The EMBO Journal vol.10 no.12 pp.3609-3619, 1991

Zebrafish *pax[zf-a]*: a paired box-containing gene expressed in the neural tube

Stefan Krauss, Terje Johansen[1], Vladimir Korzh, Ugo Moens[2], Johanna U.Ericson and Anders Fjose[3]

Molecular Genetics Group, [1]Biotechnology Group and [2]Virology Group, Institute of Medical Biology, University of Tromsø, N-9000 Tromsø, Norway

[3]Corresponding author

Communicated by M.Noll

Murine and human sequences homologous to the paired box of the *Drosophila* segmentation gene *paired* have been reported previously. Here we describe a zebrafish (*Brachydanio rerio*) paired box-containing clone, *pax[zf-a]*, which is clearly distinct from reported vertebrate *Pax* genes. The putative protein encoded by *pax[zf-a]* contains a paired box and a paired-type homeobox separated by a glycine-rich, acidic linker and a carboxy-terminal end which is remarkably rich in serine, threonine and proline residues. By *in situ* hybridization to embryonic tissue sections and whole mount embryos, *pax[zf-a]* transcripts were found within restricted regions of the central nervous system and the eye. In contrast to the murine *Pax* genes recently characterized, *pax[zf-a]* is not expressed in the segmented mesoderm. At the 17 h stage, *pax[zf-a]* expression is detected in a defined area of the diencephalon which circumscribes the presumptive thalamus. This suggests an involvement of *pax[zf-a]* in pattern formation in the rostral brain. The *pax[zf-a]* gene is also expressed throughout the hindbrain and spinal cord. This hybridization signal is restricted to a longitudinal column which includes the basal plate. Later in development, at 36 h post-fertilization, *pax[zf-a]* transcripts are no longer restricted to a specific region of the diencephalon, but are distributed over the entire developing brain.

Key words: eye development/neurogenesis/paired box gene/segmentation/zebrafish embryogenesis

Introduction

The homeotic selector genes included in the regulatory network responsible for specifying regional differences along the body axis of vertebrates and insects are highly conserved (Gehring, 1987; Graham *et al.*, 1989; Malicki *et al.*, 1990). These genes encode closely related transcription factors and are organized in gene complexes which seem to be derived from a similar complex already present in the common ancestor of insects and vertebrates (Graham *et al.*, 1989). In *Drosophila* the molecular mechanisms responsible for generating body segments have been extensively characterized (Ingham, 1988; Carroll, 1990), but it is presently unclear to what extent similar genetic elements are involved in vertebrate pattern formation.

Many of the *Drosophila* developmental regulatory genes encode transcription factors with conserved motifs like the homeobox (McGinnis *et al.*, 1984; Scott *et al.*, 1989) or the paired box (Bopp *et al.*, 1986), which are both DNA binding domains (Goulding *et al.*, 1991; Treisman *et al.*, 1991). Several paired box-containing genes in *Drosophila* are known to be involved in the segmentation process (Bopp *et al.*, 1986; Baumgartner *et al.*, 1987; Bopp *et al.*, 1989). In addition, some of the members of the paired box gene family seem to have important regulatory functions during neurogenesis at later developmental stages (Patel *et al.*, 1989).

In relation to the different segmental features of vertebrate and insect embryos, it is of considerable interest to investigate the structural and functional conservation of the different categories of segmentation genes. Among the vertebrate homologues of *Drosophila* genes analysed so far, the paired box-containing *Pax* genes seem to be most closely linked to segmentation (Kessel and Gruss, 1990). Thus, all of the *Pax* genes which have been studied in mouse (*Pax1, Pax2, Pax3, Pax7* and *Pax8*) show expression in segmented mesodermal tissues during embryogenesis (Balling *et al.*, 1988; Deutsch *et al.*, 1988; Dressler *et al.*, 1990; Nornes *et al.*, 1990; Plachov *et al.*, 1990; Jostes *et al.*, 1991; Goulding *et al.*, 1991). In addition, four of the murine *Pax* genes are expressed in cell subsets of the developing nervous system, indicating an involvement in neuronal maturation or determination. Similar paired box-containing genes have been reported for humans, and there is evidence for homologues in a variety of other species (Burri *et al.*, 1989).

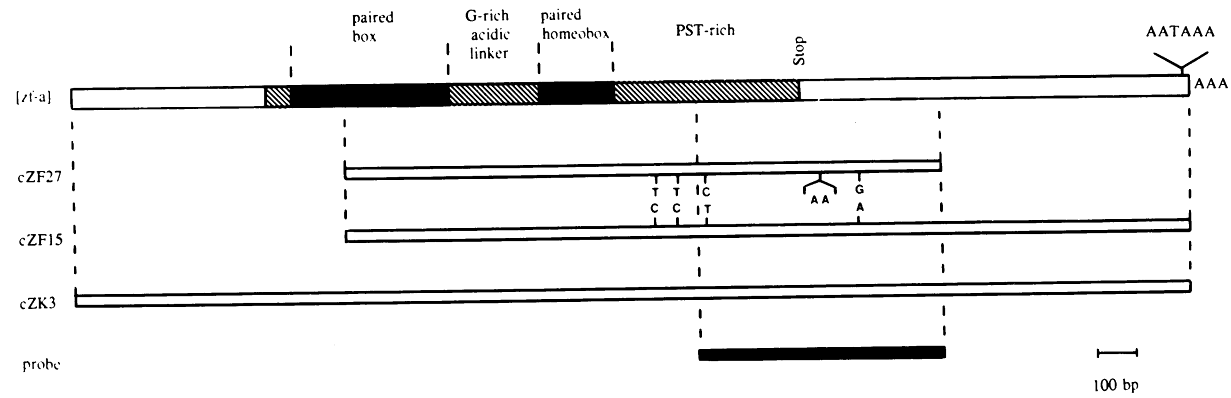
To gain further understanding about pattern formation and neuronal differentiation in the vertebrate central nervous system, it may be useful to study these aspects in a lower vertebrate, such as zebrafish (*Brachydanio rerio*), where morphological features are more simple and easier to distinguish (see Hanneman *et al.*, 1988; Trevarrow *et al.*, 1990). In a search for zebrafish genes involved in neural development, we have cloned a paired box-containing gene, *pax[zf-a]*, which in contrast to all the murine *Pax* genes described so far, is not expressed in segmented tissues of the mesoderm. Instead, transcripts of the gene are detected within a longitudinal column of the hindbrain and spinal cord which maps to the basal plate and intermediate plate. In addition, *pax[zf-a]* seems to be involved in early regionalization of the rostral brain.

Results

Structure of pax[zf-a] cDNA clones

In order to isolate zebrafish cDNA clones containing paired box sequences, a cDNA library in λ*gt11* prepared from embryos at the 33 h stage was screened by plaque hybridization at low stringency using both murine *Pax1* (Deutsch *et al.*, 1988) and zebrafish *pax[zf-d]* (H.G.Eiken, P.Villand and A.Fjose, unpublished) paired box probes.

