

Prox 1, a *prospero*-related homeobox gene expressed during mouse development

Guillermo Oliver ^{a,*}, Beatriz Sosa-Pineda ^a, Sabine Geisendorf ^a, Eric P. Spana ^b,
Chris Q. Doe ^b and Peter Gruss ^a

^a Department of Molecular Cell Biology, Max-Planck Institute of Biophysical Chemistry, 37077 Göttingen, Germany, and

^b Department of Cell and Structural Biology, University of Illinois, Urbana, IL 61801, USA

(Received 28 June 1993; revision received and accepted 15 July 1993)

Prox 1, a likely mouse homologue of the *Drosophila* homeobox gene *prospero* has been cloned and its expression pattern analyzed during development. In *Drosophila*, *prospero* is expressed in the developing CNS, lens-secreting cone cells of the eye, and midgut. In the mouse, Prox 1 is expressed in many of the same tissues: young neurons of the subventricular region of the CNS, developing eye lens and pancreas. Expression is also detected in the developing liver and heart, as well as transiently in the skeletal muscles. The similarities in protein sequence and expression patterns between the mouse and fly cognate genes suggest that Prox 1 may play, among others, a fundamental role in early development of the murine CNS.

Prox 1; *Prospero*; Central nervous system; *Drosophila*; Eye; Homeobox gene; Neuroblast; Pancreas; Muscle

Introduction

The differentiation of the nervous system is a central process in the early development of most organisms, and the molecular mechanisms underlying the generation of neuronal diversity are mostly unknown. In vertebrates, initially the embryonic neural tube is composed of an undifferentiated neuroepithelium. According to the position along the antero-posterior (A-P) and dorso-ventral (D-V) axes, the cells will differentiate and produce distinct neuronal structures. After closure, the neural tube differentiates along the A-P axis into spinal cord and three brain vesicles: forebrain, midbrain and hindbrain. The mechanisms that control the specification of the different regions in the vertebrate body, including the nervous system, are probably established by a complex network of regulatory factors, similar to those that are involved in *Drosophila* development. For example, many vertebrate homologues of the *Drosophila* homeotic genes (Hox) are expressed in different domains along the A-P and D-V axes (for

reviews see McGinnis and Krumlauf, 1992; Kessel and Gruss, 1990; Gruss and Walther, 1992). However, very little is known about the mechanisms that control the production of the different neuronal and glial populations from the undifferentiated embryonic neuroepithelium. The accepted criterion is that cells will migrate from the mitotically active ventricular neuroepithelium towards the peripheric flanking population where they will become mostly non-dividing neurons and glia cells. These will differentiate and produce all the different neuronal populations of the grey matter and glia of the white matter in a process which in mammals ends before or soon after birth. The genes involved in this precise and complex mechanism that generates neuronal diversity are unknown. The identification in the vertebrate genome of genes similar to those that are known to regulate CNS development in *Drosophila* has been one of the approaches used towards answering this question.

The *Drosophila* CNS originates from a monolayer of ventral ectodermal cells called the neurogenic region (for review see Campos-Ortega and Hartenstein, 1985). Neurogenesis begins when individual cells within the ventral neurogenic region delaminate into the embryo

* Corresponding author. Tel: 551-201507. Fax: 551-201504.

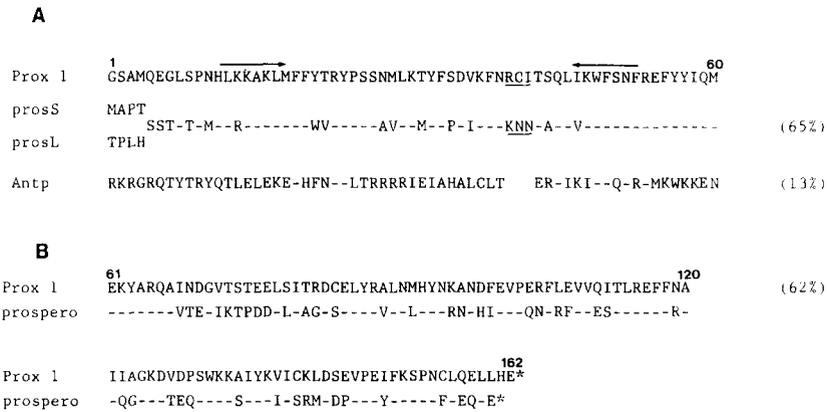


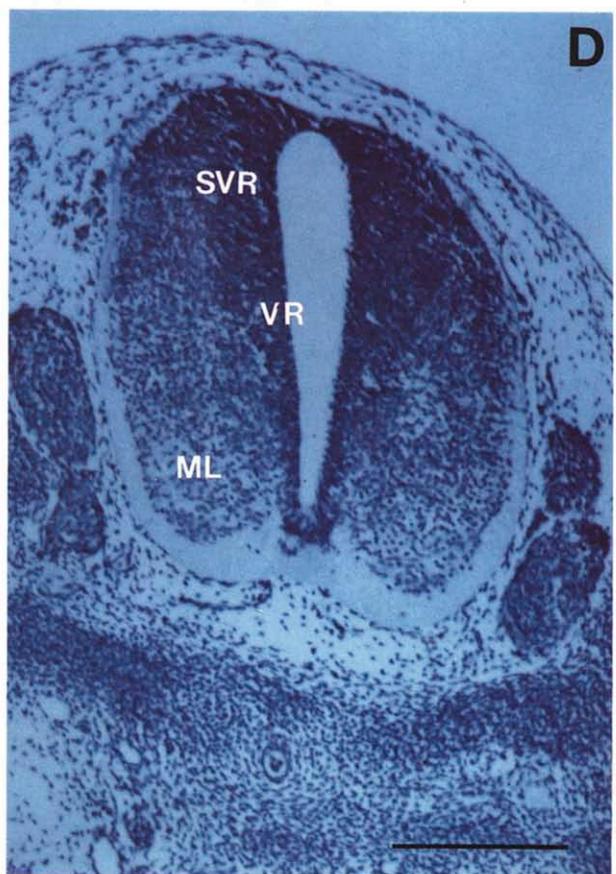
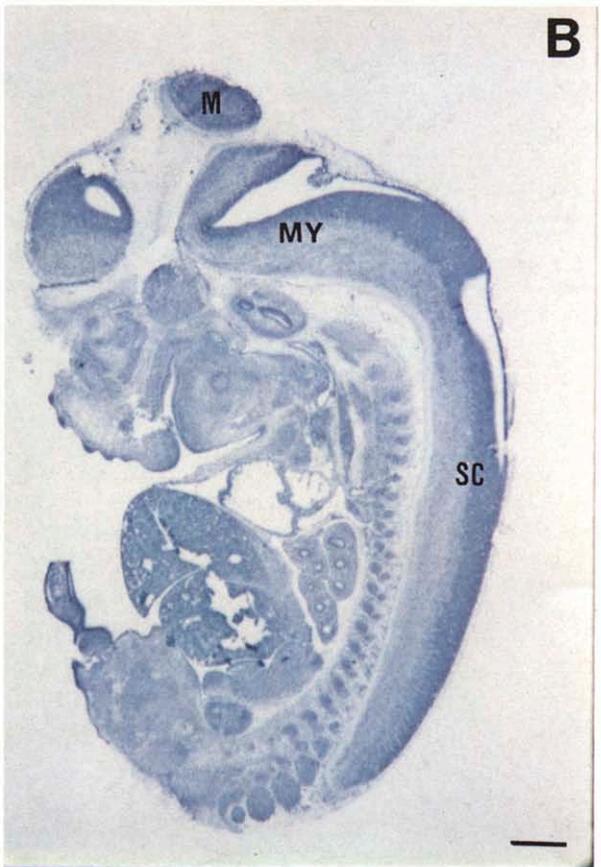
Fig. 1. PCR amplification strategy and amino acid sequence of the Prox 1 homeodomain and the conserved 3' coding flanking sequence. (A) The amino acid sequence of the Prox 1 homeodomain is shown on the top, compared with those of *prospero* (with its two different spliced variants) and *Antennapedia*. Dashes indicate similar residues. Underlined sequences on Prox 1 and *prospero* indicate those 3 extra amino acids found in the turn between helix-2 and -3. On the Antp sequence a space was left in that region. Arrows above Prox 1 sequence shows the regions against which the degenerate deoxyoligonucleotides used in the initial PCR amplification were designed. Such an amplified fragment was used as probe in the in situ hybridization analysis. (B) Amino acid sequence comparison on the conserved 3' flanking coding sequence between Prox 1 and *prospero*. Percentage amino acid homology is indicated on the right and numbers over the sequences indicate amino acids.

