Expression of the zebrafish paired box gene pax[zf-b] during early neurogenesis

STEFAN KRAUSS1, TERJE JOHANSEN2, VLADIMIR KORZH1 and ANDERS FJIOSE1

1Molecular Genetics Group and 2Biotechnology Group, Department of Microbiology, Institute of Medical Biology, University of Tromsø, 9000 Tromsø, Norway

*Corresponding author

Summary

The paired box-containing (pax) gene family encodes a group of putative transcription factors differentially expressed during embryonic development. In this study, we describe the cloning and expression of a zebrafish gene pax[zf-b], which most probably is a direct homologue to the mouse Pax2 gene. The putative protein encoded by pax[zf-b] contains a paired box, an octapeptide, but no homeobox. However, a region of homology to the N-terminal half of paired-type homeoboxes is detected C-terminal to the pax[zf-b] paired domain. In zebrafish embryos, pax[zf-b] transcripts are first seen during the formation of the neural keel. At 9-10 h of development, two laterally located transverse stripes of cells expressing the gene appear in the rostral 1/3 of the embryo. The two areas subsequently move towards the midline and form the posterior portion of the midbrain. In the following stages of development, at 10–12 h, transcripts are detected in the otic placode, the Wolfian duct including the nephritic primordium and in the optic stalk. At a later time point, beginning at 14–15 h, single cells along the spinal cord, presumably interneurons, start to express the gene. The characteristic expression pattern of pax[zf-b] in the neural tube suggests an involvement of this gene in the regionalization of the midbrain as well as in the specification of neuronal cell fates at early embryonic stages.

Key words: paired box, neurogenesis, kidney development, zebrafish embryology.

Introduction

The paired box was first described as a domain of approximately 130 amino acids conserved among the products of the Drosophila paired (prd) gene and two closely linked genes at the gooseberry locus (Bopp et al. 1986). The Drosophila gene family encoding this domain includes members that are involved in establishing segment periodicity and specific cell fates during later stages of embryogenesis (Baumgartner et al. 1987; Bopp et al. 1989). Antibody staining has demonstrated that the products of the Drosophila paired box-containing genes Pax meso and Pax neuro are located in the nucleus (Bopp et al. 1989), and recent studies by Treisman et al. (1991) and Goulding et al. (1991) have provided direct evidence for the assumption that the paired box mediates specific DNA binding. Thus, a protein containing this element might act as a transcription factor.

The paired box was first described as a domain of approximately 130 amino acids conserved among the products of the Drosophila paired (prd) gene and two closely linked genes at the gooseberry locus (Bopp et al. 1986). The Drosophila gene family encoding this domain includes members that are involved in establishing segment periodicity and specific cell fates during later stages of embryogenesis (Baumgartner et al. 1987; Bopp et al. 1989). Antibody staining has demonstrated that the products of the Drosophila paired box-containing genes Pax meso and Pax neuro are located in the nucleus (Bopp et al. 1989), and recent studies by Treisman et al. (1991) and Goulding et al. (1991) have provided direct evidence for the assumption that the paired box mediates specific DNA binding. Thus, a protein containing this element might act as a transcription factor.

The paired box was not only present in Drosophila genes, but has been conserved throughout evolution in a variety of organisms. In particular, it has been described in several human (Burri et al. 1989) and murine genes (Deutsch et al. 1988; Dressler et al. 1990; Jostes et al. 1991; Plachov et al. 1990), but it is also present in the genomes of nematodes, cockroach, zebrafish, chick and rat (Burri et al. 1989).

During vertebrate embryogenesis, members of the paired box-containing (Pax) gene family show a remarkable variety of expression patterns, and consequently a diversity of suggested functions (for review, see Kessel and Grass 1990). The murine Pax1 transcripts, for example, are seen in the segmentally organized vertebral column (Deutsch et al. 1988). Studies on the undulataed phenotype have shown that a point mutation at a highly conserved position of the Pax1 paired box correlates with severe distortions along the entire vertebral column suggesting an important role for Pax1 in the normal development of the column (Balling et al. 1988). Pax transcripts are also detected in the developing excretory system. Both Pax2 and Pax8 are expressed in the mesenchymal condensations, which are induced by the nephritic duct. Pax2 transcripts are also detected in the pronephric tubules and the extending nephritic duct (Dressler et al. 1990; Plachov et al. 1990). The expression patterns of both genes have been correlated with induction of epithelial structures forming the kidney. Transcripts of Pax8 have
in addition been detected in the invaginating thyroid gland (Plachov et al. 1990).

In the epichordal division of the neural tube, transcripts of several Pax genes have been detected in adjacent and overlapping longitudinal areas along the anterior–posterior axis. The closely related genes Pax2 and Pax8 are expressed in two parallel regions on either side of the intermediate plate (Nornes et al. 1990; Plachov et al. 1990). Pax7 is active in dorsal portions of the ventricular zone along the anterior–posterior axis of the neural tube (Jostes et al. 1990), whereas Pax3 transcripts are restricted to the dorsal portion of the neural tube (Goulding et al. 1991). In contrast to the early embryonic expression of Pax2, Pax7 and Pax8, the expression of Pax3 extends further anteriorly to the diencephalon. It has been suggested that members of the Pax gene family may, in concert with other factors, specify cell fates in the hindbrain and spinal cord (e.g. Jostes et al. 1991).

More anteriorly, in the prechordal division of the neural tube, Pax2 transcripts have been detected in the optic stalk whereas Pax7 is expressed in the forebrain and, at later developmental stages, in the midbrain (Nornes et al. 1990; Jostes et al. 1990).