A



B

```
CACGAGCAACACGGTTAACGTCCGAGGTTTGCGTGTCTCGTAAGAAAGAAGAGGGAGAGTCCAATATTCGACCTGATTTCATCTGTTAAACAGGCTAGTATTTTCGTCGACTTCACAAAG 120
               ***
GGGTTTGCAATCTCTCACAACCTTTTCCGAGGTTCTCTTTTCGAGGTTCCCTTGTTGGACTGGATATTATTTACTTTGGGACTAGTTTTGTGATTCGGATCCGGAGGGGCTCCATAGACA 240
                                                                                          ***                       ***
CTCATTACTCTTCACTTGAAACTTGGAACCGTGCGTCTCATAACGAAATCCATTACGAATGTTTTGCTATCGGAACGGGTTTATTGAGGGAATTTAAATTTTCCACCCGAGATCAGTTGG 360
                                                                                     ***
AAACTATAAAAGACAAACTGTTTGAGGACGATTTTCTAGTACTCCAGAGTCTTCTCGTTATTGTAAACGAAGAACTTCAGCGAGGATACAAAGGCTGTTGGAACTATGCCTCAAAAAGAA 480
      ***                          ***                        ***                                             M  P  Q  K  E     5
                                                                               Paired box

TACTATAACCGGGCCACGTGGGAGTCTGGTGTCGCGTCCATGATGCAAAACAGTCACAGTGGAGTGAACCAGCTCGGCGGTGTGTTCGTCAACGGCAGACCGCTACCCGACTCCACGAGA 600
 Y  Y  N  R  A  T  W  E  S  G  V  A  S  M  M  Q  N  S  H  S  G  V  N  Q  L  G  G  V  F  V  N  G  R  P  L  P  D  S  T  R    45
CAGAAAATAGTTGAACTCGCACACAGTGGCGCGCGGCCGTGTGACATATCAAGAATTCTGCAGGTGTCCAACGGCTGCGTGAGTAAAATCTTGGGTAGATACTATGAAACAGGCTCCATC 720
 Q  K  I  V  E  L  A  H  S  G  A  R  P  C  D  I  S  R  I  L  Q  V  S  N  G  C  V  S  K  I  L  G  R  Y  Y  E  T  G  S  I    85
AGACCCAGGGCGATCGGAGGAAGTAAACCACGAGTAGCGACTCCCGAGGTGGTCGGCAAAATTGCCCAGTACAAGAGGGAGTGTCCGTCAATCTTCGCGTGGGAAATCCGAGACAGGCTG 840
 R  P  R  A  I  G  G  S  K  P  R  V  A  T  P  E  V  V  G  K  I  A  Q  Y  K  R  E  C  P  S  I  F  A  W  E  I  R  D  R  L   125
CTATCAGAGGGGGTCTGCACAAACGATAATATACCCAGTGTGTCATCGATAAACAGAGTACTGCGCAACCTGGCTAGCGAAAAGCAACAGATGGGCGCAGATGGCATGTATGAAAAGCTG 960
 L  S  E  G  V  C  T  N  D  N  I  P  S  V  S  S  I  N  R  V  L  R  N  L  A  S  E  K  Q  Q  M  G  A  D  G  M  Y  E  K  L   165
AGGATGCTGAACGGTCAGACCGGCACGTGGGGACCCGGCCGGGCTGGTACCCCGGAACCTCGGTGCCAGGACAGCCCAATCAAGATGGTTGCCAACAGTCAGACGGAGGCGGTGAGAAC 1080
 R  M  L  N  G  Q  T  G  T  W  G  T  R  P  G  W  Y  P  G  T  S  V  P  G  Q  P  N  Q  D  G  C  Q  Q  S  D  G  G  G  E  N   205
                                                                                  Paired homeobox

ACAAACTCAATAAGCTCCAATGGCGAGGACTCAGATGAGACCCAAATGAGGCTTCAGCTTAAACGAAAACTGCAAAGGAATCGCACTTCTTTCACACAAGAACAAATAGAAGCACTTGAA 1200
 T  N  S  I  S  S  N  G  E  D  S  D  E  T  Q  M  R  L  Q  L  K  R  K  L  Q  R  N  R  T  S  F  T  Q  E  Q  I  E  A  L  E   245
AAAGAGTTTGAAAGAACGCACTACCCTGACGTTTTTGCACGAGAAAGACTTGCTGCAAAAATAGATTTACCAGAAGCAAGAATACAGGTCTGGTTCTCAAACAGAAGAGCGAAATGGAGG 1320
 K  E  F  E  R  T  H  Y  P  D  V  F  A  R  E  R  L  A  A  K  I  D  L  P  E  A  R  I  Q  V  W  F  S  N  R  R  A  K  W  R   285
AGGGAGGAAAAGTTAAGAAATCAAAGAAGACAAGCCAGTAATTCCTCAAGTCACATACCCATCAGCAGCAGCTTCAGCACAAGCGTCTATCAACCAATCCCTCAGCCCACAACGCCAGTA 1440
 R  E  E  K  L  R  N  Q  R  R  Q  A  S  N  S  S  S  H  I  P  I  S  S  S  F  S  T  S  V  Y  Q  P  I  P  Q  P  T  T  P  V   325
TCCTTCACGTCAGGCTCCATGTTGGGCAGATCAGACACAGCTCTTACGAACACATACAGCGCCCTGCCACCAATGCCAAGCTTTACCATGGCCAACAACCTTCCTATGCAACCCAGCCAG 1560
 S  F  T  S  G  S  M  L  G  R  S  D  T  A  L  T  N  T  Y  S  A  L  P  P  M  P  S  F  T  M  A  N  N  L  P  M  Q  P  S  Q   365
ACCTCATCCTACTCCTGCATGTTGCCCACTAGTCCTTCAGTAAACGGGCGGAGCTATGACACATACACACCCCCGCACATGCAGGCGCATATGAACAGCCAATCAATGGCCGCCTCGGGC 1680
 T  S  S  Y  S  C  M  L  P  T  S  P  S  V  N  G  R  S  Y  D  T  Y  T  P  P  H  M  Q  A  H  M  N  S  Q  S  M  A  A  S  G   405
ACAACCTCAACGGGTTTAATCTCACCTGGAGTGTCTGTACCCGTTCAAGTGCCAGGCAGTGAACCAGACATGTCCCAATACTGGCCCAGACTACAGTGAGAACGCGACGCACAAGAACAA 1800
 T  T  S  T  G  L  I  S  P  G  V  S  V  P  V  Q  V  P  G  S  E  P  D  M  S  Q  Y  W  P  R  L  Q  End                       437
AAAAGGAAATCAGAGGAGAGAAAAAAAAAAAAGAGAAGCTCCTTCACCTCCTGATGTCTCTCGGCTACAAGACAGGGGTGTTCAGCAGTATTTCACCACAGAAGGAAAGAAGGAGACTCT 1920
AAAGGACCTCTTTTTTTGTACGGGGATAGTGCCACTTCTCTATCATTCTTTGGACACAAAGACTTGAAGAATAAAAGAGAACAGACTTCTGTAAAGTGCCCGGTATTATATCGTAAAAAA 2040
ATCTGTTTTTCAGTTTCCAACCTAAGTCATTTTGATGTATTTGTACATGTAATGGCCAATTGTATGTTATGACAAAAAAAGGAAAAAAAAATTTTTTTTTTTTGAACAACCGTGGATGGA 2160
CTGTCAGTAGACCATATTGATGACTCATCCTGCACTGCAAAGATTTTTATCCAATCAGACGCTCTCTAGAACGAGAATGTCCTGGCCACTGAAAACACCTGCAGCTGACTATCAAATTGC 2280
TGAATACATTTGGCTTTTATAAGACCAAACTAAAAAAGGAAAAGAAACAACATTGTAATTTGGCCTTGAATCGGTTAATGGAGGAACATTTGTGCAGGTTTTACTCATCATTTTCTCATT 2400
TGCCGCTTATATCAAGAACTTCTGCCCATTTCTGTTTCAAGCGTGGACCTGTAGTAACAAAACGATCTCTGTAATCTTTTCGAGATGAACAAACACAGCAACAGAGTTTCAGAGGAAGCC 2520
AGTCCAGAATTAGCATTTTTTTTTCAAAGCTTGCAGGTGAAAGTGAGCTGTTTTTGTTGCTGAACCGGAGGAGGAAAATGTTTGGTAGCTTAAAAACGGTAGATTGTGTCTTCGATATAA 2640
TTCAGTTTTGTTTATGTCAAAATGTAAGTATTTGTCTTCCCTAGAAGTCTCGCAGAACAATTTCTATAATAAAATTAATTCCATTTTAAAAAAAAAAA 2738
```

Among the positive clones obtained, three (cZF1, cZF15 and cZF27) were shown by restriction analysis to be derived from the same mRNA. The coding gene was designated *pax*[*zf-a*]. Sequencing of both strands of cZF15 and cZF27 showed that the cDNAs contained a long open reading frame (ORF) with regions clearly homologous to both the paired box and the paired homeobox (Bopp *et al.*, 1986; Frigerio *et al.*, 1986). However, the cDNA clones are truncated at the 5′ end in the paired box due to an internal *Eco*RI site left unprotected during the construction of the cDNA library. We obtained a clone, cZK3, harbouring the missing 5′ end by screening another 20−28 h cDNA library in λZAP-II using a 234 bp *Eco*RI−*Cla*I 5′ fragment from cZF27 as a probe. cZK3 contains a 652 bp 5′ extension from the *Eco*RI site. The composite sequence of cZK3 and cZF15, designated *pax*[*zf-a*] cDNA, is shown in Figure 1. The *pax*[*zf-a*] cDNA is 2738 bp long and contains a consensus polyadenylation signal (Proudfoot and Brownlee, 1976; Fitzgerald and Shenk, 1981) 14 nucleotides upstream of an 11 nucleotide long poly(A) tail. An ORF capable of encoding a putative protein of 437 amino acids (mol. wt = 48 425) including both a paired box and a paired homeobox is located at positions 466−1776 (see Figures 1 and 2). There are seven in-frame stop codons upstream of the first ATG of the ORF and two additional methionine codons just upstream of the paired box (see Figure 1B). Based on criteria for optimal sequence context for efficient initiation of translation (Kozak, 1986, 1987) and length of the N-terminal sequence upstream of the paired box compared with that found for other paired box-containing genes (Dressler *et al.*, 1990 and references therein) we favour the first ATG as the start codon. As seen from the Northern blot in Figure 3, the size of the *pax*[*zf-a*] transcript is ~3.0 kb. Assuming a poly(A) tract of ~200 bases, this suggests that the *pax*[*zf-a*] cDNA sequence shown in Figure 1 is close to full length.

