage 15 embryo, showing staining in many CNS neurons. The probe used was a 3′ fragment of kek2 that did not contain the regions to which it hybridizes to kek1 at low stringency. Anterior is to the left in all panels.
not hybridize to the genomic DNA cloned from the 15A6 locus, was used to isolate nine additional cDNAs, the long-
est of which, NB7, is 4.2 kb in size. Because partial se-
quence analysis of NB7 revealed sequence similarity with
NB1, we pursued the characterization of this class of cDNA.
NB7 identifies a 5.9-kb transcript on Northern blots of 9-
to 13-hr poly(A) RNA and corresponds to a single copy
sequence on Southern blots of genomic DNA (data not
shown). The gene encoding NB7 resides in the genome out-
side of the extent of genomic DNA cloned at the 15A6
locus.

To determine whether the transcript encoded by NB7 is
expressed in a pattern similar to that of NB1 during em-
byonic development, RNA in situ hybridization to embryos
was done using a unique fragment of NB7. Like NB1, N
B7 is first expressed in a small, segmentally repeated group
of cells at the dorsal midline of the CNS at stage 11 (not
shown); it is then expressed in a ventral group of midline
cells as it expands to many cells in the CNS. The pattern
of this expansion appears slightly different than that of N
B1, with several prominent dorsolateral cells staining at earlier
stages (Figs. 4A and 4B). However, by stage 15 it is expressed
in many cells of the CNS in a pattern similar to NB1 (Fig.
4C) although NB7 appears to be expressed in a much larger
subset of neurons than NB1. In the CNS, a small group of
lateral cell clusters and a group of ventral cell clusters also
express NB7 (data not shown).

NB7 transcript expression differs from the NB1 pattern
in its tissue specificity. Notably, at early stages (11 to early
12) it is transiently expressed in a lateral patch in the head
(data not shown), and at stage 14 it is expressed in segmen-
tally repeated ventrolateral patches outside the CNS (Figs.
4A and 4B), which correspond to the position of extending
ventrolateral muscles. We did not examine the expression
of NB7 at other developmental stages.

Both NB1 and NB7 cDNAs Encode Transmembrane Proteins of the Ig Superfamily
with LRR Motifs

We sequenced the NB1 and NB7 cDNA inserts in their
entirety. We also sequenced genomic DNA from the 15A6/
NB1 locus. An interesting feature of the gene organiza-
tion at this locus is the absence of introns; the genomic DNA
sequence is entirely collinear with the sequence of NB1.

A conceptual translation of the NB1 nucleic acid se-
quence reveals a long open reading frame (ORF) of 880
amino acids (aa). The initiator methionine is preceded by
a three out of four match to the consensus for Drosophila
translation start sites (C/A AA A C ATG; Cavener, 1987)
and by multiple stop codons in all three frames in the previ-
ous 172 nucleotides. The ORF starts with a potential signal
sequence, as deduced by a Kyte-Doolittle hydrophathy ana-
lysis (Kyte and Doolittle, 1982), which also predicts a single
transmembrane segment separating the presumed extracel-
lular (428 a.a. mature) and intracellular (413 a.a.) regions. A
signal peptide cleavage site is predicted between residues
20 and 21 by the method of von Heijne (1983). The extracel-

lular portion contains motifs for N-linked glycosylation at
three sites and a single glycosaminoglycan attachment site;
three sites and a single glycosaminoglycan attachment site;
together these data suggest that NB1 encodes a processed
transmembrane glycoprotein of predicted M, 92.3 x 10^3,
unprocessed.

The sequence of the NB7 cDNA indicates that, like NB1,
it encodes a processed transmembrane glycoprotein. NB7
contains a long ORF of 892 a.a., beginning immediately at
the 5’ end of the cDNA. No initiator methionine is present,
suggesting that the cDNA is truncated at its 5’ end. How-
ever, hydropathy analysis indicates a 17-a.a. hydrophobic
stretch ending in a predicted cleavage site, consistent with a
partial signal sequence, as well as a transmembrane stretch
separating a mature extracellular domain of 362 a.a. and an
intracellular domain of 492 a.a. Thus, we believe that NB7
contains the full sequence of the processed protein it en-
codes. As with the predicted protein from NB1, there are
three N-linked glycosylation sites in the extracellular
domain. The entire mature protein has a predicted M, of 95.5
x 10^3, unprocessed.