In the present study, we report the cDNA sequence and the expression pattern of the zebrafish *pax*zf-b* gene. We show that, similar to mouse Pax2, transcripts of the zebrafish gene are detected in the optic stalk, the otic vesicle, the excretory system and along either sides of the hindbrain and the spinal cord. In addition, we demonstrate that at earlier developmental stages *pax*zf-b is also expressed within a transverse stripe in the posterior midbrain adjacent to the furrow separating midbrain and hindbrain. Furthermore, transcripts of the gene are detected in single cells, presumably interneurons, along the anterior–posterior axis of the neural tube. This suggests a fundamental role for *pax*zf-b in the regionalization of the rostral brain and in the specification of neuronal cell fates.

### Materials and methods

#### Cloning, sequencing and polymerase chain reaction

Approximately 1.5 x 10⁸ plaques of a lambda gt11 cDNA library derived from 33 h zebrafish embryos were screened at low stringency (McGinnis et al. 1984) using a mix of two paired box sequences as probe: the 313 bp HindIII–EcoRI fragment from murine Pax1 and a cloned PCR product from zebrafish *pax*zf-d. A 2409 bp long cDNA clone contained the entire *pax*zf-b coding sequence. The cDNA was subcloned into pGEM-3Zf(+) (Promega Biotech), mapped with restriction enzymes and both strands were sequenced on overlapping fragments in M13mp18 and M13mp19 vectors using Sequenase (US Biochemicals). Polymerase chain reaction was performed directly on 2.3 x 10⁸ pfu of the cDNA library as described by Friedman et al. (1990) with the modification that the following PCR profile was used: 94°C, 1 min; 58°C, 1 min; 72°C 3 min for 30 cycles. The primers used were derived from positions 1144 to 1164 (5' primer) and from 1326 to 1346 (3' primer) of the *pax*zf-b cDNA sequence. PCR products were cloned using the TA cloning kit (Invitrogen) and sequenced. DNA sequences and derived amino acid sequences were analyzed on a VAX/VMS computer using the GCG software package (version 6.2; Devereux et al. 1984).

#### In situ hybridization on tissue sections

Zebrafish embryos after 12, 18, 24 and 36 h of development at 28.5°C were manually dechorionated and subsequently fixed overnight in 4% paraformaldehyde in Fix-buffer (4% sucrose, 0.12 M CaCl₂, 0.1 M PO₄ pH 7.3) at 4°C. After fixation the embryos were washed 3×5 min in Fix-buffer and embedded in agar–sucrose (1.5% agar, 5% sucrose) at 4°C. 16 μm sections were cut at −20°C, placed on gelatinized slides and dried at room temperature. 100 ng probe was labelled by nick-translation (BRL) using [35S]dATP. The size of the probe was adjusted to an average of 40–120 nucleotides by the addition of Ro Q DNAase (Promega) to the reaction. The probe was then adjusted to a concentration of 0.1 ng μl⁻¹ in hybridization buffer (50% denatured formamide, 15% dextran sulfate, 10 mM DTT, 1 mM Denhardt's, 0.1 mg ml⁻¹ herring sperm DNA, 0.1 mg ml⁻¹ yeast (RNA and 1 μg thio S-ATP) and heat denatured.

The tissue sections were dehydrated in ethanol (50%, 75%, 95% and 100%; 4 min each), the probe was applied to the sections, covered with parafilm, sealed with rubber cement and hybridized overnight at 45°C in a wet chamber. After hybridization the parafilm was removed and slides were washed twice at 45°C in 50% formamide, 4×SSC, 0.1% 2-mercaptoethanol for 30 min, and once in 2×SSC for 30 min at room temperature. Slides were then dehydrated in ethanol-0.3 M ammonium acetate (50%, 75%, 95% and 100%; 4 min each), and dipped in undiluted LM-1 photoemulsion (Amersham). After exposure (7−10 days), slides were developed for 3 min in Kodak D19, fixed for 6 min in Kodak Unifix, washed 15 min in H₂O, stained 2 min in 5% Giemsa (in 1×PBS), washed briefly in H₂O, dried and mounted in Difco D.P.X 8711.

#### In situ hybridization on whole-mount embryos

The protocol used is a modification of the procedure described for *Drosophila* by Tautz and Pfeifle (1989). Zebrafish embryos were fixed as described above. The embryos were then washed 4×15 min in PBT (1×PBS, 0.1% Tween-20, 0.2% BSA), dipped 10 s in H₂O, transferred to acetone (7 min at −20°C) and washed again 2×15 min in PBT. PBT was then stepwise changed to hybridization buffer HB (15 min 1:1 PBT/ HB, 15 min HB; HB is 50% formamide, 5× SSC, 100 μg ml⁻¹ salmon sperm DNA, 50 μg ml⁻¹ heparin and 0.1% Tween-20). The embryos were subsequently prehybridized for 30 min at 45°C in HB. The desired fragment was labelled for 1 h at 15°C in the following reaction mix: 250 ng DNA, 5 μl 10×buffer (500 mM Tris 7.8, 50 mM MgCl₂, 100 mM 2-mercaptoethanol), 1 μl DNTP-mix (BMB, Geneus kit), 4 μl 1/10 diluted Ro Q DNAase (Promega) and 5 μl DNAase/polymerase mix (BRL nick translation system). The reaction was stopped by addition of 5 μl Stop-mix (BRL nick translation system) and heated for 10 min at 68°C. The probe was subsequently precipitated, added at a concentration of 5 μg ml⁻¹ to an appropriate volume of HB, boiled for 2 min and put on ice. The prehybridization solution was removed, and the HB was added to the embryos. The embryos were hybridized overnight at 45°C. After the hybridization, the embryos were washed for 1 h in fresh HB at 45°C, for 15 min at 45°C in 1:1 HB: PBT and twice for 15 min at room temperature in PBT. Embryos were refixed for 30 min in 4% paraformaldehyde in Fix-buffer (see above) at room temperature, washed 2×15 min in PBT and left for 1 h in PBT. (The refixation step is not essential.) After the last washing steps,
50% of the embryos were transferred into an extra tube and incubated for 1 h with anti-digoxigenin antibodies (BMB) diluted 1:2000 in PBT (700 μl antibody solution per 100 embryos). Preincubated antibodies were then added to the remaining embryos and left at 4°C overnight. Embryos were again washed 4×15 min in PBT and rinsed subsequently 3×5 min in pH 9.5 buffer (100 mM NaCl, 50 mM MgCl2, 100 mM Tris–HCl pH 9.5, 1 mM Levamisol, 0.1% Tween-20). The staining was carried out in staining solution (4.5 μl NBT (BMB) and 3.5 μl X-Phosphate (BMB) in 1 ml pH 9.5 buffer) for 20 min to 3 h. After a suitable signal-to-background relation was reached, the reaction was stopped by washing the embryos 3×5 min in PBT. Embryos were finally dehydrated in ethanol (30%, 50%, 75%, 90%, 95%, 100%, 3 min each), put for 1 h into methyl salicylate and mounted in Permount.