Comparison of the derived amino acid sequence of pax[zf-a] with that of other known paired box-containing genes

As shown in Figure 1, the *pax*[*zf-a*] cDNA encodes a 22 amino acid N-terminus, the 125 amino acid paired box, a 'linker' of 80 residues, the 61 residue long paired homeobox and a C-terminal region of 149 amino acids. In Figure 2 the paired box and the paired homeobox of *pax*[*zf-a*] are aligned with all published paired box- and paired homeobox-containing sequences. The extent of the paired box is defined according to Bopp *et al.* (1986) and the extent of the homeobox is as defined by Scott *et al.* (1989). Since there is no strong sequence conservation between all sequences beyond position 126 (see Figure 2A), we propose that the C-terminal end of the paired box should be defined at the conserved arginine residue at this position. *Pax*[*zf-a*] clearly represents a novel *Pax* gene which, based on paired box homology, is most related to the murine paired box genes *Pax8* and *Pax2*. *Pax*[*zf-a*] differs at four positions from the consensus where all the 12 other published paired box sequences are completely conserved, thus reducing the number of absolutely conserved positions to 58 out of 126 (46%).

Pax[*zf-a*] contains the most divergent paired homeobox of the hitherto published sequences with only 34 of the 61 residues conserved. From sequence comparisons with all other classes of homeoboxes it is clear that *pax*[*zf-a*] nevertheless contains a paired-type homeobox (see Figure 2B). The critical residues for positioning in the major groove (Trp49, Phe50, Asn52 and Arg54) and base-specific contacts in the recognition helix (Val48 and Ser51) (helix III in Figure 2B; Kissinger *et al.*, 1990) are conserved. For the *Drosophila* genes *prd*, *gsb-p* and *gsb-d* the homology also extends 17 amino acids upstream of the paired homeobox (Bopp *et al.*, 1986) and for the murine *Pax7* and *Pax3* genes it is extended at least 10 amino acids beyond the amino end of the 61 amino acids homeobox (Jostes *et al.*, 1991) which is similar to the situation with zebrafish *pax*[*zf-a*]. For all five genes the three amino acids immediately upstream of the homeobox (Leu-Lys-Arg) are completely conserved, supporting the notion of an N-terminally extended paired-type homeobox (Bopp *et al.*, 1986). One possible reason for the absolute conservation of the three amino acids preceding the homeobox (as defined by Scott *et al.*, 1989) is that they may form part of a nuclear localization signal; LKRKnnRnR (for a review see Roberts, 1989).

The linker region between the paired box and the homeobox is rather long in *pax*[*zf-a*] (80 amino acids) compared with the *Drosophila* and murine genes (40−60 amino acids). Interestingly, upon comparison of the amino acid composition and charge distribution, scored as theoretical isoelectric point (pI) when all side chains are solvent accessible, we find that all five sequences contain a glycine-rich (15−20%), acidic linker region with theoretical pIs varying from 3.9 to 4.6. Glycine is frequently found in hinge regions and is often used in protein structures to facilitate movement (Richardson and Richardson, 1989). This suggests that the basic paired box and homeobox domains are separated by a flexible, acidic linker (see Discussion).

Burri *et al.* (1989) described a conserved octapeptide sequence (5/8−8/8 identities) downstream of the paired box in *gsb-p*, *gsb-d*, and the human *HuP1* and *HuP2* genes. The octapeptide is also present at similar positions in the murine *Pax2*, *Pax7*, *Pax8* and *Pax3* genes (Dressler *et al.*, 1990; Plachov *et al.*, 1990; Jostes *et al.*, 1991; Goulding *et al.*, 1991). However, similar to *Drosophila prd*, no such sequence is found in the linker region of zebrafish *pax*[*zf-a*].

Fig. 1. (**A**) Structure of *pax*[*zf-a*] cDNA. The extent of the cDNA clones cZK3, cZF15 and cZF27 and the organization of *pax*[*zf-a*] cDNA into a coding region (hatched and solid boxes) which can be divided into five domains [N-terminal region, paired box, glycine-rich, acidic linker, homeobox and C-terminal rich in proline, serine and threonine (PST-rich)] and a 3′ untranslated sequence containing a polyadenylation signal (AATAAA) and a poly(A) tail (AAA_n) are shown. Four silent nucleotide substitutions between cZF15 and cZF27 and a 2 bp insertion in cZF27, are indicated. The location of the *Hin*dIII−*Eco*RI fragment from cZF27 used as a probe for *in situ* hybridizations is shown. (**B**) The nucleotide and derived amino acid sequences of *pax*[*zf-a*] cDNA (from cZF3 and cZF15) are shown. Glycine residues in the linker and the polyadenylation signal are underlined. Proline, serine and threonine residues in the C-terminal region are indicated by double underlining. The 3′ end of cZF27 at position 2113 is indicated (underlined). The *pax*[*zf-a*] sequence has been assigned the accession number X61389 in the EMBL Data Library.

The 149 amino acids C-terminal region of *pax[zf-a]* is particularly rich in serine residues (20.1%). Proline (12.8%) and threonine residues (10.1%) are also over-represented in this region (see Figure 1). Despite the high number of serine and threonine residues, there are no potential phosphorylation sites for protein kinase C, cAMP- and cGMP-dependent protein kinases. However, the serine at 361 in the derived amino acid sequence may represent a phosphorylation site for casein kinase II (Krebs *et al.*, 1988; Kemp and Pearson, 1990).

Expression of pax[zf-a] during embryogenesis

To test if *pax[zf-a]* is expressed during embryogenesis, Northern blot analysis was done on poly(A)$^+$ RNA derived from different embryonal stages from 3 h after fertilization to the time of hatching. The cZF27 probe (see Figure 1A)

