The EMBO Journal vol.10 no.5 pp.1135-1147, 1991

Pax-3, a novel murine DNA binding protein expressed during early neurogenesis

Martyn D.Goulding, Georges Chalepakis, Urban Deutsch, Jörn R.Erselius and Peter Gruss

Max Planck Institute for Biophysical Chemistry, Göttingen, Germany

Communicated by P.Gruss

We describe the isolation and characterization of *Pax-3*, a novel murine paired box gene expressed exclusively during embryogenesis. *Pax-3* encodes a 479 amino acid protein with an M_r of 56 kd containing both a paired domain and a paired-type homeodomain. The Pax-3 protein is a DNA binding protein that specifically recognizes the e5 sequence present upstream of the *Drosophila even-skipped* gene. *Pax-3* transcripts are first detected in 8.5 day mouse embryos where they are restricted to the dorsal part of the neuroepithelium and to the adjacent segmented dermomyotome. During early neurogenesis, *Pax-3* expression is limited to mitotic cells in the ventricular zone of the developing spinal cord and to distinct regions in the hindbrain, midbrain and diencephalon. In 10-12 day embryos, expression of *Pax-3* is also seen in neural crest cells of the developing spinal ganglia, the craniofacial mesectoderm and in limb mesenchyme of 10 and 11 day embryos.

Key words: DNA binding/homeodomain/paired domain/neurogenesis

Introduction

During embryogenesis each organism utilizes a complex genetic program as a blueprint for constructing its body plan. This program directs the combinatorial action of particular genes to specify positional information and regulate morphogenesis during the development of the embryo. The fruit fly *Drosophila* has been a powerful model system for dissecting the genetic program underpinning its development, resulting in the isolation and characterization of many of the genes that regulate *Drosophila* embryogenesis. These genes form a hierarchy from the early acting maternal effect genes down to the later acting homeotic genes that operate sequentially to organize the body plan of the early embryo (Scott and Carroll, 1987; Akam, 1987; Ingham, 1988). A striking feature of many of the developmental control genes isolated from *Drosophila* is the presence of conserved sequences that encode protein motifs such as the homeodomain (Gehring, 1987), the helix-loop-helix domain (Caudy *et al.*, 1988; Murre *et al.*, 1989), the zinc-finger motif (Rosenberg *et al.*, 1986) and the paired domain (Bopp *et al.*, 1986; Frigerio *et al.*, 1986). Each of these domains is present in a variety of transcription factors, implying that transcriptional regulation forms a fundamental component of an organism's developmental program.

The isolation and analysis of vertebrate genes with homology to *Drosophila* homeobox-containing genes has been a useful approach to dissecting the genetic elements underlying vertebrate development. Using this approach >30 vertebrate homeobox-containing genes have been identified, each of which exhibits a restricted pattern of expression during vertebrate embryogenesis consistent with a role in establishing positional cues within the embryo (Holland and Hogan, 1988; Graham *et al.*, 1989; Wright *et al.*, 1989b). A further highly conserved sequence, the paired box, that encodes a 128 amino acid domain, has been identified in five genes expressed during early *Drosophila* development. These are the segmentation genes, *paired, gooseberry-distal* and *gooseberry-proximal*, as well as the two developmentally expressed tissue specific genes *Pox-meso* and *Pox-neuro* (Baumgartner *et al.*, 1987; Bopp *et al.*, 1986, 1989). In addition, a number of paired box genes from mouse (Deutsch *et al.*, 1988; Dressler *et al.*, 1990; Walther,C., Guenet,J.L., Simon,D, Deutsch,U., Jostes,P., Goulding,M.D., Plachov,D. Balling,R. and Gruss,P., submitted), *Xenopus* (P.Turner, personal communication) and human (Burri *et al.*, 1989) have been isolated on the basis of homology to *Drosophila* paired box sequences. Many of these genes also exhibit a temporally and spatially restricted pattern of expression during embryogenesis that implies a regulatory role in vertebrate development. Expression of the murine *Pax-1* gene in the mouse embryo is restricted to the sclerotome cells contributing to segmented structures of the developing vertebral column and sternebrae, the thymus (Deutsch *et al.*, 1988) and the developing limb bud (P.Timmins, personal communication.). *Pax-2*, the second member of the murine paired box gene family, is also developmentally regulated during murine embryogenesis. Transcripts of *Pax-2* are first detected in the nephric cord and Wolffian duct and then later in the ureter and metanephric mesenchyme (Dressler *et al.*, 1990). *Pax-2* transcripts are also detected in primitive spinal cord, in two discrete populations of early neuroblast cells and in the optic and otic vesicles (Nornes *et al.*, 1990). *Pax-8*, a gene that is highly homologous to *Pax-2*, exhibits a very similar pattern to *Pax-2* expression in the mouse. *Pax-8* is expressed in the developing excretory system and in the developing nervous system where it coincides with *Pax-2* expression in the spinal cord of 11-12 day embryos (Plachov *et al.*, 1990). Preliminary characterization of a further murine paired box gene *Pax-7*, reveals expression in the developing nervous system and in early myoblasts (Jostes *et al.*, 1990).

A number of reports have demonstrated that altering the endogenous pattern of homeobox gene expression can disturb the normal genesis of body structures in vertebrates (Kessel *et al.*, 1989; Ruiz i Altaba and Melton, 1989; Wright *et al.*, 1989a). There is now evidence that paired box-containing genes play an important role in vertebrate pattern formation. It has been shown that mutations in the *Pax-1* gene are

responsible for the mouse mutant phenotype *undulated* (Balling *et al.*, 1988; R.Balling, personal communication). *Undulated* mice exhibit vertebral malformations possibly involving incorrect allocation or differentiation of sclerotome cells during vertebral column formation (Grüneberg, 1954). The demonstration that mutations in *Pax-1* influence morphogenetic events in the mouse highlights the potential importance of paired box-containing genes in the developmental program of vertebrates.

In this paper we present a detailed analysis of a new member of the murine *Pax* gene family, whose structure and expression reveal a number of interesting features. The *Pax-3* gene encodes a 56 kd protein that contains a paired-type homeodomain in addition to the paired box domain. The protein encoded by the *Pax-3* gene specifically binds to the e5 sequence upstream of the *Drosophila even-skipped* gene that is recognized *in vitro* by the *Drosophila paired* protein. *Pax-3* transcripts are detected only during embryogenesis consistent with a role in pattern formation in the mouse embryo. Furthermore, *Pax-3* is expressed in a unique pattern during early neurogenesis, with transcripts restricted to the dorsal half of the neural tube. Expression is also observed in a population of migrating neural crest cells which contribute to the dorsal root ganglia and the cephalic mesenchyme. The pattern of *Pax-3* in the developing nervous system suggests that *Pax-3* plays an important role in the early regionalization of the vertebrate nervous system. *Pax-3* transcripts are also found in the segmented mesoderm and limb bud mesenchyme of the mouse embryo.

Results

Complementary DNA clones encoding *Pax-3* were isolated from a mouse 8.5 day embryonic cDNA library (Fahrner *et al.*, 1987) under low stringency conditions using a 313 bp *Hin*cII–*Sac*I fragment from the mouse *Pax-1* paired box as a probe. The two longest cDNA clones isolated were compared and found to be colinear with one cDNA (pBH3.2) containing a little more 5′ sequence than the other (pMG-33). The complete sequence of the longest cDNA was determined and the protein sequence deduced. The 2347 bp of sequence of this cDNA is shown in Figure 1 along with the derived protein sequence of the longest open reading frame. The longest *Pax-3* cDNA clone contains 1437 nucleotides that code for the Pax-3 protein, and a further 607 bp of 3′ untranslated sequence containing a polyadenylation signal AATAAA (Fitzgerald and Shenk, 1981) 18 nucleotides from the 3′ terminus of the cDNA. Northern blots to *Pax-3* reveal two transcripts of 3.3 and 3.6 kb (Figure 6B), whereas the longest cDNA clone isolated to date contains 297 bp of 5′ untranslated sequence. Consequently these cDNA clones may contain only part of the 5′ untranslated region. Analysis of the *Pax-3* cDNA revealed a long open reading frame encoding a protein of 479 amino acids containing a paired domain and a paired-type homeodomain. Three in-frame translation start sites are present (Figure 1), each in the context of a moderately good concensus start sequence. It seems probable the first of these is used, as this provides the best fit (CCAGGAUGA) with the derived concensus sequence for translational initiation (CCPuCCAUGG; Kozak, 1986 and 1987). Furthermore, a stop signal TAA present upstream of the first ATG should favour the initiation of translation at the next translation start signal encountered by the ribosomal complex.

The Pax-3 protein contains three of the conserved protein motifs found in other members of this gene family, the paired domain, octapeptide and paired-type homeodomain. As with all the paired box containing genes isolated to date, the highly conserved paired domain resides close to the amino terminus of the protein. A weakly basic domain of 32 amino acids precedes the 128 amino acid paired domain of *Pax-3*. The first 74 amino acids of the paired domain are basic in nature and are highly conserved between all members of the paired gene family, however the carboxyl 54 amino acids are more divergent. The *Pax-3* gene belongs to a subclass of paired box-containing genes that includes the murine *Pax-7* gene and the human *HuP1* and *HuP2* genes (Figure 2A). The paired boxes of these genes are also closely related to those of the *paired* (*prd*), *gooseberry-distal* (*gsb-d*) and *gooseberry-proximal* (*gsb-p*) genes (Burri *et al.*, 1989; Walther *et al.*, submitted). Within the *gooseberry-Pax-3-Pax-7* subclass of paired box genes the paired domain exhibits an extremely high level of conservation. For example, the first 74 amino acids of *Pax-3, HuP2, Pax-7* and *HuP1* are identical and show only a few changes with respect to their *Drosophila* paired box gene counterparts (Figure 2A). Moreover, the carboxyl 54 amino acids of *Pax-3* contains only one change with respect to *HuP2* and seven amino acids different from *Pax-7*. The exon encoding the carboxyl region of the *Pax-3* paired domain encompasses a further conserved domain, the octapeptide. The octapeptide and paired domains of *Pax-3* are separated by a highly charged domain mainly consisting of basic and acidic residues. The octapeptide domain consists of the consensus sequence HSIAGILG (Figure 1) and is found in the *Drosphila gsb-p* and *gsb-d* genes, in two human genes *HuP1* and *HuP2* (Burri *et al.*, 1989) and in the murine *Pax-7* gene (Jostes *et al.*, 1990). Interestingly, the octapeptide of *Pax-3* is identical to the human *HuP2* octapeptide.