Results

Structure of pax[zf-b] cDNA

A cDNA clone containing the entire pax[zf-b] coding sequence was isolated by low-stringency screening of a cDNA library derived from 33 h zebrafish embryos using a mix of murine Pax1 (Deutsch et al. 1988) and zebrafish pax[zf-d] (H. G. Eiken, P. Villand and A. Fjose, unpublished data) paired box sequences as probe. The complete nucleotide sequence and the derived amino acid sequence of the pax[zf-b] cDNA are shown in Fig. 1A. The cDNA clone is 2409 bp long and contains an open reading frame (ORF) capable of encoding a putative protein of 391 amino acids (aa) with a predicted relative molecular mass of 420.8 kDa. The ORF, beginning at nucleotide 627 and ending with an AUG stop codon at position 1800, includes the paired domain. There is another possible start codon at position 663 just upstream of the paired domain. However, the AUG at position 627 is most probably the start codon; it has a good match to the Kozak core consensus (PuCAUGG) for translational initiation sites (Kozak, 1986 and 1987), and the downstream ATG is too close to the paired box compared to other known paired box genes. A 130 aa paired domain is located at aa positions 15 to 144 (Bopp et al. 1986). The 5’ untranslated leader sequence is 626 nucleotides long and contains a short (44 aa) upstream ORF (positions 15 to 146) in the same reading frame as the long ORF encoding the paired domain. The two ORFs are separated by eleven in frame stop codons (see Fig. 1A). The 3’ untranslated trailer sequence is 610 nucleotides long and does not contain a consensus polyadenylation sequence (Proudfoot and Brownlee, 1976; Wickens and Stephenson, 1984) upstream of the poly (A) tail at the 3’ end of the cDNA. Thus, the 3’ end of the cDNA clone is most probably the result of priming of cDNA synthesis by oligo-(dT) at an internal poly (A) sequence. A ‘CA repeat’ (Hamada et al. 1982) is present in the 3’ untranslated sequence.

The proteins encoded by zebrafish pax[zf-b] and mouse Pax2 are highly homologous

The paired domain sequence of pax[zf-b] contains only 5 amino acid substitutions compared to mouse Pax2 (96% identity), two of which represent changes to chemically similar amino acids. The homology to the mouse Pax8 paired domain is also high (8 substitutions; 93.8% identity). Similar to the mouse Pax1 (Deutsch et al., 1988), Pax2 (Dressier et al., 1990), Pax8 (Plachov et al. 1990) and Drosophila Pax meso and Pax neuro (Bopp et al. 1989) encoded proteins, pax[zf-b] does not contain a paired-type homeobox. Another amino acid sequence motif that has been reported to be conserved among several paired box-containing proteins, the octapeptide (Burri et al. 1989), is located 40 aa C-terminal to the paired box. A comparison of the complete amino acid sequence of pax[zf-b] to those of published genes encoding paired box-containing proteins revealed that Pax[zf-b] is highly homologous to mouse Pax2 (see Fig. 1A). The two amino acid sequences are 87.2% identical, and if changes to chemically similar amino acids (scored as in Johansen et al. 1989) are considered, the two putative proteins show a sequence similarity of 96.7%. There is no significant homology at the DNA sequence level between the 5’- and 3’-untranslated regions of pax[zf-b] and Pax2. Thus, the homology is confined to the coding sequence (the nucleic acid sequence identity here is 81.5%). As seen from Fig. 1A, pax[zf-b] contains a 6 bp deletion compared to Pax2 located between positions 1512 and 1513 and a 3 bp insertion at position 1764.

