Int. J. Dev. Biol. 44: 349-359 (2000)

Original Article

Characterization of two novel lipocalins expressed in the *Drosophila* embryonic nervous system

DIEGO SÁNCHEZ*[1], MARÍA D. GANFORNINA[1], SONIA TORRES-SCHUMANN, SEAN D. SPEESE, JOSÉ M. LORA, and MICHAEL J. BASTIANI

Department of Biology, University of Utah, Salt Lake City, Utah, U.S.A.

ABSTRACT We have found two novel lipocalins in the fruit fly *Drosophila melanogaster* that are homologous to the grasshopper *Lazarillo*, a singular lipocalin within this protein family which functions in axon guidance during nervous system development. Sequence analysis suggests that the two *Drosophila* proteins are secreted and possess peptide regions unique in the lipocalin family. The mRNAs of DNLaz (for *Drosophila* neural *Lazarillo*) and DGLaz (for *Drosophila* glial *Lazarillo*) are expressed with different temporal patterns during embryogenesis. They show low levels of larval expression and are highly expressed in pupa and adult flies. DNLaz mRNA is transcribed in a subset of neurons and neuronal precursors in the embryonic CNS. DGLaz mRNA is found in a subset of glial cells of the CNS: the longitudinal glia and the medial cell body glia. Both lipocalins are also expressed outside the nervous system in the developing gut, fat body and amnioserosa. The DNLaz protein is detected in a subset of axons in the developing CNS. Treatment with a secretion blocker enhances the antibody labeling, indicating the DNLaz secreted nature. These findings make the embryonic nervous system expression of lipocalins a feature more widespread than previously thought. We propose that DNLaz and DGLaz may have a role in axonal outgrowth and pathfinding, although other putative functions are also discussed.

KEY WORDS: *longitudinal glia, neural development, lipocalin, medial cell body glia, Lazarillo.*

Introduction

The lipocalins form a family of small (~20 kDa) soluble extracellular proteins present in prokaryotes, protists, and complex metazoans (Flower, 1996; Ganfornina *et al*., 2000). The lipocalin folding motif (Cowan *et al*., 1990; Flower, 1995) is an eight-stranded antiparallel β-barrel, open at one side and enclosing a binding pocket. There are three short structurally conserved regions (SCRs) (Flower *et al*., 1993) that are located at the bottom of the β-barrel. A broad set of hydrophobic molecules (e.g., heme metabolites, retinoids, fatty acids, and small odorants) has been shown to bind to different lipocalins. Some lipocalins have an exquisite specificity for a given ligand, but most bind a variety of ligands of very different nature (Flower, 1995). Although the transport of hydrophobic molecules is the generally accepted role for lipocalins and abundant information is available about the structural and biochemical features of lipocalins, our knowledge of their physiological roles is still fragmentary.

Our work has focused on the functional role of *Lazarillo*, a grasshopper lipocalin that is heavily glycosylated, and is attached to the extracellular side of the neuronal plasma membrane through a glycosyl-phosphatidylinositol (GPI) tail (Ganfornina *et al*., 1995). *Lazarillo* is expressed by a subset of neuroblasts, ganglion mother cells and neurons of the central nervous system (CNS), by all sensory neurons of the peripheral nervous system (PNS), and by a subset of neurons of the enteric nervous system (ENS). It is also present in non-neuronal cells associated mainly with the excretory system:

Abbreviations used in this paper: aCC, anterior corner cell; AcP, anterior commissure pioneers; ApoD, apolipoprotein D; BBP, bilin-binding protein; CDS, coding sequence; CG, channel glia; CNS, central nervous system; DGLaz, *Drosophila* glial Lazarillo; DNLaz, *Drosophila* neural Lazarillo; Dmel.DGLaz, *Drosophila melanogaster* DGLaz; Dmel.DNLaz, *Drosophila melanogaster* DNLaz; ENS, enteric nervous system; ER, endoplasmic reticulum; Gmel.Gall, *Galleria mellonella* Gallerin; GPI, glycosyl-phosphatidylinositol; Hgam.CRC1, *Homarus gammarus* Crustacyanin 1; Hgam.CRC2, *Homarus gammarus* Crustacyanin 2; HRP, horseradish peroxydase; IG, interface glia or longitudinal glia; M-CBG, medial cell body glia; mAb, monoclonal antibody; Msex.IcyA, *Manduca sexta* Insecticyanin A; Msex.IcyB, *Manduca sexta* Insecticyanin B; ORF, open reading frame; Pbra.Bbp, *Pieris brassicae* Bilin-binding protein; PNS, peripheral nervous system; Same.Laz, *Schistocerca americana* Lazarillo; SCR, structurally conserved region; UTR, untranslated region.

***Address correspondence to:** D. Sánchez. Department of Biology, University of Utah, 257 South 1400 East, Salt Lake City, UT 84112-0840, USA. Tel.: 801 581 8616. Fax: 801 581 4668. e-mail: sanchez@bioscience.utah.edu

[1] D.S. and M.D.G. contributed equally to this work

0214-6282/2000/$20.00
© UBC Press
Printed in Spain
www.ehu.es/ijdb

cells of the Malpighian tubules, putative sessile haemocytes located under the embryonic amnioserosa, and cells of the subesophageal body (Sánchez *et al.*, 1995). We have shown that *Lazarillo* has a unique role within the lipocalin family: it is involved in the guidance of developing axons. This role has been demonstrated for a particular pair of commissural neurons, the AcP cells, that are misrouted when *Lazarillo* function is specifically perturbed by the anti-*Lazarillo* monoclonal antibody (mAb) 10E6 (Sánchez *et al.*, 1995).

We report here the discovery and molecular characterization of two new lipocalins, found in the fruitfly *Drosophila melanogaster*, which like Lazarillo, are expressed in the developing nervous system. The amino acid sequence of these new lipocalins makes them the closest relatives to the grasshopper *Lazarillo*. Using whole-mount *in situ* hybridization and immunocytochemistry we have demonstrated that the two *Drosophila* lipocalins are expressed in subsets of developing neurons and glial cells during embryogenesis. Because of their expression pattern and their homologous relationship to *Lazarillo*, we have called these lipocalins DNLaz (for *Drosophila* neural *Lazarillo*) and DGLaz (for *Drosophila* glial *Lazarillo*).

Results

Routine database searches of protein sequences released by the Berkeley *Drosophila* Genome Project (BDGP) uncovered significant similarities between *Lazarillo* and two different *Drosophila* sequences derived from: 1) a region of the P1 clone DS08613, and 2) an EST clone (GH09946) and its corresponding genomic sequence (P1 DS01087). We identified the complete sequence of the potential lipocalin in the P1 DS08613 sequence by using an open reading frame (ORF) prediction program available at the BDGP web site, along with the knowledge of the stereotypic lipocalin protein sequence. Similarly, the partial sequence of the EST GH09946 and predictions from the P1 DS01087 helped us to deduce the entire sequence of the second lipocalin.

