Dorsal and neural expression of a tyrosine kinase-related Drosophila gene during embryonic development

Jochen Haller, Serge Côté, Günter Brönner, and Herbert Jäckle

Max-Planck-Institut für Entwicklungbiologie, Abt. Biochemie, D-74 Tübingen, FRG

Sequence analysis of an embryonic transcript of Drosophila predicts a tyrosine protein kinase-related gene. The prediction is based on several protein domains that are homologous to the functional domains of kinase-related oncogenes and several serine, threonine, and tyrosine protein kinases. For this reason, we named this gene Drosophila tyrosine kinase related (dTKR). dTKR maps into chromosome band 2R 60F1. It is initially expressed at blastoderm stage, showing transient transcript accumulations at dorso-lateral positions of the embryo and differences along its longitudinal axis. At later stages of embryogenesis, dTKR transcripts are found exclusively in neural anlagen. Both the region-specific pattern of expression and the putative kinase function are consistent with the suggestion of a regulatory role for this gene during development, which remains to be elucidated.

[Key Words: Drosophila melanogaster; in situ hybridization; sequence analysis; transcript pattern; tyrosine kinase-related]

Received July 1, 1987; revised version accepted July 31, 1987.

During the past years, genetic and molecular analysis of Drosophila development has identified a number of genes that are active only in a subset of cells or tissues [for review, see Gehring and Hiromi 1986; Scott and O’Farrell 1986]. The spatial and temporal patterns of expression of several of these genes have been revealed by in situ hybridization of molecular probes to tissue sections [for review, see Hafen et al. 1983; Scott and O’Farrell 1986]. Aside from confirming the spatial patterns of expression of genes for which the domains of action are already characterized by mutant phenotypes, in situ hybridization allows identification of genes that are expressed in a limited number of cells or tissues, but for which no mutant phenotype has yet been identified. The value of this approach has recently been demonstrated with the caudal gene. It was first identified on the basis of an exciting pattern of transcript accumulation: a gradient along the longitudinal axis of the developing embryo [Mlodzik et al. 1986]. Genetic analysis then proved that the caudal gene provides positional information for the developing embryo, as deduced from both the graded distribution of the protein along the longitudinal axis of the embryo and from the phenotype of caudal mutant embryos [Macdonald and Struhl 1986].

Analysis of the chromosome region 60 E9-F1 at the tip of the right arm of the second Drosophila chromosome revealed five transcripts that are expressed during early embryogenesis [Côté et al. 1987]. One of them, arbitrarily termed “g2”, maps into the 60 F1 chromosome band. This anonymous gene is first expressed as early as blastoderm stage [Côté et al. 1987]. Here we report its temporal and spatial patterns of expression during embryogenesis. Sequence comparison with protein kinases reveals conserved motifs found in tyrosine kinase. Thus, we named this gene Drosophila tyrosine kinase related (dTKR).

Results

Changes of the spatial and temporal patterns of dTKR gene expression

Initial expression of the dTKR gene occurs at blastoderm stage [Fig. 1]. As described recently [Côté et al. 1987], this expression is restricted to the dorsal and lateral region of the embryo. During the blastoderm stage (stage 14, Foe and Alberts 1983). dTKR transcripts decrease in the anterior and posterior-most region of the embryo, and they gradually disappear from lateral to dorsal [Fig. 1a–d]. At the end of blastoderm stage, the transcripts are restricted to the dorsal region of the embryo covering about 20–85% of the egg length (posterior pole is 0%). Within this domain of expression, transcripts accumulate unevenly. Between 20% and 50%, rather low levels of dTKR transcripts can be detected. They rise to high levels between 50% and 85% egg length, forming a wedge-shaped stripe on the dorsal surface [Fig. 1e,f].

During gastrulation, dTKR transcripts accumulate in the dorsal region of the cephalic furrow and cover the procephalic neurogenic ectoderm [Campos-Ortega and Hartenstein 1985]. During the extended germ band stage, transcripts begin to disappear gradually from the procephalic neurogenic region toward posterior. They remain last in the amnion serosa [Fig. 2a–c]. When para-segmental furrows appear prior to germ band retraction.
Tyrosine kinase-related Drosophila gene