A search of the sequence databases with the ORFs of each
of these cDNAs reveals that they both have two distinct
regions of homology in their extracellular domains: to
the LRR and to the Ig superfamilies. The two predicted proteins
are strikingly similar in their extracellular domain, having
the same length and number of these motifs. The structure
of their extracellular domains differs only by a 69-a.a. ser-
ine/proline-rich stretch at the amino terminus, present in
the NB1 ORF and not in the NB7 ORF, which contains a
potential glycosaminoglycan attachment site. Their overall
structure, as transmembrane proteins containing both mo-
tifs, is shared only by the Trk family of neurotrophin recep-
tors. We have named the putative proteins Kek1 and Kek2,
after the Japanese word for marriage, to reflect the unique
joining of these motifs in these proteins.

Structure of the LRR Domains of Kek1 and Kek2

Toward the amino end, both Kek1 and Kek2 contain six
t contiguous LRRs. These are flanked by the cysteine-rich
amino- and carboxy-flanking consensus regions often asso-
ciated with the repeats (Fig. 5A, B). The repeats are charac-
terized by the specific spacing of leucine and asparagine
residues (Rothberg et al., 1990; Kobe and Deisenhofer, 1994;
Nose et al., 1992). The cysteine-rich amino- and the car-
boxy-flanking regions are both characterized by the four
conserved cysteines found in these motifs.

Between them, the Keks share repeat-specific homologies
beyond the basic LRR framework. The longest stretch of
sequence identity between the two molecules occurs in the
largely basic amino-flanking cysteine-rich region and is
followed by significant homologies spanning the amino-end
boundaries of the first and second repeats and also in the
last repeat. These are regions known to adopt a beta-strand
structure in the homologous portions of ribonuclease inhib-
itor, a protein entirely composed of LRRs whose three-di-

Copyright © 1996 by Academic Press, Inc. All rights of reproduction in any form reserved.
Drosophila. Acid sequence identity (Fig. 7). Both proteins have a hydro-
several short regions of similarity with a 19% overall amino
Drosophila. Acid sequence identity (Fig. 7). Both proteins have a hydro-
taminine-rich stretches similar to opa repeats found in the imaginal discs, and ovarian follicle cells. Since the enhancer
taminine-rich stretches similar to opa repeats found in the imaginal discs, and ovarian follicle cells. Since the enhancer
taminine-rich stretches similar to opa repeats found in the imaginal discs, and ovarian follicle cells. Since the enhancer
Structure of the Ig Domain of Kek1 and Kek2
A single C2-type Ig loop begins immediately following the LRR region. It has the characteristic spacing of the
taminine-rich stretches similar to opa repeats found in the imaginal discs, and ovarian follicle cells. Since the enhancer
Structure of the Ig Domain of Kek1 and Kek2
A single C2-type Ig loop begins immediately following the LRR region. It has the characteristic spacing of the
taminine-rich stretches similar to opa repeats found in the imaginal discs, and ovarian follicle cells. Since the enhancer
Structure of the Ig Domain of Kek1 and Kek2
A single C2-type Ig loop begins immediately following the LRR region. It has the characteristic spacing of the
Structure of the Intracellular Domains of Kek1 and Kek2
The putative intracellular domains of both proteins are large (413 and 492 a.a. for Kek1 and Kek2, respectively) and more divergent than their extracellular domains, containing several short regions of similarity with a 19% overall amino
Structure of the Intracellular Domains of Kek1 and Kek2
The putative intracellular domains of both proteins are large (413 and 492 a.a. for Kek1 and Kek2, respectively) and more divergent than their extracellular domains, containing several short regions of similarity with a 19% overall amino
Structure of the Intracellular Domains of Kek1 and Kek2
The putative intracellular domains of both proteins are large (413 and 492 a.a. for Kek1 and Kek2, respectively) and more divergent than their extracellular domains, containing several short regions of similarity with a 19% overall amino
Structure of the Intracellular Domains of Kek1 and Kek2
The putative intracellular domains of both proteins are large (413 and 492 a.a. for Kek1 and Kek2, respectively) and more divergent than their extracellular domains, containing several short regions of similarity with a 19% overall amino
The spacing of leucines in the first two repeats diverges from the consensus yet is conserved between Kek1 and Kek2.
Overt Phenotypes
To examine the functions of kek1 during development, we isolated mutations that affect the kek1 transcription unit. The ry" marker within the P[IArB] transposon provided a convenient tool to generate mutations in the gene since it can be excised following hybrid dysgenesis or rearranged with X-ray mutagenesis. Two independently derived chromosomes recovered in these screens, RA5 and RM2 (Table 1), were crossed to generate a transheterozygote that completely deletes the kek1 gene, as determined by Southern analysis of genomic DNA (Fig. 8). kek1 mutant animals (of genotype RA5/RM2) are viable, lay fertile eggs with normal morphology, and do not exhibit any obvious visible phenotype.
DISCUSSION
kek Genes Are Expressed in Restricted Patterns during CNS and Disc Development
kek1 was identified adjacent to an enhancer trap inserted at 33F which is expressed in a specific pattern in the CNS, imaginal discs, and ovarian follicle cells. Since the enhancer trap expression closely parallels kek1 expression in the embryonic CNS and in follicle cells, and because of the proximity of the insertion to the kek1 transcription unit (within 0.7 kb of the putative transcription start, see Materials and Methods), we believe that the enhancer trap is a reliable indicator of kek1 transcript expression throughout development. kek2 was found by hybridization homology to kek1 and corresponds to a separate gene.
The similar, restricted expression patterns of the two genes suggest that they are involved in the same events in the CNS during embryonic development. Neurons must distinguish from among many cell interactions to direct axon growth and to find targets as they differentiate; the expression of kek1 and kek2 in many CNS neurons may have the same general distribution of charged and hydrophobic residues.
Mutations in kek1 Are Not Associated with Any Overt Phenotypes
To examine the functions of kek1 during development, we isolated mutations that affect the kek1 transcription unit. The ry" marker within the P[IArB] transposon provided a convenient tool to generate mutations in the gene since it can be excised following hybrid dysgenesis or rearranged with X-ray mutagenesis. Two independently derived chromosomes recovered in these screens, RA5 and RM2 (Table 1), were crossed to generate a transheterozygote that completely deletes the kek1 gene, as determined by Southern analysis of genomic DNA (Fig. 8). kek1 mutant animals (of genotype RA5/RM2) are viable, lay fertile eggs with normal morphology, and do not exhibit any obvious visible phenotype.
To determine whether kek1 plays a general role in many neurons during axon outgrowth, the overall pattern of CNS axon tracts in kek1 mutant embryos between stages 9 and 16 were examined with the BP102 antibody, which stains all CNS axons (see Seeger et al., 1993). Axon tracts in the kek1 CNS appear wildtype at all stages examined, and CNS neurons that normally express kek1, such as the ventral midline clusters, and PNS structures, such as the chordotonal cap cells and the antenno-maxillary complex, are present and in their normal positions, as shown by 22C10 and lac2 antibody staining (data not shown). These results indicate that kek1 is not essential for viability and that removal of kek1 alone does not grossly affect the morphological development of the cells that express it.