to form a subepidermal array of neuronal precursor cells called neuroblasts; the remaining superficial cells differentiate into ventral epidermis (Jimenez and Campos-Ortega, 1990). The mitotically active neuroblasts divide asymmetrically to produce a second order precursor, called the ganglion mother cell (GMC) with each GMC dividing to generate a characteristic pair of neurons (reviewed in Doe, 1992). Recently much has been learned about the early development of the *Drosophila* CNS. Proneural genes (e.g. the achaete-scute complex, AS-C) are expressed in subsets of neuroectodermal cells where they promote neuroblast formation; in contrast, the neurogenic genes (e.g. *Notch*) promote epidermal development within the neurogenic region (reviewed in Campos-Ortega and Jan, 1991). Following neuroblast formation, the *prospero* gene is transcribed in all neuroblasts (Doe et al., 1991), but only translated in GMCs and undifferentiated neurons (Vaessin et al., 1991; Matsuzaki et al., 1992; E.P.S and C.Q.D, unpublished results). Thus, *prospero* protein is not detected in dividing neuroblasts or mature differentiated neurons; it is highly expressed in GMCs and young neurons where it plays a fundamental role in regulating gene expression. Lack of *prospero* function results in altered expression of many genes (e.g. *even*



Fig. 2. Prox 1 early expression. Whole mount in situ hybridization was done on a E9.5 mouse embryo. Expression at this early stage is found in the anlage of the dorsal pancreas (arrow) and liver (arrowhead).

Fig. 3. Prox 1 is expressed in the dorsal part of the E12.5 mouse spinal cord. In situ hybridization analysis with Prox 1 antisense probe of E12.5 mouse embryos on dark-field (upper panels) and bright field (bottom panels) pictures. (A) Parasagittal section showing the restricted dorsal expression in the spinal cord. (C) A higher magnification of an oblique section at the level of the anterior spinal cord is shown. Expression is observed in the postmitotic subventricular neurons. SC, spinal cord; M, mesencephalon; ML, mantle layer; MY, myelencephalon; SVR, subventricular region; VR, ventricular region. Scale bar (A) 0.5 mm; (C) 0.25 mm.



skipped and *engrailed*) in GMCs, which leads to a profoundly defective CNS and embryonic lethality (Doe et al., 1991; Vaessin et al., 1991).

Homologues to the genes of the *Drosophila* AS-C have been cloned in rat and mouse (MASH 1 and MASH 2) (Johnson et al., 1990; Guillemot and Joyner, 1993). Similar to the fly, the expression pattern of one of these genes (MASH 1) is confined to populations of neural precursor cells and its expression disappears upon neuronal differentiation (Lo et al., 1991). In P19 cells differentiated in vitro, MASH 1 induction has been shown to precede the earliest detectable expression of neuronal differentiation markers (Johnson et al., 1992). This data supports the idea that the MASH genes are the functional homologues of the AS-C. Homologues of the *Drosophila Notch* gene have been also isolated from the mouse genome, and its expression has been found in the embryonic CNS restricted to the ventricular region (Franco Del Amo, 1992; Reaume et al., 1992).

Taking into account all these previous data, we decided to search for a murine homologue of the *prospero* gene. We expected that this gene may perform at least, during CNS development, some of the same functions as the *Drosophila* gene. In this paper, we report the cloning of a likely murine homologue of the *Drosophila prospero* gene (Prox 1) and we particularly focus on its expression in the CNS. Partial sequence analysis shows that the sequence homology includes the homeodomain and 3' flanking coding region. The analysis of its expression pattern reveals some interesting similarities with the one reported for *Drosophila prospero*. Expression in the developing CNS is found mainly restricted to the subventricular region containing undifferentiated neurons. In *Drosophila*, *prospero* is co-expressed in neuroblasts with AS-C genes, and it is required for the normal expression of *even skipped* and *engrailed* in GMCs. We find Prox 1 to be expressed in the developing CNS in a pattern overlapping that of MASH 1 as well as partially those reported for *evx 1* (Bastian and Gruss, 1990) and *en-2* (Davis et al., 1988). It is also expressed in the developing eye lens and pancreas. In the fly, *prospero* is not only expressed in the developing CNS, but also in the lens secreting cone cells of the developing eye and in cells associated with the midgut (E.P.S and C.Q.D, unpublished results). In addition to these similarities in the regions of expression of Prox 1 and *prospero*, we found Prox 1 to be also expressed in the heart, liver

and skeletal muscle. Therefore, Prox 1 is a homeobox-containing gene in which the conservation with its putative *Drosophila* cognate includes partial protein sequence and expression pattern. These data suggest that similarly to *prospero*, Prox 1 may play a fundamental role in the early development of the CNS.

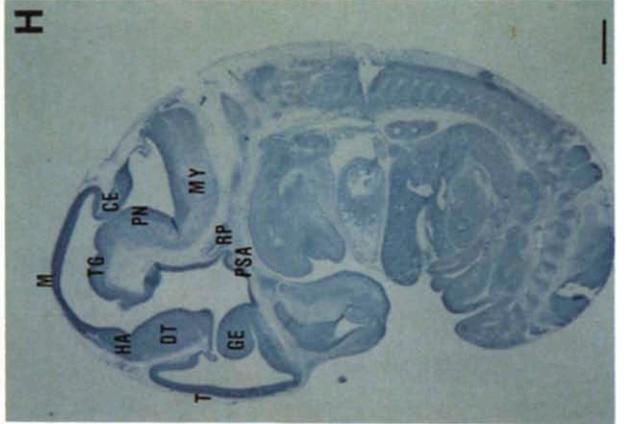
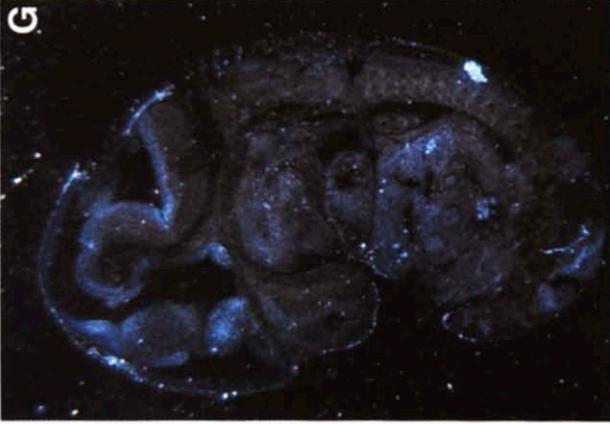
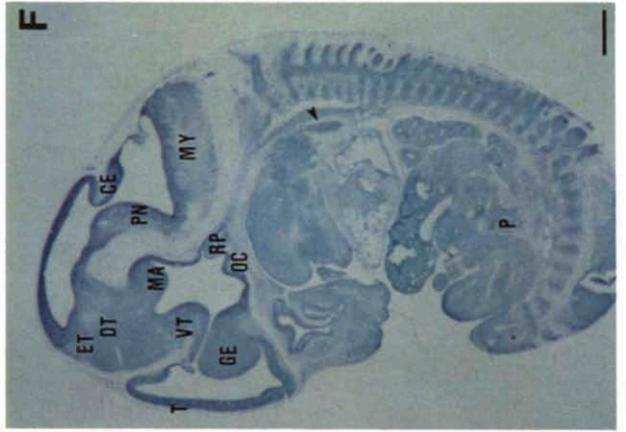
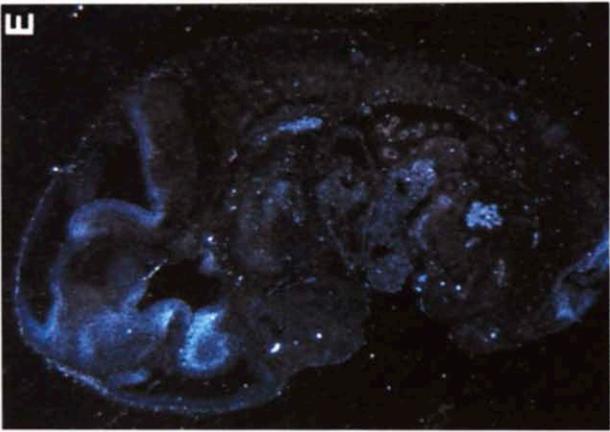
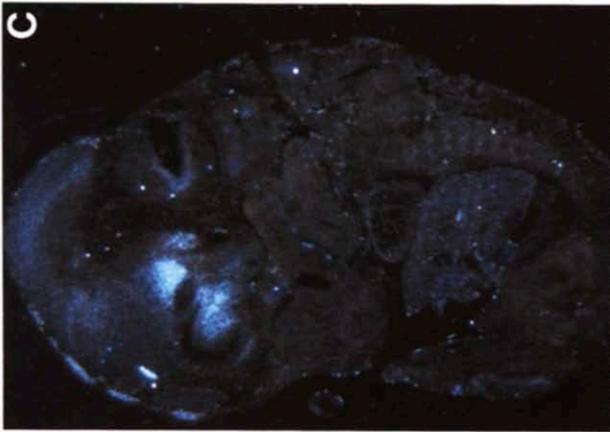
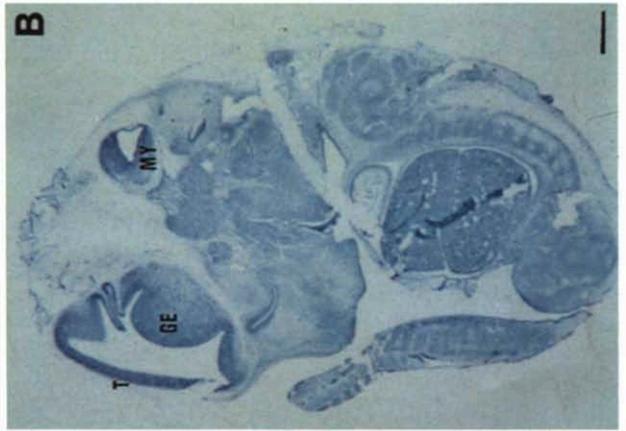
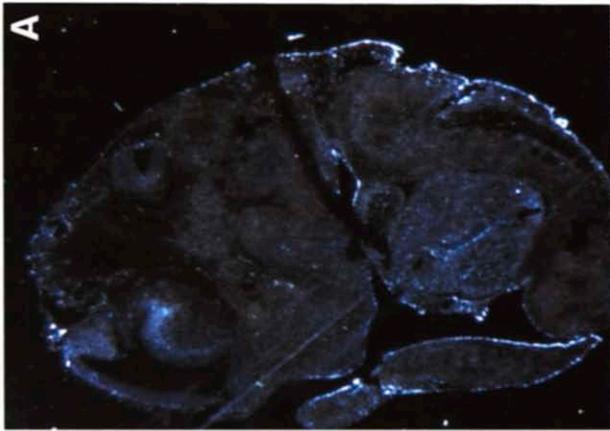
Results