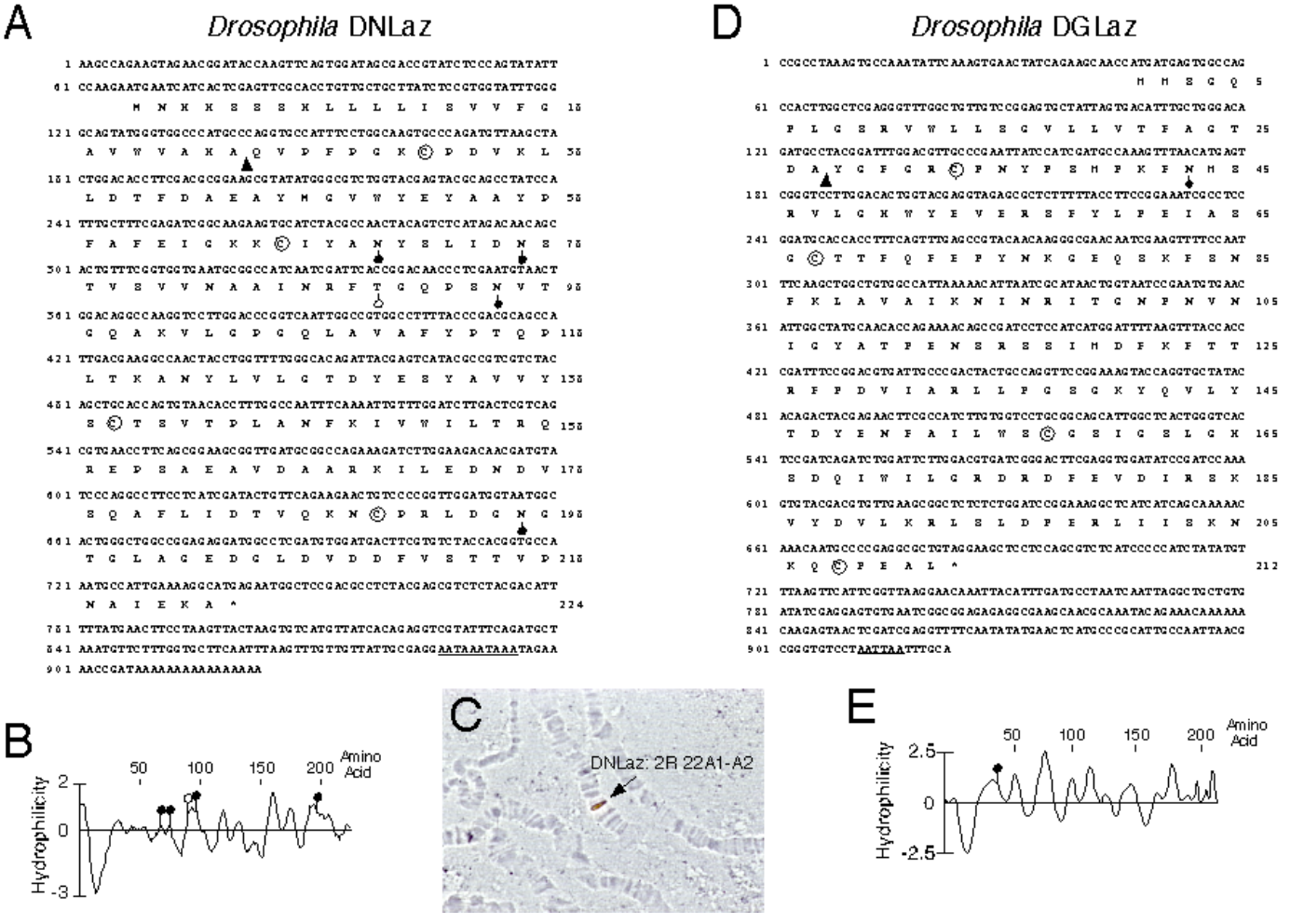

Fig. 1. The sequence of DNLaz and DGLaz cDNAs code for two putatively secreted lipocalins whose genes are located on chromosome two. (A,D) *cDNAs and deduced amino acid sequences of DNLaz and DGLaz. Nucleotide numbers are on the left and amino acid numbers are on the right (with respect to the first methionine). The predicted cleavage sites of the signal peptides are indicated by arrowheads. The potential N- and O-linked glycosylation sites are indicated by black and white dots respectively. Cysteine residues are circled. The polyadenylation sites are underlined.* **(B,E)** *Hydropathy plots of the predicted protein sequences determined by Kyte and Doolittle's method using a window of 9 residues. The dots represent the predicted glycosylation sites.* **(C)** *The DNLaz gene was mapped to the 22A1-2 band by in situ hybridization on larval polytene chromosomes. The hybridization signal is indicated by an arrow.*

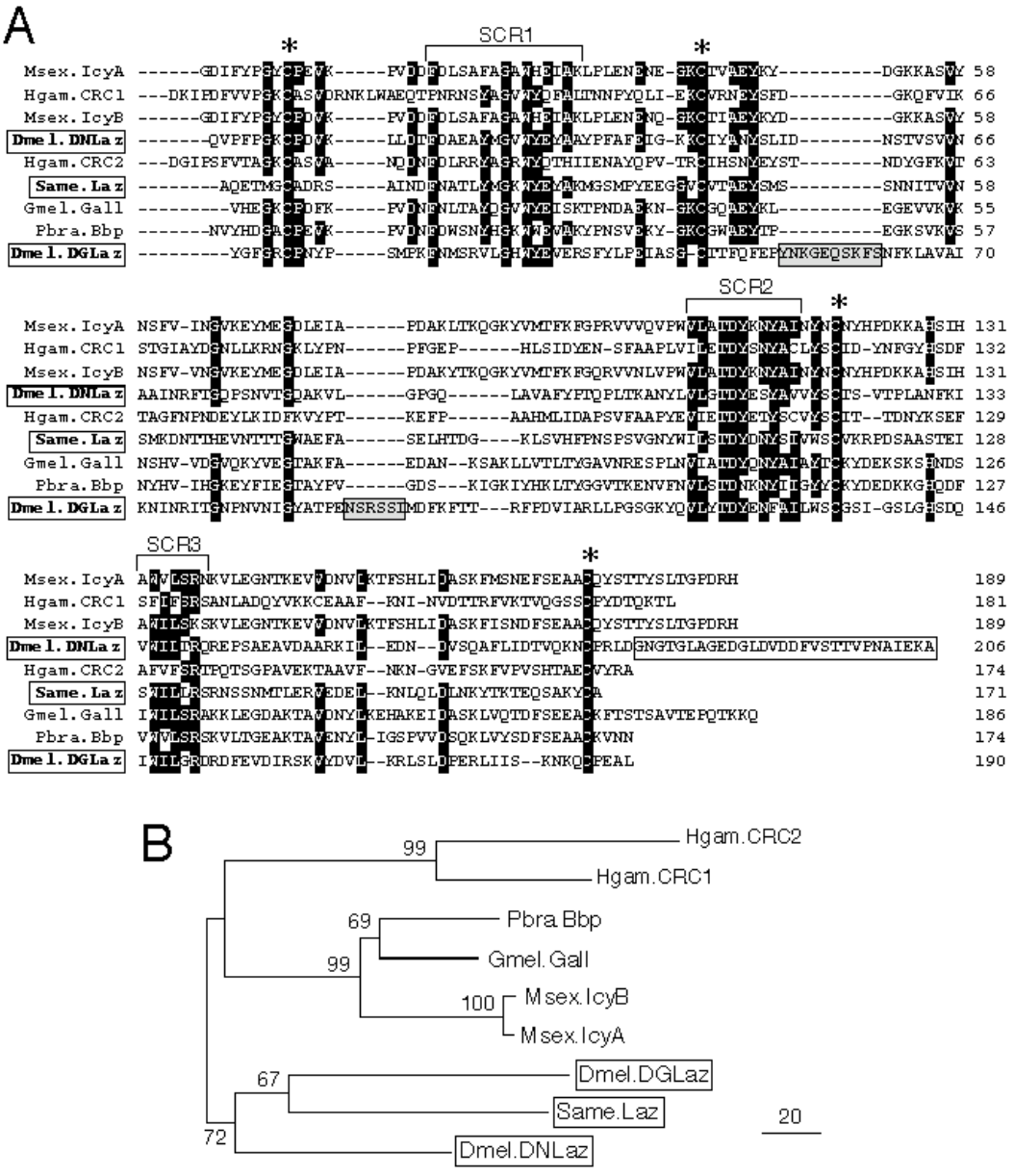

Fig. 2. Relationship of the *Drosophila* lipocalins to other arthropodan lipocalins. (**A**) *Sequence alignment of the predicted mature proteins of the novel Drosophila lipocalins (Dmel.DNLaz and Dmel.DGLaz) with Lazarillo (Same.Laz) and other arthropod lipocalins: Hgam.CRC1 and 2, Homarus gammarus Crustacyanins 1 and 2; Pbra.Bbp, Pieris brassicae Bilin-binding protein; Gmel.Gall, Galleria mellonella Gallerin; Msex.IcyA and B, Manduca sexta Insecticyanins A and B. The sequences were aligned with CLUSTAL X. Black boxes show residue identities present in six or more of the selected proteins. The three lipocalin structurally conserved regions (SCRs) are highlighted in the figure. The conserved cysteine residues are noted by asterisks. Two uniquely long loops present in DGLaz are boxed in gray, and the unique DNLaz C-terminus is highlighted by a white box.* (**B**) *Phylogenetic tree of the arthropodan lipocalins based on the alignment shown in A. Numbers at nodes are bootstrap values from 1000 replicates. The scale bar represents branch length (number of amino acid substitutions/100 residues).*

We have demonstrated by RT-PCR amplification from 6-18h embryonic total RNA, that these two novel genes are in fact transcribed during *Drosophila* embryogenesis. We have named *DNLaz* the gene found in P1 DS08613 and *DGLaz* the one corresponding to the EST GH09946.

Molecular characterization of the *Drosophila* genes *DNLaz* and *DGLaz*

The putative 5'-end of the DNLaz transcript was identified by RT-PCR with unique oligonucleotides designed from the region predicted to be the transcription initiation. The 3'-end of the DNLaz mRNA was obtained by 3'-RACE, and assembled to the 5'-end fragment. The minimal cDNA clone coding for DNLaz (Fig. 1A) has 907 bp, with a 66 bp 5'-untranslated region (UTR), and a 166 bp 3'-UTR flanking a predicted ORF. The DNLaz 3'-UTR shows several consensus polyadenylation sites.

A DGLaz cDNA clone was generated by using the sequence of the EST GH09946 clone, which includes the 5'-end of the mRNA, and by amplifying with RT-PCR a 3'-end fragment (based on predictions of polyadenylation sites) from *Drosophila* embryonic RNA. The minimal DGLaz cDNA (Fig. 1D) has 922 bp, and contains a single ORF, a 45 bp 5'-UTR, and a 237 bp 3'-UTR. The first AUG codons encountered in both lipocalin ORFs are in appropriate nucleotide contexts to be translation initiation sites.

The predicted DNLaz protein has a molecular mass of 21.6 kDa and an isoelectric point of 4.3, while the values for DGLaz are 21.2 kDa and 8.6 respectively. Both proteins have four cysteine residues, circled in Figure 1 A,D that are well conserved in the lipocalin family (Flower *et al.*, 1993). The hydrophilicity plots for DNLaz and DGLaz (Fig. 1 B,E) and the signal peptide predictions suggest that the *Drosophila* lipocalins have signal peptides at their N-termini. The expected cleavage site of the signal peptides are shown for each lipocalin sequence with black triangles in Fig. 1A,D. No other

hydrophobic or highly charged regions are found in the proteins, suggesting that these lipocalins are secreted proteins. The deduced DNLaz protein sequence shows one potential O-linked and four potential N-linked glycosylation sites (white and black dots respectively in Fig. 1A), and the DGLaz protein shows only one N-linked glycosylation site (black dot in Fig. 1D).