FIG. 5. Extracellular domains of Kek1 and Kek2. (A) The putative extracellular and transmembrane portions of the deduced amino acid sequences from the N B1 and N B7 cDNAs are labeled Kek1 and Kek2, respectively. The sequences were aligned with the Lipman–Pearson algorithm (Lipman and Pearson, 1985) and are shown arranged on different lines according to their structural features. Amino acid identities are highlighted in black background, and relative gaps are indicated by dashes. Hydrophobic regions identified by the Kyte–Doolittle (1982) hydrophobicity analysis are underlined. The potential glycosaminoglycan attachment site in Kek1 is indicated by a double underline, potential N-linked glycosylation sites are boxed in dashed lines, and the conserved cysteines within the Ig loop (boxed) are indicated with a black dot. The conserved positions of the 24-amino-acid leucine-rich repeat consensus xLxxLxLxxNxLxxaFxxLx found in many Drosophila LRR proteins (where x is any residue and a is A, V, L, I, F, Y, or M) are indicated by vertical gray bands. Note the unusual substitution of C for N at position 10 of the fifth LRR repeat. A single, full repeat is counted as LxxLxLxxN/C, which comprises the beta-strand portion of the repeat (Kobe and Deisenhofer, 1994). (B) Entire Kek1 and Kek2 predicted proteins shown schematically, lined up at their putative transmembrane domains. Amino acid residue numbers are shown above each diagram. Percentage similarities between homologous regions (like-shaded boxes) are based on amino acid identity within the sequence alignments shown in (A) and in Fig. 6. The position of the SPDEGY motif at the C-terminus (see text) is indicated by vertical bars.
FIG. 6. Sequence comparison of Kek1 and Kek2 Ig loops with related loops of other Ig superfamily members. Alignment of the single C-2 type Ig domain of Kek1 and Kek2 with Ig domains from the GenBank database that show the closest relationship:

<table>
<thead>
<tr>
<th>Species</th>
<th>Ig Domain</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>Drosophila</td>
<td>Fasciclin II</td>
<td></td>
</tr>
<tr>
<td>rat NCAM</td>
<td>Ig loop 3</td>
<td></td>
</tr>
<tr>
<td>rat TAG1</td>
<td>Ig loop 5</td>
<td></td>
</tr>
<tr>
<td>Drosophila</td>
<td>Neuroglian</td>
<td></td>
</tr>
<tr>
<td>rat NRG</td>
<td>Ig loop 1</td>
<td></td>
</tr>
<tr>
<td>rat TRK1</td>
<td>Ig loop 1</td>
<td></td>
</tr>
<tr>
<td>human FGFR</td>
<td>Ig loop 3</td>
<td></td>
</tr>
</tbody>
</table>

Sequence comparison between Ig sequences indicates that they tend to share a glycine seven residues before the first cysteine and an aspartic acid within seven residues before the second cysteine. Ig domains share some homologies among them that are not present in the Keks, such as tyrosine two residues 5° to the second cysteine. Amino acid identities in common with both Kek proteins are shown in black background; identities with either Kek1 or Kek2 are shown in dark gray and light gray backgrounds, respectively. No attempt has been made to show identities among the other Ig domains. The position of conserved cysteine residues is marked with a black dot above the alignments, and potential N-linked glycosylation sites are within dashed boxes. Predicted beta-strand structures are overlined with dashes and labeled after Williams and Barclay (1988).
and intracellular proteins in species ranging from yeast to man. Drosophila LRR-containing genes often have a critical role in directing cell interactions; these include Toll (Keith and Gay, 1990), slit (Rothberg et al., 1988), contactin (Nose et al., 1992), chaoptin (Reinke et al., 1988; Krantz and Zipur et al., 1990), and 18 wheeler (Eldon et al., 1994). Proteins that contain Ig motifs and that are expressed in the CNS include centuly, and intracellular family members NCAM, L1, and TAG-1 (see review by Keynes and Cook, 1995). Further, members of each superfamily have been shown to operate as receptors in cell signaling events; i.e., the Toll protein (Hashimoto et al., 1991). The activity of the Trk family of kinases is controlled by their neurotrophin ligands in stimulating axon growth (Soppet et al., 1991). Numerous other receptor tyrosine kinases, such as FGFR and PDGFR, contain extracellular Ig domains (Williams and Barclay, 1988), as do transmembrane phosphatases expressed in the CN S (Zinn, 1993). Thus these structural motifs have been conserved in evolution and adapted to diverse cell-cell interaction processes. The similarity in the structure of the kek genes to these proteins suggests that they are likely to encode related functions.

Individually, LRR and Ig motifs have been implicated in direct protein interactions, although presumably by different mechanisms. The LRR motif can form a pocket that surrounds other protein moieties (Kobe and Deisenhofer, 1995), whereas the Ig domain is globular and has been shown to interact with carbohydrate side chains. It may not be surprising, therefore, that many members of these superfamilies in Drosophila can act as adhesion molecules in both heterophilic and homophilic assays in vitro. Recently, however, more complex activities, such as chemotraction (FasIII; Chiba et al., 1995) and repulsion (connectin; Neuroglian (Bieber et al., 1989), Fasciculins II and III (Snow et al., 1989), Dtrk (Pulido et al., 1992), and Amalgam (Seeger et al., 1990), have been exposed in elegant expression experiments in vivo. It is increasingly apparent that neither LRR nor Ig motifs confer a specific function themselves and that the function of molecules containing them is dependent on their cellular context (see review by Keynes and Cook, 1995).

Both LRR and Ig motifs can be adapted to specific functions by their variation and repetition, and by combination with other structural motifs. The Ig domain has been conserved in evolution and adapted to diverse cell-cell interaction processes. The similarity in the structure of the kek genes to these proteins suggests that they are likely to encode related functions.

Individually, LRR and Ig motifs have been implicated in direct protein interactions, although presumably by different mechanisms. The LRR motif can form a pocket that surrounds other protein moieties (Kobe and Deisenhofer, 1995), whereas the Ig domain is globular and has been shown to interact with carbohydrate side chains. It may not be surprising, therefore, that many members of these superfamilies in Drosophila can act as adhesion molecules in both heterophilic and homophilic assays in vitro. Recently, however, more complex activities, such as chemotraction (FasIII; Chiba et al., 1995) and repulsion (connectin; Neuroglian (Bieber et al., 1989), Fasciculins II and III (Snow et al., 1989), Dtrk (Pulido et al., 1992), and Amalgam (Seeger et al., 1990), have been exposed in elegant expression experiments in vivo. It is increasingly apparent that neither LRR nor Ig motifs confer a specific function themselves and that the function of molecules containing them is dependent on their cellular context (see review by Keynes and Cook, 1995).