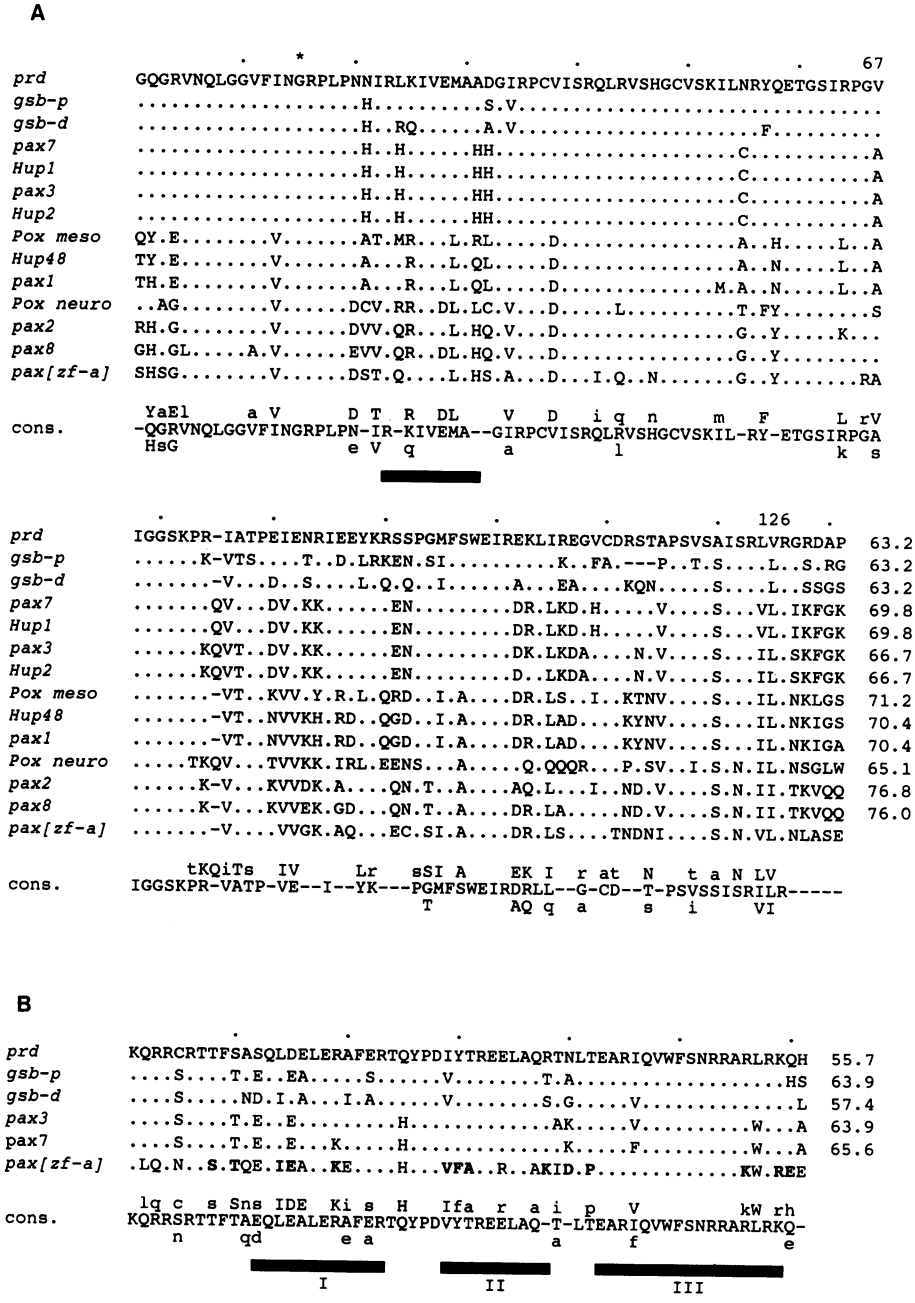


Fig. 2. (**A**) Comparison of the paired box sequence of *pax[zf-a]* with known paired box sequences from *Drosophila* (Bopp *et al.*, 1986; Baumgartner *et al.*, 1987; Bopp *et al.*, 1989), mouse (Deutsch *et al.*, 1988; Dressler *et al.*, 1990; Plachov *et al.*, 1990; Jostes *et al.*, 1991; Goulding *et al.*, 1991) and humans (Burri *et al.*, 1989). Amino acids identical to those at corresponding positions of the *prd* sequence are indicated by dots. Gaps introduced are represented by dashes. In the consensus line positions varying by more than three amino acids are indicated by dashes whereas amino acids found in only one sequence are shown in small letters. The α-helix proposed by Burri *et al.* (1989) and shown by Treisman *et al.* (1991) to be essential for DNA binding, is indicated by the bar below the consensus line. The percentage identity to *pax[zf-a]* (calculated from 126 aligned positions) is shown to the right of each sequence. The location of the *undulated* mutation is indicated by a star. (**B**) Comparison of the known paired homeobox sequences with *pax[zf-a]*. The extent of the three α-helices determined from the *engrailed* homeodomain crystal structure (Kissinger *et al.*, 1990) is shown below the consensus line. Percentage identity to *pax[zf-a]* is shown to the right of each sequence. Changes to chemically similar amino acids in *pax[zf-a]* compared with *prd* are shown in bold. If these changes are considered, the *pax[zf-a]* homeobox is 78.7% and 80.3% similar, respectively, to the *Drosophila* and mouse sequences shown.

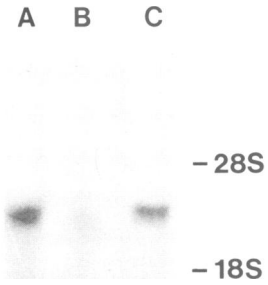


Fig. 3. Expression of *pax[zf-a]* in total embryonic tissue. Northern blot analysis of poly(A)$^+$ RNA derived from (**A**) zebrafish larvae shortly after hatching, (**B**) embryos after 33 h of development and (**C**) embryos after 24 h of development. For high stringency hybridization the cZF27 fragment shown in Figure 1A was used as a probe. The blot was exposed for 6 days.

hybridized to a 3.0 kb transcript which was detected in RNA of 24 and 36 h embryos as well as in RNA derived from zebrafish larvae shortly after hatching (Figure 3). However, no *pax[zf-a]* expression was apparent at 3 and 10 h of development (data not shown). These results are entirely consistent with the data obtained from the *in situ* hybridization experiments. *In situ* hybridizations were performed on tissue sections and whole mount embryos of different developmental stages. Using a 600 bp *Hin*dIII−*Eco*RI cDNA fragment as a probe (see Figure 1A), transcripts of the *pax[zf-a]* gene are first detected in two separate regions of the neural tube at 12 h of development (Figure 5a and b). While a relatively high expression level is seen in the diencephalon, only faint hybridization signals are detected in the hindbrain region. At 17 h of development the expression is substantially increased in both regions (Figure 5c and d). A rather intense hybridization signal is now seen in a clearly circumscribed area of the diencephalon and the other expression domain extends from a sharp border

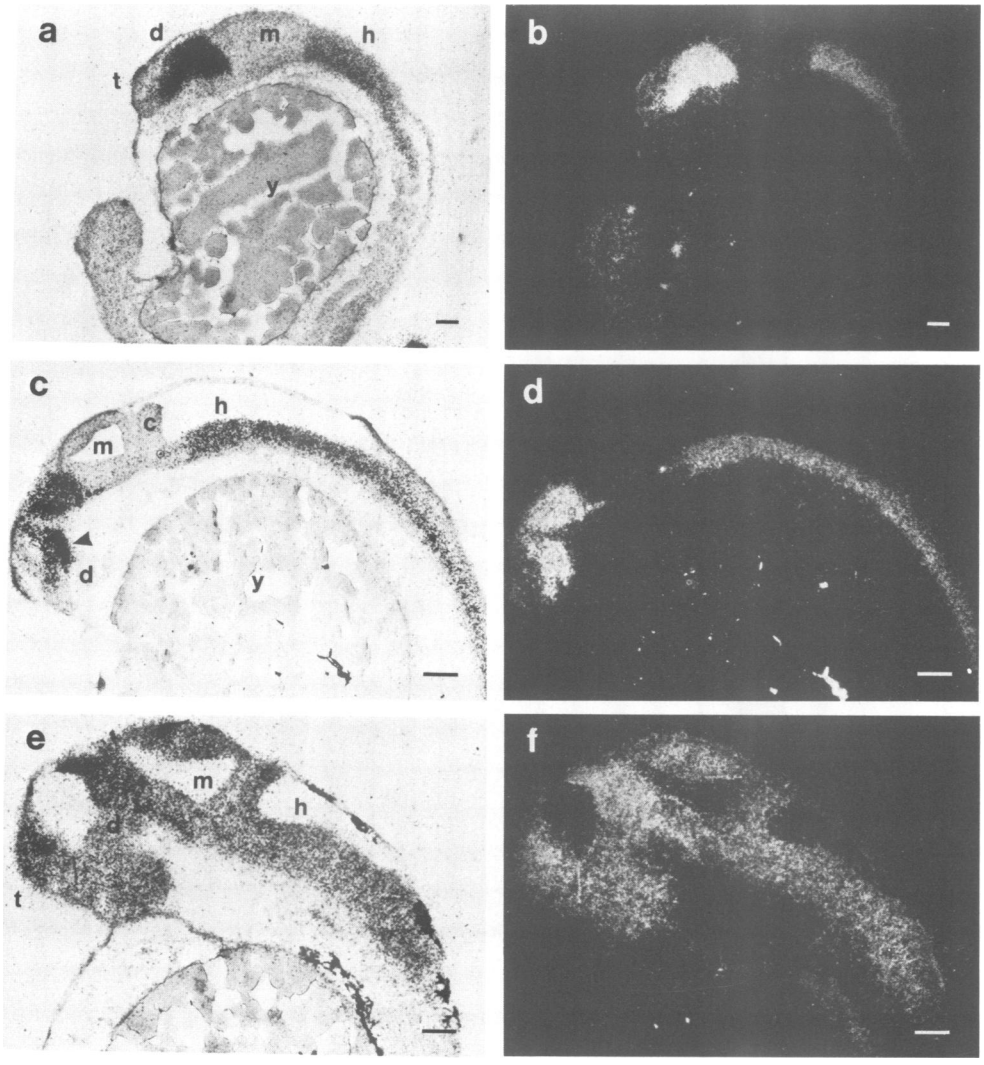


Fig. 4. Localization of *pax[zf-a]* transcripts in tissue sections of zebrafish embryos of different developmental stages by *in situ* hybridization. Sagittal sections are shown for embryos after 17 h (**a** and **b**), 24 h (**c** and **d**) and 36 h (**e** and **f**) of development. Brightfield and darkfield images of each section are shown side by side. The embryos are oriented with their anterior end to the left. Note the circumscribed area expressing *pax[zf-a]* in the diencephalon of the 17 h embryo (a and b). Note also the region of stronger expression in the diencephalon of the 24 h embryo which is marked by an arrowhead in (c). Due to the section level, the region of *pax[zf-a]* expression in the telencephalon at this stage cannot be seen (for comparison see Figure 5l and o). Abbreviations: c, cerebellum; d, diencephalon; h, hindbrain; m, midbrain; t, telencephalon; y, yolk. Bars, 50 μm.