The homeodomain has been identified in proteins encoded by a number of developmental control genes and contains a helix–turn–helix motif that is distantly related to the helix–turn–helix motif present in prokaryote DNA binding proteins and in certain eukaryote transcription factors (Goulding and Gruss, 1989). The Pax-3 protein contains in addition to the paired domain a paired-type homeodomain, found in the *Drosophila paired, gsb-d* and *gsb-p* proteins (Figure 2B). The homeodomain of *Pax-3* shares an extended homology over 78 amino acids with the homeodomains of all three *Drosophila* genes with the Pax-3 homeodomain being somewhat more divergent from *gsb-p* (74% homology) than *prd* and *gsb-d* (84% homology). The third helix of the paired-type homeodomain is highly conserved between all four genes with a single conservative change of valine (*Pax-3*) for isoleucine (*prd, gsb-d, gsb-p*) at the fourth residue. Most importantly, the *Pax-3* homeodomain contains a serine at position nine in the third helix and that confers sequence specific recognition by the homeodomain (Treisman *et al.*, 1989). Helix 2 of the homeodomain is also highly conserved except for the two carboxy amino acids which have both undergone non-conservative changes with respect to *gsb-d* and *gsb-p*.

The Pax-3 protein contains a further 201 amino acids at its carboxyl terminus in addition to the paired domain, homeodomain and octapeptide motifs. The presence of proline and glycine residues scattered throughout this domain indicates the relative paucity of any helical structure within this part of the Pax-3 protein. This part of the carboxyl portion of Pax-3 is very Pro-Ser-Thr rich. No extended

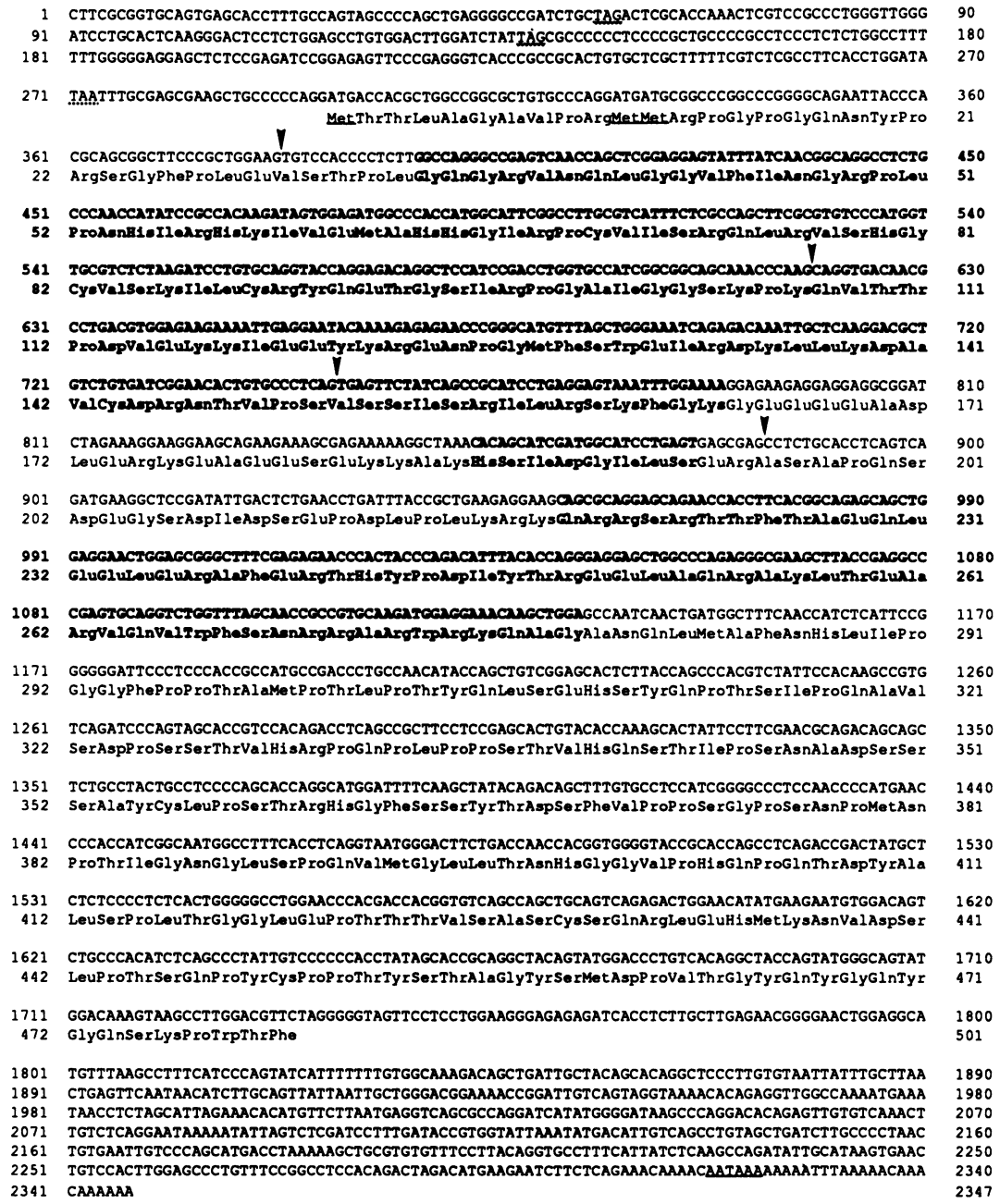

Fig. 1. Nucleotide and predicted protein sequence of the *Pax-3* gene. The paired domain octapeptide and homeodomain sequences are shown in bold type. The translation initiation codons (ATG) and polyadenylation signal (AATAAA) are underlined. Splice sites are indicated by arrows.

similarity was detected between the carboxyl region of the Pax-3 protein and those of other members of the paired box gene family. The His−Pro paired-repeat sequence identified in the *Drosophila paired* and *bicoid* proteins (Frigerio *et al.*, 1986) is not present in Pax-3. However, the three amino acids Tyr-Gly-Gln appear three times in an imperfect repeat adjacent to the carboxyl terminus of the Pax-3 protein (Figure 1).

In order to analyse the size of the protein encoded for by the *Pax-3* cDNA, pBH3.2 plasmid DNA was linearized, purified and then used to prepare RNA by *in vitro* transcription. An aliquot of the reaction mix was then translated *in vitro* in rabbit reticulocyte lysates. Translation of the pBH3.2 transcripts revealed a single major protein band of 56 kd specific for the pBH3.2 cDNA (Figure 3). The *in vivo* protein product of the *Pax-3* cDNA was also analysed following expression of *Pax-3* in embryonal carcinoma cells. An 1820 bp *Bam*HI−*Bal*I fragment containing the coding region of *Pax-3* was inserted into the pHSβ vector downstream of the mouse hsp68 promotor and used to stably transform P19 cells. Polyclonal antisera to a 14-mer peptide derived from the carboxyl terminus of Pax-3 (AbM-4) was used to immunoprecipitate Pax-3 protein from whole cell lysates of [^{35}S]methionine-labelled P19 stem cells and P19 cells overexpressing Pax-3. A 56 kd protein was observed in P19 cells only after heat shock which resulted in the overexpression of *Pax-3* mRNA (Figure 3). The M_r of 56 kd for the Pax-3 protein in SDS−polyacrylamide gels is consistent with the use of the first ATG in the cDNA to give a protein of 479 amino acids with a predicted M_r of 55 kd (cf. Figure 1). The AbM-4 antisera also immunoprecipitated a 32 kd protein specifically from extracts

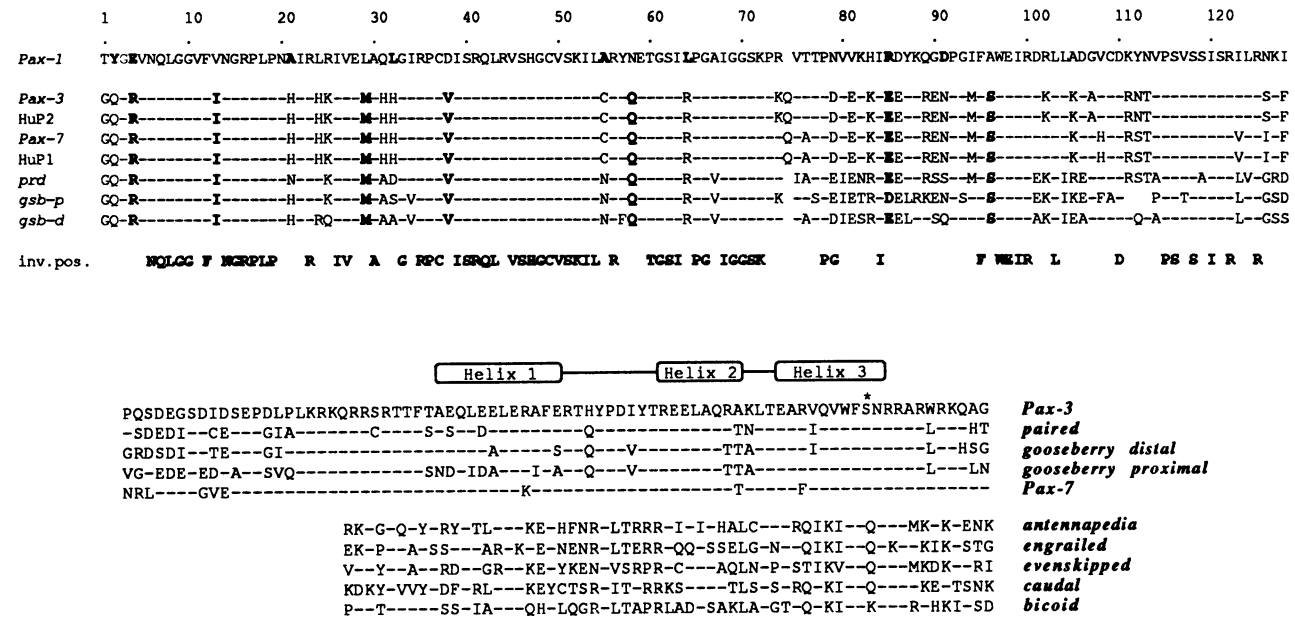

Fig. 2. (**A**) Protein sequence comparison of the *Pax-3* subclass of paired domains. A comparison with the murine *Pax-1* paired domain sequence is also shown (*Pax-1*). The *Pax-3* subclass-specific amino acids are denoted by bold type. The invariant amino acids for the paired domain are also shown ('inv. pos.'). The paired domains of other members of the *Pax-3* subclass are also shown. (**B**) Comparison of the Pax-3 homeodomain with other extended paired-type homeodomains. The homeodomains of the *Drosophila Antennapaedia, engrailed, evenskipped, caudal* and *bicoid* proteins are also shown. Bars indicate conserved amino acids. The structure of the *Antennapaedia* homeodomain has been determined (Gehring *et al.*, 1990) and the three helices of the homeodomain are indicated by open boxes. The ninth residue of the recognition helix that is responsible for sequence specific binding is indicated (*). This is a serine in the Pax-3 and the other paired-type homeodomains, a lysine in the *bicoid* homeodomain and a glutamine in the *Antennapaedia, engrailed, even-skipped* and *caudal* homeodomains.

of cells that overexpressed the Pax-3 protein (data not shown). This probably represents a proteolytic cleavage product of the Pax-3 protein, since it was only observed after inducing Pax-3 protein expression.