Cloning of murine Prox 1 gene

Using a polymerase chain reaction (PCR) protocol, a murine homologue of the *Drosophila prospero* homeobox was isolated. Two degenerate oligonucleotides were designed (see Materials and Methods) encoding two stretches of amino acids of the helix-2 and -3 of the *prospero* homeodomain (Chu-Lagraff et al., 1991). These oligonucleotides were used as primers in reactions with cDNA produced from 10.5-day old (E10.5) mouse embryos. A single amplified fragment with homology to the *prospero*-homeobox gene was obtained. The PCR fragment was used to screen an amplified E12.5 mouse cDNA library under high-stringency conditions. Only one single cDNA clone of 2.2 kb in length was isolated (Prox 1). We have not succeeded in detecting the size of the transcript on Northern blot analysis.

The partial Prox 1 DNA sequence obtained shows some interesting features (Fig. 1). As in *Drosophila prospero* (Chu-Lagraff et al., 1991; Matsuzaki et al., 1992), when Prox 1 homeodomain is compared with the one from *Antennapedia*, it is clearly seen that it is a highly diverged one. Sequence comparison of the homeodomain shows 65% homology at the amino acid level with *prospero*, whereas with any other homeodomain it only shares approximately 13% homology (Fig. 1). It contains many of the highly conserved amino acids found in most homeodomains, including the typical helix-3 region which is necessary for sequence-specific DNA recognition (for a review see Scott et al., 1989). The DNA binding amino acid at position 50, considered to be important for the DNA binding specificity is a serine, similar to the paired-type homeodomains. Also, as in the fly *prospero* homeodomain, Prox 1 has six amino acids in the turn between helix-2 and -3, unlike most homeodomains which contain three amino acids. Although most amino acids in the homeodomain of Prox 1 and *prospero* are conserved, the first seven of the N-terminal region are

Fig. 4. Prox 1 transcripts in mouse brain. RNA in situ hybridization on sagittal sections of an E12.5 embryo. From Panel A to G, sections are from parasagittal to midsagittal. Dark-field is on the top and bright field below. Labelling is found in the myelencephalon, pons, mesencephalon and forebrain and is mostly in the SVR. Expression in the developing pancreas is also seen in (E). Some non-specific labelling is found over the dorsal aorta (arrow in (F)). CE, cerebellum; DT, dorsal thalamus; ET, epithalamus; GE, ganglionic eminence; HA, habenula; M, mesencephalon; MA, mammillary area; MY, myelencephalon; OC, region of the optic chiasm; P, pancreas; PN, pons; PSA, postoptic area; RP, Rathke's pouch; T, telencephalon; TG, tectum; VT, ventral thalamus. Scale bar, 0.5 mm.



totally different. One of the characteristics of the *prospero* gene is that it produces two alternative-spliced transcripts giving rise to different N-terminal amino acids of the homeodomain (Chu-Lagraff et al., 1991). It is important to note that contrary to most other homeobox-containing genes, Prox 1 and *prospero* amino acid sequences are conserved not only in the homeobox but also along the 306 bp of the 3' flanking coding region, with 62% homology at the amino acid level (Fig. 1). These sequence similarities argue in favor of Prox 1 being the real cognate of the *Drosophila prospero* gene. Partial DNA sequence of the homeodomain 5' flanking region of the isolated cDNA clone have not shown any significant homology. Genomic Southern hybridization of total genomic mouse DNA suggests that there are no additional Prox related genes (data not shown).

Expression of Prox 1 during development

Various approaches have been used during this work in order to follow Prox 1 expression. It was evident that despite the wide distribution of Prox 1 transcripts in the embryo, gene expression is very low, such that even the *in situ* hybridization analysis performed here with the more sensitive α -³³P requires more than 20 days of exposure to detect a clear signal. A developmental PCR analysis was initially used in order to determine the temporal expression pattern of Prox 1. We found that expression commences at 8.5 day post-coitus embryos (E8.5), and becomes almost undetectable at E18.0 (data not shown). The major regions of Prox 1 expression are described below.

Prox 1 expression pattern in the developing CNS

A spatio-temporal analysis of Prox 1 expression pattern was initially performed using whole-mount RNA *in situ* analysis on mouse embryos between E7.5 and E9.5. We did not observe any expression at E7.5, and it was extremely weak at E8.5. Prox 1 expression was initially observed at E9.5. At this stage no expression can be detected in the developing CNS and labelling is clearly seen in a region that may correspond to the developing dorsal pancreatic anlage and liver anlage (Fig. 2). By *in situ* hybridization analysis using paraffin sections, expression in the CNS was first seen around E10–E10.5 with weak labelling detected over the spinal cord (not shown).