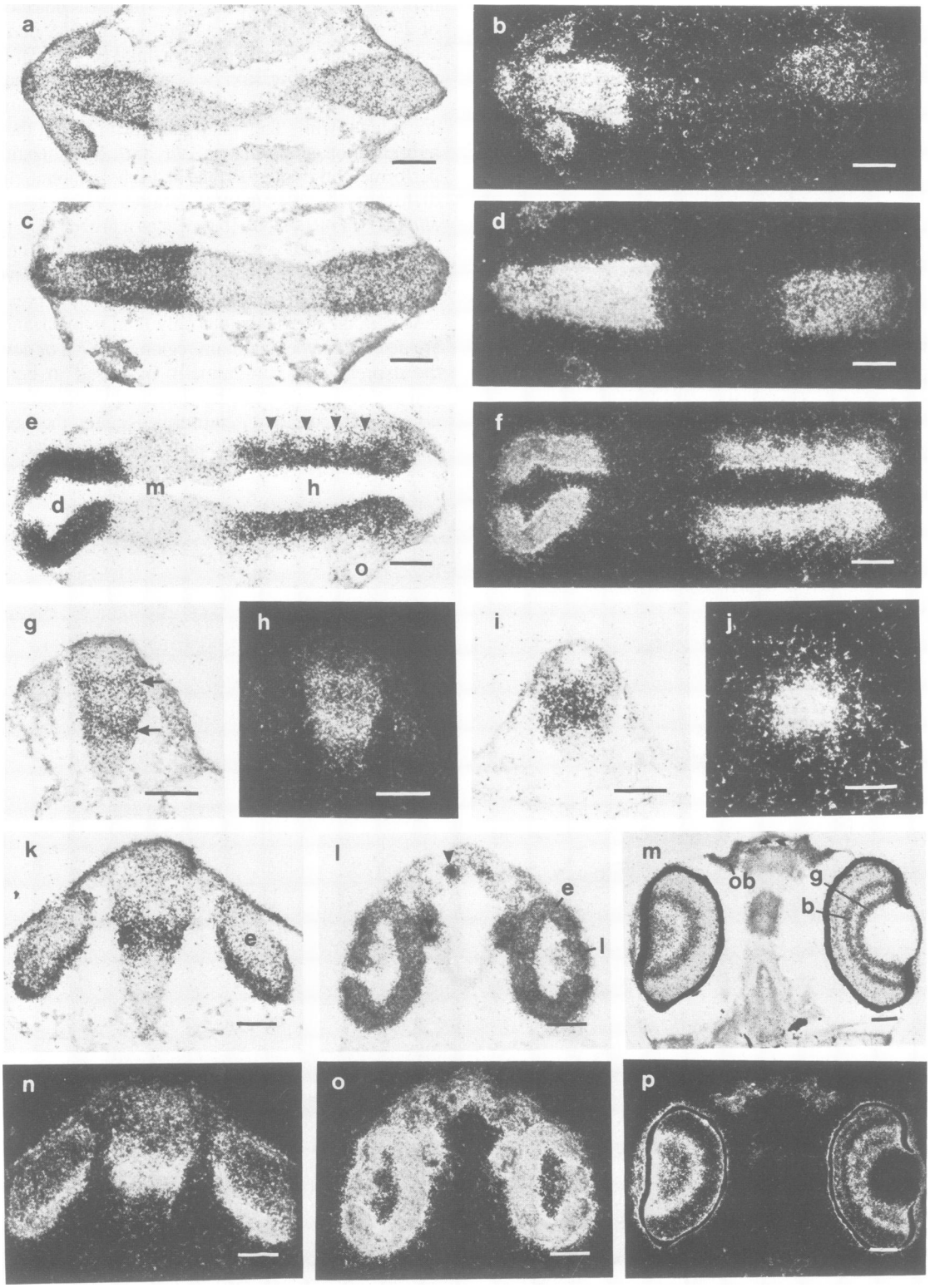


Fig. 5. Distribution of *pax[zf-a]* transcripts in horizontal sections of embryos at 12 h (**a** and **b**), 17 h (**c** and **d**) and 24 h (**e** and **f**) of development. (**g**–**j**) Details of expression in the spinal cord of 17 h (g and h) and 24 h (i and j). (**k**–**p**) Details of the *pax[zf-a]* expression are also shown for the eye at 17 h (k and n), 24 h (l and o) and one week after hatching (m and p). The embryos are oriented with their anterior end towards the left (a–f) or the top (k–p). Note the arrowheads in (e) indicating the rhombomeres in the hindbrain. The arrowheads in (l) show the areas of increased signal intensity in the diencephalon and telencephalon at 24 h. These areas are not seen at earlier stages and seem to be located at the periphery rather than the lumen of the neural tube. The two arrows in (g) indicate the two horizontal stripes of increased labelling in the posterior hindbrain at 17 h. Abbreviations: b, bipolar nerve layer; d, diencephalon; e, eye; g, ganglion cell layer; h, hindbrain; l, lens; m, midbrain; o, otic vesicle; ob, olfactory bulb; t, telencephalon. Bars, 50 μm.

in the anterior hindbrain being progressively reduced towards the spinal cord (Figure 4a and b). In the following stages of development major changes in the expression pattern of *pax*[*zf-a*] are mainly observed in the rostral brain where the signal intensity becomes less uniform and expression is detected in additional areas (Figure 4c–f). Thus, by 36 h post-fertilization, *pax*[*zf-a*] expression is no longer restricted to specific regions of the rostral brain (Figure 4e and f).

Expression of pax[zf-a] in the diencephalon

Transcripts derived from the *pax*[*zf-a*] gene are first detected within the rostral portions of the neural tube in 12 h embryos (Figure 5a and b). This *pax*[*zf-a*] expressing region extends anteriorly to the edge of the optic stalk, while posteriorly it delineates a clear cut border in the middle of the rostral brain. At this time no morphological landmarks are visible in the rostral portion of the neural tube. A very similar rostral expression pattern was revealed in whole mount *in situ* hybridization experiments on 15 h embryos (see Figure 6a and b). The rostral component of the *pax*[*zf-a*] expression is most clearly defined in 17 h embryos (Figure 4a and b and Figure 5c and d; see also Figure 7a). At this stage, it is limited caudally by a sharp, transverse border in the region between diencephalon and midbrain. Ventrally, the expression border has the shape of a convex line that coincides roughly with the boundary separating the presumptive thalamic and hypothalamic areas of the diencephalon (Figure 4a and b). However, since clear morphological landmarks are not yet visible in this region, we are not able to specify the location of this expression border further. Rostrally, the ventral border extends to the level of the optic stalk, and an additional restriction of *pax*[*zf-a*] expression is seen in the dorsal part of the rostral brain where transcripts are not detected in the presumptive telencephalon (Figure 4a and b). The rather uniform signal intensity observed in the rostral brain at the 15–17 h stage, is a transient phenomenon. Two additional strongly labelled areas appear at each side of the third ventricle which has been formed in the 24 h embryo. Whereas one symmetrical pair is located at the ventral border of the rostral expression domain described for the earlier stages (Figure 4c and d), the other pair can be seen in the central part of the developing forebrain (Figure 5l and o). These two areas seem to map to the regions of the neural tube underlying the eyes and the olfactory bulbs respectively (Figure 5l and o). In 36 h embryos, hybridization signals are detected in additional regions of the forebrain and midbrain (Figure 4e and f). Thus, at this developmental stage, a relatively uniform level of *pax*[*zf-a*] transcripts appears to be present throughout the entire prechordal division. Further modulation of the rostral expression pattern occurs later in development and one week after hatching of the zebrafish larva, *pax*[*zf-a*] transcripts are detected only in specific groups of cells at different positions in the brain such as the anterior part of the olfactory bulb and in the eye (Figure 5m and p; data for further brain regions not shown).

Expression of pax[zf-a] in the hindbrain and the spinal cord

In the hindbrain and spinal cord, the spatial distribution of *pax*[*zf-a*] transcripts is clearly different from the expression pattern observed in the diencephalon. The intensity of the

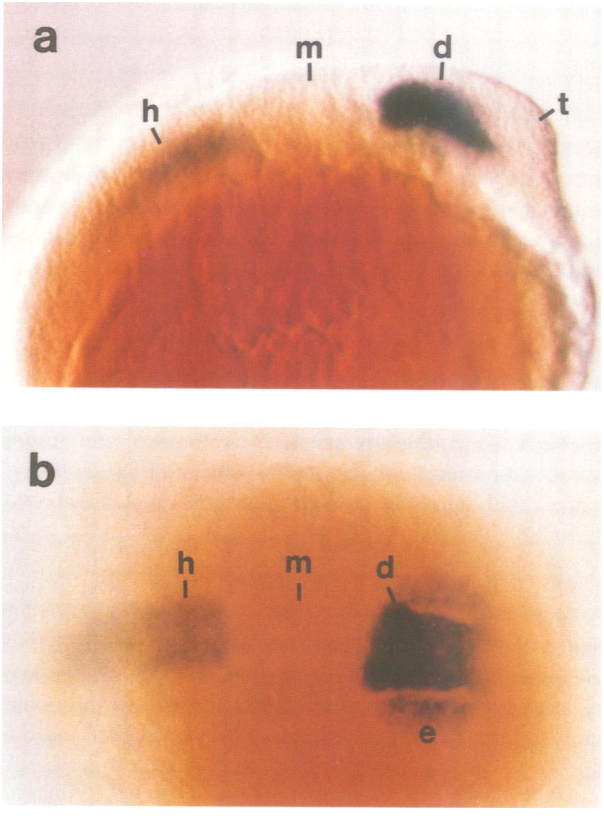


Fig. 6. *In situ* hybridization of *pax*[*zf-a*] on whole mount embryos at 15 h of development. A lateral (**a**) and a dorsal (**b**) view of the embryo are shown. The embryos are oriented with the anterior end to the right. Abbreviations: d, diencephalon; e, eye; h, hindbrain; m, midbrain; t, telencephalon.