Dressler et al. (1990) described two overlapping cDNA clones for Pax2, one of which contained a 69 nucleotides insertion (23 aa) in the ORF downstream of the octapeptide. The pax[zf-b] cDNA clone described here lacks this insertion in the coding sequence. To investigate whether zebrafish also contain pax[zf-b] transcripts with such an insertion, we performed polymerase chain reaction (PCR) directly on 2.3×10^6 pfu of the embryonic cDNA library using primers flanking the insertion site. Two PCR products were obtained, cloned and sequenced. The shorter product of 203 bp corresponds exactly to the product expected from the cDNA sequence shown in Fig. 1A. The longer 248 bp product contains an insertion of 45 bp (15 aa) compared to the pax[zf-b] cDNA sequence (see Fig. 1C). In addition to the introduction of 15 aa, the insertion also interrupts a codon thus leading to an aspartate to histidine substitution. The 161 bp sequence between the primers flanking the insertion site is identical to the pax[zf-b] cDNA sequence (data not shown). This result indicates that two versions of the Pax[zf-b] protein differing only in 16 aa are expressed. However, both the length and the sequence of the insertion are different from that of mouse Pax2. Interestingly, there is no homology between the 15 aa insertion in zebrafish Pax[zf-b] and the 23 aa insertion reported for mouse Pax2.

Analysis of the derived amino acid sequence of the C-terminal region downstream of the paired domain of pax[zf-b] reveals several interesting features (see Fig. 1B). This region is rich in serine residues (16.6%), a feature shared with other paired box-containing proteins (Goulding et al. 1991; Krauss et al. 1991). Stretches with local overrepresentation of either serine and proline residues or both are evident. Furthermore,
Fig. 1. Sequence of the zebrafish pax[zf-b] cDNA. (A) The nucleotide and derived amino acid sequence of pax[zf-b] cDNA. The paired domain is boxed. Amino acid changes found in mouse Pax2 relative to Pax[zf-b] are shown below the Pax[zf-b] sequence. Changes to chemically different residues are indicated by double underlining. The octapeptide and a 'CA repeat' are underlined. The downward arrow indicates the position of the 45 bp (15 aa) insertion found in some transcripts and the upward arrow points to a 6bp deletion in pax[zf-b] relative to Pax2. A 3bp insertion in pax[zf-b] is shown (bold, underlined). # indicates an upstream in frame start codon, and stars are used to indicate in frame stop codons. (B) Schematic representation of the putative Pax[zf-b] protein showing the location of the paired domain, the octapeptide (solid box), Ser- and Pro-rich regions, a charged/acidic region, a putative amphipathic alpha-helix and the 15 aa insertion found in some pax[zf-b] transcripts. (C) DNA and derived amino acid sequence of the 45 bp (15 aa) insertion (boxed) and flanking regions found in some pax[zf-b] transcripts. The numbers refer to the position in the cDNA sequence shown in A.
from amino acid position 194 to 280 a charged region with an acidic theoretical isoelectric point (=4.7) is found. The N-terminal part of this region is especially rich in charged residues (e.g. 10/15 from 194 to 208). Secondary structure predictions using both Garnier-Robson and Chou-Fasman algorithms (GCG software package version 6.2) show the C-terminal region to be largely composed of random coils, some short beta sheets and many turns. Interestingly, a 28-residue-long alpha-helix is strongly predicted from position 219 to 246 in the charged region. Helical wheel analysis suggests this to be an amphipathic helix with both acidic, basic and aliphatic surfaces.

Pax[zf-b] is homologous to the mouse Pax8 protein in the C-terminal region downstream of the paired domain

It has previously been shown that the paired domains of Pax2 and Pax8 are highly homologous (Plachov et al. 1990). We therefore asked whether Pax[zf-b] show any significant homology to Pax8 outside the paired box. Fig. 2A shows that the derived amino acid sequence of Pax[zf-b] is clearly homologous to Pax8 and Pax2b (the variant of Pax2 lacking the 23 aa insertion) in the region C-terminal to the paired box. Pax8 contains a 69 aa insertion relative to Pax[zf-b]/Pax2b located 101 aa C-termally to the 23 aa insertion in Pax2. The C-terminal region downstream of the paired box show 52% identity and 78% similarity between Pax8 and Pax[zf-b]. The most conserved part of the C-terminal region is located downstream of the insertion in Pax8. Here, the identity and similarity scores are 66% and 95%, respectively (see Fig. 2B).

Homology to paired-type homeoboxes in pax[zf-b]/Pax2 and Pax8

By comparing the derived amino acid sequence of pax[zf-b] to those of known paired box-containing genes, we found that a region located about 80 amino acids downstream of the paired box in Pax[zf-b] is homologous to the N-terminal half of the paired-type homeobox (see Fig. 3). In the region comprising amino acids 8 to 31 in the paired homedomain (as defined by Scott et al. 1989), the identity between Pax[zf-b]/Pax2 and 6 paired homeodomain sequences varies from 46% to 58% and the similarity (including changes to Pax2b) 50 100

---

Fig. 2. Homology between Pax[zf-b]/Pax2 and Pax8. (A) Alignment of the Pax[zf-b] and Pax2 amino acid sequences to that of Pax8. Two dots indicate identity and one dot indicates a change to a chemically similar amino acid. Gaps introduced are shown as dashes. (B) Identity and similarity scores (within brackets) between Pax[zf-b]/Pax2 and Pax8.
chemically similar amino acids) varies from 79% to 92%. This homology is also found in Pax8 (data not shown). We do not presently understand the functional significance of this homology, but it is interesting that the predicted alpha-helix mentioned above ends exactly at the same position within this homology region as the first alpha-helix of the homeobox determined from the X-ray crystal structure of the engrailed homeobox (Kissinger et al. 1990; see Fig. 3).