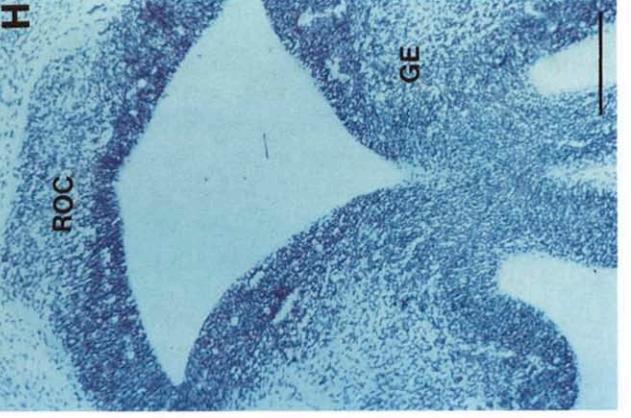
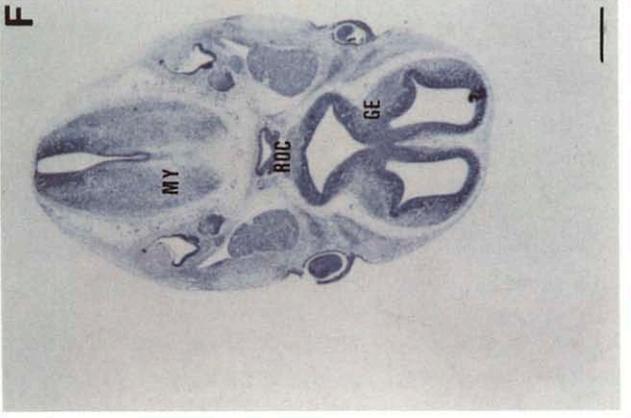
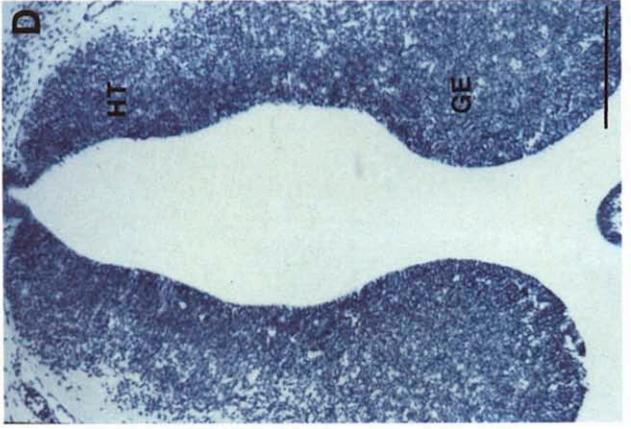
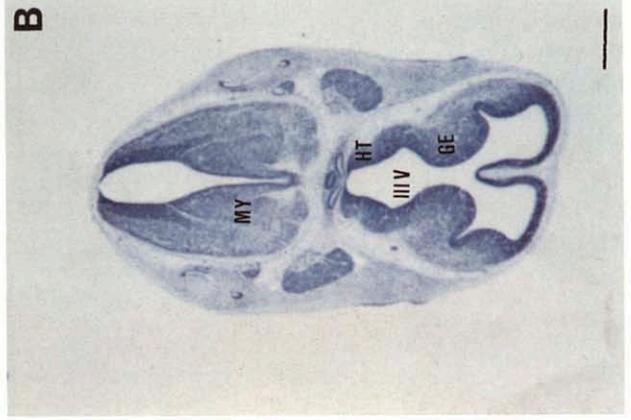
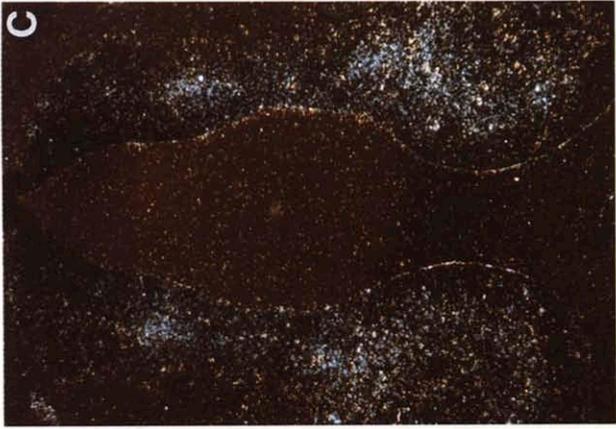
After this stage, expression starts to increase, and by E12.5 Prox 1 transcripts are found all along the an-

tero-posterior axis of the developing neural tube with a restriction to the dorsal alar plate of the spinal cord (Fig. 3A). As documented by transverse sections, the transcripts are located in the subventricular region (SVR) of the tube, and maybe a few labelled cells are also present in the more compacted cells of the ventricular region (VR) (Fig. 3C). At this stage, the spinal cord has a mitotically active VR containing the neuroepithelial precursors. Once mitosis is completed, the newly formed young neuronal precursors migrate away from the VR and differentiate into the different neuronal populations of the mantle layer (ML, Fig. 3D). In addition, a ventro-dorsal neuronal differentiation wave is known to occur in the neural tube at this stage. Prox 1 expression follows a graded dorso-ventral expression into the differentiating SVR of the spinal cord, being mostly absent from the most ventral and marginal areas containing differentiated neurons.

In E12.5, additional sites of expression of Prox 1 are observed in the rhombencephalon, midbrain and forebrain and they are also mainly restricted to the SVR. On the most parasagittal sections of Fig. 4, positive cells are found in the developing telencephalon over the basal region (GE) and hippocampus, in the myelencephalon (MY), and weakly in the roof of the mesencephalon (M) (Fig. 4A, C). In Fig. 4C, expression is observed in the GE of the telencephalon and in the diencephalon. In this last structure, following Herrick's nomenclature, label is seen mostly in the ventral thalamus (VT) and also weakly in the mammillary area (MA). On the more midsagittal sections (Fig. 4E, G), the expression delineates the entire shape of the hindbrain (myelencephalon and pons), mesencephalon and tegmentum. No expression is found in the cerebellum. In the forebrain, expression is faint in the epithalamic region (ET), moderate in the habenula (HA) and dorsal thalamus (DT), and strong in the ventral thalamus (VT) (Fig. 4C, E, G). In the hypothalamic area, Prox 1 is also found in the MA, postoptic area (PSA) and region of the optic chiasm (OC) (Fig. 4E, G).

Transverse sections at higher magnifications are included in order to better illustrate the main expression patterns in the brain, and especially to clearly distinguish the negative ventricular region from the strongly labeled subventricular region (Fig. 5). Expression in the SVR of myelencephalon (MY), hypothalamus (HT) and base of telencephalon (GE) is shown in Fig. 5A. This is better seen using the higher magnification shown in panel C. In Fig. 5E and G, similar expression is seen in the SVR of myelencephalon and base of telen-

Fig. 5. Distribution of Prox 1 transcripts over the subventricular region of the brain. Transverse sections across a E12.5 mouse brain. Dark field is on the top row and bright below (A) and its higher magnification (C) show the expression over the myelencephalon, hypothalamic region (HT) of diencephalon and base of the telencephalon (GE) restricted to the SVR. (E) and its higher magnification (G) show Prox 1 expression over the SVR of the region of the optic chiasm and ganglionic eminence (GE). GE, ganglionic eminence; HT, hypothalamic region; MY, myelencephalon; ROC, region of the optic chiasm; III V, third ventricle. Scale bar, (A,E) 0.5 mm; (C,G) 0.25 mm.



cephalon, and also in the region of the optic chiasm (ROC).

The time of origin of the ventral thalamic neurons is from E10 to E13, with a differentiation of the germinal and mantle layers around E11 (Angevine, 1970). This agrees with the expression we find for *Prox 1* in the diencephalon at E12.5, which is mainly located in the ventral area and may well correspond to already postmitotic neurons still in the process of differentiation. In the diencephalon, expression spreads beyond the SVR. Angevine (1970), has previously shown that in the thalamus, the sites of neuron origin also include internal areas of the thalamic wall. The more diffuse labelling we see in this area could also be related to this other populations of young postmitotic neurons.

As development progresses, *Prox 1* expression in the CNS slowly decreases, reaching very low levels at E18.5. When compared with E12.5, expression at E14.5 in the spinal cord is clearly weaker, concomitant with the end of the neuronal differentiation stage (Fig. 6). At this stage, expression in the brain is also decreasing in intensity and up to birth the main remaining labelling is seen outside of the VR in the base of the telencephalon, hippocampus (HI) and thalamus (Fig. 6). In this last region, the expression profile previously observed at E12.5 has now changed. Meanwhile, as already shown in Fig. 4C, the main labelling at E12.5 is found over the ventral thalamic region (VT), at E14.5 the strongest expression is detected over the dorsal thalamus (DT) and much weaker over the ventral part (Fig. 6). On this figure, *Prox 1* expression can be also observed in the heart (HE), liver (LI) and in the intercostal (arrow) and tongue (T) muscles.

The strong expression of *Prox 1* in the DT at these later stages could be related to the fact that the dorsal thalamus differentiates later than the rest of the thalamic areas, and around E14.5–E16.0 correspond to the timing of maximum development of the habenular nuclei (Angevine, 1970). Therefore, *Prox 1* expression in the CNS seems to be mostly restricted to those young neuronal precursors of the subventricular region, which in the multistep process leading of neuronal differentiation have already stopped dividing in the adjacent ventricular area. Its expression pattern seems to coincide with the time of neuronal differentiation itself.

Other major regions of expression of *Prox 1* outside the CNS can be seen during mouse development.

Prox 1 in developing muscles

In vertebrates, both cardiac and skeletal striated muscles are mesoderm derivatives, and the heart muscle is the first to be formed followed by somitogenesis and skeletal myogenesis. One of the last steps during myogenesis is muscle differentiation in which cell division stops and cells fuse. Finally the skeletal muscle will mature giving rise to the different muscle fibers (Ott and Buckingham, 1992).