Both LRR and Ig motifs can be adapted to specific functions by their variation and repetition, and by combination with other structural motifs. The Ig domain has been conserved in evolution and adapted to diverse cell-cell interaction processes. The similarity in the structure of the kek genes to these proteins suggests that they are likely to encode related functions.
FIG. 8. Identification of a kek1 null mutation. (a) Southern blot analysis of genomic DNA digested with BamHI and hybridized with the NB1 cDNA. The genotypes of the DNAs are indicated above the lanes. Note the absence of fragments detected in the lanes of the RA5/RM2 lane. (b) The same blot probed with a 3′ EcoRI subclone of phase 6-2, indicating the respective loading of DNA in each lane.

Genetic Analysis of kek1 Reveals No Overt Phenotype

Despite its conserved protein-interaction motifs and restricted expression pattern, a deletion of kek1 has no obvious effect on the development and function of the tissues in which it is expressed. This situation has been found for several genes related to cell recognition and expressed in Drosophila neurons, nrg, fasIII, and connexin null mutants show no major abnormalities in the CNS or in their peripheral projections (Biebler et al., 1989; Elkins et al., 1990a; Nose et al., 1994). Neuroupsinulin, expressed in the PNS and a subset of target muscles, does not affect these cells when mutated (Kania et al., 1993), and there have been no lethal mutations found for amalgam, despite saturation screens (Seeger et al., 1988).

There are two possible interpretations for this result. One is that kek1 acts alone in a simple manner in subtle aspects of outgrowth or guidance specificity in most neurons, or in general cohesion or cell physiology that our analysis has failed to reveal. A second interpretation is that kek1 participates with other genes in a complex, partially redundant function, such that its phenotype may be observed in only a few neurons. Similar alternatives have been suggested with respect to the examples noted above (also, see reviews by Goodman and Shatz, 1993; Keynes and Cook, 1995). Detailed developmental analysis of identified axon pathways will be necessary to distinguish these possibilities.

If the function of kek1 is redundant, kek2 is an likely candidate gene to overlap in function with kek1 at least in the CNS. Identification of kek2 mutants should help in the
investigation of this hypothesis. Another possibility is that structurally dissimilar proteins may contribute to the function of \textit{kek1}. For example, mutation of the transmembrane protein \textit{FasI} alone has no effect on CNS viability; however, in combination with a mutation in the \textit{Abl} tyrosine kinase, \textit{fasI} mutations cause specific defects in axonal pathfinding (Elkins et al., 1990b).

Finally, it may be that, in higher metameric organisms, the mechanism of cell recognition is encoded in various overlapping functions, such that disruption of a single function often is not critical for axon guidance in all neurons. As has been noted before, the use of additional genetic methods along with loss-of-function analysis has defined intriguing functions not exposed by mutation or in vitro analysis. Preliminary data on misexpression of \textit{kek1} show defects in the eye, wing, and chorion, demonstrating that \textit{kek1} can interact with other components in tissues where its expression is normally regulated (J. Duffy and Z. Wills, unpublished observations). The characterization of a new class of potential adhesion molecules provides a basis for more directed misexpression studies aimed at revealing their role in CNS morphogenesis.

ACKNOWLEDGMENTS

We are indebted to Roham Zamanian for help with the sequencing of \textit{kek2}, Joe Duffy for identifying the gradient of expression in follicle cells, Ray Hsu for confirmation of sequence near the 15A6 insertion site, Nipam Patel and the Goodman lab for sending the BP102 and 22C10 antibodies, Trudi Schupbach for sharing unpublished information on the AN296 insertion line, Bob Nelson for \textit{kek1}, Marisa DiNardo for \textit{KEK2}, Joe Duffy for \textit{KEK2}, Trudi Schupbach for \textit{KEK2}, and Randy Johnson for the suggestion during neuronal development. We are indebted to Roham Zamanian for help with the sequencing of \textit{kek2}, Joe Duffy for identifying the gradient of expression in follicle cells, Ray Hsu for confirmation of sequence near the 15A6 insertion site, Nipam Patel and the Goodman lab for sending the BP102 and 22C10 antibodies, Trudi Schupbach for sharing unpublished information on the AN296 insertion line, Bob Nelson for insights of motif alignments, and Randy Johnson for the suggestion of the name \textit{kekkon}. We are grateful for the discussions and advice of Bob Nelson, Ken Yoffe, Andrea Brand, Ed Lauffer, Doug Harrison, and Joe Duffy. We thankWelcome Bender, Chris Hughes, and Doug Harrison for critical reading of the manuscript. M. M. was supported in part by a NIH Genetics training grant. N. P. is an investigator of the Howard Hughes Medical Institute.

REFERENCES


Received for publication January 11, 1996
Accepted April 24, 1996