Figure 1. Spatial and temporal patterns of dTKR transcripts during blastoderm stage. Sections of blastoderm stage embryos were hybridized in situ using the dTKR cDNA probe. Orientation of embryos is anterior left [a,c,e], dorsal uppermost [a–d,f]. Photomicrographs [left] and corresponding dark-field images [right] are shown. (a) Sagittal section of an early-stage 14 embryo [Foe and Alberts 1983]. Transcripts accumulate between 20% and 85% of the egg length [posterior pole is 0%]. Note the absence of transcripts at the ventral region as can be seen in the cross section of an embryo at the same stage shown in (b). During stage 14, transcripts become restricted to the dorsal site [see cross sections d,f] and accumulate more heavily in the 50–85% region of the embryo [see parasagittal section c]. At late-stage 14, transcripts show a wedge shape accumulation pattern of high intensity in the dorsal region [dorsal-horizontal, tangential section in e] and low levels of transcripts posterior to 50% egg length. (Martinez-Arias and Lawrence 1985), transcripts disappear from the dorsal position [Fig. 2d], and new dTKR expression was observed in the developing nervous system. From then on, the transcripts remain in the developing nervous system, including the brain [Fig. 2e,f], until they disappear completely at the end of embryogenesis. Based on Northern blot analysis, major dTKR expression is unlikely to occur at later stages of the Drosophila life cycle [Côté et al. 1987].

Structure and function of dTKR

Several cDNA clones were isolated from a library prepared from poly[A]+ RNA of 3- to 12 hr-old embryos as described in Côté et al. [1987]. The longest cDNA clone was used to map most of the dTKR transcript to genomic DNA [Fig. 3a]. It extends over more than 8 kb of genomic DNA, suggesting several introns. To analyze the dTKR gene structure and possibly to reveal the biochemical function of the dTKR gene product, we sequenced the cDNA clone, the corresponding genomic DNA, and the genomic DNA upstream of the 5' end of the cDNA clone.

Most of the dTKR transcript was recovered in a cDNA clone which, however, is close to full size [see below and Fig. 3b]. The transcription and translation initiation sites are not unambiguously identified. However, likely candidates for a CAAT and the TATA box as well as the presumptive translation start site were detected in the genomic DNA upstream from the 5’ end of the cDNA clone [Fig. 3b]. A single open reading frame extends from the first ATG codon to position 2259 [Fig. 3b], which would encode a 753-amino-acid polypeptide [see below]. The open reading frame is constructed from three exons: a 5’ exon of 1357 bp followed by a small 132-bp exon and by a 1895-bp 3’ exon. They are separated by two introns of about 2.5 and 3.8 kb, respectively [Fig. 3b,c]. The 3’ region of the cDNA contains a single polyadenylation signal [AATAAA] followed by two such signals in genomic DNA. This suggests that the first polyadenylation signal is the one being used at least in the transcript represented by the cDNA that we sequenced [Fig. 3b]. From these data, differential termination, initiation, and additional splicing cannot be excluded, although Northern blot analysis revealed only a single major transcript, the size of which would be consistent with the length of the transcript of about 4 kb [Côté et al. 1987].

The dTKR sequence predicts a 753-amino-acid polypeptide with putative protein kinase function. We observed two portions of the dTKR protein that are weakly homologous to the protein kinase domain of the avian ves oncogene [Fig. 4]; one is conserved in other avian oncogenes as well as in several homologous serine, threonine, and tyrosine kinases [Fig. 4b,c]. The degree of homology with these known kinases is low, but it suggests that dTKR may encode a protein kinase. Further evidence for this suggestion comes from the notion of
Figure 2. Patterns of dTKR transcripts during gastrulation and neurogenesis. Parasagittal sections through embryos are oriented anterior left and dorsal uppermost; [left] photomicrographs; [right] dark-field images. (a) Early gastrulation. Labeling appears in the amnion-serosa precursors extending into the procephalic neurogenic region. [b,c] Extended germ band stages. Note the accumulation of transcripts over amnion serosa cells. (d) Advanced extended germ band stage. Note transcripts appearing in the nervous system. (e) Germ band retraction; (f) Shortened germ band. Note that transcripts accumulate heavily in the ventral nervous system and brain.