We mapped the *DNLaz* gene to the 22A1-2 genomic interval by *in situ* hybridization to polytene chromosomes of larval salivary glands with a digoxigenin-DNA probe derived from the DNLaz cDNA. The specific hybridization is shown in Fig. 1C. The *DGLaz* gene is located in the 49F12-13 region, according to the chromosomal position of two molecularly characterized genes, a *CNS-specific receptor tyrosine kinase* (Oishi *et al.*, 1997), and a *tripeptidyl peptidase II* (Renn *et al.*, 1998), that are adjacent to *DGLaz*. We did not find mutations or deficiencies that could help us to explore the role of the *Drosophila* lipocalin genes. However, there are nearby P-elements located in the vicinity of the loci of interest that can be mobilized to generate *DNLaz* and *DGLaz* mutants.

When the DNLaz and DGLaz mature protein sequences are compared to other lipocalins, they show global identity values ranging from 15-40%. Fig. 2A shows an alignment of DNLaz and DGLaz with a group of arthropodan lipocalins. Sequence conservation is high in the structurally conserved regions of the lipocalins (SCRs 1-3 in Fig. 2A) and in additional residues such as the four cysteines involved in forming two alternate disulfide bonds (asterisks in Fig. 2A). Sequence gaps are present in predicted loops according to the known tertiary structure of several lipocalins of this group. It is important to note the presence of two extended loops in DGLaz (gray boxes in Fig. 2A), which are in fact a unique feature in the entire lipocalin family. The two loops are located in the closed end of the predicted lipocalin calyx, opposite to the binding pocket entrance. The C-terminal region of DGLaz is very similar to a standard lipocalin, while that of DNLaz is rather long, exceeding by 29 residues (boxed in Fig. 2A) the C-termini of other arthropodan lipocalins. Interestingly, *Lazarillo* also has a long C-terminal region (not shown in Fig. 2A) that represents a GPI anchoring signal, a type of membrane attachment found in *Lazarillo* and unique within the lipocalin family (Ganfornina *et al.*, 1995). However, the DNLaz C-terminus is hydrophilic and does not resemble a GPI signal.

Based on the sequence alignment shown in Fig. 2A we carried out a phylogenetic analysis of the arthropodan lipocalins. The resulting unrooted tree (Fig. 2B) separates these proteins into two well-supported groups. One of the groups associates two crustacean lipocalins (Hgam.CRC1 and 2) and four lepidopteran lipocalins, while the two *Drosophila* lipocalins associate monophyletically with the grasshopper *Lazarillo* protein. Moreover, DGLaz groups monophyletically with *Lazarillo* in a well supported node. Therefore DGLaz seems to be more closely related to *Lazarillo* than DNLaz.

Temporal expression pattern of DNLaz and DGLaz

We characterized the temporal expression pattern of the *DNLaz* and *DGLaz* genes by Northern analysis of total RNA from different embryonic and larval stages, pupa, and adult flies. We used radiolabeled probes generated from the DNLaz and DGLaz cDNAs. The 18S rRNA band was assessed as a loading standard. The DNLaz probe hybridizes to a single band of 0.9 kb (Fig. 3), which matches the expected size predicted by the sequence in Fig. 1A. The DNLaz transcript is first detected in 12-16 hour embryos (embryonic stages 14-16), and increases as embryogenesis proceeds. DNLaz mRNA decreases during larval stages, increases during pupariation, and has its highest expression in adulthood.

A distinct band of 1 kb is detected with the DGLaz probe. DGLaz is expressed at high levels in 0-2 hour embryos. Subsequently, the DGLaz message decreases from 2-8 hours and then gradually increases after stages 11-12 (8-12 hours), reaching a peak at the end of embryogenesis. The DGLaz transcript is nearly absent during larval stages, and then increases dramatically in pupal stage and adulthood.

The *Drosophila* lipocalins are expressed in subsets of cells in the developing CNS

The expression patterns of the *Drosophila* lipocalins were studied by *in situ* hybridization on whole mount embryos with DNLaz and DGLaz digoxigenin-labeled RNA probes. The DNLaz mRNA is detected in the developing CNS from stage 12 until stage 16, after which the presence of the cuticle limits access of the probe. No labeling was obtained either in the peripheral or the enteric nervous system. At stage 14-15 the DNLaz signal is present in a subset of 12-15 cells/hemisegment (Fig. 4A). The changing patterns of labeling observed when comparing hemisegments of the same embryo may be the result of a transient expression of DNLaz within each cell, and suggests that DNLaz expression has a very fine temporal regulation. Most cells expressing DNLaz mRNA are located ventral to the neuropil. We show the focal plane of the axonal scaffold in Fig. 4B (commissures are marked by arrows). The location of some of these cells with respect to the neuropil makes possible to infer their identities (e.g. the aCC and the MP1 neurons shown in Fig. 4C, Thomas *et al.*, 1984) In addition to neurons, a subset of neuronal precursors seem to express DNLaz as well. A lateral view of the CNS reveals large DNLaz-expressing cells (arrows in Fig. 4D) at the ventral-most side of the nerve cord (the dorsal neuropil side is indicated by an arrowhead in Fig. 4D). These ventral cells (1 per neuromere) are located at the

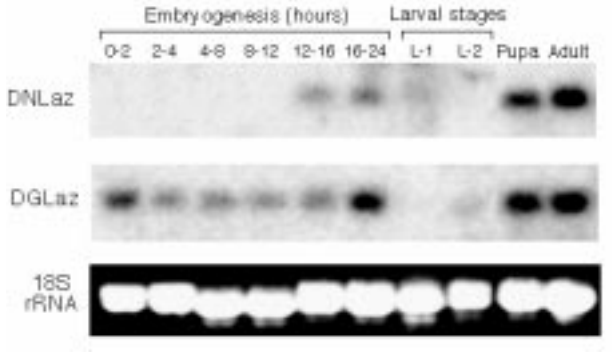

Fig. 3. Developmental Northern analysis of DNLaz and DGLaz transcripts. *Total RNA was extracted from different embryonic and larval stages, pupa and adult flies, and hybridized with the DNLaz and DGLaz cDNA probes. The 0.9 kb DNLaz mRNA is present during the second half of embryogenesis, decreases in amount during larval stages, and increases to a maximum in pupae and adult flies. The 1 kb DGLaz mRNA is highly transcribed during early embryogenesis, decreases at mid-embryogenesis, and increases at the end of the embryonic period. DGLaz mRNA decreases during larval stages, and increases in pupae and adult tissues. The ethidium bromide-stained 18S rRNA bands are shown below.*

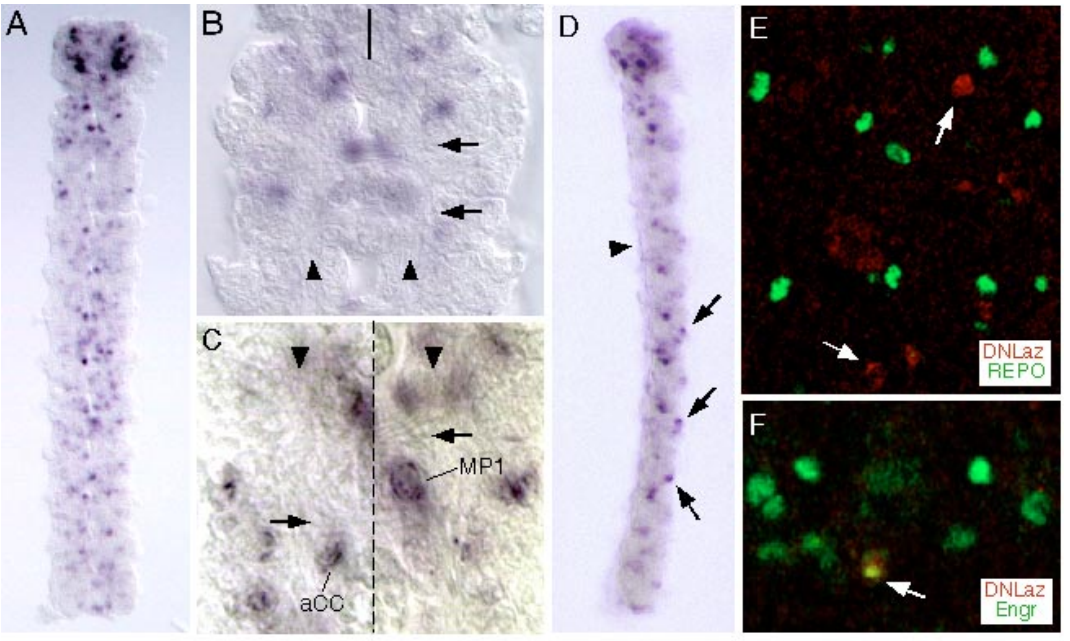

Fig. 4. CNS Expression pattern of DNLaz as revealed by *in situ* hybridization using an antisense DNLaz riboprobe. (**A**) *Dorsal view of a Drosophila CNS dissected out of a stage 15 embryo, showing the localization of the DNLaz mRNA in a restricted subset of cells.* **(B)** *A closer view of a dorsal focal plane of a thoracic neuromere showing the two commissures (arrows), the two longitudinal connectives (arrowheads), along with several out of focus DNLaz-positive cells. The midline is marked with a solid line.* **(C)** *Some of the cells expressing DNLaz that can be identified by position are shown in this composite panel of a dorsal view of two hemineuromeres of the same segment at a slightly different focal plane (midline marked by dashed line). Commissures and longitudinal connectives are indicated by arrows and arrowheads, respectively.* **(D)** *Lateral view of a nerve cord (stage 15) showing a neuroblast expressing DNLaz (arrows). The arrowhead points to the dorsal side of the nerve cord.* **(E,F)** *Confocal dorsal sections of embryonic nerve cords (stage 14) double-labeled with the DNLaz riboprobe and* (E) the REPO antiserum or (F) mAb 4D9 anti-engrailed. Some neurons (REP*O-negative, arrows in E) and an engrailed-positive cell (arrow in F) express DNLaz. Anterior is up in all panels. No labeling was observed with the DNLaz sense probe (not shown).*