Pax-3 encodes a DNA binding protein

Proteins containing a homeodomain are known to bind specifically to their cognate DNA recognition sequences to regulate gene expression. The presence of a homeodomain within the Pax-3 protein suggested that it might function as a DNA binding factor. Previously the *Drosophila paired* protein has been shown to bind and footprint *in vitro* a region upstream of the *Drosophila even-skipped* gene known as the e5 sequence (Hoey and Levine, 1988; Treisman *et al.* 1989). To ascertain whether the Pax-3 protein was able to specifically recognise the e5 sequence, electrophoretic mobility shift assays (EMSA) were performed using cell extracts derived from P19 and PCC7 teratocarcinoma cells that overexpress the Pax-3 protein under the control of the mouse hsp68 promoter (Figure 4). Cell extracts from undifferentiated P19 cells were seen to contain very little endogenous e5-binding activity either before or after heat shock treatment (Figure 4, lanes a and c). Extracts from cells containing the *Pax-3* gene cloned in the reverse orientation showed no binding following heat shock (Figure 4, lane b). Heat shock induction of *Pax-3* expression in P19 and PCC7 cells resulted in an increase of three e5 DNA binding complexes (Figure 4, lanes d and e). These complexes are denoted by e5-I, e5-II and e5-III. It is unclear whether the e5-I complex represents a Pax-3 protein-specific complex since in some experiments high levels of this complex were observed prior to heat shock and did not increase upon induction of Pax-3 protein.

The e5 sequence consists of two distinct recognition sequences for the paired domain and homeodomain of the *Drosophila paired* gene (C.Desplan, personal communication, see Materials and methods). When each of these specific recognition sequences was used separately to compete for Pax-3 binding, both significantly reduced complex formation when used in a 100-fold excess of unlabelled oligonucleotide (Figure 4, lanes f and g). In contrast, a 100-fold excess of an olignucleotide containing the consensus recognition sequence for the *Antp*-type homeodomain (TCAATTAAAT) derived from the *Drosophila engrailed* gene (Hoey and Levine, 1988) failed to compete with the e5 sequence for Pax-3 binding (Figure 4, lane h). To confirm that these complexes were specific Pax-3−e5 complexes, cell extracts from P19 and PCC7 cells were incubated with an affinity purified antibody to Pax-3 (AbM-3). Incubation of the e5 oligonucleotide with the AbM-3 antisera alone did not produce any specific complexes (Figure 4, lane j). Following incubation of PCC7 cell extracts with AbM-3, no complex II was formed and the e5-III complex formation decreased (Figure 4, lane k and l). Incubation of cell extracts with the AbM-3 antibody also resulted in the appearance of two novel low mobility complexes, presumably corresponding to the antibody−Pax-3−e5 complex (Figure 4, lanes k and l). The two low mobility complexes were specific for the Pax-3 antibody complex, since incubating PCC7 cell extracts with preimmune sera from the same animal did not give these two low mobility complexes (Figure 4, lane i). The reason for increased e5-1 complex with higher concentrations of antibody is unclear.

Mutational analysis of the e5 DNA binding sequence was undertaken to ascertain the nature of the Pax-3 protein interaction with the e5 sequence. Pax-3 protein was able to

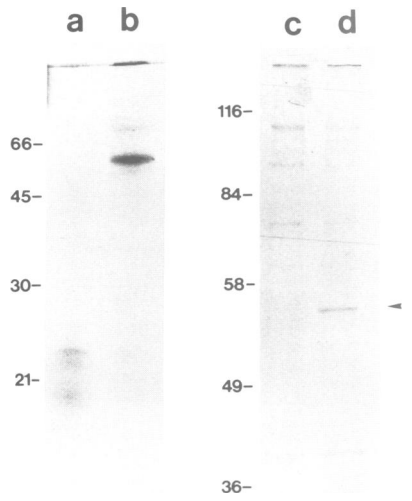

Fig. 3. Lanes **a** and **b** *In vitro* translated Pax-3 protein. The pBH3.2 cDNA encoding Pax-3 was transcribed *in vitro* and translated in a rabbit reticulocyte lysate. Translated proteins were separated on a 14% SDS−polyacylamide gel. (**Lane a**) No RNA (**Lane b**) *in vitro* transcribed Pax-3 RNA. **Lanes c** and **d:** Radioimmunoprecipitation of Pax-3 protein. P19 cells containing the *Pax-3* gene under the control of the mouse hsp68 promotor (P19hsPax cells) were incubated at 42°C for 2 h to induce *Pax-3* expression. Cell lysates were prepared from P19-hsPax cells labelled with [^{35}S]methionine and Pax-3 protein was immunoprecipitated using AbM-4 antisera and protein A−agarose radioimmunoprecipitated proteins were analysed on a 10% SDS−polyacrylamide gel. (**Lane c**) 37°C (**Lane d**) 42°C. The relative mobilities of protein molecular weight markers are shown. The immunoprecipitated Pax-3 protein is marked by an arrow.

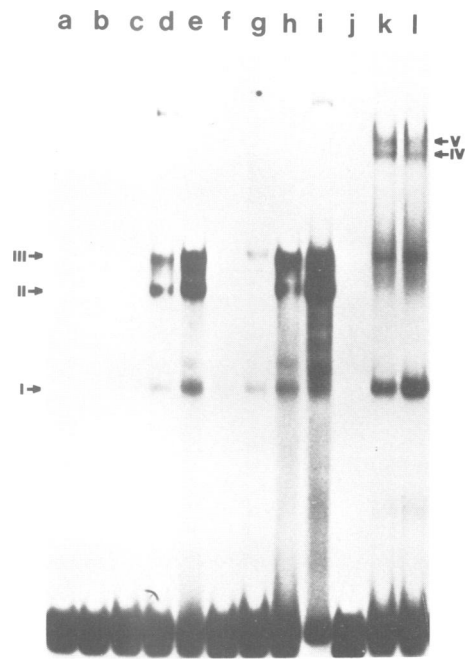

Fig. 4. Pax-3 binding to the *evenskipped* e5 sequence. Pax-3 protein was ectopically expressed in P19 and PCC7 cells using the mouse hsp68 promoter. P19hsPax and PCC7hsPax cells contain the *Pax-3* coding region in the correct orientation. P19hsPaxRO cells contain the *Pax-3* coding region in the reverse orientation. Cell extracts were prepared from cells following heat stock (42°C) or in the case of control extracts, from cells grown at 37°C. Cell extracts containing 1.5 μg of protein and a ^{32}P-labelled double-stranded oligonucleotide (oligonucleotide 110) were used for each EMSA reaction. **Lane a**: P19 stem cells. **Lane b**: P19hsPaxRO cells, 42°C. **Lane c**: P19 stem cells, 42°C. **Lane d**: P19hsPax cells. Extracts from PCC7hsPax cells that had been heat shocked to 42°C were used in **lanes e, f, g, h, i, k** and **l**. Binding to oligonucleotide 110 was competed with the oligonucleotide 45 encompassing the paired domain recognition sequence (**lane f**), the homeodomain recognition sequence (oligonucleotide 89, **lane g**) and the consensus *Antennapaedia* homeodomain recognition sequence (oligonucleotide Antp-1, **lane h**). In **lane i** the binding reaction was performed in the presence of pre-immune antisera. In **lanes j, k** and **l**, assays contained antisera to Pax-3 (AbM-3). **Lane j** contained AbM-3 antisera alone, **lanes k** and **l** contained cell extract and AbM-3 antisera.

bind the complete e5 sequence (oligonucleotide 110) that is footprinted by the *paired* protein or to a partial e5 sequence (oligonucleotide 45) containing the complete paired domain binding sequence and the core of the homeodomain binding sequence (Figure 5). Oligonucleotides containing the paired domain recognition sequence alone (oligonucleotide 22) exhibited only weak binding, while the homeodomain recognition sequence alone (oligonucleotide 89) gave no specific complexes with the Pax-3 protein. Oligonucleotide 14 contained a point mutation in the paired domain recognition sequence of oligonucleotide 45 that (CACCGTTCCGCTCC to CACCTTTCCGCTCC) significantly reduced binding by the Pax-3 protein (Figure 5, oligonucleotide 14). A non-specific oligonucleotide from the thyroglobulin promoter region partially homologous to the e5 sequence was unable to bind Pax-3 protein (oligonucleotide 24).

Pax-3 is expressed only during embryogenesis

The pattern of *Pax-3* expression during embryonic development was analysed by Northern hybridization of mRNA derived from embryos from 9 to 17 days *post coitum* (p.c.) (Figure 6A). Expression of *Pax-3* in mouse embryos was highest during mid-gestation with two transcripts of M_r 3.3 and 3.6 kb present from day 9 to day 12. From day 13 onwards, the level of both *Pax-3* transcripts declined and no expression of *Pax-3* was apparent on day 17. A similar temporal pattern of *Pax-3* expression was observed using RNase protection analysis of total RNA from embryos, where expression was also seen in 8 day embryos (data not shown). When adult tissues were examined for *Pax-3* transcripts, no tissues including the brain and spinal cord contained detectable levels of *Pax-3* mRNA (data not shown). The expression of *Pax-3* was also examined in two murine teratocarcinoma cell lines. No expression of *Pax-3* was seen in either undifferentiated F9 cells or P19 cells (Figure 6B). Differentiation of F9 cells into parietal endoderm with 10^{-7} M retinoic acid and 1 mM DBcAMP resulted in a sharp increase in two *Pax-3*-specific transcripts of 3.3 and 3.6 kb. Similarly, differentiation of P19 cells into neuroectoderm-like cells with retinoic acid resulted in the induction of these two *Pax-3* mRNAs.