Figure 3. Physical map and sequence of the dTKR gene. (a) Restriction map of dTKR DNA (top) and dTKR cDNA (below). For location and orientation of the genomic DNA in chromosome region 2R 60E9-F1 see Figs. 2 and 3c in Cote et al. (1987). Exons are indicated by black boxes in cDNA, transcription is right to left (indicated by arrow at the 3' end). Restriction enzymes used were AccI (A), BamHI (B), HindIII (H), PstI (P), PvuII (I), SacI (S). (b) Nucleotide sequence and predicted amino acids of the dTKR gene. Sequence of genomic DNA [light letters] and cDNA [heavy letters]. A single base difference was found in the open reading frame [position 2156; C in genomic DNA vs. G in cDNA; position values for bases are given on the left side above the sequence], (c) exon–intron boundaries. Three base pair differences were observed in the 3' untranslated sequence in positions 2766, 2797, and 3039 as A,G,G in genomic DNA, respectively. The position of putative CAAT and TATA boxes as well as the polyadenylation signals are underlined. The amino acid sequence is given below the nucleotide sequence; positions are indicated at the right side. Amino acids encoding the consensus sequence of the putative ATP-binding site are printed dark, the putative autophosphorylation site is marked *. Putative glycosylation sites [Asn-X-Set/Thr] are underlined, as is the direct repeat Lys-Glu-[Leu or Phe]-Gly-Lys-Glu-Phe-Gly. Note that this repeat contains a subrepeat Lys-Glu-[Leu or Phe]-Gly. See Fig. 4 for comparison with other kinases.
Figure 3. [See facing page for legend.]
protein sequences are printed in the one-letter code, amino acids are indicated by dots. Islamist amino acid conservation of amino acid residues within the catalytic kinase do-

dTKR

| dTKR | 354 | LLGPHONMPFGG |
| v-yes | 337 | GP GM PPFGG |

| dTKR | 377 | GNGCPCQAEHRLKSHALKEL |
| v-yes | 559 | G GC K K K K  |
| dPYC | 349 | G G G S S  |
| c-erec | 273 | G GC K K  |

| dTKR | 544 | PYNGLMIVAGKESLEKLEGREDC |
| v-yes | 627 | SLL G KE L Q  |
| dPYC | 423 | L Y K S K E K  |
| v-aro | 343 | SLL G G K L L Q  |
| v-fgr | 361 | SLL G G K L L Q  |
| EGF R | 279 | P L Y VA K  |

Figure 4. Homology between the dTKR protein and other serine, threonine, and tyrosine kinases and related oncogenes. [a] Amino-terminal homology with the avian v-yes oncogene. [b] Homology within the ATP-binding domain. [*] Gly-X-Gly-X-Gly motif. [+] putative ATP-binding lysine residue. [c] Conservation of amino acid residues within the catalytic kinase domain of several kinases and kinase-related oncogenes as described in detail by Hunter and Cooper (1985) and by Rosenthal et al. (1987) for the Drosophila protein kinase C homolog gene (dPKC). Numbers and abbreviations refer to the positions of sequences within the different proteins as described in the above references. Corresponding amino acid residues in the different protein sequences are printed in the one-letter code, exchanged amino acids are indicated by dots. [%] Percent amino acid conservation.

strong candidates for both an ATP-binding lysine residue [Kamps et al. 1984] and a tyrosine autophosphorylation site [Hunter and Cooper 1985] within the dTKR amino acid sequence (see below).

The presumptive ATP-binding lysine residue within the dTKR sequence [position 398, Figs. 3b, 4] is preceded by the amino-terminal consensus sequence Gly-X-Gly-X-Gly, which is characteristic for an ATP-binding domain in protein kinases [Wierenga and Hol 1983]. This lysine, however, is part of an eight-amino-acid direct repeat. Thus, ATP-binding to this second lysine (position 406) within this repeat or even the ATP-binding to both residues at the same time cannot be excluded. As observed with several tyrosine kinases, the carboxy-terminal tyrosine of the dTKR sequence [position 744] contains a basic amino acid residue seven positions upstream of the tyrosine [Fig. 3b]. This tyrosine is adjacent to an acidic amino acid, thus, it fulfills the criteria for representing the site of highest autophosphorylation in tyrosine kinases [reviewed in Hunter and Cooper 1985]. The presence of putative sites for both ATP-binding and autophosphorylation suggests that the dTKR gene product may have tyrosine kinase function. Hydropathy plots [Kyte and Doolittle 1982] with the dTKR sequence did not show any evidence for a membrane-spanning domain. This may argue for dTKR being a cellular rather than a membrane-associated tyrosine kinase; in this case, the putative glycosylation sites [Fig. 3] would not be glycosylated.

Discussion

We report the identification of a tyrosine kinase-related Drosophila gene that is active during early stages of embryonic development. The domains of dTKR expression suggest specific and transient activity of this gene in a dorsal location and, later, in the developing nervous system. The conservation of an ATP-binding site and a tyrosine autophosphorylation site, as well as homology within the catalytic domains of several other protein kinases, suggests a similar, if not identical, biochemical function for this gene as known for well-characterized tyrosine kinases [reviewed in Hunter and Cooper 1985], although the invariant sequence His-Arg-Asp-Leu-X-Ala-X-Asn-X-Leu-Val [located just downstream of the sequences shown in Fig. 4c] of all tyrosine kinases is absent from the dTKR sequence.