midline, and might be the midline precursors, a specific set of neuronal and glial precursors of the *Drosophila* CNS (Bossing and Technau, 1994). In addition, DNLaz is expressed by a subset of unidentified cells in the CNS during the development of third instar larvae (not shown). In order to ascertain whether DNLaz is mainly expressed in one cell type within the CNS (i.e., neurons and neuronal precursors), we double-labeled *Drosophila* embryos with the DNLaz riboprobe and different antibodies that mark distinct cell types. In Figure 4E we show the DNLaz mRNA signal (red cells, some pointed by arrows) and glial cells expressing REPO, a nuclear protein expressed by all glial cells except the midline glia (Campbell *et al.*, 1994; Halter *et al.*, 1995). No co-localization of DNLaz and REPO is observed (Fig. 4E), demonstrating that DNLaz is expressed by a subset of postmitotic neurons. Similarly, we used a monoclonal antibody raised against the *Drosophila* engrailed protein (Patel *et al.*, 1989). The stripe of cells expressing engrailed includes a specific type of glial cells, the channel glia (CG) (Ito *et al.*, 1995). Double-labeling experiments showed that the DNLaz and engrailed signals co-localize in cells that, by position, could be the CG cells (arrow in Fig. 4F). However, given the results obtained in the double labeling with REPO, these cells are most probably engrailed-expressing neurons. The combined results of double-labeling experiments indicate that DNLaz is expressed mostly by a subset of neurons, but also by a precursor cell at the ventral midline in the embryonic *Drosophila* CNS.

The DGLaz transcript is also detected in the CNS, and like DNLaz, it is not present in the PNS or the ENS. The CNS pattern is first detected early in stage 12 and is evident until stage 16. The DGLaz signal appears mainly confined to the neuropil region (Fig. 5A), and more specifically to a symmetrical chain of cells (Fig. 5B) that is aligned with the axonal longitudinal connectives. There are 8 DGLaz positive cells/hemineuromere. Because of their position and morphology, the DGLaz-expressing cells can be identified as the interface glia (IG, also known as the longitudinal glia, Jacobs *et al.*, 1989). The IG cells continue to express DGLaz without a significant change in its spatiotemporal pattern until stage 16 (Fig. 5C). Figure 5D shows that the DGLaz transcript (blue labeling in Fig. 5D, black arrowheads) is located in the cytoplasm of cells surrounding the longitudinal connectives. The latter are labeled in brown with a HRP antiserum (Snow *et al.*, 1987) that reveals the axonal scaffold in a stage 14 nerve cord (commissures marked by white arrowheads in Fig. 5D). The identity of the longitudinal connective DGLaz-positive cells was determined by a double-labeling experiment with the DGLaz riboprobe (red cells in Fig. 5E) and the REPO antiserum (green nuclei in Fig. 5E). The co-localization of the DGLaz transcript and the REPO protein (yellow nuclei in Fig. 5E) demonstrates that the DGLaz expressing cells are glial in nature and confirms that they are the IG cells. Another DGLaz positive cell appears in each hemineuromere of embryos from stages 13 to 16 (arrowheads in Fig. 5F). These cells are located in the ventral half of the neuroectoderm, and occupy a lateral and posterior position (close to the neuroblast 6-4). The glial identity of these cells was demonstrated by the co-localization of DGLaz and REPO in a double-labeling experiment (arrows in Fig. 5G). Therefore, a subset of glia, and no other cell type, express DGLaz within the CNS.

Non-neural embryonic expression of DNLaz and DGLaz

The *Drosophila* lipocalins are also expressed in other tissues outside the nervous system during embryogenesis. DNLaz is detected at stage 11 in some scattered cells in the developing amnioserosa (white arrowheads in Fig. 6A) and some cells that seem to be located in the visceral mesoderm (black arrowheads in

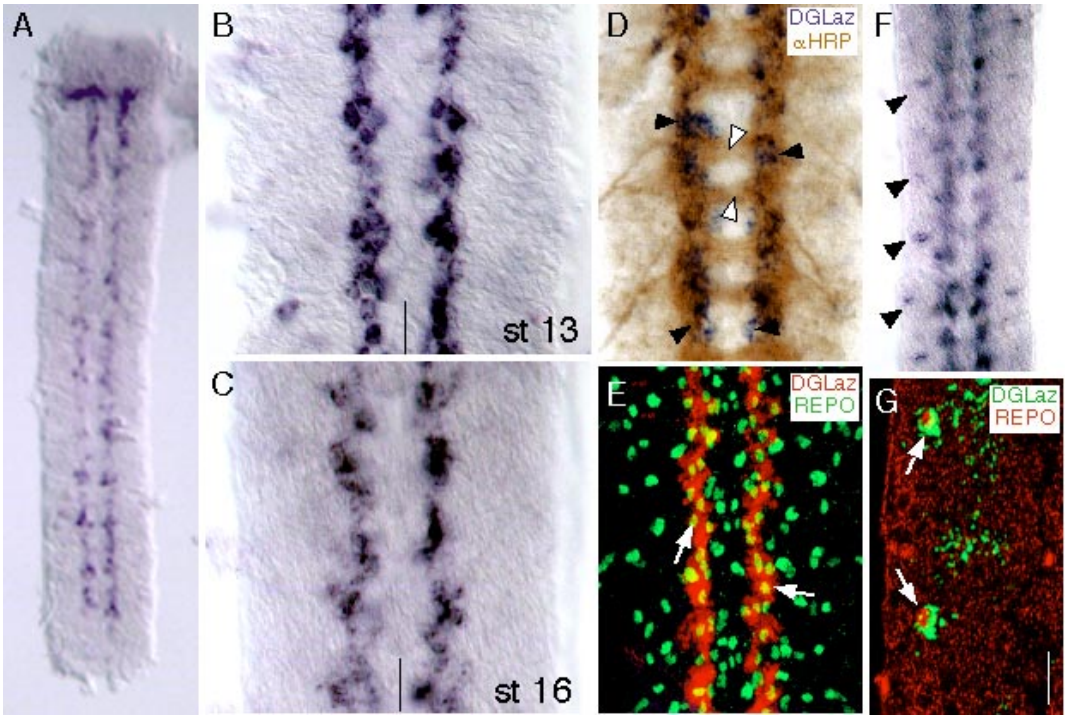

Fig. 5. Expression pattern of DGLaz during CNS development. *A DGLaz riboprobe was used in whole-mount in situ hybridization. (**A**) Dorsal view of a Drosophila CNS showing the localization of DGLaz mRNA in a subset of cells. (**B**) Dorsal focal plane of two abdominal neuromeres of a stage 13 embryo showing the DGLaz-positive cells. (**C**) These cells maintain DGLaz expression during the succeeding stages of embryogenesis. (**D**) The DGLaz cells (labeled by the DGLaz riboprobe, some marked by black arrowheads) are located along the longitudinal connectives of the axonal scaffold, as revealed by double-labeling with HRP antiserum (commissures of a neuromere marked by white arrowheads). (**E**) Z-series of confocal sections covering three neuromeres (stage 14) double-labeled with the DGLaz riboprobe and REPO antiserum. The co-localization of REPO nuclei and the DGLaz cytoplasmic labeling (some indicated by arrows) identifies the interface glia as the cells expressing DGLaz. (**F**) Dorsal view of several neuromeres of a stage 14 embryo showing another DGLaz positive cell located laterally at the ventral side of each neuromere (arrowheads). (**G**) Confocal dorsal section of the cells shown in F double-labeled with the DGLaz riboprobe and the REPO antiserum. The co-localization (arrows) identifies the M-CBG glial cells. The midline is marked in B, C, and G with a solid line. Anterior is up. No labeling was observed with the DGLaz sense probe (not shown).*

Fig. 6A). The identity and final fate of these cells are unknown. A group of mesodermal cells, anteriorly and dorsally located (arrow in Fig. 6A), start expressing DNLaz at stage 10-11. These cells seem to migrate along a dorsomedial pathway (large black arrow in Fig. 6B) and invade the developing amnioserosa by stages 12-13 (white arrowhead in Fig. 6B). Another tissue expressing DNLaz is the fat body, where the signal is first detected at stages 12-13 (white arrow in Fig. 6B). A group of cells of the midgut primordium at the foregut-midgut boundary also express DNLaz at stage 13. These cells (not shown) keep expressing DNLaz throughout embryogenesis, and appear more intensely labeled in the gastric caeci at stage 15. Some cells of the hindgut (black arrowhead in Fig. 6B) also express DNLaz after stage 13. Numerous gut cells expressing DNLaz are seen at later stages of embryogenesis (arrowheads in Fig. 6C). These cells appear to be sessile cells on the surface of the gut epithelium, and not the epithelial cells proper. DGLaz is also expressed in the embryo in a dynamic and tissue-specific pattern. It is detected early at the blastoderm stage (not shown), and is later expressed in the visceral mesoderm at stage 9-10 (black arrowheads in Fig. 6D). A weak DGLaz labeling is also observed in the amnioserosa (not shown). The DGLaz transcript is very abundant in cells of the developing gut. The epithelial cells of the hindgut primordium start expressing DGLaz at stage 9 (white arrowhead in Fig. 6D). The cells of the midgut primordium express DGLaz at the foregut- and hindgut-midgut boundaries at stages 13-14 (large black arrows in Fig. 6 E,F). The midgut epithelial cells located at the prospective sites of the midgut constriction, start expressing DGLaz before any morphological change is evidenced in the gut epithelium (stage 14, not shown). This expression is maintained while the midgut constrictions form (white arrows in Fig. 6F, stage 15) and it continues during stage 16 (not shown). Other places of DGLaz expression are the salivary glands primordia, a subset of cells located in the clipeolabrum, and the primordia of the anal pad and the posterior spiracles (data not shown). See Bate and Martínez-Arias (1993) for a general description of the morphological structures mentioned.

