MOLECULAR AND CELLULAR BIOLOGY, May 1995, p. 2858–2871 Vol. 15, No. 5
0270-7306/95/$04.00+0

DNA-Binding and Transactivation Properties of Pax-6: Three Amino Acids in the Paired Domain Are Responsible for the Different Sequence Recognition of Pax-6 and BSAP (Pax-5)

THOMAS CZERNY AND MEINRAD BUSSLINGER*

Research Institute of Molecular Pathology, A-1030 Vienna, Austria

Received 28 September 1994/Returned for modification 7 November 1994/Accepted 17 February 1995

Pax-6 is known to be a key regulator of vertebrate eye development. We have now isolated cDNA for an invertebrate Pax-6 protein from sea urchin embryos. Transcripts of this gene first appear during development at the gastrula stage and are later expressed at high levels in the tube foot of the adult sea urchin. The sea urchin Pax-6 protein is highly homologous throughout the whole protein to its vertebrate counterpart with the paired domain and homeodomain being virtually identical. Consequently, we found that the DNA-binding and transactivation properties of the sea urchin and mouse Pax-6 proteins are very similar, if not identical. A potent activation domain capable of stimulating transcription from proximal promoter and distal enhancer positions was localized within the C-terminal sequences of both the sea urchin and mouse Pax-6 proteins. The homeodomain of Pax-6 was shown to cooperatively dimerize on DNA sequences consisting of an inverted repeat of the TAAT motif with a preferred spacing of 3 nucleotides. The consensus recognition sequence of the Pax-6 paired domain deviates primarily only at one position from that of BSAP (Pax-5), and yet the two proteins exhibit largely different binding specificities for individual, naturally occurring sites. By creating Pax-6-BSAP fusion proteins, we were able to identify a short amino acid stretch in the N-terminal part of the paired domain which is responsible for these differences in DNA-binding specificity. Mutation of three Pax-6-specific residues in this region (at positions 42, 44, and 47 of the paired domain) to the corresponding amino acids of BSAP resulted in a complete switch of the DNA-binding specificity from Pax-6 to BSAP. These three amino acids were furthermore shown to discriminate between the Pax-6- and BSAP-specific nucleotide at the divergent position of the two consensus recognition sequences.

The *Pax* genes constitute a family of developmental control genes that were initially identified in *Drosophila melanogaster* (39). Nine members of the vertebrate paired box-containing (*Pax*) gene family have since been isolated by homology to these *Drosophila* genes (48, 56). Their distinct spatiotemporal expression pattern in the vertebrate embryo has implicated these genes in the control of morphogenesis and pattern formation (reviewed in reference 22). Indeed, mutations in *Pax-1*, *Pax-3*, and *Pax-6* have been associated with *undulated* (3), *Splotch* (14), and *Small eye* (26) mouse developmental mutants, respectively. Moreover, the *Pax-5* gene coding for the transcription factor BSAP has recently been shown to play a key role in early B lymphopoiesis and patterning of the midbrain (53). In humans, genetic lesions in the *PAX-3* gene are known to generate Waardenburg's syndrome (2, 50), while mutations in the *PAX-6* gene cause the disorders aniridia and Peters' anomaly (24, 51). *PAX-3* and *PAX-7* have, furthermore, been implicated in the genesis of alveolar rhabdomyosarcomas, as these myogenic tumors are characterized by specific chromosome translocations resulting in the fusion of one of the two *PAX* genes to a gene of the fork head transcription factor family (12, 17, 46). A potential role of *Pax* genes in oncogenesis has been further suggested by ectopic expression experiments in fibroblasts (38).

The vertebrate *Pax-6* genes code for a subfamily of Pax proteins which show extraordinarily high conservation not only at the sequence level but also in their expression pattern (33, 36, 51, 55). During embryogenesis, the *Pax-6* gene is expressed in discrete regions of the forebrain and hindbrain as well as along the entire neural tube, where its dorsoventral expression pattern is determined by signals from the notochord and floor plate (21, 55). *Pax-6* expression is additionally seen in the developing eye and nasal primordia of the embryo (55). Consistent with this expression pattern, homozygous *Small eye* mice show a complete failure of lens induction, lack nasal structures, and exhibit dramatic dysgenesis of the entire forebrain (18, 19, 26).