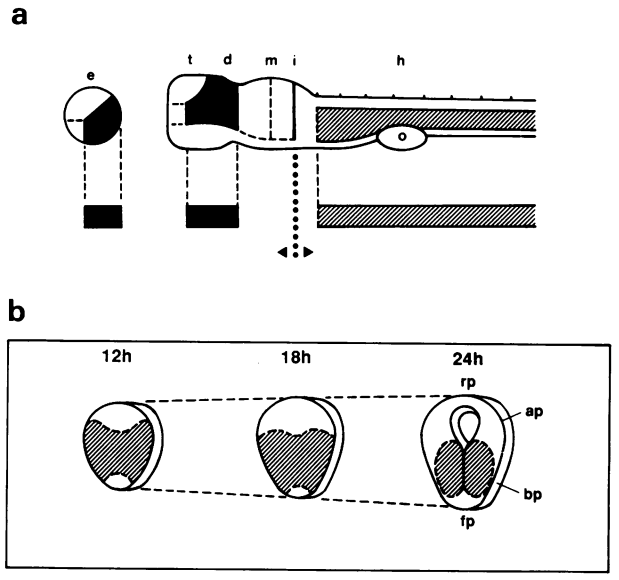


Fig. 7. Schematic illustrations showing *pax*[*zf-a*] expression in the neural tube. (**a**) Simplified scheme of the neural tube at 17 h demonstrating how the *pax*[*zf-a*] expression correlates with surrounding morphological features; (**b**) schematic illustration of *pax*[*zf-a*] expression in cross-sections of the posterior hindbrain at different developmental stages. Abbreviations: ap, alar plate; bp, basal plate; d, diencephalon; e, eye; fp, floor plate; h, hindbrain; i, isthmus; m, midbrain; o, otic vesicle; rp, roof plate; t, telencephalon.

hybridization signal is considerably weaker and the region in which the gene is transcribed extends caudally from a border in the anterior hindbrain. As seen from Figure 4c and d and Figure 5e and f, at 24 h of development the anterior limit of expression seems to map near the border between hindbrain rhombomeres ro1 and ro2. We cannot say exactly when this expression is initiated since 12 h embryos were the earliest stage analysed by *in situ* hybridization (Figure 5a and b). However, by this developmental stage, weak hybridization signals are detected in an area of the hindbrain which in the caudal direction becomes progressively more restricted. A similar cone shaped expression domain was also detected in whole mount *in situ* hybridization experiments on 15 h embryos (Figure 6). At this embryonic stage, the hybridization signal extends somewhat further posterior in the spinal cord. This extension seems to be complete at 24 h of development, when the posterior component of the *pax[zf-a]* expression has reached the tail of the embryo (not shown). At the same developmental stage, horizontal sections of the hindbrain reveal that *pax[zf-a]* transcripts are distributed in a wave-like pattern toward the luminal margin of the neural tube reflecting the segmental organization of rhombomeres (see arrowheads in Figure 5e and f). In cross sections derived from a region in the posterior hindbrain of 17 h embryos, *pax[zf-a]* transcripts are mainly detected in the ventral portion of the neural tube. Within this area, two zones of higher signal intensity are observed at medial and ventral levels, respectively (see arrows in Figure 5g and Figure 7b). However, in the most ventral region where the floor plate is located, the *pax[zf-a]* gene is not expressed. At 24 h, *pax[zf-a]* transcripts are exclusively detected within a ventral longitudinal column which includes the basal plate and most probably the intermediate plate (Figure 5i and j and Figure 7b).

Expression of pax[zf-a] in the eye

Apart from the expression described above, *pax[zf-a]* transcripts are also detected in the developing eye. Expression of *pax[zf-a]* is clearly seen in the optic vesicles of 12 h embryos (Figure 5a and b). This is shortly after the optic vesicles can first be distinguished as lateral bulges of forebrain tissue. From the optic stalk which is located at the border between the diencephalon and the telencephalon, the optic cup bends caudally and covers a major part of the lateral walls of the diencephalon. The initial *pax[zf-a]* expression is restricted to the portions of the optic cup which face towards the diencephalon, whereas the hybridization signals become more diffuse in the region of the optic stalk (see Figures 5k and n and 6b). At 24 h a uniform hybridization signal is detected throughout the eye (Figure 5l and o) excluding the anterior three-quarters of the optic stalk. A more differentiated expression pattern is seen in the eyes of one week old zebrafish larvae where *pax[zf-a]* transcripts are mainly detected in the two medial layers of the retina (see Figure 5m and p) which correspond to the ganglion cell layer and the bipolar nerve layer (Mann, 1964).

Interestingly, *pax[zf-a]* is also expressed in a tissue derived from an invagination of the ectoderm overlaying the optic vesicle (McKeehan, 1951). The lateral portions of this tissue will later form the cornea whereas the medial portions will generate the lens (Mann, 1964). As seen in 24 h embryos (Figure 5l and o), *pax[zf-a]* transcripts are restricted to the lateral tissue.

Discussion

Molecular structure of pax[zf-a]

In this report we describe the isolation and characterization of cDNAs for a novel paired box-containing gene *paz[zf-a]* from zebrafish as well as a detailed analysis of the expression pattern of this gene during embryonal development. *Pax[zf-a]* encodes a putative protein organized into five domains: an N-terminal region, a paired box separated from a paired homeobox by an 80 amino acid long glycine-rich, acidic linker and a C-terminal 149 amino acids region with a preponderance of serine, proline and threonine residues.

The paired box-containing genes are thought to encode transcription factors. *Drosophila* Prd, containing both a paired box and a homeobox, is capable of sequence-specific DNA binding (Hoey and Levine, 1988; Han *et al.*, 1989). Prd contains two distinct DNA-binding activities. The homeodomain is responsible for one of them while the other is not dependent on the helix−turn−helix (HTH) motif of the homeodomain (Treisman *et al.*, 1989). Due to its basic nature and evolutionary conservation the paired box is an obvious candidate for another DNA-binding domain outside the homeodomain (Burri *et al.*, 1989; Bopp *et al.*, 1989; Jostes *et al.*, 1991; Goulding *et al.*, 1991). In this context it is interesting to note that the *Drosophila Pox meso* and *Pox neuro* genes, which contain a paired box but no homeobox, have been shown to encode nuclear proteins (Bopp *et al.*, 1989). Very recently, Treisman *et al.* (1991), using the e_4 and e_5 sequences of the *even-skipped* promoter, showed that the paired box of Prd harbours a DNA-binding activity which is both independent of the paired homeodomain and displays a different sequence specificity. The N-terminal 90 amino acids of the paired box including an earlier proposed α-helix (Burri *et al.*, 1989, see Figure 2A) is both necessary and sufficient for DNA binding (Treisman *et al.*, 1991). In addition, Goulding *et al.* (1991) showed that murine Pax3 is also able to bind to the e_5 site. If both the paired domain and the homeodomain represent DNA-binding elements the question arises whether proteins containing both domains, like Pax[zf-a], recognize two different DNA sequences or if the two domains act in concert to specify binding to one target sequence. The latter scenario has recently been shown to be correct for transcription factors of the POU family where both the POU box and the homeobox contribute to specific binding to one target sequence (for a recent review see Ruvkun and Finney, 1991). On the other hand, in prokaryotes it is known that two autonomous DNA-binding domains may reside in the same polypeptide (Moitoso de Vargas *et al.*, 1988). The SV40 enhancer-binding factor TEF-1 is also able to bind to two different sites (Xiao *et al.*, 1991). Our observation (see Results) that the paired box is separated from the paired homeobox by a glycine-rich, acidic linker in the known sequences of this class may be significant, because this could imply that the two basic domains (the paired box and the homeobox) can be placed in close apposition on the DNA due to the flexible linker which also is subjected to charge repulsion from the DNA. In fact, Treisman *et al.* (1991) present data from binding to the e_4 site in the *even-skipped* promoter implying that the paired domain and the homeodomain can simultaneously bind adjacent sites in a co-operative manner. For another site in the same promoter (e_5) the paired domain binds independently of the homeodomain. Since the e_4 and e_5 sites most probably do not

represent important sites for regulation of transcription by Prd *in vivo* (Frasch and Levine, 1987; Treisman *et al.*, 1991) it will be important to identify such sites and use these for further studies of possible co-operativity versus independent binding of the two domains. It is likely that a double potential including both cooperative binding to one site and linking of two sites is retained by paired box-containing proteins harbouring a homeodomain since this would increase the complexity and flexibility of transcriptional regulation by these proteins.

Studies of transcriptional activation domains have revealed different structural features such as acidic helices, glutamine-rich, proline-rich and serine/threonine-rich domains (Ptashne, 1988; Courey *et al.*, 1989; Mermod *et al.*, 1989; Tanaka and Herr, 1990). Interestingly, Pax[zf-a] contains a C-terminal region particularly rich in serines, but also with a preponderance of proline and threonine residues. A high serine content (12−20%) is also present in the C-terminal regions of all the other paired box-containing genes where this region has been sequenced; *prd*, *gsb-p*, *gsb-d*, *Pax2*, *Pax3* and *Pax8*. It remains to be demonstrated whether these regions contain functional transcriptional activation domains.