Pax-3 is first expressed during neurulation

The spatial distribution of *Pax-3* transcripts in the developing mouse embryo was analysed from 7−14 days p.c. by *in situ* hybridization. No transcripts of *Pax-3* were detected in 7 day mouse embryos (data not shown). Transcripts of *Pax-3* were first seen in 8.5 and 9 day embryos in a small population of cells in the dorsal region of the neural groove and in the recently closed neural tube (Figure 7F,G; Figure 8).

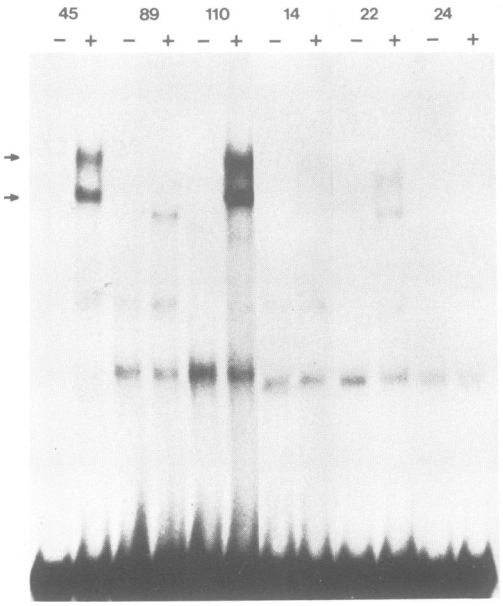

Fig. 5. Pax-3 binding to mutant e5 oligonucleotides. Cell extracts were prepared from P19hsPax cells either prior to heat shock (−) or after heat shock (+). The sequences of the oligonucleotides used for the binding reactions are given in Materials and methods. The two Pax-3-specific complexes are indicated by arrows.

In the trunk region of the embryo, expression of *Pax-3* was not observed in the neural plate and only occurred just prior to closure of the neural tube. In the region of the head fold, expression of *Pax-3* was observed in the lateral neuroepithelium of the prosencephalon (Figure 7G) and in the midbrain−hindbrain neuroepithelium (data not shown). In 8.5 day embryos no expression of *Pax-3* was detected in the primitive streak, presomitic mesoderm or the allantois (Figure 7G).

The onset of *Pax-3* expression in the neuroepithelium and mesoderm was examined in serial cross sections through the posterior neuropore region of 9 day embryos. In caudal regions of the embryo, neural development is delayed compared to more rostral regions of the embryo (Nornes and Carry, 1978). In regions where the neural groove was just closing, *Pax-3* transcripts were detected in the dorsalmost cells of the neuroepithelium but were absent from the adjacent mesoderm (Figure 8F,G). Following the closure of the neural tube, expression was restricted initially to the dorsal neural tube (Figure 8H,I), while more rostral sections showed expression of *Pax-3* in both the dorsal neural tube and in the dorso-lateral region of the adjacent somites (Figure 8J,K). These results indicate a temporal difference in the onset of *Pax-3* expression in the posterior part of the embryo, with *Pax-3* transcripts appearing first in the neuroepithelium and then in somites.

Pax-3 transcripts are restricted to the dorsal neural tube

The induction of *Pax-3* during neurulation led to a restricted pattern of *Pax-3* expression in the neural tube with transcripts present only in the dorsal neural tube. In 8.5 day embryos, *Pax-3* expression in the neuroepithelium extended almost continuously from the prosencephalon to the anterior margin of the posterior neuropore. In the primitive spinal cord of 10 day embryos, *Pax-3* transcripts were distributed

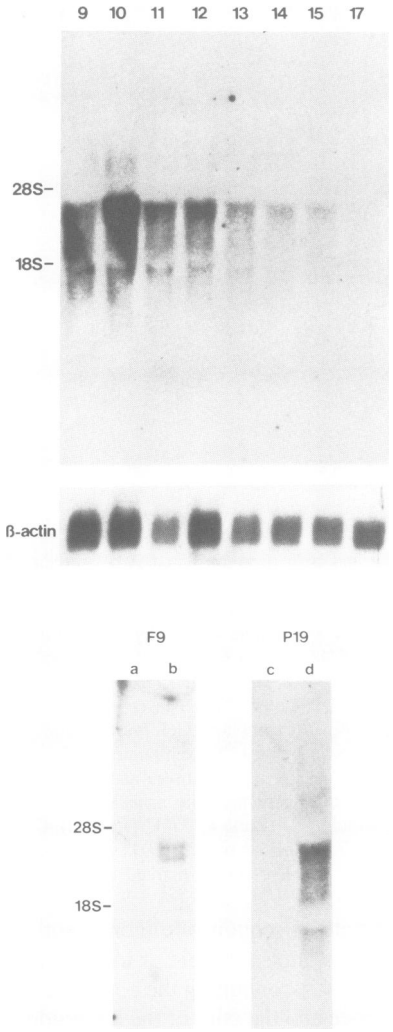

Fig. 6. (**A**) Northern blot analysis of *Pax-3* expression during murine embryogenesis. 10 μg samples of poly(A)$^+$ RNA prepared from 9−17 day embryos were loaded in each lane. The relative mobility of the 28S and 18S ribosomal RNA is indicated. The two *Pax-3* transcripts of 3.3 and 3.6 kb cannot be distinguished. As a control, the blot was hybridized with a mouse β-actin probe (β-actin). The faint band comigrating with the 18S RNA is probably due to degradation of *Pax-3* transcripts during RNA isolation. (**B**) Northern Blot analysis of *Pax-3* expression in F9 and P19 teratocarcinoma cells. (**Lane a**) 10 μg of poly(A)$^+$ RNA from undifferentiated F9 cells. (**Lane b**) 10 μg of poly(A)$^+$ RNA from F9 cells differentiated into parietal endoderm following treatment with 1 mM DBcAMP and 10^{-7} M retinoic acid. (**Lane c**) 20 μg of total RNA from undifferentiated P19 cells. (**Lane d**) 20 μg of total RNA from P19 cells differentiated into neuroectoderm with 10^{-7} M retinoic acid. Two *Pax-3* transcripts of 3.3 and 3.6 kb are clearly discernable in **lane b**.

throughout the dorsal half of the neural tube in both the alar and roof plates along the entire A−P axis of the embryo (Figure 7I). A clear ventral border of *Pax-3* expression coinciding with the sulcus limitans was observed (Figure 7D,I; Figure 8D,F). This boundary of *Pax-3* expression delineates the presumptive ventral boundary of the alar plate. The sharp ventral border of *Pax-3* expression was maintained in the spinal cord of 11, 12 and 13 day embryos (Figure 7E,J,K; Figure 8F).

The expression of *Pax-3* in the neural tube was analysed further by examining cross sections through the neural tube