At approximately E10.5 we start to detect *Prox 1* labelling over the developing heart (not shown). This expression remains and becomes stronger at later stages (see Figs. 4, 6 and 7). Expression is no longer seen after E16.5. At E14.5–E15.5, *Prox 1* transcripts are also detected in most skeletal muscles of the embryo. This includes muscles of the tongue, neck, trunk, fore and hindlimbs, and the intercostal and oculomotor muscles as well (Figs. 6 and 7). In contrast to many of the muscle specific genes whose expression is already detected at the somite stage and remains during later developmental stages (Ott and Buckingham, 1992), *Prox 1* expression is transient and not observed in any type of muscle precursor cell of the somites. During the stages in which *Prox 1* is expressed among the developing muscle masses, the mature pattern of muscle groups is formed, secondary muscle fibers appear and neuromuscular junctions are established (Ontell and Kozeka, 1984). Around E16.0 no more *Prox 1* muscle expression is detected.

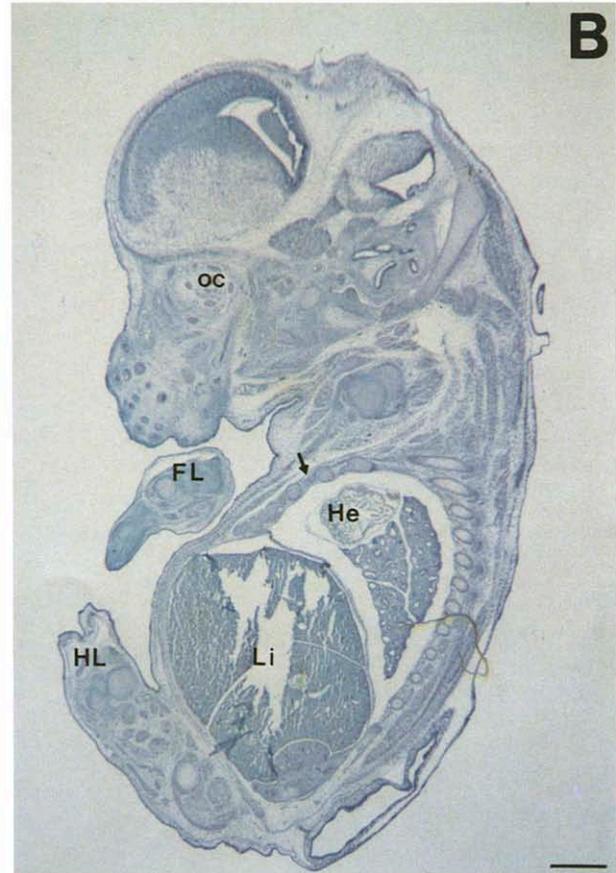
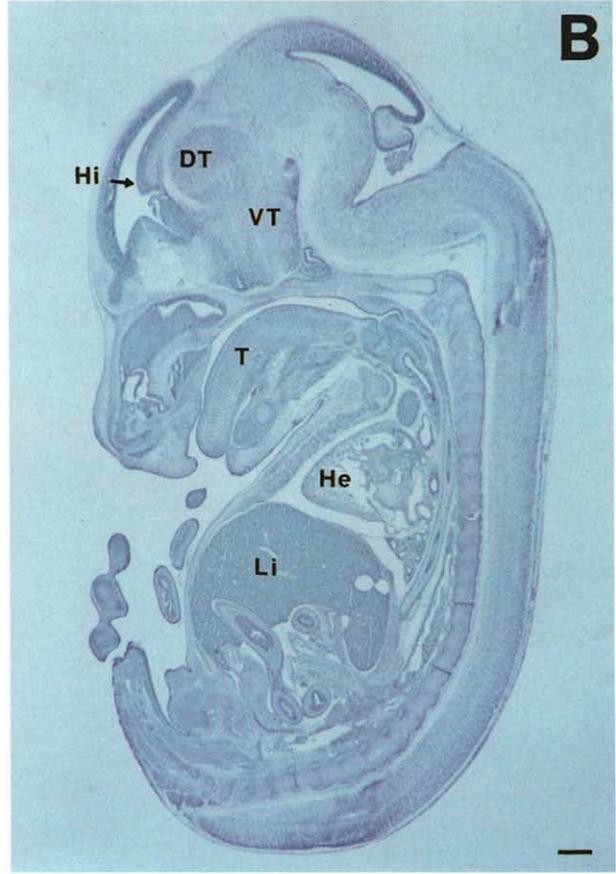
Prox 1 expression during eye development

During eye development cells in the inner portion of the lens vesicle elongate and produce the lens fibers under the influence of the neural retina. The anterior cells of the lens vesicle are mitotically active cells of the germinal epithelium. These dividing cells move toward the equator of the vesicle and once they pass through the equatorial region they start to elongate. The lens has three regions: the anterior zone of dividing epithelial cells, an equatorial zone of cellular elongation and a posterior one of crystallin-containing fiber cells.

At E10–10.5 expression in the region of the optic vesicles is first weakly detected (not shown). At E12.5 strong expression is clearly seen in the eye lens (Fig.

Fig. 6. *Prox 1* expression in later stages. Sagittal section of an E14.5 mouse embryo. On the dark-field of (A), it can be seen that *Prox 1* in the spinal cord is decreasing. The same is true for most regions of the brain with the exception of the hippocampus (Hi) and thalamus. On this last region the strongest labelling is now detected over the dorsal part (DT). *Prox 1* is also found on the liver (Li), heart (He), and intercostal (arrow) and tongue (T) muscles. (B) Bright-field of the same section. Scale bar, 0.5 mm.

Fig. 7. *Prox 1* is transiently expressed in developing skeletal muscles. (A) *Prox 1* transcripts are seen in the dark-field of an E14.5 mouse sagittal section. The labelling is present in most of the skeletal muscle fibers including muscles of the head (oculomotor, oc), neck, intercostals (arrow) and fore (FL) and hindlimb (HL). Expression is also in the heart (He) and liver (Li). Scale bar, 0.5 mm. (B) Bright-field of the same section.



8A; L), and it remains until around E16.5. Signal disappears at E18 (not shown). No labelled cells are observed in the neuronal retina (R) or in the dividing cells of the lens epithelium; only the lens fibers are positive.

According to this data, Prox 1 expression is seen mostly in the region in which the lens fibers elongate. As in the CNS, *prospero* is expressed after precursor divisions have ceased, but prior to cellular differentiation. Once again a surprising evolutionary convergence between Prox 1 and *prospero* is seen. The fly *prospero* gene is also exclusively expressed early in cone cell development in the eye imaginal disk; the cone cells secrete the lens of the *Drosophila* eye (Fig. 10B). Photoreceptor neurons do not express *prospero*.

Prox 1 expression in the developing pancreas and liver

The pancreas develops caudal to the stomach, from two separate areas of the duodenal epithelium which later become fused. One primordium grows dorsally, directly from the duodenal endoderm and the other grows ventrally from the endoderm of the hepatic diverticulum. After increasing in size the two primordia fuse (Wessells and Cohen, 1967). Both of these primordia arise from the endoderm and its differentiation is the result of a mesenchymal/endodermal induction process. It is known that the first pancreas-specific proteins are made only at the time when the initial bulge begins to form, and none of them can be detected before (Rutter et al., 1978). As early as E9.5 Prox 1 expression in an area that could correspond to the dorsal pancreas primordium is detected (Fig. 2). Its expression increases at E10.5 in the enlarged endodermal portion of the dorsal pancreas (not shown). At E12.5 the pancreas-specific expression of Prox 1 is strong as shown in Figs. 4C and 9A. Expression is seen in the excretory ducts (PD). Labelling is still observed at E14.5 and at later stages expression disappears, concomitant with the completion of pancreatic differentiation and branching. Interestingly, as can be observed in Fig. 10A, in *Drosophila* *prospero* is expressed

not only in the CNS, but also in some small cells associated with the midgut endoderm. Whether this region has a secretory function similar to the vertebrate pancreas is not known.

According to these data, Prox 1 is a marker for the developing pancreas, from the appearance of its anlage to its completed morphogenesis.

At E9.5, the liver anlage, which as the pancreas is also originated by a mesenchyme induction on the endoderm, is positive for Prox 1 (Fig. 2). At this stage the liver diverticulum is already formed and cells are rapidly proliferating. Expression continuous during development until around E15.0 (see Figs. 4 and 6) and it becomes almost extinguished afterwards.