Pax proteins are known to be transcriptional regulators which recognize their target genes via the DNA-binding function of the paired domain (1, 6, 31, 52, 61). The paired domain is a highly conserved motif of 128 amino acids which does not have any obvious sequence homology with other known protein domains. Detailed DNA-binding studies of the transcription factor BSAP (Pax-5) revealed a bipartite structure of the paired domain and its binding site and led to the definition of a consensus recognition sequence which is bound by all Pax proteins analyzed so far (10). The functional importance of the paired domain is illustrated by the fact that missense mutations interfering with DNA binding of the paired domain are associated with mouse developmental mutants and human disease syndromes (reviewed in reference 49).

Here, we report the cDNA cloning and characterization of the sea urchin Pax-6 protein. Although echinoderms and vertebrates have diverged long ago in evolution, the sea urchin Pax-6 protein exhibits extensive sequence homology to its mouse counterpart. Consequently, no differences between the two Pax-6 proteins could be detected in DNA-binding assays. Pax-6 and BSAP (Pax-5), which are members of two distinct subfamilies of Pax proteins (56), differ in the DNA-binding specificities of their paired domains. On the basis of mutational

* Corresponding author. Mailing address: Research Institute of Molecular Pathology, Dr. Bohr-Gasse 7, A-1030 Vienna, Austria. Phone: (43/1) 797 30-452. Fax: (43/1) 798 71 53.

analyses, these DNA-binding differences could be attributed to 3 amino acid residues within a 7-amino-acid stretch which is responsible for discriminating between a Pax-6- and a BSAP-specific nucleotide at the primary divergent position of the two consensus recognition sequences. Moreover, we identified a potent transactivation function in the C-terminal region of the sea urchin and mouse Pax-6 proteins which is capable of stimulating transcription from proximal promoter and distal enhancer positions.

MATERIALS AND METHODS

PCR screening and cDNA cloning. Poly(A)$^+$ RNA (5 to 10 μg) isolated from gastrula embryos of the sea urchin *Paracentrotus lividus* was transcribed into double-stranded cDNA by priming with $P(dT)_{15}$ as described elsewhere (23). After *Eco*RI linker addition, the cDNA was cloned into bacteriophage λgt10. Recombinant phages ($\sim 10^6$) were screened with the ^{32}P-labelled suPax-6 DNA probe A (Fig. 1A). Positive clones were rescreened and plaque purified. The *Eco*RI inserts were subcloned into the expression vector pKW10 (1), and DNA sequencing was then performed. The DNA probe A was obtained from gastrula cDNA (250 ng) by PCR amplification (1 min at 94°C; 2 min at 55°C; 4 min at 72°C; 35 cycles) with degenerate primers. The resulting fragment was gel purified, cloned into the *Eco*RI and *Hin*dIII sites of plasmid pSP64, and sequenced. The degenerate primers 5′GCGGAATTCGTNAA(C/T)CA(G/A)(C/T)TNG GNGGNGTNTT3′ and 5′CGCAAGCTT(C/T)(T/G)NGG(C/T)TTN(C/G)(T/A)NCCNCC(G/A/T)AT3′ were used. PCR and reverse transcription (RT)-PCR analyses of the *suPax-6* paired box region were performed with genomic DNA (100 ng) of *P. lividus* and total RNA (100 ng) from gastrula embryos, respectively, with the primers 5′GCGGAATTCGTGAATGGTCGTCCCTTGCC3′ and 5′CGCAAGCTTGGCTCGGGACGTATGCTCCC3′.

RNA preparation and analysis. *P. lividus* sea urchins were collected along the Côte d'Azur. Spawning, fertilization, and embryo culturing were carried out in Millipore-filtered sea water at 18°C. Embryos were collected at various stages, and RNA was prepared by the guanidinium thiocyanate method (8). RNase protection analysis was carried out according to Vitelli et al. (54). The *suPax-6* riboprobe was generated by subcloning a 166-bp PCR fragment (nucleotides 750 to 915 [Fig. 1B]) from the cDNA fragment A (Fig. 1A) in the antisense orientation into pSP64. In situ hybridization onto sections of *P. lividus* tube feet was performed as described elsewhere (1). Sense and antisense RNA probes were obtained by cloning of a 700-bp *Nhe*I-*Sca*I fragment from clone λ3 (Fig. 1A) into the polylinkers of pSP64 and pSP65, respectively.