The DNLaz protein is secreted from developing axons

Two different antibodies were raised against selected peptides of the DNLaz sequence. These peptides are located in protein regions predicted to be hydrophilic, exposed, and not glycosylated. Both anti-DNLaz antibodies recognize the DNLaz protein in immunocytochemistry experiments, but the results we present here were obtained with the anti-peptide A antiserum (see Materials and Methods). The DNLaz protein is detected in embryonic amnioserosa (arrows in Fig. 7A-C), and fat body cells (arrowheads in Fig. 7B). Moreover, the DNLaz protein appears in the nervous system mainly restricted to the dorsal aspect of the embryonic nerve cord and brain (arrow and arrowhead respectively in Fig. 7D). This labeling is specific for the DNLaz antiserum, since it is absent in embryos exposed to preimmune serum (not shown), and to immune serum preincubated with excess peptide A (Fig. 7E).

The CNS labeling obtained with anti-DNLaz sera is due to the presence of the protein in the embryonic neuropil (Fig. 7F). The labeling is diffuse and therefore it is difficult to determine the precise DNLaz cellular localization from these results. When *Drosophila* embryos are treated with monensin, a potent inhibitor of secretion (Tartakoff, 1983), the labeling obtained with the anti-DNLaz antibody changed dramatically (Fig. 7G-I). The monensin

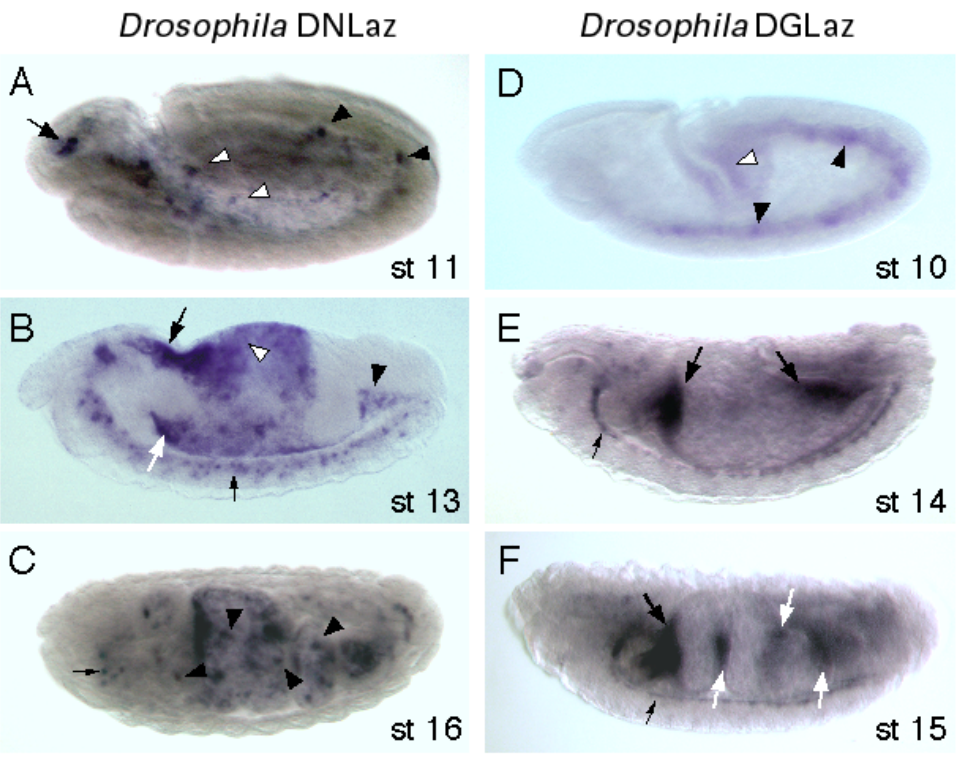

Fig. 6. Non-neural embryonic expression pattern of DNLaz and DGLaz as revealed by whole mount *in situ* hybridizations with an antisense DNLaz (A-C) or DGLaz (D-F) riboprobe. *Lateral views of embryos of different stages. Anterior is to the left, dorsal is up.* **(A)** *Arrow, black arrowheads and white arrowheads point to DNLaz-positive cells of mesodermal origin, cells in the visceral mesoderm and over the amnioserosa respectively.* **(B)** *Large black arrow, white arrow, black arrowhead and white arrowhead point to DNLaz expression in migrating mesodermal cells, the fat body, hindgut and the amnioserosa respectively.* **(C)** *Arrowheads point to DNLaz positive cells in the gut.* **(D)** *Black and white arrowheads point to DGLaz expression in the visceral mesoderm and hindgut primordium respectively.* **(E)** *Large black arrows point to DGLaz expression in the foregut- and hindgut-midgut boundaries.* **(F)** *White arrows point to DGLaz expression in the epithelium of the midgut constrictions. Black arrow indicates the expression of DNLaz in the foregut-midgut boundary. The CNS labeling of DNLaz and DGLaz is indicated by small arrows in B, C, E, and F.*

treatment results in a stronger labeling (Fig. 7G), and in a more clear delineation of the neurites contour (arrowheads in Fig. 7G), which is likely due to the sequestration of DNLaz-containing vesicles inside the developing neurites. The monensin experiments also confirm that the DNLaz antibody labels a subset of embryonic axons (Fig. 7H). At stage 16, only one of the three longitudinal axon fascicles, and a small number of commissural axons are labeled (arrows and arrowheads respectively in Fig. 7H). DNLaz is also present in axons of the medial fiber tract (Goodman and Doe, 1993), running along the CNS midline (small arrow in Fig. 7H), and around the progeny of the dorsal medial neuroblast (small arrowhead in Fig. 7H). The neurites labeled with the anti-DNLaz sera can be observed not only in the ventral nerve cord, but also in the brain neuropil (arrows in Fig. 7I). Some axons exiting the CNS are also labeled (not shown), which should be motor axons given the absence of PNS cells expressing DNLaz. These results together show that the DNLaz protein is produced by a subset of neurons, transported into vesicles along developing neurites, and secreted to the extracellular environment.

Discussion

We have identified two new lipocalins in the fruitfly *Drosophila melanogaster* that display significant sequence identity to the grasshopper lipocalin *Lazarillo*. The sequence similarity and the shared expression in the developing nervous system support the homology of these lipocalins, and set them apart in the lipocalin family. In contrast, the four other lipocalins that have been characterized in different arthropods may function in cuticle pigmentation. These are the bilin-binding proteins (BBP) of two different butterflies (Holden *et al.*, 1987; Huber *et al.*, 1987), the protein Gallerin from the moth *Galleria mellonella* (F. Sehnal, pers. comm.) and the lobster crustacyanins (Keen *et al.*, 1991a,b). The insect and crustacean lipocalins are synthesized in the fat body and epidermis, and the hepatopancreas respectively. Only Gallerin has been detected by *in situ* hybridization in glial cells of the larval nervous system (F. Sehnal, pers. comm.), but its biological role is unknown.

The molecular properties of Drosophila lipocalins

The cDNA clones of DNLaz and DGLaz and their deduced amino acid sequences clearly show the characteristics of a kernel lipocalin: 1) They have mRNAs of approximately 1 kb. 2) They contain N-terminal signal peptides, that along with the absence of other potential transmembrane-spanning domains suggests that DNLaz and DGLaz, like most lipocalins, are secreted proteins. 3) They have the three SCRs present in most lipocalins. 4) They show the conserved pattern of cysteine residues involved in two alternating disulfide bonds. However, like their homologous protein *Lazarillo*, DNLaz and DGLaz show several features uncommon for the lipocalins. DNLaz shows an unusually long C-terminal region that is relatively rich in glycine residues, which probably confers a greater flexibility to the DNLaz C-terminus. This C-terminal region also has a very acidic region that might be important for establishing electrostatic interactions with other proteins. On the other hand, DGLaz shows two rather unique polypeptide regions that are not present in any other known lipocalin. According to the sequence alignment and a homology-based three-dimensional model (not shown), these regions in DGLaz might be hydrophilic loops located at the closed end of the lipocalin calyx. This location suggests that the loops are not involved in ligand binding, but they could be important for specific protein-protein interactions, as it has been proposed for regions of the closed end of the lipocalin fold (Flower *et al.*, 1993). Finally, the potential glycosylation state of the *Drosophila* lipocalins also distinguishes them from other arthropodan lipocalins, and relates them to the heavily glycosylated