Discussion

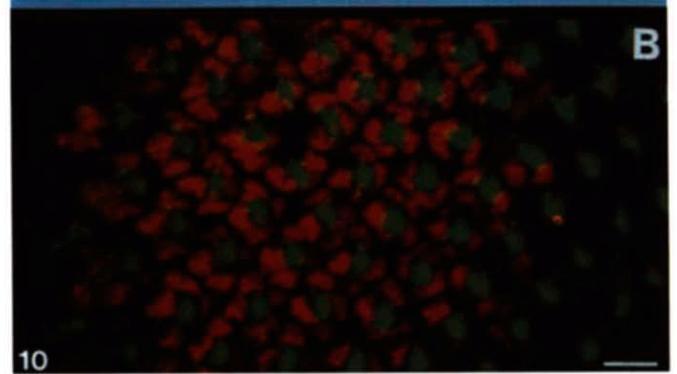
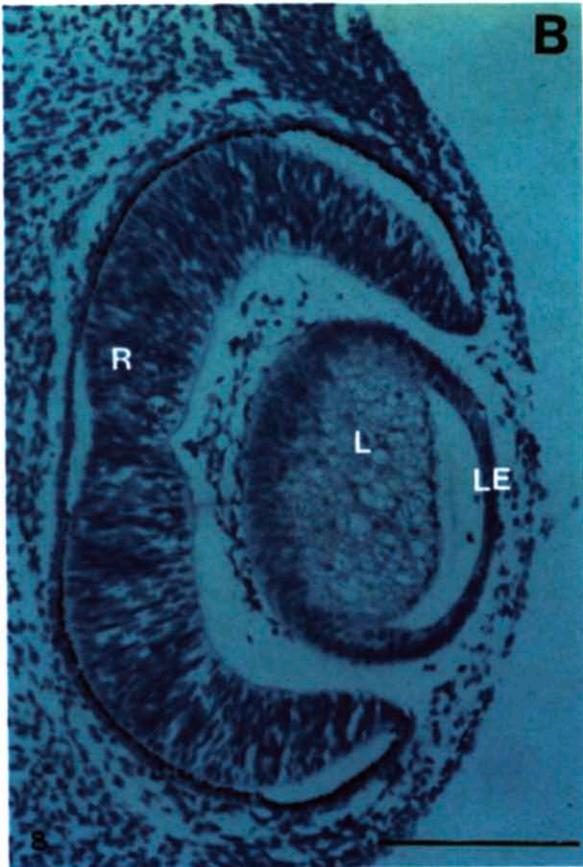
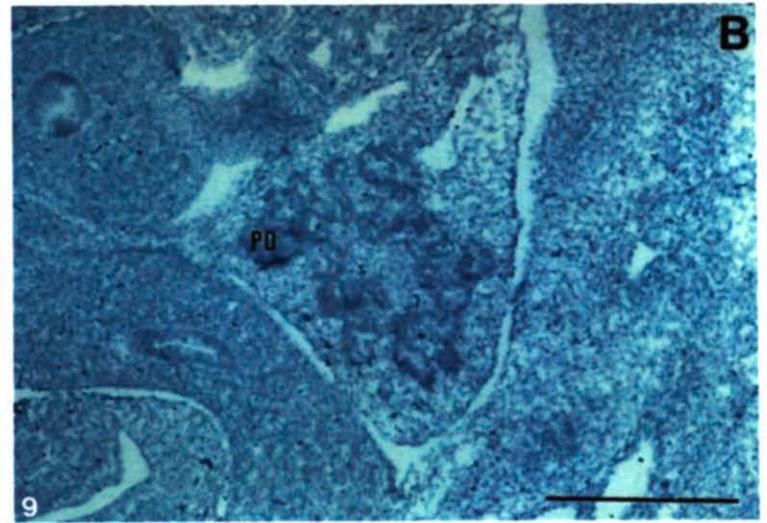
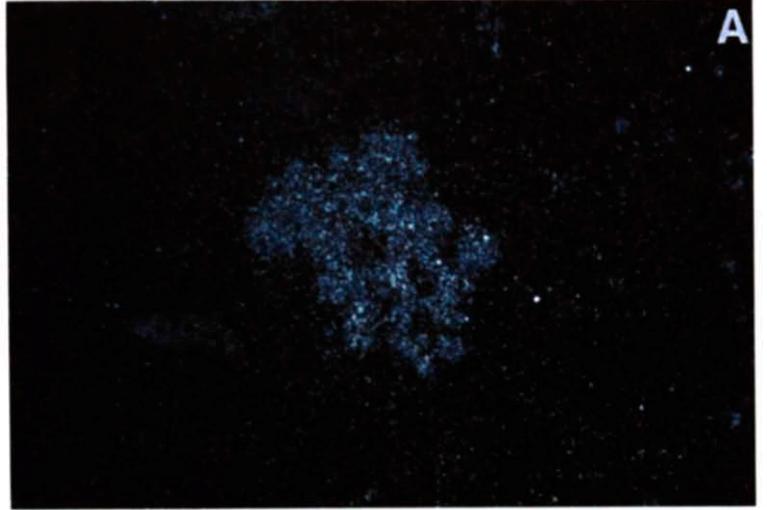
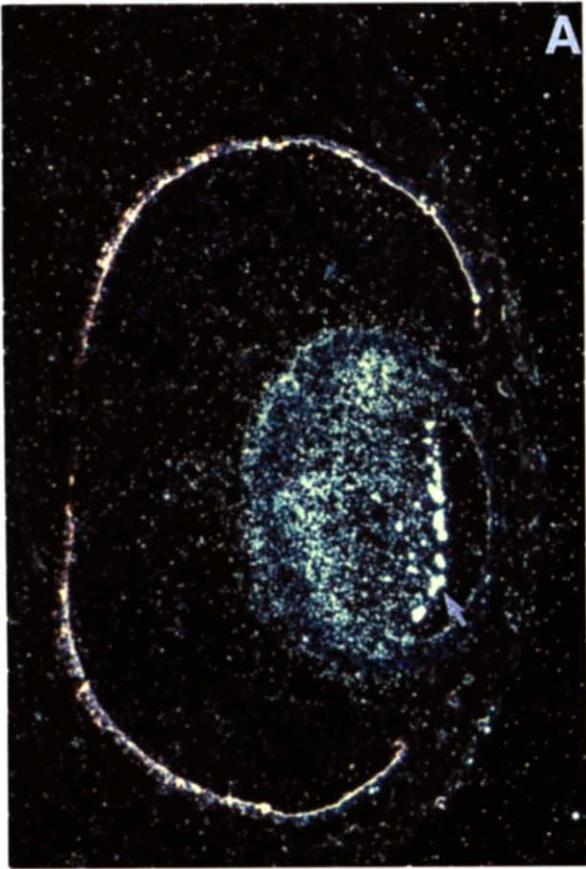
Prox 1 is likely the murine cognate of prospero

When one compares the sequence similarities between homeobox-containing genes from different organisms, particularly between mouse and *Drosophila*, the homologies are generally restricted to those regions known to be involved in sequence-specific DNA binding (Treisman et al., 1989; Kissinger et al., 1990). Those conserved regions generally correspond to the homeodomain, and in some cases also to an octapeptide found in the amino terminal region of the protein and a pentapeptide found before the homeobox (Wright et al., 1989). In some of the Pax genes (Gruss and Walther, 1992), this strong similarity extends to another binding domain, the paired-box. Another conserved region found upstream the homeodomain is named the POU-specific domain, with an extension of 75 amino acids shared by another group of transcription factors (Wright et al., 1989). With few exceptions, there is no clear evidence that vertebrate homeobox-containing genes can substitute for putative cognates of a different organism. These exceptions are the human Hox D4 (Hox-4.2) which is able to partially replace *Drosophila* *Deformed* gene (McGinnis et al., 1990) and

Fig. 8. Prox 1 is expressed in the non-neuronal cells of the eye. (A) dark-field, (B) Bright-field. In situ hybridization on a transverse section of an E12.5 mouse embryo showing the expression in the non-dividing region of the lens but not in the mitotically active lens epithelium. The strong white spots over the lens region (arrow) and surrounding the retina are non-specific labelling. L, lens; LE, lens epithelium; R, retina. Scale bar, 0.125 mm.

Fig. 9. Prox 1 is a marker for the developing pancreas. High magnification picture of an in situ hybridization analysis on the pancreas of a E12.5 mouse embryo. (A) Dark-field. (B) Bright field. Expression is found over the excretory ducts of the pancreas. PD, pancreatic ducts. Scale bar, 0.25 mm.