Recently, the involvement of a specific tyrosine kinase activity in neuronal cell specification has been demonstrated with the Drosophila gene sevenless (sev). The sev protein product, a transmembrane receptor with a tyrosine kinase domain [Hafen et al. 1987], is essential for the development of a single class of photoreceptor cells in the compound eye [Tomlinson and Ready 1986; Banerjee et al. 1987]. Furthermore, cell communication seems to be essential for developmental decisions within the “neurogenic region” of the Drosophila blastoderm stage embryo [for review, see Campos-Ortega and Hartenstein 1985]. In the neurogenic region, cells of the same pool of precursors become committed to specify either neural or epidermal cells, which allows them to develop either the ventral nervous system or epidermis. This developmental decision involves the class of neurogenic genes [Lehmann et al. 1983]. One of the neurogenic genes, Notch, encodes a transmembrane protein and epidermal growth factor (EGF)–repeat sequences [Wharton et al. 1985; Kidd et al. 1986]. This finding is consistent with the genetic and molecular analysis favoring the argument that Notch [as well as several other neurogenic genes] participates in cell communication mediated by membrane-bound molecules. At least one other neurogenic gene contains EGF-like repeat sequences [Knust et al. 1987].

The transient and region-specific expression of dTKR favors the argument that dTKR has a regulatory function, as proposed for the other tyrosine kinases. This function may be, for example, phosphorylation of cellular proteins or receptors to modulate the activities of neurogenic gene products. This proposal for dTKR function is supported by the striking resemblance of the spatial patterns of expression of dTKR to a candidate for crumbs (crb), a putative neurogenic gene of Drosophila that contains EGF-like repeats [see above; Knust et al. 1987].

As part of the effort to understand the role of tyrosine kinases and in particular the role of cellular oncogenes containing tyrosine kinase domains, a number of Drosophila genes homologous to vertebrate oncogenes or hormone receptors have been isolated, including c-src, the EGF and insulin receptors [reviewed in Shilo 1987], and the protein kinase C homolog [Rosenthal et al. 1987]. No
mutants have been reported for any of these genes. For dTKR, we obtained chromosome deficiencies that uncover at least two genes in addition to dTKR (Côté et al. 1987). Thus, the biological function of dTKR cannot be analyzed unambiguously in those mutants. On the other hand, our molecular analysis supports the suggestion of a regulatory function for dTKR. This encourages further examination of the dTKR developmental pathway involving screens for additional dTKR mutations, interaction analysis with neurogenic genes, and the identification of the substrate for phosphorylation by dTKR.

Materials and methods

DNA procedures

DNA preparation of phages and plasmids, nick-translations to prepare 32P-labeled DNA fragments, and restriction map analysis were done as described in Schuh et al. [1986]. Single-strand probes for in situ hybridization experiments were prepared from T3/T7 constructs as described by Côté et al. [1987] using 35S-labeled ATP (Amersham).

In situ hybridization of embryo sections

In situ hybridization of molecular probes to tissue sections of embryos was carried out as described by Côté et al. [1987]. Briefly, single-strand 32P-labeled RNA probes were used on tissue sections of paraffin-embedded embryos. The specific activity of the probes used was above 107 cpm/µg, and the average size was smaller than 200 bp.

Sequencing

Both cDNA and genomic DNA fragments were sequenced from single-strand phage subclones by the dideoxy chain termination method of Sanger as described previously (Schuh et al. 1986). For the data presented in Figure 3, the coding sequences were sequenced in both orientations in several subclones.

Acknowledgments

We would like to thank our colleagues in the lab for many stimulating discussions, and Dr. S.M. Cohen for critically reading the manuscript. S.C. was supported by a NSERC of Canada postdoctoral fellowship. Part of the work was supported by a DFG grant (Leibniz-Programm).

REFERENCES

Banerjee, U., P.J. Renfranz, J.A. Pollack, and S. Benzer. 1987. Molecular characterization and expression of sevenless, a gene involved in neuronal pattern formation in the Drosophila eye. Cell 49: 281–291.

Campos-Ortega, J. and V. Hartenstein. 1985. The embryonic development of Drosophila melanogaster. Springer Verlag, Berlin, Heidelberg.