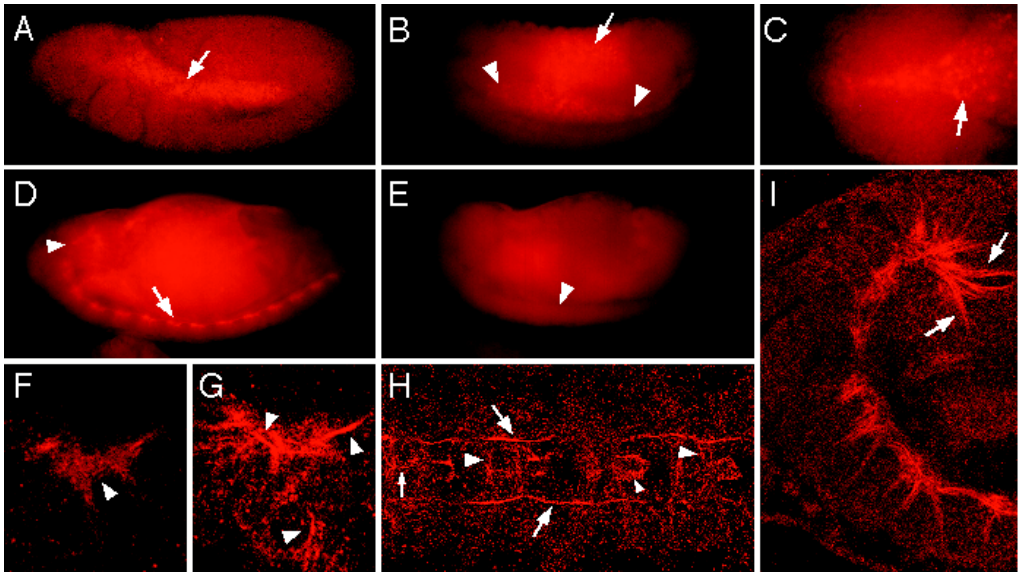

Fig. 7. Protein localization of DNLaz in whole mount *Drosophila* embryos using a DNLaz peptide antiserum and a fluorescent secondary antibody. *Anterior is to the left in all panels. Dorsal is up in all panels with the exception of the dorsal views shown in C and H.* **(A-C)** *The DNLaz protein is detected in the amnioserosa (arrows) and the fat body (arrowheads).* **(D)** *DNLaz is also detected in the ventral nerve cord (arrow) and brain (arrowhead).* **(E)** *The CNS labeling is not present if the DNLaz antiserum is preincubated with the immunizing peptide. The DNLaz labeling is different in control* **(F)** *versus monensin-treated* **(G)** *embryos.* **(H)** *A subset of developing axons is labeled with the DNLaz antiserum. A single longitudinal fascicle (arrows), a subset of commissural axons (big arrowheads), the medial fiber tract (small arrow) and some processes around the medial neuroblast (small arrowhead) are indicated.* **(I)** *Lateral view of the subesophageal ganglia and the brain showing neurites labeled with the DNLaz antiserum (arrows).*

Lazarillo (Ganfornina *et al.*, 1995). None of the lobster crustacyanins, the *Manduca* and *Pieris* BBPs, and the partial sequence of two new lipocalins of *Bombyx mori* (Acc. numbers: AU002238, AU002083) show consensus glycosylation sites. Interestingly, those arthropodan lipocalins with nervous system expression (Lazarillo, DNLaz, DGLaz and Gallerin) show potential glycosylation sites, suggesting a putative relationship between the presence of sugars in these lipocalins and their role in nervous system development and/or physiology.

The spatiotemporal expression of Drosophila lipocalins

The transcription of the two *Drosophila* lipocalin genes has a radically different temporal pattern during embryogenesis. The DNLaz mRNA is detected, and gradually increases, in the second half of embryogenesis. However, the presence of DGLaz mRNA shows a bimodal distribution during embryogenesis, with peaks at the beginning and at the end of this period. The *in situ* labeling observed in blastoderm-stage embryos and the early DGLaz signal detected by Northern analysis could be due to maternally supplied mRNA. An interesting result of our Northern analysis is the extremely low levels of DNLaz and DGLaz mRNAs detected during larval stages, in contrast with the relatively abundant lipocalin mRNAs observed in other insect larvae (Li and Riddiford, 1994; Schmidt and Skerra, 1994). In contrast, both lipocalin transcripts are very abundant during pupal stage, when the nervous system undergoes intense reorganization. The maximal mRNA production of the *Drosophila* lipocalins occurs during adulthood. This adult neural expression is also a shared characteristic with the grasshopper Lazarillo (Ganfornina *et al.*, 1995), and it could involve different roles for lipocalins in nervous system function.

Similarly to the grasshopper Lazarillo (Sánchez *et al.*, 1995), the *Drosophila* lipocalins are expressed in other organs and tissues. DGLaz is principally expressed by epithelial cells of the developing gut. Here, as well as in the salivary glands and posterior spiracle primordia, DGLaz transcription seems to be temporally associated with epithelial invagination. This suggests a potential role for DGLaz in the epithelial movements necessary for the organogenesis of the gut, trachea, and salivary glands. DNLaz is also expressed in the gut. However, expression is detected in sessile cells, but not in the gut epithelium. The expression in the fat body is worth mentioning because of the prevalence of this tissue (and its crustacean equivalent, the hepatopancreas) for lipocalin production in other arthropods. An intriguing finding is that DNLaz is also expressed in a group of mesodermal cells located in the clipeolabrum. These DNLaz positive cells seem to migrate along a narrow dorsomedial pathway into the amnioserosa, and spread from this tissue into the prospective gut. These cells will eventually condense at the foregut-midgut boundary. DNLaz seems to be a specific marker for these cells, whose fate and physiological role are still unknown (V. Hartenstein, personal communication). DNLaz is also expressed by cells that migrate on the surface of the amnioserosa, a pattern similar to that reported for Lazarillo in the grasshopper (Sánchez *et al.*, 1995).

The nervous system expression and function of Drosophila lipocalins

Both *Drosophila* lipocalins are expressed in the developing nervous system at a time when the axonal scaffold is being laid down. The DGLaz transcript is expressed by the IG cells, a group of 8-9 glial cells/hemineuromere (Ito *et al.*, 1995; Schmidt *et al.*, 1997) that arise from a special glioblast (Jacobs *et al.*, 1989; Schmidt *et al.*, 1997) and migrate toward the dorsal neuropil. Their processes enwrap axons that run along the longitudinal connectives (Jacobs and Goodman, 1989). The IG cells express DGLaz shortly after they reach their final destination at the dorsal surface of the nerve cord, a time when the first axons are forming the longitudinal connectives and commissures. Interestingly, the interaction of the IG cells with specific neurons regulates the expression of particular membrane proteins on the neuronal surface (Klaes *et al.*, 1994). Also, mutant embryos lacking the transcription factor *pnt* show defects both in longitudinal glial differentiation and axon scaffold formation (Klaes *et al.*, 1994). Moreover, ablation experiments and analysis of glial cells missing mutants show that the IG cells directly contribute to the formation

and maintenance of longitudinal axon tracts (Hidalgo *et al.*, 1995; Booth *et al.*, 2000; Hidalgo and Booth, 2000). These data support the view of the IGs being involved in axon guidance (reviewed by Klämbt *et al.*, 1996; Klämbt *et al.*, 1999). Given the specific expression of DGLaz in IG cells, it is possible that this lipocalin, like Lazarillo, plays a role in axon guidance.

The only other CNS cell that expresses DGLaz during embryonic development is located ventrally. We have identified the glial identity of this cell by double labeling with the glial marker REPO. Because of its position in each neuromere, we propose this is the medial cell body glia (M-CBG), a cell that originates from NB 6-4 (Schmidt *et al.*, 1997) and whose function is currently unknown.

We have shown the pattern of CNS expression of DNLaz mRNA and protein. The DNLaz mRNA is detected in a subset of CNS neurons during the developmental stages where the nervous system is forming. Among the 25-30 cells expressing DNLaz per neuromere, we can assign a neuronal identity to some of them (MP1 and aCC neurons). However, a specific class of glial cells, the channel glia, may also express DNLaz. The DNLaz neurons send their axons along the main axonal routes of the embryonic nerve cord: longitudinal connectives, commissures, and medial fiber tract. Moreover, some axons that exit the nerve cord are seen labeled with the DNLaz antiserum, which indicates that some motoneurons are expressing this lipocalin. An identified neuroblast, a midline precursor, also expresses DNLaz in every neuromere. The restricted expression of DNLaz to a subset of neuroblasts and neurons resembles that of Lazarillo (Sánchez *et al.*, 1995). The labeling obtained in embryos treated with monensin, an ionophore which mainly impairs the vesicular traffic from the ER to the Golgi apparatus (Tartakoff, 1983), confirms that DNLaz is secreted to the extracellular environment. Furthermore, the comparison of the control labeling with the monensin treatment suggests that, after release from the cell, DNLaz remains in the immediate surroundings of the secreting axons probably by interacting with extracellular matrix proteins or with membrane proteins in the same or neighboring axons. A similar antibody labeling behavior was observed for the growth cone-specific protein Conulin (Sánchez *et al.*, 1996). Further experiments will be needed to resolve the extracellular localization of the protein. In summary, our results show the spatially and temporally dynamic expression of DNLaz in the developing CNS, which is also consistent with a role for this lipocalin in axon outgrowth and guidance.