Fig. 10. In *Drosophila*, the *prospero* protein is expressed in the developing CNS, midgut, and eye. (A) *pros* protein is in the GMCs and undifferentiated neurons of the CNS (arrow), and in small cells associated with the internal surface of the midgut endoderm (arrowhead) in a 10 h embryo. (B) Confocal image of a third instar eye emaginal disk showing *pros* protein in the nuclei of lens-secreting cone cells (red), but not in the photoreceptor neurons (green). Anterior is to the left in A and to the right in B. Scale bar, (A) 32 μ m; (B) 6.7 μ m.



the mouse Hox B6 (Hox 2.2) and Hox A5 (Hox-1.3) which can also partially substitute for the *Drosophila Antennapedia* and *Sex combs reduced* (Malicki et al., 1992; Zhao et al., 1993). In these cases, the similarities are also mainly restricted to the domains mentioned before, so it could be that as the main binding specificity is unaltered, the whole regulatory circuit is maintained.

Prox 1 contains a highly diverged homeodomain whose amino acid sequences, similar to the *Drosophila prospero* one (Chu-Lagraff et al., 1991), are only 13% homologous to that of *Antennapedia*. We have found that in this case the amino acid homologies between Prox 1 and *prospero* include not only the homeodomain but also at least the translated 306 bp of the 3' flanking region, with an overall amino acid homology of 62%. This sequence similarity is the first argument supporting the notion that Prox 1 could be the mouse homologue of the *Drosophila prospero* gene. The reasons for an evolutionary conservation of sequences outside the homeodomain itself is unknown; however, one possibility could be that in this case the whole protein structure and not just the homeodomain is necessary to be able to fulfil its function. Perhaps Prox 1 protein needs to interact with other proteins in order to bind to specific DNA sequences. Maybe, as suggested by Hayashi and Scott (1990), the regions flanking the homeodomain could be involved in interactions with the transcriptional machinery or unknown cofactors. Moreover, Prox 1 protein could contain different functional domains aside from the homeodomain itself, and the conserved 3' flanking region may also be involved in DNA binding or transactivations. The fact that in *Drosophila* two alternative *prospero* cDNAs (Chu-Lagraff et al., 1991) producing different homeodomain amino termini have been found, and that Prox 1 has yet a third variant homeodomain, suggests that either this N-terminal homeodomain region is functionally unimportant or that these variations could be involved in cell, tissue or species-specific functions. It is also possible that in future, other variations of the homeodomain amino termini could be found. Until experiments addressing these questions are performed none of these possibilities can be excluded.

Prox 1 CNS expression pattern shares similarities with the one of Drosophila prospero

A detailed analysis of Prox 1 expression pattern, particularly at E12.5 has been done by in situ hybridization. We were originally interested in the possibility that this gene may play during mouse CNS development some roles similar to its *Drosophila* counterpart. One of the interesting aspects we found about its expression pattern is that indeed it is detected in

structures of the CNS (spinal cord and brain), but mainly restricted to the subventricular region. At this stage, in the spinal cord the SVR is formed by postmitotic young neurons. Similarly, expression in the brain is also mostly restricted to the SVR, although its distribution is more widespread, particularly in the diencephalon and ganglionic eminence. Expression of Prox 1 is almost extinguished in most of the previously expressing tissues by stage E18.5, concurrent with the termination of neuronal differentiation. According to these expression data, it seems that Prox 1 expression in the CNS is mostly restricted to postmitotic cells; maybe immature neurons. In the *Drosophila* CNS, *prospero* protein is detected only in second order precursors (GMCs) and young neurons, and is not observed in differentiated neurons (E.P.S and C.Q.D, unpublished results). Based on the similar CNS expression pattern here reported for Prox 1, at least a partial functional similarity between the two genes is possible. Our data, are in agreement with the notion that one of the Prox 1 functions could be to regulate gene expression and development of postmitotic undifferentiated young neurons, similar to the role of *prospero* in controlling GMC development in *Drosophila*. Further studies using different cell markers will be necessary in order to substantiate this assumption.

Prox 1 CNS expression correlates with that of Mash 1

As mentioned in the Introduction, the *Drosophila* AS-C genes are expressed in dividing neuroblasts, whereas *prospero* expression overlaps and extends to the GMC progeny and young neurons. Remarkably, a similar expression is observed for the most likely murine homologues of these genes Mash 1 and Prox 1 (Lo et al., 1991; Guillemot and Joyner, 1993). Mash 1 transcripts are found over most of the developing CNS (spinal cord and brain), but up to E13.5 expression is mainly restricted to the mitotic active neuronal precursors of the ventricular region. Prox 1 is found in the same regions of the CNS, but in the adjacent subventricular area containing mostly postmitotic young neurons. Using various markers for differentiated neurons, Lo et al. (1991) have clearly correlated the areas of expression of Mash 1 with subpopulations of early neuroepithelial precursor cells, showing that the induction of this gene precedes neuronal differentiation. This statement can also be applied for Prox 1, based on the fact that as shown here its expression in the CNS is found in similar regions as Mash 1. These spatial expression data suggest, but do not prove, that the early steps in neuronal differentiation could be conserved from *Drosophila* to mouse.

Such evolutionary convergence between expression of AS-C/*prospero* in *Drosophila* and Mash 1/Prox 1 in mouse is remarkable, it is equally noteworthy that

additional vertebrate cognates of *Drosophila* genes known to participate in this network leading to neural development also have overlapping expression patterns. This is the case for Notch 1, which as pointed out by Guillemot and Joyner (1993), shows extended overlapping expression patterns with Mash 1, very similar to Notch/AS-C in *Drosophila*. Perhaps more interestingly, *prospero* is co-expressed with, and regulates the genes *eve* and *en* in the *Drosophila* CNS. Murine homologues of *eve* and *en* have been cloned, and the reported expression patterns of these genes, *en-2* and *evx 1* (Davis et al., 1988; Bastian and Gruss, 1990), show once again partial overlapping expression patterns in brain and spinal cord with Prox 1. We can speculate that the well characterized gene regulatory network controlling the generation of the *Drosophila* CNS may be substantially conserved in vertebrates. The spatio-temporal parallels are striking: in *Drosophila* Notch and AS-C function first to control the formation of neuroblasts. *Prospero* controls the fate of the neuroblast progeny by regulating expression of a number of genes including *eve* and *en*. Similarly, in murine neurogenesis Mash 1 and Notch are expressed in the mitotic neurons of the neuroepithelia, Prox 1 is found in the young postmitotic neurons, while *evx 1* and *en-2* show partial overlap with Prox 1. This gene hierarchy, together with a Hox code, could produce some of the different neuronal identities found in the vertebrate body. The availability of these different molecular markers will allow us to test some of these hypotheses in the near future. Gain and loss of function mutations of these genes, and the subsequent analysis of the effect on each putative member of the cascade, may help to understand how in higher organisms such a complex mechanism as the production of neuronal diversity is accomplished. Two mutants will be particularly interesting to test according to the genetic hierarchy known in *Drosophila*. Firstly, the Mash 1 knock out may result in loss or reduction in Prox 1 expression. Secondly, Prox 1 loss of function may dramatically affect neuronal gene expression and differentiation.

Materials and Methods

PCR cloning of the Prox 1 homeodomain

Preparation of RNA from mouse embryos was performed using the LiCl procedure (Auffray and Rougeon, 1980). Single strand cDNA template was prepared as recommended by the suppliers (Pharmacia), and using an oligo(dT)-*NotI* primer-adaptor to prime the first strand synthesis. PCR amplification with degenerate primers (Wilkie and Simon, 1991) was performed in a volume of 50 μ l. Buffer containing 2

mM MgCl₂ and enzyme conditions were used as recommended by the suppliers (Promega). Primers used for the amplification were designed following the amino acids sequence of the *Drosophila prospero* homeobox and they were: (1) CAC/TCTC/GC/AGC/GAAA/GGCC/TAAA/GCTC/GATGT (5' primer for helix 1), and (2) AAG/ATTG/AG/CA/TG/AAACCAT/CTTG/CAC (3' primer for helix 3). 1.5 μ g of RNA was used in the amplification, together with 5 μ M of each primer. The conditions of the PCR reactions were: ten cycles (94°C 1 min; 45–72°C, 2 min) followed by 30 cycles (94°C, 1 min; 50°C, 1.5 min; 72°C, 2 min), and a final incubation (50°C, 1.5 min and 72°C, 10 min). Amplified products were electrophoresed on a 2% low-melting agarose gel. The fragment with the expected size was excised and the agarose plug was melted at 65°C for 5 min. 3 μ l of the DNA-containing molten agarose were reamplified with 30 cycles (94°C, 1 min; 50°C, 1.5 min; 72°C, 2 min). Amplified DNA was purified from a 2% agarose gel; termini were blunt-ended using Klenow enzyme (Promega). DNA was ligated into *SmaI*-cut pBluescript KSII + (Stratagene), transformed into competent *E. coli* DH5 α (Stratagene). Clones with the expected size insert were sequenced using T7 DNA polymerase sequencing kit (Pharmacia).