Plasmids. Expression constructs were obtained by insertion of the different *Pax* cDNAs downstream of the cytomegalovirus (CMV) promoter of pKW10 (1). The mouse Pax-6 translation initiation sequence CCAGCTCCAGCATG was introduced upstream of the last in-frame ATG codon (Fig. 1B) of the *suPax-6*, *suPax-6*(1–363), and *suPax-6*(1–145) cDNAs by the PCR method. The unique *Nhe*I site present in the overlapping region of clones λ1 and λ3 was used for assembly of the full-length *suPax-6* cDNA clone. Cloning of the mouse *Pax-6* cDNA was previously described (10). A stop codon was introduced by PCR downstream of the paired domain of the *suPax-6* and *mPax-6* cDNAs to create *suPax-6*(1–145) and *mPax-6*(1–131). Two restriction sites of the mouse BSAP cDNA (1) were used for generating chimeric *suPax6-BSAP* and *BSAP-suPax6* cDNAs: a *Bam*HI site present in the original *BSAP* cDNA and a *Xba*I site which was created by introduction of two silent mutations at codons 86 and 87 (TCC AGG → TCT AGA) by a PCR-based mutagenesis method (10). *suPax-6* cDNA fragments, which were created by PCR amplification of clone λ1 with primers containing *Bam*HI or *Xba*I sites, were then used to replace the corresponding *BSAP* cDNA fragments in the expression plasmid pKW10-mBSAP (1). The GAL4 fusion constructs were generated by introduction of appropriate *suPax-6* and *mPax-6* PCR fragments into a pKW10 plasmid containing the DNA-binding domain (positions 1 to 147) of GAL4. To create the reporter plasmid lucCD19, three copies of the CD19-2(A-ins) oligonucleotide were inserted immediately upstream of a β-globin TATA box and initiator region linked to a luciferase gene. The plasmid plGC-luc was constructed by insertion of the GAL4-responsive promoter described by Braselmann et al. (5) upstream of the luciferase gene. The plasmid lucTK-GAL/CD19 was generated by introduction of a DNA fragment containing five GAL4-binding sites together with three multimerized CD19-2(A-ins) sites at position −109 of the thymidine kinase (TK) promoter in plasmid lucTK (44). A DNA fragment containing the GAL4 and CD19-2(A-ins) sites present in lucTK-GAL/CD19 was inserted at the *Pvu*II site downstream of the luciferase gene in plasmid lucTK to generate the reporter construct lucE-GAL/CD19.

EMSA analysis of Pax proteins synthesized in COP-8 cells or by in vitro translation. Pax expression vectors were transiently transfected into COP-8 cells by the DEAE-dextran method, and whole-cell extracts were prepared 2 days later as described elsewhere (1). The same CMV expression plasmids, which contained an SP6 promoter in the 5′ region, were used as templates for in vitro transcription, and the in vitro-synthesized RNA was translated in a nuclease-treated rabbit reticulocyte lysate as described elsewhere (10). All paired domain recognition oligonucleotides used for electrophoretic mobility shift assay (EMSA) have been previously described (10) or are shown in Fig. 7. A short P3 oligonucleotide (5′TCGACCCTAATCGCATTACCC3′ annealed with 5′TC GAGGGTAATGCGATTAGGG3′) and P2 and P2* derivatives thereof were used for DNA-binding analysis of the Pax-6 homeodomain (see Fig. 8B). The following longer P3 oligonucleotide was used as probe for EMSA analysis of the full-length Pax-6 protein (see Fig. 8C): 5′TCGAGGGCATCAGGATGCTAA TGCGATTAGCATCCGATCGGG3′ annealed with 5′TCGACCCGATCGGA TGCTAATCGCATTAGCATCCTGATGCCC3′. Annealed oligonucleotides were end labelled with Klenow fragment DNA polymerase. A 5-fmol sample of labelled DNA probe was incubated together with whole-cell extract or in vitro-translated proteins at room temperature in 20 μl of a buffer containing 1 μg of poly(dI-dC), 10 mM HEPES (*N*-2-hydroxyethylpiperazine-*N*′-2-ethanesulfonic acid, pH 7.9), 100 mM KCl, 4% Ficoll, and 1 mM EDTA. Protein-DNA complexes were separated on a native 6% polyacrylamide gel (in 0.25× Tris-borate-EDTA) and detected by autoradiography.