Expression of pax[zf-b] during embryogenesis

The expression of pax[zf-b] was analyzed at different stages of zebrafish development by in situ hybridization on tissue sections and whole-mount embryos. Two different cDNA fragments were used as probes with consistent results; (i) the 715 bp EcoRI–HindIII fragment which excludes the conserved paired box (position 1694–2409), and (ii) the 2409 bp EcoRI fragment containing the complete cDNA. For maximum signal intensity only results from hybridizations with the latter probe are shown here. Expression of pax[zf-b] was first detected at 9–10 h (see Fig. 4A–H) of development and was thereafter seen throughout embryogenesis (see Fig. 5A–H). In summary, transcripts of the gene were confined to the kidney, the optic stalk, the midbrain, the otic vesicle and a row of cells along the spinal cord and the hindbrain (Fig. 9A–C).

Expression in the midbrain

The earliest detectable transcripts of pax[zf-b] appear in the rostral 1/3 of the embryo when the process of epiboly is close to completion (9–10 h of development). At this stage, two faint transversal bands expressing pax[zf-b] can be detected in whole mount in situ hybridizations. The two bands describe an angle of about 50°–60° compared to the midline, and are separated by a non-expressing area in the midline of the embryo (Fig. 4A,E). Subsequently, the two transversal bands enlarge in size and, in particular, the medial portion of the two bands moves towards the midline. In addition, the two bands broaden along the longitudinal axis of the embryo (Fig. 4B,F). This process continues and at a slightly later stage the two transverse bands join at the midline. At this stage the longitudinal enlargement continues while the two bands start to shorten along the transverse axis (Fig. 4C,G). At the time point when the rostral portion of the neural keel is clearly visible, the staining for pax[zf-b] is restricted to the posterior region of the prospective midbrain. During this phase, the shape of the area expressing the gene resembles an arrowhead (Fig. 4D,H), but at a later stage (Fig. 8A), the transcripts of the pax[zf-b] gene form a solid transverse band. In the following period of development, the expression in the midbrain seems to remain without major visible changes (compare in situ hybridization on whole-mount embryos at 15 h of development in Fig. 8A with hybridization on sections at 18 h of development in Fig. 5C,D). At 24 h of development, this region in the midbrain can be correlated with several morphological landmarks as seen in horizontal and sagittal sections of the embryo. (i) The posterior border of expression coincides with the position where the furrow separating midbrain and hindbrain is formed at 20 h of development. As the expression of pax[zf-b] precedes the appearance of this furrow by several hours, it could be suggested that the gene may be involved in specifying the position of the furrow (Figs 5E,F; 7D). (ii) The anterior border of pax[zf-b] expression coincides with the posterior end of the tectal ventricle, as seen in 24 h sagittal sections (Fig. 7D). In horizontal sections, the position of pax[zf-b] expression appears precisely in the portion of the neural tube that surrounds the connection between the tectal ventricle in the midbrain and the 4th ventricle in the hindbrain (see Fig. 6E,F). (iii) The expression of pax[zf-b] in the posterior midbrain does not include the most ventral portions of the neural tube suggesting that the floor plate located most ventrally does not contain transcripts of the gene (not shown). In this context, it is noteworthy that the floor plate serves as a pathway for axon tracts (e.g. Wilson et al. 1990).
Fig. 4. Localization of pax/zf-b transcripts by in situ hybridization on a series of whole-mount embryos of early developmental stages between 9 and 10h. Lateral views (E–H) and dorsal views (A–D) are shown. The embryos are oriented with the anterior end to the right. Bar, 100 μm.

Fig. 8. In situ hybridization of pax/zf-b on whole-mount embryos at 15h of development. The embryo is oriented as in Fig. 3. (A) Lateral view of the embryo. (B) Dorsal view of the spinal cord of the embryo. (C) Close up of the spinal cord seen from lateral. Dorsal is to the top, the two rows of cells described in Discussion are marked with arrows and arrowheads, respectively. Abbreviations: d, diencephalon; e, eye; h, hindbrain; m, midbrain; np, nephritic primordium; os, optic stalk; ov, otic vesicle; sc, spinal cord; t, telencephalon; w, Wolffian duct; y, yolk. Bars: 50 μm.
Zebrafish pax[zf-b] expression
Expression in the optic stalk

At 11–12 h of development, the wall of the neural tube begins to bulge out laterally at the region of the embryonic diencephalon to form the optic vesicle. As this process continues, the connection between the optic vesicle and the neural tube forms the optic stalk. Transcripts of \textit{pax[zf-b]} are first seen at around 12 h of development in the region of the optic stalk. The expression is stronger at the anterior portion of the stalk than more posteriorly. In addition, lower levels of transcripts are also detected in the lateral parts of the optic vesicle (see Fig. 5A,B for sagittal sections and Fig. 6A,B for horizontal sections). At 18 h of development, the distribution of transcripts within the stalk is
more homogenous (Fig. 5C,D; Fig. 6C,D) and as the stalk constricts, the staining increases considerably (Fig. 7D; see also Fig. 5E,F).

**Expression in isolated cells along the spinal cord and the hindbrain**

Surprisingly, the *pax*zf-b*] gene is also expressed in isolated cells along the anterior-posterior axis of the spinal cord and the hindbrain. As seen in Fig. 8A–C, these cells first appear at approximately 14–15 h between somites 1 and 10. The hybridization on whole-mount embryos shows that two different groups of *pax*zf-b*]-expressing cells can be distinguished along the axis of the cord (see arrows and arrowheads in Fig. 8C; for 18 h see Fig. 5C,D and Fig. 7E). As development continues, additional cells expressing *pax*zf-b*] are detected at both sides of the midline anteriorly and posteriorly to the cells that already express the gene, leading to a row of adjacent single cells that extends continuously in both rostral and caudal directions (Fig. 5C,D and Fig. 7F). Furthermore, in between the *pax*zf-b*]-expressing cells, additional cells appear to transcribe the gene. Cells staining for *pax*zf-b*] are seen at 18 h at the level of the otic vesicles (Fig. 5C,D) and, at 24 h, *pax*zf-b*]-expressing cells are detected at the anterior end of the hindbrain as well as at the posterior end of the spinal cord (Figs 5E,F, 7F).