Isolation of Prox 1 cDNA

An amplified 12.5 days post-coitum murine C3H murine λ gt10 cDNA library (kindly provided by Dr. Gail Martin, University of California at San Francisco) was hybridized with the Prox 1 homeobox probe obtained by PCR at 42°C in high stringency hybridization buffer (50% formamide, 5 \times Denhardt's, 5 \times SSC, 0.02% sodium pyrophosphate, 100 μ g/ml Torula RNA, 0.1% SDS, 125 U/ml heparin, 0.03 M sodium phosphate, pH 6.5). Filters were washed three times with 2 \times SSC, 0.1% SDS at 42°C and twice with 0.2 \times SSC, 0.1% SDS at 65°C. 10⁶ plaques were screened and 2 positives clones were obtained and purified. The isolated inserts were subcloned into pBluescript KSII + and sequenced using Pharmacia sequencing kit. Both isolated clones were the same and contained the Prox 1 homeobox plus 5' and 3' flanking sequences.

In situ hybridization

Embryos were dissected out, fixed overnight in 4% paraformaldehyde, and embedded in Paraplast (Monoject Scientific). Sections (8 μ m) were cut and dried onto chromalum-gelatin slides. All the steps of high-stringency hybridization, washings, and RNase treatment are as described in Kessel and Gruss (1991). T3 or T7 derived in vitro transcribed sense or anti-sense probes were generated from the pBluescript KSII +

subclone carrying the homeobox obtained by PCR. A different probe, 3' of the homeobox, was used for the in situ hybridization with similar results; the results presented here were done with the homeobox probe. Probes (1×10^8 cpm/ml) were used as described by Kessel and Gruss (1991) with the only difference that, α - ^{33}P (Amersham) was used instead of α - ^{35}S with a considerable enhancement of the signal. The exposure time was approximately 20 days. The sense control hybridizations gave no specific signal above background. Whole mount in situ hybridization using the same RNA probe as before but digoxigenin-labeled together with an alkaline phosphatase-coupled anti-digoxigenin antibody was made as described (Wilkinson, 1992).

Immunohistochemical localization of prospero protein in Drosophila embryos and eye disks

Monoclonal antibodies to a *prospero*-glutathione fusion protein were generated (Span and Doe, in preparation) and used to detect *pros* protein in embryos and eye imaginal disks. Standard fixation and staining methods were used (Doe, 1992). The monoclonal supernatant was used at a 1:1 dilution and detected with and HRP-conjugated secondary antibody for embryos or a lissamine rhodamine-conjugated secondary antibody for eye disks (both from Jackson Immunoresearch). Eye disks were counterstained for neuronal membranes using goat-anti-HRP conjugated to fluorescein (Cappel) and viewed on a BioRad 600 confocal microscope.

Acknowledgements

We would like to acknowledge Drs. F. Guillemot and A. Joyner for sharing data before publication, Dr. M. Kessel for critical reading of the manuscript, and especially Dr. A. Stoykova for the help and discussions. This work was supported by the Max-Planck Society. C.Q.D. is supported by grants from the NIH, NSF, PVI, and Searle Scholars Program.

References

Angevine Jr., J.B. (1970). *J. Comp. Neur.* 139, 129–188.
 Auffray, C. and Rougeon, F. (1980). *Eur. J. Biochem.* 107, 303–314.
 Bastian, H. and Gruss, P. (1990). *EMBO J.* 9, 1839–1852.
 Campos-Ortega, J.A. and Hartenstein, V. (1985). *The Embryonic*

Development of Drosophila melanogaster (Berlin: Springer-Verlag).
 Campos-Ortega, J.A. and Jan, Y.N. (1991). *Annu. Rev. Neurosci.* 14, 399–420.
 Chu-Lagraff, Q., Wright, D.M., McNeil, L.K. and Doe, C.Q. (1991). *Development Suppl.* 2, 79–85.
 Davis, C.A., Noble-Topham, S.E., Rossant, J. and Joyner, A.L. (1988). *Genes Dev.* 2, 361–371.
 Doe, C.Q., Chu-LaGriff, Q., Wright, D.M. and Scott, M.P. (1991). *Cell.* 65, 451–464.
 Doe, C.Q. (1992). In *Determinants of Neuronal Identity*, pp. 119–154. Academic Press.
 Franco Del Amo, F., Smith, D.E., Swiatek, P.J., Gendron-Maguire, M., Greenspan, R.J., McMahon, A.P. and Gridley, T. (1992). *Development* 115, 737–744.
 Gruss, P. and Walther, C. (1992). *Cell.* 69, 719–722.
 Guillemot, F. and Joyner, A.L. (1993). *Mech. Dev.* 42, 171–185.
 Hayashi, S. and Scott, M.P. (1990). *Cell.* 63, 883–894.
 Herrick, C.J. (1910). *J. Comp. Neurol.* 20, 413–547.
 Jimenez, F. and Campos-Ortega, J.A. (1990). *Neuron* 5, 81–89.
 Johnson, J.E., Birren, S.J. and Anderson, D.J. (1990). *Nature* 346, 858–861.
 Johnson, J.E., Zimmerman, K., Saito, T. and Anderson, D.J. (1992). *Development* 114, 75–87.
 Kessel, M., and Gruss, P. (1990). *Science* 249, 374–379.
 Kessel, M. and Gruss, P. (1991). *Cell* 67, 89–104.
 Kissinger, C.R., Liu, B., Martin-Blanco, E., Kornberg, T. and Pabo, C.O. (1990). *Cell* 63, 579–590.
 Lo, L.C., Johnson, J.E., Wuenschell, C.W., Saito, T. and Anderson, D.J. (1991). *Genes Dev.* 5, 1524–1537.
 Malicki, J., Schughart, K. and McGinnis W. (1992). *Cell* 63, 961–967.
 Matsuzaki, F., Koizumi, K., Hama, C., Yoshioka, T. and Nabeshima, Y. (1992). *Biochem. Biophys. Res. Commun.* 182, 1326–1332.
 McGinnis, N., Kuziora, M.A. and McGinnis, W. (1990). *Cell* 63, 969–976.
 McGinnis, W. and Krumlauf, R. (1992). *Cell* 68, 283–302.
 Ontell, M. and Kozeka, K. (1994). *Am. J. Anat.* 171, 133–148.
 Ott, M.O. and Buckingham, M. (1992). In: *Formation and Differentiation of Early Embryonic Mesoderm*. NATO ASI Series: Life Science. Vol. 231 (ed. R. Bellairs et al.) pp 211–225. New York: Plenum.
 Reaume, A.G., Conlon, R.A., Zirngibl, R., Yamaguchi, T.P. and Rossant, J. (1992). *Dev. Biol.* 154, 377–387.
 Rutter, W.J., Pictel, R.O.L., Harding, J.D., Chirgwin, J.M., MacDonald, R.J. and Przybyla, A.E. (1978). In: *Molecular Control of Proliferation and Differentiation*. Academic, New York.
 Scott, M.P., Tamkun, J.W. and Hartzell III, G.W. (1989). *Biochem. Biophys. Acta* 989, 25–48.
 Thomas, J.B., Bastiani, M.J., Bate, C.M. and Goodman, C.S. (1984). *Nature* 310, 203–207.
 Treisman, J., Gonczy, P., Vashishtha, M., Harris, E. and Desplan, C. (1989). *Cell* 59, 553–562.
 Vaessin, H., Grell, E., Wolff, E., Bier, E., Jan L.Y. and Jan, Y.N. (1991). *Cell* 67, 941–953.
 Wilkie, T.M. and Simon, M.I. (1991). *Methods* 2, 32–41.
 Wilkinson, D.G. 1992. pp. 75–83. In: *In Situ Hybridization. A Practical Approach*. IRL Press.
 Wessells, N.K. and Cohen, J.H. 1967. *Dev. Biol.* 15, 237–270.
 Wright, C.V.E., Cho, K.W.Y., Oliver, G. and De Robertis, E.M. (1989). *Trends Biol. Sci.* 14, 52–56.
 Zhao, J.J., Lazzarini, R.A. and Pick, L. (1993). *Genes Dev.* 7, 343–354.