Côté, S., A. Preiss, J. Haller, R. Schuh, A. Kienlin, E. Seifert, and H. Jackle. 1987. The gooseberry-zipper region of Drosophila: Five genes encode different spatially restricted transcripts in the embryo. EMBO J. [in press]

Foe, V.E. and B.M. Alberts. 1983. Studies of the nuclear and cytoplasmic behaviour during the five mitotic cycles that precede gastrulation in Drosophila embryoogenesis. J. Cell. Sci. 61: 31–70.

Gehring, W.J. and Y. Hiromi. 1986. Homeotic genes and the homeobox. Annu. Rev. Genet. 20: 147–173.

Hafen, E., M. Levine, R.L. Garber, and W.J. Gehring. 1983. An improved in situ hybridization method for the detection of cellular RNAs in Drosophila tissue sections and its application for localizing transcripts of the homeotic Antennapedia gene complex. EMBO J. 2: 617–623.

Hafen, E., K. Basler, J.-E. Edström, and G.M. Rubin. 1987. sevenless, a cell-specific homeotic gene of Drosophila, encodes a putative transmembrane receptor with a tyrosine kinase domain. Science 236: 55–63.

Hunter, T. and J.A. Cooper. 1985. Protein-tyrosine kinases. Annu. Rev. Biochem. 54: 897–930.

Kamps, M.P., S.S. Taylor, and B.M. Sefton. 1984. Direct evidence that oncogenic tyrosine kinases and cyclic AMP-dependent protein-kinases have homologous ATP-binding sites. Nature 310: 589–592.

Kidd, S., M.F. Kelley, and M.W. Young. 1986. Mol. Cell. Biol. 6: 3094–3108.

Knust, E., U. Dietrich, U. Tepass, K.A. Bremer, D. Weigel, H. Vässin, and J.A. Campos-Ortega. 1987. EGF homologous sequences encoded in the genome of Drosophila melanogaster, and their relation to neuronal genes. EMBO J. 6: 761–766.

Kyte, J. and R.F. Doolittle. 1982. Evaluation of protein hydropathy. J. Mol. Biol. 157: 105–132.

Lehmann, R., F. Jimenez, U. Dietrich, and J. Campos-Ortega. 1983. On the phenotype and development of mutants of early neurogenesis in Drosophila melanogaster. Wilhelm Roux’s Arch. Dev. Biol. 192: 62–74.

Macdonald, P.M. and G. Struhl. 1986. A molecular gradient in early Drosophila embryos and its role in specifying the body pattern. Nature 324: 537–545.

Martinez-Arias, A. and P. Lawrence. 1985. Parasegments and compartments in the Drosophila embryo. Nature 313: 639–642.

Mlodzik, M., A. Fjose, and W.J. Gehring. 1986. Isolation of canadal, a Drosophila homebox containing gene with maternal expression, whose transcripts form a concentration gradient at the preblastoderm stage. EMBO J. 4: 2961–2969.

Rosenthal, A., L. Rhee, Y. Yasegari, R. Paro, A. Ullrich, and D.V. Goeddel. 1987. Structure and nucleotide sequence of a Drosophila melanogaster protein kinase C gene. EMBO J. 6: 433–441.

Schuh, R., W. Aicher, U. Gaul, S. Cote, A. Preiss, D. Maier, E. Seifert, U. Nauber, C. Schröder, R. Kemmler, and H. Jäckle. 1986. A conserved family of nuclear proteins controlling structural elements of the finger protein encoded by Krüppel, a Drosophila segmentation gene. Cell 47: 1025–1032.

Scott, M.P. and P.H. O’Farrell. 1986. Spatial programming of gene expression in early Drosophila embryogenesis. Annu. Rev. Cell. Biol. 2: 49–80.

Shilo, B.-Z. 1987. Proto-oncogenes in Drosophila melanogaster. TIG 3: 69–73.

Tomlinson, A. and D.F. Ready. 1986. sevenless, a cell-specific homeotic mutant of the Drosophila eye. Science 231: 400–402.

Wharton, K.A., K.M. Johansen, T. Xu, and S. Artavanis-Tsakonas. 1985. Nucleotide sequence from the neurogenic locus Notch implies a gene product that shares homology with proteins containing EGF-repeats. Cell 43: 567–581.

Wierenga, R.K. and W.G.J. Hol. 1983. Predicted nucleotide-binding properties of p21 protein and its cancer-associated variant. Nature 302: 842–844.
Dorsal and neural expression of a tyrosine kinase-related Drosophila gene during embryonic development.

J Haller, S Côté, G Brönner, et al.

*Genes Dev.* 1987, 1:
Access the most recent version at doi:10.1101/gad.1.8.862