At their earliest appearance (14–15 h), the labelled cells are of different size and location: (i) more dorsally, there is a row of cells larger than the surrounding cells with spherical soma or soma elongated along the rostral-caudal axis (arrowheads in Fig. 8C); (ii) slightly ventrally, we detect a second row of cells showing a smaller size but a similar spherical or rostral-caudal oriented cell body (arrows in Fig. 8C). Although most of these cells are seen as bilateral pairs with an average of two cells present on each side of a spinal segment (Fig. 8B), the cells do not show a strictly repeated pattern. In cross sections on 18 and 24 h embryos at the level of the hindbrain–spinal cord junction, the *pax*zf-b*] transcripts are located laterally at the wall of the neural tube (for 24 h, see Fig. 7A). Since neurons are known to migrate from the luminal part towards the
wall of the neural tube in order to differentiate, \(\text{pax}\{zf-b\}\) is most likely expressed in neurons.

At 24 h, transcripts of \(\text{pax}\{zf-b\}\) can be detected along the entire length of the spinal cord and the hindbrain but not in rhombomere 1 and the cerebellum (Fig. 5E,F). Unfortunately, at this developmental stage, we did not obtain satisfactory whole-mount embryos and we are therefore not able to give a more detailed description of the number and shape of the \(\text{pax}\{zf-b\}\)-positive cells. Finally, at 36 h, \(\text{pax}\{zf-b\}\)-staining cells are also detected within rhombomere 1 and the cerebellum (Fig. 5G,H).

Expression in the otic vesicle and the nephritic primordium

In addition to expression in the neural tube, \(\text{pax}\{zf-b\}\) transcripts are also detected in the developing otic vesicle, the pronephric duct and the nephritic primordium. The otic vesicle is derived from epithelial tissue (the otic placode) invaginating towards the neural tube. Beginning at 11 h, \(\text{pax}\{zf-b\}\) transcripts are detected in the otic placode, an epithelial condensation which is located at this time point close to the junction between the epithelium covering the yolk and that covering the embryo (Fig. 5A,B; for horizontal sections, see Fig. 6A,B). At 18 h, \(\text{pax}\{zf-b\}\) transcripts describe an arch surrounding the otic cavity. These \(\text{pax}\{zf-b\}\)-expressing cells are located at the part of the otic vesicle that faces towards the neural tube. This tissue will at later developmental stages form the neural portions of this organ (Fig. 6C,D for horizontal sections; Fig. 7B for cross sections and Fig. 7C and F for sagittal sections).

Further posterior in the embryo, close to the described border between the embryonic body and the yolk epithelium, a narrow band of mesodermal cells start to express \(\text{pax}\{zf-b\}\) at 12–14 h (not shown). This band shows at 15 h anteriorly a distinct enlargement at the position between somite 3 and 5 (Fig. 8A). Moreover, a thickened bifurcation is seen posteriorly at the junction between the yolk and the tail (Fig. 7E, for horizontal sections, see Fig. 5E,F). Judging from the location, the staining maps to the Wolffian duct and the elongation of it, the urogenital system. Both will form the excretory system.

The urogenital system will be induced by the metanephrogenic mesenchyme to branch and will together with the metanephric condensations form the collecting duct system of the kidney. No \(\text{pax}\{zf-b\}\) hybridization is seen in the lateral mesoderm anteriorly to the urogenital system.

Discussion

Molecular structure of \(\text{Pax}\{zf-b\}\)

In this report, we have described the isolation and characterization of a cDNA derived from the zebrafish paired box-containing gene \(\text{pax}\{zf-b\}\) as well as the temporal and spatial expression pattern of this gene during embryogenesis. The derived amino acid sequence of the putative \(\text{Pax}\{zf-b\}\) protein is highly homologous to mouse \(\text{Pax}2\). We also show that there is a high degree of homology between \(\text{Pax}\{zf-b\}/\text{Pax}2\) and the murine \(\text{Pax}8\) in the C-terminal region downstream of the conserved paired box. This is interesting in view of the similarity seen in the expression pattern between the murine \(\text{Pax}2\) and \(\text{Pax}8\) genes (Dressler et al. 1990; Plachov et al. 1990). The murine \(\text{Pax}3\) and \(\text{Pax}7\) genes which similarly show high amino acid homology, also display overlapping expression patterns (Goulding et al. 1991). Moreover, we report a novel homology of \(\text{Pax}\{zf-b\}/\text{Pax}2\) and \(\text{Pax}8\) to the N-terminal half of paired-type homeobox sequences. This homology is located in a region downstream of the paired box roughly where the homeobox is found in homeobox-containing paired box genes. At present, we do not know if this homology is an evolutionary remnant of a homeobox where the C-terminal half has been lost in a process of exon shuffling or if it is a result of convergent evolution due to some functional constraint. It is possible that this region may have an important function in DNA binding (together with the paired domain) or transcriptional activation (see discussion below).

From the murine \(\text{Pax}2\) cDNA sequence, it was impossible to determine the start codon for the protein since there were no in frame stop codons in the upstream sequence and two other possible start codons upstream of the paired box in addition to the one proposed as the start codon (Dressler et al. 1990). The sequence of the zebrafish \(\text{pax}\{zf-b\}\) cDNA described here shows that the start codon predicted for \(\text{mouse} \text{Pax}2\) is most probably correct since there are many stop codons upstream of the putative start codon in \(\text{pax}\{zf-b\}\) cDNA and since the methionine codon located two codons downstream of the proposed initiator codon in \(\text{mouse} \text{Pax}2\) is substituted with an isoleucine codon in the zebrafish \(\text{Pax}8\) homologue. The short upstream ORF found in \(\text{pax}\{zf-b\}\) cDNA is probably too close to the Cap site to be translated efficiently (Sedman et al. 1990).