DNA-binding site selection. A PCR fragment containing *mPax-6* cDNA sequences from codon 207 to codon 294 (55) was inserted into the *Bam*HI site of plasmid pETH-2a (1). The resulting expression vector was used to synthesize the polyhistidine-tagged Pax-6HD protein in *Escherichia coli* BL21 (DE3, pLysS), and Pax-6HD was purified on Ni^{2+}-nitriloacetic acid (NTA) agarose as described elsewhere (1). The following 86-bp oligonucleotide was used for binding-site selection assay: 5′GCGGGATCCACTCCAGGCCGGATGCT$(N)_{35}$CACCAG GGTGTAAGGCGGATCCCGC3′. The second strand of this oligonucleotide was synthesized with Klenow fragment DNA polymerase and primer A (5′GCGGGATCCGCCTTACACCCTGGTG3′). This double-stranded oligonucleotide (1 μg) and the Pax-6HD polypeptide (0.6 μg) were incubated for 1 h at room temperature in 100 μl of a buffer containing 100 mM KCl, 20 mM HEPES (pH 7.9), and 0.5 μg of bovine serum albumin per ml, before the Ni^{2+}NTA silica resin (4 mg; Qiagen) was added. Following a further 30-min incubation, the Ni^{2+}NTA resin was washed four times with 20 mM HEPES (pH 7.9) containing increasing KCl concentrations (100, 150, and 200 mM KCl). Bound DNA was eluted with 300 mM KCl, phenol extracted, ethanol precipitated, amplified by PCR (30 cycles of 30 s at 94, at 60, and at 72°C) with primers A and B (5′GCGGGATCCACTCCAGGCCGGATGCT3′), and gel purified. After four rounds of selection, bound DNA fragments were cloned into the *Bam*HI site of pUC19 and sequenced.

Transient transactivation assays. The luciferase reporter gene (3 μg), the transactivator plasmid (0.2 to 8 μg, as indicated in Fig. 9 and 10), and the reference CMV chloramphenicol acetyltransferase (CAT) gene (0.1 μg) were transiently transfected by electroporation into J558L cells or by calcium phosphate coprecipitation into RAC65 cells (40). Two days later, the cells were washed twice with phosphate-buffered saline (PBS) and resuspended in 100 μl of 250 mM Tris (pH 7.0). Following three freeze-thaw cycles and a final centrifugation step, the supernatant was directly used for measuring luciferase and CAT activities.

Nucleotide sequence accession number. The entire DNA sequence shown in Fig. 1 has been submitted to GenBank (accession number U14621).

RESULTS

Cloning of sea urchin *Pax-6* cDNA. Previously, we characterized the sea urchin transcription factor TSAP (4), which, as a homolog of the mammalian BSAP (Pax-5) protein, proved to be a member of the Pax protein family (1). In an attempt to clone this transcription factor and to identify other sea urchin Pax proteins, we have devised a PCR screen based on degenerate primers which were derived from two highly conserved regions of the paired domain (see Fig. 2). cDNA prepared from gastrula embryos of the sea urchin *P. lividus* was used for PCR amplification and cloning of a 210-bp DNA fragment (Fig. 1A, fragment A) with an open reading frame predicting a peptide which was identical with known vertebrate Pax-6 proteins (see Fig. 2). Subsequently, we isolated several cDNA clones by screening a gastrula cDNA library with this probe (Fig. 1A). The sequence of two overlapping clones (λ1 and λ3) was determined, and the full-length Pax-6 cDNA sequence was assembled (Fig. 1B). It contains a long open reading frame which is flanked by 0.6 kb of 5′ untranslated leader and 3 kb of trailer sequences. A putative imperfect polyadenylation signal (AATATA) is located at the 3′ end 14 bp upstream of the poly(A) tail. Although three potential in-frame translation start sites precede the open reading frame, the sequence context of only the last of these sites fits the consensus sequence for optimal translation initiation well (30). The resulting pro-