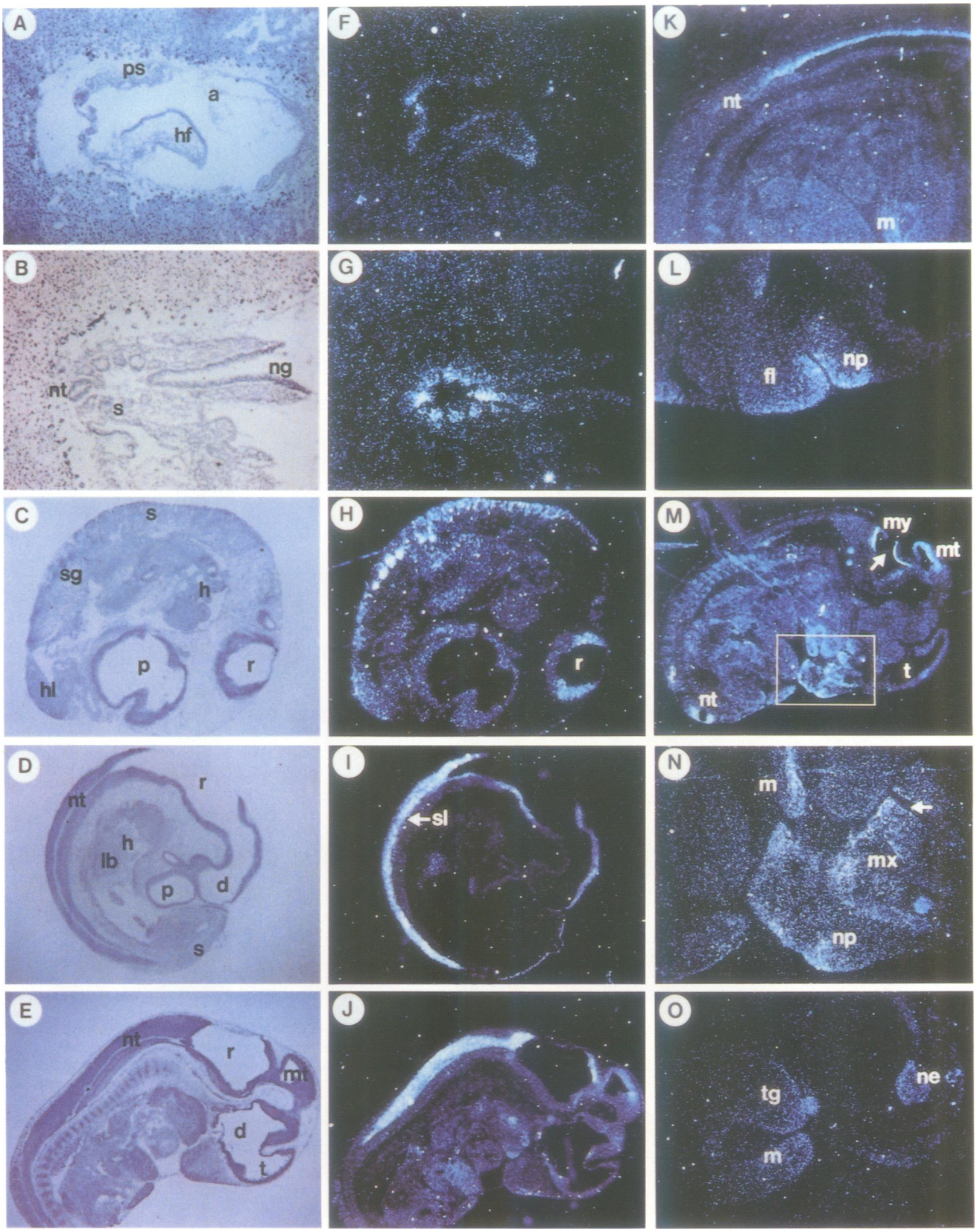

Fig. 7. Expression of *Pax-3* in 8.5–13 day mouse embyros. All sections were hybridized with an antisense *Pax-3* probe. (**A–E**) Bright field images. (**F–O**) Dark field images. (**A,F**), Sagittal section of a 8.5 day embryo. (**B,G**), Transverse section of an 8.5 day embryo. (**C,H**), Parasagittal section of a 10 day embryo. (**D,I**), Sagittal section of a 10 day embryo. Note the sharp ventral boundary of expression coinciding with the sulcus limitans (sl). (**E,J**), Sagittal section of a 11 day embryo. (**K**), Sagittal section through a 13 day embryo. (**L**) Parasagittal section through a 10 day embryo showing the nasal process and forelimb bud. (**M**) Parasagittal section of a 13 day embryo. The choroid plexus is indicated by an arrow. (**N**) Higher magnification of M. (**O**) Section of a 13 day embryo showing *Pax-3* transcripts in the tongue, mandible and olfactory neuroepithelium. (a) allantois. (d) diencephalon. (fl) forelimb bud. (h) heart anlage. (hf) head fold. (hl) hindlimb bud, (1b) limb bud mesenchyme. (m) mandible. (mt) metencephalon. (my) myelencephalon. (mx) maxilla. (ne) nasal epithelium. (np) nasal process. (nt) neural tube. (p) prosencephalon. (ps) primitive streak. (r) rhombencephalon. (s) somites. (sg) spinal ganglia. (sl) sulcus limitans. (t) telencephalon. (tg) tongue.

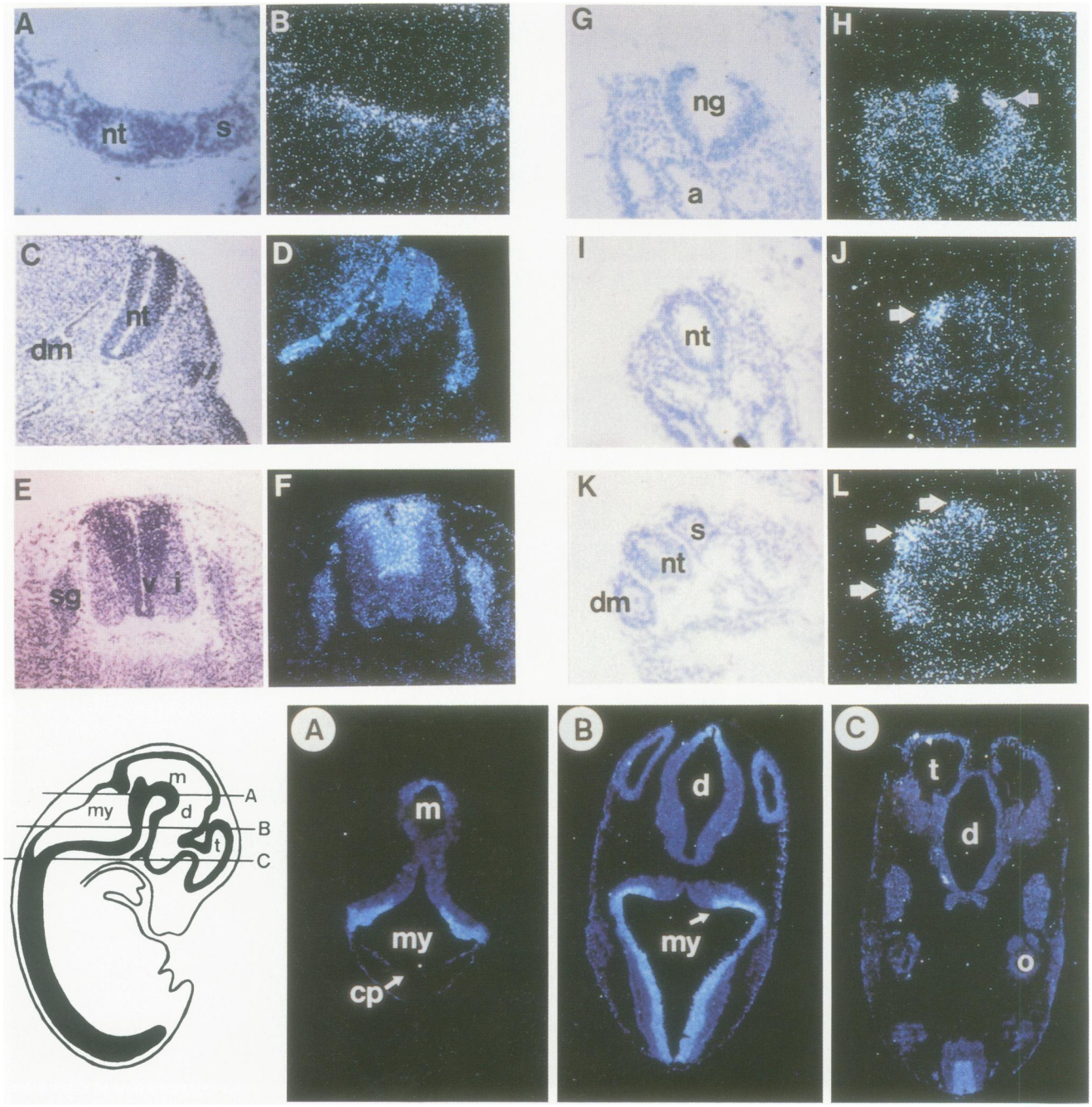

Fig. 8. Upper panel: (**A–F**) Expression of *Pax-3*: Transverse sections through the neural tube. (**A,C,E**) Bright field images. (**B,D,F**) Dark field images. (A,B), day 8.5 (C,D), day 10. (E,F), day 12. The ventricular zone (v) and the intermediate zone (i) are clearly visible at day 12 in E. (**G–L**) Transverse sections through the posterior neuropore of a 9 day mouse embryo. Sections G and H are most caudal, while sections I, J and then K, L are progressively more rostral. Bright field images. (G,I,K) Dark field images (H,J,L). *Pax-3* expression is marked by arrows. (a) dorsal aorta. (dm) dermomyotome. (ng) neural groove. (nt) neural tube. (s) somite. (sg) spinal ganglion. **Lower panel:** *Pax-3* expression in the developing brain. Transverse sections through the head region. The plan of the three sections is shown. Dark field images are shown for each section. The myencephalon and mesencephalon are indicated. The sections are orientated with the ventral side of the embryo facing upwards. (cp) anlagen of the choroid plexus. (d) diencephalon. (m) mesencephalon. (my) myelencephalon. (t) telencephalon. (o) otic vesicle. Note the sharp border of *Pax-3* expression in the fourth ventricle coinciding with the boundary of the alar and basal plates (arrow). The absence of expression in the choroid plexus is also indicated by an arrow. Expression in the spinal cord can be seen in C.

from 8–14 day of embryogenesis. In vertebrates, the early neural tube consists of a single layer of pseudostratified epithelial cells radially arranged around the lumen, the ventricular zone. In 8.5 and 10 day embryos, the neural tube consists mainly of ventricular cells with a few early motor neuroblasts appearing in the anterior basal plate of day 10 embryos (Nornes and Carry, 1978; McConnell, 1981). High levels of *Pax-3* were observed throughout the dorsal part of the neural tube on 8.5 and 10 days of gestation (Figure 8B and D). By day 11, a second more superficial zone of cells has formed in the thickening neural tube. This is the intermediate zone which initially consists of migrating and differentiating neuroblasts that in the adult animal form the dorsal, intermediate and ventral grey of the spinal cord. In

12 day embryos, *Pax-3* expression was present only in cells in the ventricular zone and absent from cells in the intermediate zone (Figure 8F). By day 14, no expression of *Pax-3* was found in anterior parts of the neural tube in either the ventricular or intermediate zones. Nevertheless, expression of *Pax-3* was still present in mitotically active cells in the caudal part of the neural tube. Throughout neural development *Pax-3* was restricted to mitotic stem cells present in the ventricular zone of the developing nervous system. The complete absence of *Pax-3* transcripts from the intermediate zone indicates that *Pax-3* transcription is rapidly down-regulated prior to the migration of early neuroblasts. Expression of *Pax-3* in the hindbrain, mesencephalon and prosencephalon of the 10 day embryo was found in the dorso-lateral walls of the third and fourth brain vesicles. In early embryos, *Pax-3* transcripts were found as far anterior as the diencephalon (Figure 7I). By day 11, expression had been further restricted to the dorso-lateral walls of the myelencephalon, metencephalon and mesencephalon where it was again present only in the ventricular layer (Figure 7J).

In 12 day embryos, transverse sections through the head region were used to analyse the distribution of *Pax-3* transcripts in the hindbrain and midbrain. Cells throughout the ventricular zone of the dorso-lateral walls of the myelencephalon showed strong *Pax-3* expression (Figure 8, lower panel). A sharp boundary of *Pax-3* expression coinciding with the margin of the pyrimidal process that later forms the pyramid of the medulla oblongata was apparent. In the myelencephalon, *Pax-3* transcripts were again found only in the ventricular zone and were completely absent from migrating neuroblasts (Figure 8, lower panel). The dorso-lateral walls of the mesencephalon were also positive for *Pax-3*, as were a few cells at the dorsal midline of the posterior part of the diencephalon that will develop into the epithalamus. Sagittal sections through 13 day embryos indicated a significant decline in the expression of *Pax-3* in the midbrain and hindbrain. Transcripts of *Pax-3* were confined to a thin layer of cells lining the ventricles of the myelencephalon and metencephalon (Figure 7M). No expression of *Pax-3* was seen in the developing choroid plexus consisting of post-mitotic cells derived from the thin roof plate of the fourth ventricle (Figure 7M; Figure 8, lower panel). The observation that *Pax-3* is not expressed in the choroid plexus parallels the absence of *Pax-3* transcripts in the ependyma following the degeneration of the ventricular zone.