A model for the molecular mechanism of function of Drosophila lipocalins

The results presented above point to a role for the *Drosophila* lipocalins in nervous system development, possibly similar to the one demonstrated for the grasshopper Lazarillo (Sánchez *et al.*, 1995). Given the general role of lipocalins as transporters of hydrophobic molecules, we could propose a molecular mechanism by which a hydrophobic ligand specific for each *Drosophila* lipocalin would be involved in helping particular axons to properly grow and decide the appropriate pathway. However, given that these proteins are secreted from neighboring cells within the CNS, an alternative mechanism based on molecular properties other than hydrophobic ligand binding could underlie their role in neural development. The analysis of homology-based DNLaz and DGLaz molecular models shows that the binding pockets of the *Drosophila* lipocalins do not greatly differ from each other or from other lipocalins (e.g., the Manduca Insecticyanin or the vertebrate ApoD) in the amino acid residues potentially interacting with a hydrophobic ligand. On the contrary, the *Drosophila* lipocalins show a unique repertoire of molecular features not shared between them (e.g., glycosylation potential, predicted pI), and some of them are even unique in the family (the long loops of DGLaz, and the C-terminal DNLaz region). Although we can not discard the hypothesis stated above that different hydrophobic ligands could specifically bind to either DNLaz or DGLaz, our results support better the alternative hypothesis that functional specificity is based on protein-protein interactions, with the hydrophobic ligand (the same or different) being a regulator of those interactions between proteins.

The evolutionary history of nervous system lipocalins

In spite of sequence and expression pattern similarities, the precise evolutionary relationships between Lazarillo, DNLaz, and DGLaz remains uncertain. The discovery of more lipocalins in *Drosophila* might help to resolve these relationships. However, the existence of two lipocalins in lobsters (Acc. #: P80029, P80007) and in the moths *Manduca sexta* (Acc. #: P00305, Q00630), *Bombyx mori* (Acc. #: AU002238, AU002083) and *Samia cynthia* (Saito, 1998), together with our phylogenetic analysis of lipocalins (Ganfornina *et al.*, 2000) suggest that two lipocalins are indeed the expected number in the arthropod lineage.

We have shown that the two *Drosophila* lipocalins are expressed in subsets of developing cells during nervous system embryogenesis, a characteristic that was so far unique to the grasshopper Lazarillo. These three genes are part of a monophyletic group of lipocalins that arise from the most ancestral branch of the lipocalin tree within the metazoan lineage (Ganfornina *et al.*, 2000). It is an intriguing question whether this expression and potential function in the development of the nervous system is in fact an ancestral character for the family, and whether they were subsequently co-opted (Ganfornina and Sánchez, 1999) for many other roles (including the general one of hydrophobic ligand transporters). The potential for multifunctionality might have been present since their origin because of their expression in other developing cells outside the nervous system, some of which are shared with other lipocalins. Besides the common features between Lazarillo, DNLaz and DGLaz, they show interesting differences, both in their molecular properties and in their specific expression sites. Within the nervous system, Lazarillo is specific of neurons, DNLaz is mainly in neurons, and DGLaz is glial-specific. Now, with *Drosophila* genetic techniques we will be able to test whether these lipocalins share a common function in axonal pathfinding, even when produced from different cellular sources, or whether their functions have diverged concomitantly with their changes in expression pattern.

Materials and Methods

Molecular cloning and sequencing of the *Drosophila* lipocalins DNLaz and DGLaz

Total RNA used for RT-PCR was obtained from *Drosophila* embryos as described (Ganfornina *et al.*, 1995). Reverse transcription was carried out using AMV reverse transcriptase (USB) at 45 ºC for 30 min. The transcription reaction was followed by phenol extraction and filtration through a microcon-100 (Amicon). PCR experiments were performed in a Perkin Elmer machine using unique primers. The PCR-amplified fragments were subcloned into pCR-II using TOPO-TA cloning (Invitrogen), and sequenced on an ABI Prism 377 automated DNA sequencer using Taq FS DNA Polymerase.

Sequence and phylogenetic analyses

Similarity searches with DNA and protein sequences were analyzed with the BLAST service available at the NCBI web site. Promoter and ORF predictions were performed with programs available at the BDGP web site. The *Drosophila* lipocalin sequences were aligned with CLUSTAL X (Thompson *et al.*, 1997) using a Gonnet series scoring matrix. Phylogenetic analysis was carried out using the PHYLIP (3.5) (Felsenstein, 1993) package.

Analysis of DNLaz and DGLaz mRNA expression

Northern analyses. Total RNA was extracted from embryos and larvae of different stages, pupae, and adult flies using Trizol. The RNA (10 μg of each developmental stage) was electrophoresed in a formaldehyde 1.2% agarose gel, and blotted to a nylon membrane. The blotted RNAs were then hybridized in Ultrahyb solution (Ambion) for 18 hours at 65 ºC to either DNLaz or DGLaz radiolabeled probes generated from the entire DNLaz or DGLaz cDNAs. The membranes were then washed at 65 ºC, and exposed to film.

Whole-mount in situ hybridizations. Digoxigenin-11-dUTP labeled RNA probes were synthesized (Genius-4 kit, Boehringer-Mannheim) using the DNLaz and DGLaz cDNA clones as templates. Embryos were dechorionated in 50% clorox for 5 min, fixed in equal volumes of heptane and PEM-formaldehyde (4% formaldehyde in 0.1 M PIPES pH 6.9, 2 mM EGTA, 1 mM $MgSO_4$) for 45 minutes, and devitellinized in a heptane/methanol mixture. The hybridization solutions and protocol were previously described (Ganfornina *et al.*, 1995). In order to detect the labeled RNA, either alkaline phosphatase-conjugated or rhodamine-conjugated anti-digoxigenin antibodies were used (Boehringer-Mannheim). Alkaline phosphatase-reacted embryos were mounted in 90% glycerol, and fluorescently labeled embryos were mounted in Vectashield (Vector).

Chromosomal localization of DNLaz

Squashes of polytene chromosomes were obtained from third instar larvae. A digoxigenin-labeled DNA probe was prepared by PCR amplification of a 763 bp fragment. Hybridization to denatured chromosomes was performed at 65 ºC for 18 hours, and washes were performed at 60 ºC. The hybridization reaction was detected with an alkaline phosphatase-conjugated anti-digoxigenin antibody (Boehringer-Mannheim), the chromosomes were visualized with a phase-contrast Zeiss Microscope.

DNLaz peptides used to generate antibodies

Two different peptides were chosen for immunizations based on their antigenicity index and their exposed nature, as predicted by a DNLaz three-dimensional model generated by the Swissmodel program. The sequence of peptide A (AGEDGLDVDDFV) corresponds to residues 202-213 of the DNLaz CDS, and that of peptide B (QREPSAEAVDAARK) represents residues 158-171 of DNLaz CDS. The synthetic DNLaz peptides were coupled to ovoalbumin.

Generation and purification of polyclonal sera against DNLaz

Swiss-Webster female mice were immunized according to the protocol by Ou *et al.*, (1993). 40 μg of each ovoalbumin-coupled DNLaz peptides, were emulsified (1:1) in MPL+TDM adjuvant (Sigma) and injected intraperitoneally. Injections were repeated at 14-day intervals. Reactive antibodies were used for whole mount *Drosophila* embryo immunocytochemistry. Antibodies generated against DNLaz peptides were affinity purified using the peptides linked to a HiTrap column (Amersham). The affinity purified antibodies were concentrated in a Centricon-50 (Amicon) and their concentration measured by gel electrophoresis with standard proteins.

Immunocytochemistry

Dechorionated *Drosophila* embryos were fixed in equal volumes of heptane and 4% paraformaldehyde in BBS (55 mM NaCl, 40 mM KCl, 7 mM $MgCl_2$, 5 mM $CaCl_2$, 20 mM glucose, 50 mM sucrose, 10 mM Tris pH 7.5) for 35 min at room temperature, and devitellinized in a heptane/methanol mixture. After washing in BBT (BBS, 0.1% Tween-20, 1.5 mg/ml BSA (Sigma), embryos were incubated for 30 min in blocking solution (BBT-5% normal goat serum). Anti-DNLaz antibodies were used at 10 μg/ml in blocking solution. In control experiments, the anti-DNLaz antibodies were incubated with the corresponding DNLaz peptide (1.5 mg/ml) at 20 ºC for 2 hours before being added to embryos. Other primary antibodies were used at concentrations recommended by the supplying laboratories. Embryos were incubated with the primary antibody at 20 ºC for 8-12 hours, then washed with BBT, blocked again and exposed to the secondary antibody of choice diluted in blocking solution for 2 hours at 20 ºC. Labeling was reacted with 0.5 mg/ml 3,3'-diaminobenzidine and 0.03% H_2O_2 when using a HRP-conjugated secondary antibody. The embryos were then washed and mounted either in glycerol 90%, for HRP-labeling, or in Vectashield when using a fluorescently-conjugated secondary antibody.

Monensin treatment

Treatment of live embryos with monensin was carried out as described (Gonzalez *et al.*, 1991). Dechorionated embryos were transferred to a 1:1 mixture of octane and Schneider's *Drosophila* medium (Gibco BRL) for 2 min, and then incubated for 30 min in medium containing 70 μM monensin (Sigma). Control embryos were incubated in medium for the same amount of time. Embryos were subsequently fixed and processed as described above.

Acknowledgments

We want to thank Tim Beagle for help with 3'-RACE, J.P. Couso and F. González for help with antibody labeling, and Virginie Pomies for technical assistance. We thank G. Technau and N. Patel for the gift of antibodies anti-REPO, and mAb 4D9 anti-Engrailed, respectively. We also thank V. Hartenstein and G. Technau for comments on the DNLaz expression pattern. We thank the DNA Sequencing and Peptide Synthesis Facilities at the University of Utah. This work was supported by NIH grant (2 R01 NS25387-10A1) to M.J.B.

Note added in proof. The accession numbers for the cDNA sequences described are: DNLaz, AF276505; DGLaz, AF276506.

References

BATE, M. and MARTÍNEZ-ARIAS, A. (Eds.) (1993). *The development of Drosophila melanogaster*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

BOOTH, G.E., KINRADE, E.F. and HIDALGO, A. (2000). Glia maintain follower neuron survival during *Drosophila* CNS development. *Development* 127: 237-244.