Expression of pax[zf-a] in the neural tube

At early developmental stages, *pax*[*zf-a*] is expressed in two areas of the neural tube, anteriorly in the portion of the diencephalon which will later form the thalamus, and posteriorly in the basal plate and intermediate plate of the hindbrain and spinal cord. Both areas are separated by a non-expressing region extending from the midbrain to rhombomere 2. Morphological studies and cell lineage analyses have been shown that the developing neural tube of vertebrates can be partitioned in the transverse plane at the midbrain−hindbrain isthmus to form a prechordal (rostral) and an epichordal (caudal) division (see Jacobson, 1985). The early expression patterns of many vertebrate homologues to *Drosophila* developmental regulatory genes seem to reflect the two divisions. One major example includes the large number of homeobox genes of the murine *Hox* clusters which are expressed posterior to different anterior limits in the hindbrain and spinal cord (Wilkinson *et al.*, 1989; Murphy *et al.*, 1989; Kessel and Gruss, 1990). Some vertebrate genes which are closely related to *Drosophila* segmentation genes are known to be expressed in both the epichordal and prechordal divisions during early embryonic stages. Similar to *pax*[*zf-a*] the features of their expression patterns seem to differ on either side of the midbrain−hindbrain isthmus (Wilkinson *et al.*, 1987; Jostes *et al.*, 1991).

We observe that the early embryonic expression of the *pax*[*zf-a*] gene in the hindbrain and spinal cord is restricted to the basal plate and possibly the intermediate plate (for definition see Altman and Bayer, 1984). Thus, similar to several of the murine *Pax* genes which show overlapping or complementary patterns of expression along the epichordal division of the neural tube (Nornes *et al.*, 1990; Jostes *et al.*, 1991; Goulding *et al.*, 1991), *pax*[*zf-a*] may be involved in the specification of neuronal cells or cell groups in the hindbrain and spinal cord.

During the early stages of neural development *pax*[*zf-a*] is also expressed in the diencephalic region of the prechordal division. In 12 h embryos and more pronouncedly at 17 h of development, before morphological subdivisions are visible in the rostral brain, this area of *pax*[*zf-a*] expression has features which suggest an involvement in early regionalization of the diencephalon. One of these is that most of the expression boundaries seem to be colonized by axons (Wilson *et al.*, 1990; Krauss *et al.*, 1991a). Furthermore, the area of expression is related to the neurogenic pattern known for the developing rostral brain (Bergquist and Kallen, 1954; Keyser, 1972). Similarly, another zebrafish paired box gene, *pax*[*zf-b*], is regionally expressed in the optic stalk and the posterior midbrain (Krauss *et al.*, 1991a,b). In addition, also the murine *Dlx* gene and several members of the *Wnt* family appear to be involved in early regionalization of the rostral brain (Roelink and Nusse, 1991; Price *et al.*, 1991).

The increased complexity of *pax*[*zf-a*] expression observed in the prechordal division at later developmental stages probably reflects that the gene exhibits additional functions in the rostral brain. One clear illustration of this is seen in 24 h embryos where the strongest hybridization signals are detected in restricted regions directly underlying the eye and the olfactory bulb. Within these regions the highest expression levels are observed laterally in the wall of the neural tube. The locations of these sites coincide with the areas in which high levels of acetylcholinesterase (AChE) are found at this stage of zebrafish embryogenesis (Wilson *et al.*, 1990). Since AChE activity appears in postmitotic neurons which migrate to the periphery of the neural tube (Layer *et al.*, 1983), the *pax*[*zf-a*] expression detected at these sites indicates an involvement in neuronal maturation and/or differentiation.

Materials and methods

Cloning and sequencing

Approximately 1.5×10^6 plaques of a 33 h zebrafish embryonic λ*gt11* cDNA library and 1×10^6 plaques of a 20−28 h embryonic λZAP-II library were screened by plaque hybridization at low stringency (McGinnis *et al.*, 1984) using a mix of two paired box sequences as probe: a 313 bp *Hin*dIII−*Eco*RI from murine *pax1* and a cloned genomic PCR product from zebrafish *pax*[*zf-d*]. Four clones thus isolated contained *pax*[*zf-a*] cDNA sequences (see Results). cDNA inserts were subcloned into pGEM-3Zf(+) (Promega Biotec) or subcloned by helper phage cotransfection (Stratagene) and mapped with restriction enzymes. Both strands of overlapping fragments were sequenced after subcloning in M13mp18 and M13mp19 using Sequenase (US Biochemicals). 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).

RNA isolation and Northern blotting

Total RNA from zebrafish embryos was isolated by the guanidine thiocyanate−CsCl method (MacDonald *et al.*, 1987). Poly(A)$^+$ RNA was purified (mRNA purification kit; Pharmacia) and RNA was size fractionated on a 1% agarose gel containing 2.2 M formaldehyde. Capillary transfer was performed in 10 × SSC (1 × SSC is 0.15 M NaCl, 0.015 M sodium citrate, pH 7.0). The DNA for the hybridization was ^{32}P-labelled using an oligolabelling kit (Amersham). Prehybridization and hyridization were carried out in 50% formamide, 5 × SSPE, 2 × Denhardt's reagent, 100 μg/ml denatured salmon sperm DNA and 200 μg/ml yeast tRNA. The filters were washed twice for 15 min in 1 × SSPE, 0.1% SDS at 50°C and once for 15 min at 60°C in the same buffer.

In situ hybridization on tissue sections

Zebrafish embryos after 12 h, 17 h, 24 h and 36 h of development at 28.5°C were manually dechorionated and subsequently fixed overnight in 4% paraformaldehyde in fixation buffer (4% sucrose, 0.12 mM $CaCl_2$, 0.1 M PO_4, pH 7.3) at 4°C. The embryos were then washed (3 × 5 min) in fixation buffer and embedded in agar−sucrose solution (1.5% agar, 5% sucrose) at 40°C. The hardened agar−sucrose block was trimmed according to the desired plane of sectioning and left overnight in 30% sucrose. 16 μm sections were cut at −20°C, placed on gelatinized slides and air-dried for 2 h at room temperature. Tissue sections were dehydrated in ethanol (50%,

75%, 95% and 100%; 4 min each). 100 ng probe was labelled by nick translation (BRL nick translation system) using [^{32}S]dATP. The size of the fragments was adjusted to an average of 40–120 nucleotides by the addition of RQ DNase (Promega) to the reaction. Probes were passed twice over a spin column, ethanol precipitated, and diluted to a concentration of 0.1 μg/ml in the hybridization buffer (50% formamide, 15% dextran sulphate, 0.5 M NaCl, 10 mM Tris–HCl, pH 7.5, 2.5 mM EDTA, 10 mM DTT, 1 × Denhardt's, 0.1 mg/ml herring sperm DNA, 0.1 mg/ml yeast tRNA and 1 μM thio S-ATP). The probe was then heat denatured and cooled on ice. Tissue sections with the probe were 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 for 30 min at 45°C in 50% formamide, 4 × SSC, 0.1% 2-mercaptoethanol, and once for 15 min in 2 × SSC at room temperature. Slides were dehydrated stepwise in ethanol and 0.3 M ammonium acetate (50%, 75%, 95% and 100%; 4 min each). For exposure they were 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 for 15 min in H_2O, stained for 2 min in 5% Giemsa in 1 × PBS, dried and mounted in Difco DPX 8711 mounting medium.

In situ hybridization on whole mount embryos

The whole mount *in situ* hybridization protocol is a modification of the procedure described for *Drosophila* by Tautz and Pfeifle (1989). Zebrafish embryos were manually dechorionated and fixed overnight in 4% paraformaldehyde in fixation buffer (see above) at 4°C. The embryos were washed for 4 × 15 min in PBT (1 × PBS, 0.1% Tween-20, 0.2% BSA), 10 s in H_2O, and transferred for 8 min to acetone at −20°C and washed again 2 × 15 min in PBT. PBT was then stepwise changed to hybridization buffer as follows: 15 min in 1:1 PBT/hybridization buffer and 15 min 100% hybridization buffer (50% formamide, 5 × SSC, 100 μg/ml salmon sperm DNA, 50 μg/ml heparin and 0.1% Tween-20). The embryos were then prehybridized for 30 min at 45°C in hybridization buffer. 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 pH 7.8, 50 mM $MgCl_2$, 100 mM 2-mercaptoethanol), 1 μl dNTP mix (BMB Geneus kit), 4 μl 1/10 diluted RQ DNase (Promega) and 5 μl DNase–polymerase mix (BRL nick translation system). The reaction was stopped by addition of 5 μl Stop mix (BRL nick translation system) and heating for 10 min at 68°C. The probe was then run through a Sephadex G50 column, precipitated and taken up in an appropriate volume of hybridization buffer. After prehybridization, the digoxygenin labelled DNA probe was added at a concentration of 5 μg/ml to an appropriate volume of hybridization buffer, boiled for 2 min and put on ice. Prehybridization solution was removed and hybridization buffer was added to the embryos. Embryos were hybridized overnight at 45°C. After hybridization, embryos were washed for 1 h at 45°C in hybridization buffer, for 15 min at 45°C in 1:1 hybridization buffer and PBT, and twice for 15 min at room temperature in PBT. Embryos were refixed for 30 min in 4% paraformaldehyde fixation buffer, 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 a new tube and incubated for 1 h with anti-digoxygenin antibodies (BMB) diluted 1:2000 in PBT. 700 μl antibody–PBT solution were used per 100 embryos. Pretreated antibodies were then added to the remaining embryos and left at 4°C overnight. Embryos were again washed (4 × 15 min) in PBT, and washed subsequently (3 × 5 min) in pH 9.5 buffer (100 mM NaCl, 50 mM $MgCl_2$, 100 mM Tris–HCl pH 9.5, 1 mM Levamisol, 0.1% Tween-20). The staining reaction 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 2 h. After a suitable signal: background ratio was reached, the reaction was stopped by washing the embryos twice for 5 min in PBT. Embryos were finally dehydrated in ethanol (30%, 50%, 75%, 85%, 95% and 100%; 3 min each), put for 1 h in methyl salicylate and mounted in Permount.