The neural folds and dorsal neural tube give rise to a population of cells, the neural crest, that generates the peripheral nervous system and skull of the vertebrate embryo. In the early embryo, *Pax-3* expression clearly overlaps with those regions of the neural tube from which neural crest cells are known to migrate. In 8.5 day embryos, transcripts of *Pax-3* were clearly visible in regions of the recently closed neural tube from which neural crest cells migrate (Figure 8B). While it was difficult to visualize migrating neural crest cells during early mouse embryogenesis, *Pax-3* transcripts were clearly discernable in the spinal ganglia of 10−12 day embryos. Parasagittal sections through 10 day embryos revealed strong expression of *Pax-3* in the developing dorsal root ganglia positioned between the neural tube and the somitic mesoderm (Figure 7H; Figure 8F). From days 10 to 13, expression of *Pax-3* was also observed in cranio-facial neural crest derivatives. Cells expressing *Pax-3* were also present in cephalic neural crest cells in the nasal process and in structures derived from the first and second branchial arches. In 12−13 day embryos, *Pax-3*-expressing cells were found in the tongue, the mandible and maxilla (Figure 7M and N). Not all cells of the facial mesoectoderm were positive suggesting that only a subpopulation of cranial neural crest cells express *Pax-3*. Embryos were also examined for expression in other structures known to be derived from the neural crest. At no time was *Pax-3* expression detected in melanocytes, chromaffin granule cells, the developing heart or sympathetic ganglia, indicating that *Pax-3* may not be expressed in these neural crest lineages.

Expression of Pax-3 in somitic mesoderm

Transcripts of *Pax-3* were detected in the segmented mesoderm from 8.5 to 11 days of mouse embryogenesis. Transcripts of *Pax-3* in newly formed somites of an 8.5 day embryo can be seen clearly in Figure 7. Expression was observed in dorso-laterally located mesoderm soon after the closure of the neural tube and prior to the formation of the dermomyotome caps (Figure 8A,B,J,K). In 10 day embryos, parasagittal sections revealed a segmented pattern of *Pax-3* expression in the dermomyotome, in both anterior and posterior halves of the somite (Figure 7C,H). This metameric pattern of *Pax-3* expression in somites was present along the entire A−P axis of the embryo. The onset of *Pax-3* expression appeared to coincide closely with the division of the presegmental mesoderm into discrete somites and preceded the formation of the dermomyotome and sclerotome. Beginning at day 11, the expression of *Pax-3* in the anterior-most somites in the mouse began to disappear coinciding with the dissociation of somites and the migration of constituent cells that will form the dermis, the musculature and skeletal structures. This study was unable to determine whether *Pax-3* is expressed in the dermatome or myotome, however no expression of *Pax-3* was seen in migrating myoblasts which are known to express *Pax-7* (Jostes *et al.*, 1990), while expression in migrating dermatome cells would be difficult to visualize. In 12−14 day embryos, expression of *Pax-3* was still seen in the caudal-most somites reflecting the delayed development of these structures in the posterior part of the embryo (data not shown). *Pax-3* transcripts were detected in two further structures in the mouse embryo. From day 10 to day 11, expression of *Pax-3* was observed in the undifferentiated mesenchyme of both the forelimb and hindlimb. In 10 day embryos, diffuse expressions was present throughout the distal limb bud mesenchyme (Figure 7I and L). *Pax-3* expression was also visible in the epithelial cells of the nasal pit in 11 and 13 day embryos (Figure 7O).

Discussion

The structure of the Pax-3 protein

In this paper we describe the isolation and characterization of a new member of the murine paired box gene family, *Pax-3*, which contains a paired-type homeodomain as well as the paired domain. Comparison of *Pax-3* with other known paired proteins reveals three conserved domains within the Pax-3 protein, the paired domain, the octapeptide and the homeodomain. Aside from these three motifs there are no further conserved sequences in the protein when

compared with the *Drosophila* prd, gsb-p and gsb-d proteins. In particular, a fourth conserved domain, the paired repeat, found in the *Drosophila paired* and *bicoid* proteins (Frigerio *et al.*, 1986) is not present in Pax-3.

Comparison of the *Pax-3* sequence with the recently reported sequences for three human paired box genes indicates that *Pax-3* is the murine homologue of the human *HuP2* gene. No other murine paired box gene isolated to date is as homologous to *HuP2* as *Pax 3*. The Pax-3 protein is identical to the HuP2 protein in both the paired and octapeptide domains (Burri *et al.*, 1989) and in the highly charged region of 21 amino acids spanning these two domains. Furthermore, the intron−exon boundaries in the paired box of both genes are identical, as are the sizes of the intervening introns (M.Goulding, unpublished results; Burri *et al.*, 1989). The human *HuP2* gene has been characterized only on the basis of genomic sequences that encompass part of the *HuP2* coding region (Burri *et al.*, 1989). In view of the extended homology to *Pax-3* over this region, we would predict the presence of a paired-type homeodomain in the human HuP2 protein. We have recently mapped the murine *Pax-3* gene to mouse chromosome 1 between the *leaden* (*ln*) and *tumbler* (*tb*) loci (Walther *et al.*, submitted). Since this region of mouse chromosome 1 is syntenic with a region 2q on human chromosome 2 bounded by the isocitrate dehydrogenase and γ-crystallin genes (Nadeau, 1989), it seems likely that *HuP2* will also map to this region of human chromosome 2. It should be noted that *Pax-3* maps very closely to two known mouse developmental mutants, *Splotch* (*Sp*) and *Double foot* (*Dbf*) (M.Lyon, personal communication) both of which exhibit abnormalities in structures that express *Pax-3* during embryogenesis.

Pax-3 shows extensive homology with the murine *Pax-7* gene (Jostes *et al.*, 1990). This homology at the protein level extends from the paired domain through to the homeodomain of both proteins. The paired domains of Pax-3 and Pax-7 are very similar with only a few conservative exchanges restricted to the 54 amino acid carboxyl region of the paired domain (Figure 2, Jostes *et al.*, 1990). The paired-type homeodomain encoded by both genes is also highly conserved (Figure 2). Consequently, the murine *Pax-3* and *Pax-7* genes along with the human *HuP2* and *HuP1* genes constitute a family of closely related genes that may have been derived from a common ancestral gene. Interestingly, the murine *Pax-3* and *Pax-7* genes show similar overlapping patterns of expression during embryogenesis (Goulding *et al.*, in preparation). This phenomenon has been seen with two other closely related genes, *Pax-2* and *Pax-8*, both of which exhibit overlapping patterns of expression in the developing excretory and nervous system (Dressler *et al.*, 1990; Plachov *et al.*, 1990).

Pax-3 expression during embryogenesis

The temporally and spatially restricted pattern of *Pax-3* expression in the mouse embryo indicates that *Pax-3* may be an important morphogenetic determinant during embryonic development. The observation that *Pax-3* is not expressed in the adult organisms and is only expressed during a discrete period during embryogenesis supports this hypothesis. Although expression of *Pax-3* was not analysed prior to day 7 of gestation, a number of observations indicate that *Pax-3* may not be expressed prior to neurulation. In particular, *Pax 3* mRNA was not detected in embryonic stem cells (S.Sterrer, personal communication) or in two undifferentiated embryonal carcinoma cell lines (Figure 6B), which are considered to be representative of the preimplantation embryo (Graham, 1977). In F9 and P19 teratocarcinoma cells and in PCC7 cells (data not shown) the induction of differentiation by retinoic acid led to a strong increase in *Pax-3* mRNA levels, demonstrating that during embryogenesis *Pax-3* expression may be responsive to the morphogen, retinoic acid.

The vertebrate nervous system arises from a pseudostratified neuroepthelium that closes to form the neural tube (Altman and Bayer, 1984). The primitive neural tube consists of radially orientated mitotic cells that are attached to the luminal surface by tight junction (Holley, 1982). During the cell cycle, cell bodies migrate between the ventricular surface during mitosis and the pial surface during S phase (Sauer, 1935, 1959). When cells in the ventricular zone become postmitotic, they migrate radially to specific positions in the intermediate zone and marginal zones where they differentiate. The complex structure of the adult nervous system is generated from discrete regional populations of cells that arise during early neurogenesis (Jacobson, 1985). Although very little is known of the mechanisms that direct this regionalization, it seems likely to involve differential gene expression in response to inductive signals from the chorda-mesoderm. A dorsal−ventral polarity is evident in the developing neural tube. Cells in the dorsal neural tube give rise to the neural crest, commissural neurons and the microneurons of the substantia gelatinosa, while motor neurons arise from ventrally located cells (McConnell, 1981; Altman and Bayer, 1984; Leber *et al.*, 1990). There is evidence that the notochord and floorplate play a primary role in generating this dorsal−ventral gradient within the neural tube (van Straaten *et al.*, 1988; Placzek *et al.*, 1990). *Pax-3* transcripts are detected in the dorsal neuroepithelium just prior to closure of the neural tube and remain restricted to dorsally located mitotic neuroepithelial cells throughout early neurogenesis (Figure 8). This expression pattern may be a primary response to inductive signals from the notochord and floorplate that regionalize the neural tube and as a result *Pax-3* could play an active role in generating the dorsal−ventral polarity within the neural tube.

An important landmark in the early spinal cord is the sulcus limitans, a longitudinal widening of the lumen that coincides with the boundary of the alar and basal plates (His, 1888). The alar plate gives rise to the dorsal grey laminae including the substantia gelatinosa, while the basal plate generates the ventral and intermediate grey areas that contain somatic and autonomic motor neurons. Recently, cell lineage studies in the embryonic chick spinal cord indicate that the movement of cells between the alar and basal plates is restricted (Leber *et al.*, 1990). Consequently, the sulcus limitans delineates a boundary between two distinct populations of neuronal precursor cells, one in the alar and roof plates, the other in the basal plates. Our analysis of *Pax-3* expression during early neurogenesis reveals that *Pax-3* exhibits a sharp ventral border of expression in the ventricular zone of the spinal cord that coincides with the sulcus limitans (Figures 7 and 8). Expression of *Pax-3* is found throughout the alar plate and roof plate, whilst being completely absent from the basal and floor plate regions. While, the present study reveals *Pax-3* as a genetic marker for the cells in the alar and roof plate, the fact that *Pax-3*

expression is restricted to the dorsal neural tube prior to the formation of the alar and basal plates raises the possibility that *Pax-3* has an active role in the genesis of these compartments.

It should also be noted that *Pax-3* expression in the developing nervous system appears to co-localize with populations of neuronal stem cells contributing to neural structures having a major somatosensory input. In the early embryo *Pax-3* expression is found in the dorsal neural tube, spinal ganglia and the primordia of the medulla oblongata, cerebellum, superior colliculi and thalamus. Expression of *Pax-3* is also observed in primordia of target tissues for peripheral neurons, namely the dermomyotome, the nasal process where vibrissae will form and in the nasal pit epithelium (Figure 8). Nevertheless, the functional significance of this expression pattern still needs to be demonstrated.