BOSSING, T. and TECHNAU, G.M. (1994). The fate of the CNS midline progenitors in *Drosophila* as revealed by a new method for single cell labeling. *Development* 120: 1895-1906.

CAMPBELL, G., GÖRING, H., LIN, T., SPANA, E., ANDERSSON, S., DOE, C.Q. and TOMLINSON, A. (1994). RK2, a glial-specific homeodomain protein required for embryonic nerve cord condensation and viability in *Drosophila. Development* 120: 2957-2966.

COWAN, S.W., NEWCOMER, M.E. and JONES, T.A. (1990). Crystallographic refinement of human serum retinol binding protein at 2 Å resolution. *Proteins. Struct. Funct. Genet.* 8: 44-61.

FELSENSTEIN, J. (1993). *PHYLIP: Phylogeny inference package.* Department of Genetics. University of Washington, Seattle.

FLOWER, D.R. (1995). Multiple molecular recognition properties of the lipocalin protein family. *J. Mol. Recogn.* 8: 185-195.

FLOWER, D.R. (1996). The lipocalin protein family: structure and function. *Biochem. J.* 318: 1-14.

FLOWER, D.R., NORTH, A.C.T. and ATTWOOD, T.K. (1993). Structure and sequence relationships in the lipocalins and related proteins. *Protein Sci.* 2: 753-761.

GANFORNINA, M.D. and SÁNCHEZ, D. (1999). Generation of evolutionary novelties by functional shift. *Bioessays* 21: 432-439.

GANFORNINA, M.D., GUTIÉRREZ, G., BASTIANI, M.J. and SÁNCHEZ, D. (2000). A phylogenetic analysis of the lipocalin protein family. *Mol. Biol. Evol.* 17: 114-126.

GANFORNINA, M.D., SÁNCHEZ, D. and BASTIANI, M.J. (1995). *Lazarillo*, a new GPI-linked surface lipocalin, is restricted to a subset of neurons in the grasshopper embryo. *Development* 121: 123-134.

GONZALEZ, F., SWALES, L., BEJSOVEC, A., SKAER, H. and MARTINEZ ARIAS, A. (1991). Secretion and movement of *wingless* protein in the epidermis of the *Drosophila* embryo. *Mech. Dev.* 35: 43-54.

GOODMAN, C.S. and DOE, C.Q. (1993). Embryonic development of the *Drosophila* central nervous system. In *The development of Drosophila melanogaster*, (Eds. Bate, M. and Martínez-Arias, A.). Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y., pp. 1131-1206.

HALTER, D.A., URBAN, J., RICKERT, C., NER, S.S., ITO, K., TRAVERS, A.A. and TECHNAU, G.M. (1995). The homeobox gene *repo* is required for the differentiation and maintenance of glia function in the embryonic nervous system of *Drosophila melanogaster. Development* 121: 317-332.

HIDALGO, A. and BOOTH, G.E. (2000). Glia dictate pioneer axon trajectories in the *Drosophila* embryonic CNS. *Development* 127: 393-402.

HIDALGO, A., URBAN, J. and BRAND, A.H. (1995). Targeted ablation of glia disrupts axon tract formation in the *Drosophila* CNS. *Development* 121: 3703-3712.

HOLDEN, H.M., RYPNIEWSKI, W.R., LAW, J.H. and RAYMENT, I. (1987). The molecular structure of insecticyanin from the tobacco hornworm *Manduca sexta* L. at 2.6 Å resolution. *EMBO J.* 6: 1565-1570.

HUBER, R., SCHNEIDER, M., EPP, O., MAYR, I., MESSERSCHMIDT, A., PFLUGRATH, J. and KAYSER, H. (1987). Crystallization, crystal structure analysis and preliminary molecular model of the bilin-binding protein from the insect *Pieris brassicae. J. Mol. Biol.* 198: 423-434.

ITO, K., URBAN, J. and TECHNAU, G.M. (1995). Distribution, classification, and development of *Drosophila* glial cells in the late embryonic and early larval ventral nerve cord. *Roux's Arch. Dev. Biol.* 204: 284-307.

JACOBS, J.R. and GOODMAN, C.S. (1989). Embryonic development of axon pathways in the *Drosophila* CNS. I. A glial scaffold appears before the first growth cones. *J. Neurosci.* 9: 2402-2411.

JACOBS, J.R., HIROMI, Y., PATEL, N.H. and GOODMAN, C.S. (1989). Lineage, migration, and morphogenesis of longitudinal glia in the *Drosophila* CNS as revealed by a molecular lineage marker. *Neuron* 2: 1625-1631.

KEEN, J.N., CACERES, I., ELIOPOULOS, E.E., ZAGALSKY, P.F. and FINDLAY, J.B.C. (1991a). Complete sequence and model for the A_2 subunit of the carotenoid pigment complex, crustacyanin. *Eur. J. Biochem.* 197: 407-417.

KEEN, J.N., CACERES, I., ELIOPOULOS, E.E., ZAGALSKY, P.F. and FINDLAY, J.B.C. (1991b). Complete sequence and model for the C_1 subunit of the carotenoprotein, crustacyanin, and model for the dimer, β-crustacyanin, formed from the C_1 and A_2 subunits with astaxanthin. *Eur. J. Biochem.* 202: 31-40.

KLAES, A., MENNE, T., STOLLEWERK, A., SCHOLZ, H. and KLÄMBT, C. (1994). The ets transcription factors encoded by the *Drosophila* gene *pointed* direct glial cell differentiation in the embryonic CNS. *Cell* 76: 149-160.

KLÄMBT, C., HUMMEL, T., MENNE, T., SADLOWSKI, E., SCHOLZ, H. and STOLLEWERK, A. (1996). Development and function of embryonic central nervous system glial cells in *Drosophila. Dev. Genet.* 18: 40-49.

KLÄMBT, C., SCHIMMELPFENG, K. and HUMMEL, T. (1999). Glia development in the embryonic CNS of *Drosophila. Adv. Exp. Med. Biol.* 468: 23-32.

LI, W.C. and RIDDIFORD, L. (1994). The two duplicated insecticyanin genes, *ins-a* and *ins-b* are differentially expressed in the tobacco hornworm, *Manduca sexta*. *Nucleic Acids Res.* 22: 2945-2950.

OISHI, I., SUGIYAMA, S., LIU, Z.-J., YAMAMURA, H., NISHIDA, Y. and MINAMI, Y. (1997). A novel *Drosophila* receptor tyrosine kinase expressed specifically in the nervous system. *J. Biol. Chem.* 272: 11916-11923.

OU, S.K., HWANG, J.M. and PATTERSON, P.H. (1993). A modified method for obtaining large amounts of high titer polyclonal ascites fluid. *J. Immunol. Methods* 165: 75-80.

PATEL, N.H., MARTÍN-BLANCO, E., COLEMAN, K.G., POOLE, S.J., ELLIS, M.C., KORNBERG, T.B. and GOODMAN, C.S. (1989). Expression of *engrailed* proteins in arthropods, annelids, and chordates. *Cell* 58: 955-968.

RENN, S.C.P., TOMKINSON, B. and TAGHERT, P.H. (1998). Characterization and cloning of tripeptidyl peptidase II from the fruit fly, *Drosophila melanogaster*. *J. Biol. Chem.* 273: 19173-19182.

SAITO, H. (1998). Purification and characterization of two insecticyanin-type proteins from the larval hemolymph of the Eri-silkworm, Samia cynthia ricini. *Biochim. Biophys. Acta* 1380: 141-150.

SÁNCHEZ, D., GANFORNINA, M.D. and BASTIANI, M.J. (1995). Developmental expression of the lipocalin *Lazarillo* and its role in axonal pathfinding in the grasshopper embryo. *Development* 121: 135-147.

SÁNCHEZ, D., GANFORNINA, M.D. and BASTIANI, M.J. (1996). Developmental expression and biochemical analysis of Conulin, a protein secreted from a subset of nerve growth cones. *J. Neurosci.* 16: 663-674.

SCHMIDT, F.S. and SKERRA, A. (1994). The bilin-binding protein of *Pieris brassicae*. cDNA sequence and regulation of expression reveal distinct features of this insect pigment protein. *Eur. J. Biochem.* 219: 855-863.

SCHMIDT, H., RICKERT, C., BOSSING, T., VEF, O., URBAN, J. and TECHNAU, G.M. (1997). The embryonic central nervous system lineages of *Drosophila melanogaster*. II Neuroblast lineages derived from the dorsal part of the neuroectoderm. *Dev. Biol.* 189: 186-204.

SNOW, P., PATEL, N., HARRELSON, A. and GOODMAN, C.S. (1987). Neural-specific carbohydrate moiety shared by many surface glycoproteins in *Drosophila* and grasshopper embryos. *J. Neurosci.* 7: 4137-4144.

TARTAKOFF, A. (1983). Perturbation of vesicular traffic with the carboxylic ionophore monensin. *Cell* 32: 1026-1028.

THOMAS, J.B., BASTIANI, M.J., BATE, M. and GOODMAN, C.S. (1984). From grasshopper to *Drosophila*: a common plan for neuronal development. *Nature* 310: 203-207.

THOMPSON, J.D., GIBSON, T.J., PLEWNIAK, F., JEANMOUGIN, F. and HIGGINS, D.G. (1997). The ClustalX windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tool. *Nucleic Acids Res.* 24: 4876-4882.

Received: March 2000
Accepted for publication: May 2000