Dressler et al. (1990) showed that two variants of the murine \(\text{Pax}2\) protein can be deduced from cDNA sequences. One form containing a 23 aa insertion relative to the other. These authors suggest that this variation is due to differential splicing. For zebrafish \(\text{pax}\{zf-b\}\), we also find two types of transcripts, but, surprisingly, the insertion is both different in length and sequence from that of \(\text{mouse} \text{Pax}2\). It is presently unclear if the two variants of \(\text{Pax}\{zf-b\}\), presumably generated by differential splicing, show any difference in their spatial and temporal expression patterns.

During the past few years, evidence indicating that paired box-containing genes encode transcription factors have accumulated. The homeodomain is a well-established DNA-binding element (for review see Hayashi and Scott, 1990). The \(\text{Drosophila} \text{Paired} \text{(Prd)}\) protein, which contains both a paired domain and a homeodomain, binds to specific DNA sequences (Hoey and Levine, 1988; Han et al. 1989). Prd contains two specific DNA-binding activities (Treisman et al. 1989), one mediated by the homeodomain and one mediated
by the paired domain (Treisman et al. 1991). Goulding et al. (1991) have also shown for the murine Pax3 protein that the paired box is capable of sequence-specific DNA binding. This strongly suggests that both paired box-containing gene products with and without homeodomains are transcription factors. It has also been shown that the *Drosophila* *Pox meso* and *Pox neuro* genes, which contain a paired box but no homeobox, encode nuclear proteins (Bopp et al. 1989). Since the N-terminal region upstream of the paired box generally is rather short, the C-terminal region downstream of the paired box should harbor transcriptional activating elements. Such elements are relatively poorly characterized, but different structural features such as acidic helices, glutamine-rich, proline-rich, and serine/threonine-rich domains have been shown to be involved in activation of transcription (Ptashne, 1988; Courey and Tjian, 1988; Mermod et al. 1989; Tanaka and Herr, 1990). Interestingly, the *Pax*[zf-b] protein contains a C-terminal region particularly rich in serines with short stretches containing a high density of serine-, proline- or both residues. In addition, we find that an amphipathic alpha-helix with one acidic surface is predicted in a rather long charged/acidic region. All these are features typical for transcriptional activating elements.

**Fig. 9.** Camera-lucida reconstruction of the embryo at 18 h of development and comparison of the *pax*[zf-b] expression at 18 h with the expression of other transcription factors mapping to the zebrafish brain. (A) Lateral view of the embryo. (B) Three-dimensional reconstruction of the embryo seen from a dorsolateral angle. The scheme is based on a series of in situ on tissue sections. (C) Location of *pax*[zf-a], *pax*[zf-b] and *enraged* transcripts in a simplified neural tube. Dots mark individual neurons staining for *pax*[zf-b]. Abbreviations: d, diencephalon; e, eye; h, hindbrain; m, midbrain; np, nephritic primordium; o, otic vesicle; s, somites; t, telencephalon; y, yolk.
Expression of pax\(\text{zf-b}\) in the spinal cord

The zebrafish spinal cord is organized in an extremely simple manner and therefore provides a suitable system for studying neuronal differentiation in vertebrates (Eisen et al. 1986; Kuwada, 1986; see also Kuwada et al. 1990). The zebrafish cord contains at early developmental stages (18–20 h) approximately 18 postmitotic neurons per hemisegment (segments 5–15) which are identifiable according to their form, size and the stereotyped, cell-specific pathway of their axonal outgrowth (for review, see Kuwada and Bernhardt, 1990).

\(\text{Paxzf-b}\) shows, at 15 h of development, a remarkable expression pattern along the cord, which is restricted to an average of two single cells on the lateral surface of each hemisegment. From the intermediate position of the positive cells, we can exclude that \(\text{paxzf-b}\) is expressed within the primary motoneurons and the mechanosensory Rohon-Beard cells (Myers et al. 1986). Thus, at early embryonic stages, the \(\text{paxzf-b}\) gene is most probably expressed in interneurons. However, a firm identification of the exact class of interneurons that stain for \(\text{paxzf-b}\) would require more detailed analyses such as double staining with antibodies.

The presumptive murine homologue to the zebrafish \(\text{paxzf-b}\) gene, \(\text{Pax2}\) is at a relatively early developmental stage (day 10 p.c.) expressed in two clearly distinguishable longitudinal columns in the spinal cord, the pars ventralis and the pars dorsalis of the intermediate grey (Nornes et al. 1990). Since these two genes encode very similar proteins and share comparable expression patterns in other parts of the embryonic body (optic stalk, kidney), we would assume that they are both involved in the determination of single interneurons within the cord.

Importantly, the observed restriction of \(\text{paxzf-b}\) transcripts to single cells (neurons) seems to be unique among the vertebrate paired box-containing genes analyzed so far. The transcripts of at least three other zebrafish \(\text{Pax}\) genes (\(\text{zf-a}\), \(\text{zf-e}\) and \(\text{zf-f}\)) are never restricted to specific cells. Instead they show a dorsoventral restricted, homogenous staining along the neural tube at early embryonic stages (S. Krauss et al. unpublished data).