Comparison with Drosophila developmental control genes

The murine *Pax* genes were isolated on the basis of their homology to the paired domain present in the *Drosophila* pair-rule and segment identity genes *paired*, *gooseberry*, *Pox-meso* and *Pox-neuro*. These genes are part of the genetic hierarchy which establishes the segmented body plan of *Drosophila*. Anatomical structures that are reminiscent of a segmented body plan are also present in vertebrates, in the form of somites, blocks of mesoderm tissue that are present transiently in the early vertebrate embryo. In higher vertebrates, the remnants of this segmented body plan present during embryogenesis are seen in the vertebral column, intercostal musculature, segmented peripheral nervous system and dermatomes that are innervated by the peripheral nervous system. During murine embryogenesis, genes containing a paired box are also expressed in segmented structures like their *Drosophila* segmentation gene counterparts. *Pax-1* is expressed in the sclerotome cells that will form the intervertebral disc anlagen (Deutsch *et al.*, 1988), while *Pax-2* and *Pax-8* are expressed in the segmentally arranged mesonephric tubules that are phylogenetically derived from the nephrotomes (Dressler *et al.*, 1990; Plachov *et al.*, 1990). *Pax-3* also exhibits a metameric pattern of expression in somites, where expression is confined to the dermomyotome. While *Pax-3* expression in somites was detected as early as day 8.5, expression appears to be absent from the presomitic mesoderm suggesting that *Pax-3* is not involved in the primary segmentation of the somitic mesoderm. *Pax-3* may in part regulate dermomyotome formation, since *Pax-3* transcripts appear in the dorsolaterally located cells of each somite (i.e. presumptive dermomyotome) prior to the formation of the sclerotome and dermomyotome.

Many of the genes that are involved in *Drosophila* segmentation are also expressed during *Drosophila* neurogenesis. Altered neuronal phenotypes have been observed when the pair-rule genes *fushi tarazu* (Doe *et al.*, 1988a), *evenskipped* (Doe *et al.*, 1988b) and the segment polarity genes *gooseberry, patched* and *wingless* (Patel *et al.*, 1989) are mutated, while *Pox-neuro* is expressed in neuroblasts at the late germ band stage of *Drosophila* development (Bopp *et al.*, 1989). As such these genes may have a dual role in shaping the segmented nervous system of invertebrates by regulating primary segmentation and then the neuronal identity of differentiating neuroblasts within these segments. A number of vertebrate homologues of these *Drosophila* genes are also expressed in the developing nervous system. While many of the *Antp*-like *Hox* genes exhibit restricted expression along the A–P axis (Holland and Hogan, 1988; Graham *et al.*, 1989; Wilkinson *et al.*, 1989), analysis of the expression of the vertebrate homologues of *even-skipped* (*(Evx-1)*, *wingless* (*Wnt-1*), *En-1*, *Pax-2* and *Pax-3* demonstrate that these genes are expressed in transversely defined longitudinal columns of cells along the entire spinal cord (Wilkinson *et al.*, 1987; Davidson *et al.*, 1988; Davis *et al.*, 1988; Bastian and Gruss, 1990; Nornes *et al.*, 1990). This is not surprising because apart from the spinal ganglia, the vertebrate spinal cord shows little in the way of segmental organization. The longitudinal organization of the spinal cord appears to be a necessary development in higher vertebrates, allowing the relay of information to and from the brain. In this respect, the pattern of expression of *Pax-3* in the developing nervous system is particularly interesting, since a segmented pattern of expression has been maintained in the peripheral nervous system but not in the central nervous system. The pattern of *Pax-3* expression in the neural tube, neural crest and somitic mesoderm agree with the hypothesis that compartmentalization of the spinal cord is mainly longitudinal and any segmentation present arises from the peripheral nervous system. As a result, a gene specifying a particular neuronal compartment would be expressed in a continuous pattern along the A–P axis of the spinal cord. Consequently, proteins sharing conserved domains may regulate similar developmental processes within the context of the differing strategies utilized for organizing the invertebrate and vertebrate nervous system.

Is Pax-3 a transcription factor?

The presence of a homeodomain in the Pax-3 protein suggests that *Pax-3* may regulate gene expression, since *Drosophila* and vertebrate homeodomain-containing proteins can modulate transcription (Jaynes and O'Farrell, 1988; Sturm *et al.*, 1988; Bodner *et al.*, 1988). However, until recently it was not known if the paired domain was also involved in protein interactions with DNA. Although the *Drosophila Pox-meso* and *Pox-neuro* proteins that contain a paired domain alone are localized in the nucleus of *Drosophila* embryos, their ability to bind DNA has not been demonstrated (Bobb *et al.*, 1989). The *paired* protein is able to bind *in vitro* to two recognition sequences upstream of the *Drosophila even-skipped* gene known as e4 and e5 (Hoey and Levine, 1988; Treisman *et al.*, 1989). The e5 sequences consists of two overlapping recognition elements, one that is recognized by the paired domain and the other by the homeodomain (C.Desplan, personal communication; M.Goulding, unpublished results). We have now demonstrated that Pax-3 protein is also able to specifically recognize the e5 recognition sequence and both recognition elements are required for high affinity Pax-3 protein binding (Figure 5). This indicates that both the paired domain and the homeodomain are involved in DNA binding by the Pax-3 protein. Presently, it is unclear whether the paired domain of Pax-3 alone can bind with high affinity to DNA.

We have been unable to demonstrate any activation of transcription by the Pax-3 protein in P19 cells using the *Drosophila even-skipped* e5 recognition sequence as a target for Pax-3 protein binding. This is not surprising since the *Drosphila even-skipped* gene is not regulated by paired

(Frasch and Levine, 1987). Nevertheless, the observation that transcriptional activation can be conferred on the Gal4 DNA-binding domain when sequences encompassing the carboxyl domain of Pax-3 were fused to the DNA-binding domain of Gal4 (G.Chalepakis and M.Goulding, unpublished results) indicates that Pax-3 is able to activate gene transcription. Interestingly, the carboxyl terminus of the Pax-3 protein contains a proline-serine-threonine rich domain rather like a domain in the Oct-2 protein that is obligatory for transcriptional activation (Tanaka and Herr, 1990). Further target sequences for Pax-3 binding need to be identified, since it is likely that Pax-3 binding to the e5 sequence is unable to confer transcriptional regulation.

Materials and methods

Screenings of an embryonic mouse cDNA library

A first cDNA clone encoding *Pax-3* designated pBH3.2 was initially isolated from an 8.5 day embryonic mouse γgt10 cDNA library (Fahrner *et al.*, 1987). After transfer to Hybond-N filters (Amersham), 8×10^5 clones were hybridized under low stringency conditions (7 × SSC, 60°C) with a 313 bp *Hin*cII–*Sac*I *Pax 1* paired box probe, labelled by random priming (Feinberg and Vogelstein, 1983). The filters were washed three times in 2 × SSC/0.2%SDS at 42°C. A second overlapping *Pax-3* cDNA clone pMG-33 was isolated using a 350 bp *Hin*dIII–*Sac*I genomic fragment encompassing the 5′ most sequences of pBH3.2 1.6×10^6 phages of the 8.5 day embryonic mouse cDNA library were hybridized under high stringency conditions (4 × SSC, 68°C). As a final step, filters were washed in 0.2 × SSC, 0.1% SDS at 68°C. Genomic *Pax-3* clones were isolated by screening a C57B6 mouse DNA library made in EMBL-3, kindly provided by K.Imai.

DNA sequencing

All nucleotide sequences were determined using the dideoxy chain termination method (Sanger *et al.*, 1977). Use was made of the USB Sequenase kit for all sequencing reactions.

Isolation of RNA from cells, tissues and embryos

Embryos were obtained from naturally mated NMRI mice with the day of the vaginal plug designated day 0. Mouse embryos were collected from 9–17 days of gestation, free of extra-embryonic tissue and stored at −70°C prior to use. Tissues were dissected out of adult animals using standard techniques. RNA was prepared from embryos and tissues by homogenization in guanidinium-thiocyanate (Chirgwin *et al.*, 1979) followed by centrifugation through a 5.7 M caesium chloride, 25 mM sodium acetate (pH 5.0) cushion for 24 h in a Beckman SW 40 rotor at 31 000 r.p.m. Poly(A)$^+$ RNA was prepared by oligo(dT) cellulose chromatography (Maniatis *et al.*, 1982). RNA was prepared from tissue culture cells, lysed with NP-40 in the presence of vanadyl-ribonucleoside complexes. Nuclei were removed by centrifugation (1000 *g* for 5 min at 4°C). RNA was ethanol precipitated after proteinase K-treatment followed by phenol–chloroform extraction (Maniatis *et al.*, 1982).

Northern hybridization

RNA samples containing either 20 μg or poly(A)$^+$ RNA or 30 μg of total RNA were electrophoresed through 1.2% agarose gels in the presence of 3.7% of formaldehyde and 1 × MOPS buffer (20 mM morpholinopropane sulphonic acid, pH 7.0, 50 mM sodium acetate, 10 mM EDTA) and transferred to Genescreen plus membranes. Membranes were hybridized overnight with a [^{32}P]-oligolabelled 516 bp *Hin*dIII–*Pst*I fragment of *Pax-3* in 50% formamide, 7.5% dextran sulphate, 5 × SSC, 1% SDS, 10 mM sodium phosphate pH 6.8 and 100 μg/ml denatured salmon sperm DNA. Blots were washed twice in 2 × SSC, 1% SDS at 65°C, then twice in 0.1 × SSC, 0.1% SDS at 65°C for 20 min.

In situ hybridization

A Bluescript vector containing a 516 bp *Hin*d–*Pst*I fragment from the 3′ end of the *Pax-3* gene was linearized and single-stranded RNA probes were transcribed *in vitro* using 100 μCi[^{35}S]UTP and either T3 or T7 polymerase (Promega Biotech). Details of the *in situ* hybridization procedure are described in Erselius *et al.* (1990).

In vitro transcription and translation

Plasmid pBH3.2 was linearized using the *Bam*HI and *Sal*I restriction sites in the Bluescript polylinker and the sense and anti-sense RNA strands were synthesized with the T3 and T7 RNA (Promega) polymerases, respectively. The reaction conditions were 40 mM Tris–HCl pH 8.0, 8 mM $MgCl_2$, 25 mM NaCl, 2 mM spermidine, 5 mM DTT, 2 mM dNTPs, 2 mM m7G(5′)ppp5′G, 2 μg template and 2 μl enzyme. 1/50 of the RNA template was translated in a rabbit reticulocyte lysate following the manufacturer's specifications (BRL).