Acknowledgements

We are indebted to K.Zinn, D.Grunwald and B.Riggleman for providing the cDNA libraries and to P.Gruss for providing the murine *Pax1* probe. We would like to thank C.Kimmel, A. Lumsden, A.Molven, M.Westerfield and S.Wilson for their helpful discussions. We are grateful to Inna Korzh and Turid Holm for excellent technical assistance and to Sarah El-Gewely for help with the manuscript. S.K. and V.K. are fellows of the Aakre Cancer Foundation and the Norwegian Research Council for Science and the Humanities, respectively. Financial support was also obtained from the Norwegian Cancer Society, Hasselblad Foundation and Nansen Foundation.

References

Altman,J. and Bayer,S. (1984) In Beck,F., Hild,W., van Limborgh,J., Ortmann,R., Pauly,J.E. and Schiebler,T.H. (eds) *Advances in Anatomy, Embryology and Cell Biology*. Vol. 85, Springer-Verlag, New York.
Balling,R., Deutsch,U. and Gruss,P. (1988) *Cell*, **55**, 531–535.
Baumgartner,S., Bopp,D., Burri,M. and Noll,M. (1987) *Genes Dev.*, **1**, 1247–1267.
Bergquist,H. and Kallen,B. (1954) *J. Comp. Neurol.*, **100**, 627–660.
Bopp,D., Burri,M., Baumgartner,S., Frigerio,G. and Noll,M. (1986) *Cell*, **47**, 1033–1040.
Bopp,D., Jamet,E., Baumgartner,S., Burri,M. and Noll,M. (1989) *EMBO J.*, **8**, 3447–3457.
Burri,M., Tromvoukis,Y., Bopp,D., Frigerio,G. and Noll,M. (1989) *EMBO J.* **8**, 1183–1190.
Carroll,S.B. (1990) *Cell*, **60**, 9–16.
Courey,A.J., Holtzman,D.A., Jackson,S.P. and Tjian,R. (1989) *Cell*, **59**, 827–836.
Deutsch,U., Dressler,G.R. and Gruss,P. (1988) *Cell*, **53**, 617–625.
Devereux,J., Haeberli,P. and Smithies,O. (1984) *Nucleic Acids Res.*, **12**, 387–397.
Dressler,G.R., Deutsch,U., Chowdhury,K., Nornes,H.O. and Gruss,P. (1990) *Development*, **109**, 787–795.
Fitzgerald,M. and Shenk,T. (1981) *Cell*, **24**, 251–260.
Frasch,M. and Levine,M. (1987) *Genes Dev.*, **1**, 981–995.
Frigerio,G., Burri,M., Bopp,D., Baumgartner,S. and Noll,M. (1986) *Cell*, **47**, 735–746.
Gehring,W.J. (1987) *Science*, **236**, 1245–1252.
Goulding,M.D., Chalepakis,G., Deutsch,U., Erselius,J.R. and Gruss,P. (1991) *EMBO J.*, **10**, 1135–1147.
Graham,A., Papalopulu,N. and Krumlauf,R. (1989) *Cell*, **57**, 367–378.
Han,K., Levine,M.S. and Manley,J.L. (1989) *Cell*, **56**, 573–583.
Hanneman,E., Trevarrow,B., Metcalfe,W.K., Kimmel,C.B. and Westerfield,M. (1988) *Development*, **103**, 49–58.
Hoey,T. and Levine,M. (1988) *Nature*, **332**, 858–861.
Ingham,P.W. (1988) *Nature*, **335**, 25–34.
Jacobson,M. (1985) *Annu. Rev. Neurosci.*, **8**, 71–102.
Jostes,B., Walther,C. and Gruss,P. (1991) *Mech. Dev.*, **33**, 27–38.
Kemp,B.E. and Pearson,R.B. (1990) *Trends Biochem. Sci.*, **15**, 342–346.
Kessel,M. and Gruss,P. (1990) *Science*, **249**, 374–379.
Keyser,A. (1972) *Acta Anat.*, **83** (suppl. 59), 1–178.
Kissinger,C.R., Liu,B., Martin-Blanco,E., Kornberg,T.B. and Pabo,C.O. (1990) *Cell*, **63**, 579–590.
Kozak,M. (1986) *Cell*, **44**, 283–292.
Kozak,M. (1987) *Nucleic Acids Res.*, **15**, 8125–8148.
Krauss,S., Johansen,T., Korzh,V. and Fjose,A. (1991a) *Nature*, **353**, 267–270.
Krauss,S., Johansen,T., Korzh,V. and Fjose,A. (1991b) *Development*, in press.
Krebs,E.G., Eisenman,R.N., Kuenzel,E.A., Litchfield,D.W., Lozeman, F.J., Luscher,B. and Sommercorn,J. (1988) *Cold Spring Harbor Symp. Quant. Biol.*, **53**, 78–84.
Layer,P.H., Rommel,S., Bulthoff,H. and Hengstenberger,R. (1983) *Cell Tiss. Res.*, **251**, 587–595.
MacDonald,R.J., Swift,G.H., Przybyla,A.E. and Chirgwin,J.M. (1987) *Methods Enzymol.*, **152**, 219–227.
Malicki,J., Schughart,K. and McGinnis,W. (1990) *Cell*, **63**, 961–967.
Mann,J. (1964) *The Development of the Human Eye*. Grune and Stratton, New York.
McGinnis,W., Garber,R.L., Wirz,J., Kuroiwa,A. and Gehring,W.J. (1984) *Cell*, **37**, 403–408.
McKeehan,M. (1951) *J. Exp. Zool.*, **117**, 31–64.
Mermod,N., O'Neill,E.A., Kelly,T.J. and Tjian,R. (1989) *Cell*, **58**, 741–753.
Moitoso de Vargas,L., Pargellis,C.A., Hasan,N.M., Bushman,E.W. and Landy,A. (1988) *Cell*, **54**, 923–929.
Murphy,P., Davidson,D.R. and Hill,R.E. (1989) *Nature*, **341**, 156–159.
Nornes,H.O., Dressler,G.R., Knapik,E.W., Deutsch,U. and Gruss,P. (1990) *Development*, **109**, 797–809.
Patel,N.H., Schafer,B., Goodman,C.S. and Holmgren,R. (1989) *Genes Dev.*, **3**, 890–904.
Plachov,D., Chowdhury,K., Walther,C., Simon,D., Guenet,J.-L. and Gruss,P. (1990) *Development*, **110**, 643–651.
Price,M., Lemaistre,M., Pischetola,M., Di Lauro,R. and Duboule,D. (1991) *Nature*, **351**, 748–751.
Proudfoot,N.J. and Brownlee,G.G. (1976) *Nature*, **263**, 211–214.
Ptashne,M. (1988) *Nature*, **335**, 683–689.

Richardson,J.S. and Richardson,D.C. (1989) In Fasman,G.D. (ed.), *Prediction of Protein Structure and the Principles of Protein Conformation*. Plenum Press, New York, pp. 1–98.
Roberts,B. (1989) *Biochim. Biophys. Acta*, **1008**, 263–280.
Roelink,H. and Nusse,R. (1991) *Genes Dev.*, **5**, 381–388.
Ruvkun,G. and Finney,M. (1991) *Cell*, **64**, 475–478.
Scott,M.P., Tamkun,J.W. and Hartzell,III,G.W. (1989) *Biochim. Biophys. Acta*, **989**, 25–48.
Tanaka,M. and Herr,W. (1990) *Cell*, **60**, 375–386.
Tautz,D. and Pfeifle,C. (1989) *Chromosoma*, **98**, 81–85.
Treisman,J., Gönczy,P., Vashishtha,M., Harris,E. and Desplan,C. (1989) *Cell*, **59**, 553–562.
Treisman,J., Harris,E. and Desplan,C. (1991) *Genes Dev.*, **5**, 594–604.
Trevarrow,B., Marks,D.L. and Kimmel,C.B. (1990) *Neuron*, **4**, 669–679.
Xiao,J.H., Davidson,I., Matthes,H., Garnier,J.-M. and Chambon,P. (1991) *Cell*, **65**, 551–568.
Wilkinson,D.G., Bailes,J.A. and McMahon,A.P. (1987) *Cell*, **50**, 79–88.
Wilkinson,D.G., Bhatt,S., Cook,M., Boncinelli,E. and Krumlauf,R. (1989) *Nature*, **341**, 405–409.
Wilson,S.W., Ross,L.S., Parrett,T. and Easter,S.S., Jr (1990) *Development*, **108**, 121–145.

Received on July 23, 1991; revised on August 23, 1991