Immunological studies on zebrafish embryos using an antibody that detects epitopes of the frog homeobox \(\text{XIIBox 1}\), revealed an expression pattern confined to different cell populations within the spinal cord (Molven et al. 1990). However, the features of the \(\text{XIIBox 1}\) staining within the spinal cord differs fundamentally from the expression of \(\text{paxzf-b}\) as it includes a wide variety of cell types and shows a graded expression extending posteriorly from a defined anterior boundary. Another example has recently been demonstrated with the antibodies \(\text{4D9}\) and \(\text{alphaEnhhb-1}\) (which most probably detect \(\text{engrailed}\) homeoproteins). Whereas \(\text{4D9}\)-positive cells are strictly confined to the anterior three rhombomeres, \(\text{alphaEnhhb-1}\) labels cells throughout the length of the hindbrain and spinal cord (Hatta et al. 1991). The specific identities of these cells remain unknown but, similar to the \(\text{paxzf-b}\)-expressing cells, the \(\text{engrailed}\) staining cells are located in the lateral mantle zone of the neural tube. Several classes of \(\text{Drosophila}\) segmentation genes are expressed in the nervous system and several of them are transcribed in a specific subset of neuronal precursor cells, neurons and glial cells. Such expression has for instance been observed for the pair-rule genes \(\text{ftz}\) and \(\text{eve}\) (Carroll and Scott, 1985; MacDonald et al. 1986). Subsequently, it has been shown by Doe and coworkers (1988a, b) that \(\text{ftz}\) and \(\text{eve}\) are directly involved in specifying the identity of individual neurons. Another class of genes that are expressed in the developing nervous system of the fruit fly are the paired box-containing segmentation genes \(\text{Pox neuro}\) and \(\text{gsb-p}\) (BSH4). Their expression appear in neuroblasts and are detected in the presumably clonally related neural precursors (Burgmartert et al. 1987; Bopp et al. 1989; see also Patel et al. 1989). Hence, it is possible that the vertebrate paired box-containing gene \(\text{paxzf-b}\) is also involved in a network specifying neuronal identities.

Expression of pax\(\text{zf-b}\) in the midbrain

Nornes and coworkers (1990) described the rostral limit of the \(\text{Pax2}\) expression as located at the midbrain–hindbrain border. They did not detect transcripts in the posterior midbrain. In contrast, we observe \(\text{paxzf-b}\) transcripts in a region in front of the furrow separating the midbrain and the hindbrain. This is surprising as the zebrafish \(\text{paxzf-b}\) probe detects transcripts at a very early developmental stage. In fact, \(\text{paxzf-b}\) staining can already be seen at a time point when the process of epiboly is completed. At this stage (early neurula), cells of the embryonic shield converge towards the dorsal midline, moving from lateral to medial positions in order to form the embryonic axis. Warga and Kimmel have demonstrated that most gastrula cell lineages are already tissue specific (Kimmel and Warga, 1986; see also Warga and Kimmel, 1990). However, it remains unclear at what time point and in which way cells get committed to subdivisions within the neural keel. Judging from the time-point of appearance and the molecular structure, the product of the \(\text{paxzf-b}\) gene might play a fundamental role in this process.

At later stages of development, as the tectal ventricle enlarges in zebrafish, the tissue between the furrow separating hindbrain and midbrain gets considerably thinner. Since cells or cell groups within the hindbrain start to express \(\text{paxzf-b}\), the expression in the posterior midbrain could be obscured at the stages that the murine \(\text{Pax2}\) was analyzed.

Recent studies have identified two more genes that share a comparable expression border anteriorly to the furrow; the homeobox-containing gene \(\text{en-2}\) (Joyner et al. 1985; for zebrafish, see Hatta et al. 1990) and the protooncogene and secreted signal protein \(\text{Wnt-1}\) (Wilkinson et al. 1987; for zebrafish, see Molven et al. 1991). Whereas \(\text{Wnt-1}\) shows a similar clear cut border at the furrow as \(\text{paxzf-b}\), the expression of \(\text{en-2}\) shares
this border only ventrally and extends dorsally to the cerebellum. Zebrafish pax\(zf-b\) is therefore the only transcription factor known thus far that shows a posterior border that coincides with the midbrain–hindbrain border.

**Expression in the optic stalk, the otic vesicle and the nephric primordium**

The expression of pax\(zf-b\) in the optic stalk, the otic vesicle and the nephric primordium is comparable to the expression reported for the presumptive murine homolog Pax2. In parallel to what has been observed for homeoproteins, it seems that the expression patterns of genes encoding paired box-containing proteins are maintained largely unaltered during vertebrate evolution. This leads to the suggestion that the Pax2/pax\(zf-b\) gene product performs fundamental and probably highly conserved functions during vertebrate embryogenesis.

We would like to thank Kai Zinn for providing the cDNA library and Peter Gruss for providing the murine\(Pax\)1 probe. We are grateful to John Kuwada, Monte Westerfield and Charles Kimmel for their helpful discussions. We like to thank Stephen Wilson for critical reading of the manuscript. We are grateful to Inna Korzh and Turid Holm for excellent technical assistance, to Nils Peder Willansen for help with Fig. 1. S.K. and V.K are fellows of the Aakre Cancer Foundation and the Nansen Foundation. Financial support was also obtained from the Norwegian Cancer Society, Hasselblad Foundation and Nansen Foundation.

**References**


SEDMAN, S. A., GELEMBIUK, G. W. AND MERTZ, J. E. (1990). Translation initiation at a downstream AUG occurs with increased efficiency when the upstream AUG is located close to the 5' Cap. J. Virol. 64, 453–457.


(Accepted 6 September 1991)