Preparation of antibodies

Peptides from selected regions of the Pax-3 protein was synthesized by FMOC synthesis on a Milligen-Biosearch 9050 peptide synthesizer. The peptides used to produce antibodies were as follows, M-3 (QSQKKAKHSIDGIC*) corresponding to amino acids 170–191 of Pax-3 and M-4 (C*PVTGYQYGQYGQS) corresponding to amino acids 462–474. Peptides were sequenced by Edman degradation prior to coupling. Peptides were coupled via a cysteine residue (*) to bovine serum albumin (BSA) using maleimido-benzoyl-*N*-hydroxysuccinimide ester. Rabbits were immunized with 500 μg of coupled peptide in Freund's complete adjuvant using standard procedures. Rabbits were boosted twice at four weekly intervals and blood was collected 10 days after the final boost. Antibodies to M-3 (AbM-3) were purified by ammonium sulphate precipitation followed by affinity purification to the M-3 peptide coupled to Sepharose–EAH. Antibodies to M-4 were purified by ammonium sulphate precipitation and then antibodies to BSA were removed by affinity chromatography with BSA coupled to Sepharose-4B. All procedures used are described in Harlow and Lane (1988).

Nuclear extract preparation and gel shift assays

P19 cells containing *Pax-3* coding sequences under the control of the mouse heat shock promoter were induced by heat shocking cells to 42°C for 2 h. Cells were then left for 3 h at 37°C to recover before extracts were prepared. Cell extracts were prepared as described previously (Dignam *et al.*, 1983). DNA binding and gel electrophoresis conditions were carried out as described before, with the exception that binding was performed at 30°C (Schöler *et al.*, 1989). A double-stranded DNA fragment (CTCAGCAC CGCACGATTAG**CACCGTTCCGCTTC**) encompassing the e5 *paired* protein recognition sequence from the upstream region of the *Drosophila even-skipped* gene was used for binding assays (Hoey and Levine, 1988; C.Desplan, personal communication). The recognition sequence bound by the homeodomain is underlined and the paired domain recognition sequence is shown in bold print.

Oligonucleotides

45. TCGACTAGTCACGATTAGCACCGTTCCGCTCTAGATATC
89. TCGACTAGTTCAGCACCGCACGATTAGCATCTAGATATC
110. TCGACTCTCAGCACCGCACGATTAGCACCGTTCCGCTCC
14. TCGACTAGTCACGATTAGCACCTTTCCGCTCTAGATATC
22. TCGACTAGTCACTGGGCTCACCGTTCCGCTCTAGATATC
24. TCGACGTTGACTAGCAGAGAAAACAAAGTGAGTCTAGAC
ANTP-1 TCGACTCAATTAAATGTCAATTAAATAGATCTC

Acknowledgements

We would like to thank H.Bastian for providing the embryonic mRNA, H.-P.Geithe for synthesizing oligonucleotides and peptides, G.Kristjansson for his advice on generating antibodies, R.Altschäffel for the photography and B.Fohring for help with the computer analysis. Many thanks to A.Püschel for his helpful comments on the manuscript. This work was supported by the Max Planck Society.

References

Akam,M. (1987) *Development*, **101**, 1–22.
Altman,J. and Bayer,S.A. (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.
Bastian,H. Gruss,P. (1989) *EMBO J.*, **9**, 1839–1852.
Baumgartner,S., Bobb,D., Burri,M. and Noll,M. (1987) *Genes Dev.*, **1**, 1247–1267.

Bodner,M., Castrillo,J-L., Theill,L.E., Derrinck,T., Ellisman,M. and Karin,K. (1988) *Cell*, **55**, 505–518.
Bobb,D., Burri,M., Baumgartner,S., Frigerio,G. and Noll,M. (1986) *Cell*, **47**, 1003–1040.
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.
Caudy,M., Vaessin,H., Brand,M., Tuma,R., Jan,L.Y. and Jan,Y.N. (1988) *Cell*, **55**, 1061–1067.
Chirgwin,J.M., Przybyla,A.E., MacDonald,R.J. and Rutter,W.J. (1979) *Biochemistry*, **18**, 5294–5299.
Davidson,D., Graham,E., Sime,C. and Hill,R. (1988) *Development*, **104**, 305–316.
Davis,C.A., Noble-Topham,S.E., Rossant,J. and Joyner,A.L. (1988), *Genes Dev.*, **2**, 361–371.
Deutsch,U., Dressler,G.R. and Gruss,P. (1988) *Cell*, **53**, 617–625.
Dignam,J.D., LeBowitz,R.M. and Roeder,R.G. (1983) *Nucleic Acids Res.*, **11**, 1475–1489.
Doe,C.W., Hiromi,Y., Gehring,W.J. and Goodman,C.S. (1988a) *Science*, **239**, 170–175.
Doe,C.Q., Smouse,D. and Goodman,C.S. (1988b) *Nature*, **333**, 376–378.
Dressler,G.R., Deutsch,U., Balling,R., Simon,D., Guenet,J.L. and Gruss,P. (1988) *Development*, **104** (Suppl.) 181–186.
Dressler,G.R., Deutsch,U., Chowdhury,K., Nornes,H.O. and Gruss,P. (1990) *Development*, **109**, 787–795.
Erselius,J.R., Goulding,M.D. and Gruss,P. (1990) *Development*, **110**, 629–642.
Fahrner,K., Hogan,B.L.M. and Flavell,R.A. (1987) *EMBO J.*, **6**, 1269–1271.
Feinberg,A.P. and Vogelstein,B. (1983) *Anal. Biochem.*, **132**, 6–13.
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.
Gehring,W.J., Müller,M., Affolter,M., Percival-Smith,A., Billeter,M., Qian,Y.Q., Otting,G. and Wüthrich,K. (1990) *Trends Genet.*, **6**, 323–329.
Goulding,M.D. and Gruss,P. (1989) *Current Opinion Cell Biol.*, **1**, 1088–1093.
Graham,C.F. (1977) *Embryonic and Fetal Development*. p. 159–180. Cambridge University Press, Cambridge.
Graham,A., Papaplopolu,N. and Krumlauf,R. (1989) *Cell*, **57**, 367–378.
Grüneberg,H. (1954) *J. Genet.*, **52**, 441–455.
Harlow,E. and Lane,D. (1988) *Antibodies. A Laboratory Manual.* Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
His,W. (1988) *Abh. K. Sachs. Ges. Wiss. Math.-Phys. Kl.*, **14**, 341–372.
Hoey,T. and Levine,M. (1988) *Nature*, **332**, 858–861.
Hogan,B., Constantini,F. and Lacy,E. (1986) *Manipulating the Mouse Embryo: A Laboratory Manual.* Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY. pp. 228–235.
Holley,J.A. (1982) *J. Comp. Neurol.*, **205**, 371–382.
Holland,P.W.H. and Hogan,B.L.M. (1988) *Genes Dev.*, **2**, 773–782.
Ingham,P.W. (1988) *Nature*, **335**, 25–34.
Jacobson,M. (1985) *Annu. Rev. Neurosci.*, **8**, 71–102.
Jaynes,J.B. and O'Farrell, (1988) *Nature*, **336**, 744–749.
Jostes,B., Walther,C. and Gruss,P. (1990) *Mech. Dev.*, **33**, 27–37.
Kessel,M., Balling,R. and Gruss,P. (1989) *Cell*, **61**, 301–308.
Kozak,M. (1986) *Cell*, **44**, 283–292.
Kozak,M. (1987) *Nucleic Acids Res.*, **15**, 8125–8146.
Leber,S., Breedlove,S.M. and Sanes,J.R. (1990) *J. Neurosci.*, **10**, 2451–2482.
Maniatis,T., Fritsch,E.F. and Sambrook,J. (1982) *Molecular Cloning: A Laboratory Manual.* Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
McConnell,J.A. (1981) *J. Comp. Neurol.*, **200**, 273–288.
Murre,C., McCaw,P.S. and Baltimore,D. (1989) *Cell*, **56**, 777–783.
Nadeau,J. (1989) *Trends Genet.*, **5**, 83–86.
Nornes,H.O. and Carry,M. (1978) *Brain Res.*, **159**, 1–16.
Nornes,H.O., Dressler,G.R., Knapik,E.W. and Gruss,P. (1989) *Development*, **109**, 797–805.
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.
Placzek,M., Tessier-Lavigne,M., Yamada,T., Jessell,T. and Dodd,J. (1990) *Science*, **250**, 985–988.
Rosenberg,U.B., Schröder,C., Preiss,A., Kienlen,A., Cote,S., Riede,I. and Jäckle,H. (1986) *Nature*, **319**, 336–339.
Ruiz i Altaba,A. and Melton,D (1988) *Cell*, **57**, 317–326.
Sanger,F., Nicklen,S. and Coulson,A.R. (1977) *Proc. Natl. Acad. Sci. USA*, **74**, 5463–5467.
Sauer,F.C. (1935) *J. Comp Neurol.*, **62**, 377–405.
Sauer,M.E. (1959) *Anat. Record*, **133**, 456.
Schöler,H.R., Hatzopoulos,A.K., Balling,R., Suzuki,N. and Gruss,P. (1989) *EMBO J.*, **8**, 2543–2550.
Scott,M.P. and Carroll,S.B. (1987) *Cell*, **51**, 689–698.
Sturm,R.A., Das,G. and Herr,W. (1988) *Genes Dev.*, **2**, 1582–1599.
Tanaka,M. and Herr,W. (1990) *Cell*, **60**, 375–386.
Treisman,J., Gönczy,P., Vashintha,M., Harris,E. and Desplan,C. (1989) *Cell*, **59**, 553–562.
van Straaten,H.W.M., Hekking,J.W.M., Beursgens,J.P.W.M., Terwindt-Rouwenhoest,E. and Drukker,J. (1989) *Development*, **107**, 793–803.
Wilkinson,D.G., Bailes,J.A. and McMahon,A. (1987) *Cell*, **50**, 79–86.
Wilkinson,D.G., Bhatt,S., Cook,M., Boncinelli,E. and Krumlauf,R. (1989) *Nature*, **341**, 405–409.
Wright,C.V.E., Cho,K.W.Y., Hardwicke,J., Collins,R. and De Robertis,E.H. (1989a) *Cell*, **59**, 81–93.
Wright,C.V.E., Cho,K.W.Y., Oliver,G. and De Robertis,E.M. (1989b) *Trends Biochem. Sci.*, **14**, 52–56.

Received on December 5, 1990; revised on February 7, 1991