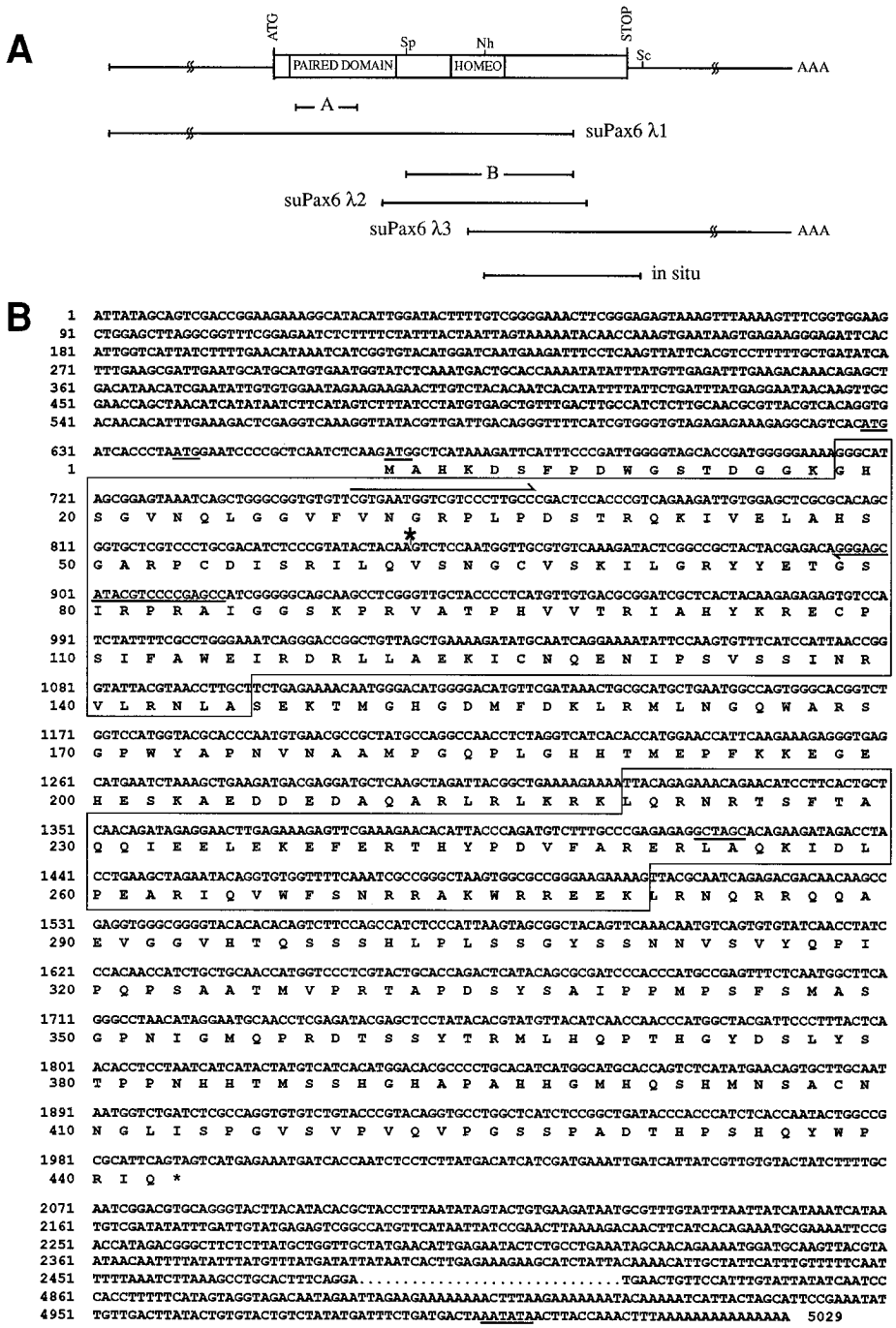


FIG. 1. Cloning and sequencing of sea urchin *Pax-6* cDNA. (A) Schematic diagram of the structural organization of the *suPax-6* cDNA. 5′ and 3′ untranslated sequences are indicated by a line, and the open reading frame containing the paired domain and homeodomain is shown by a boxed region. DNA fragment A was isolated by PCR amplification from gastrula cDNA and was subsequently used for screening of a gastrula cDNA library resulting in clone λ1. The 3′ part (fragment B) of this cDNA clone was used for a second screen of the cDNA library, which yielded clones λ2 and λ3. The *Nhe*I-*Sca*I DNA fragment used for in situ hybridization analysis (Fig. 3B) is indicated below (in situ). Restriction sites are as follows: Sp, *SpeI*; Sc, *Sca*I, and Nh, *Nhe*I. (B) Nucleotide and deduced amino acid sequence of the sea urchin *Pax-6* cDNA. The sequence was assembled from the two overlapping cDNA clones λ1 and λ3 by joining the sequences at the *Nhe*I site (positions 1419 to 1424, underlined). Several silent nucleotide substitutions were present in the overlapping region of the two cDNA clones, which merely reflects the high degree of sequence polymorphism within sea urchin populations. The three potential translation initiation sites and the putative poly(A) addition signal are underlined. The paired domain and homeodomain are boxed. Arrows indicate the primers which were used for RT-PCR assay of gastrula mRNA, PCR analysis of genomic DNA, and PCR cloning of the RNase protection probe. A large asterisk above the line marks the position where all vertebrate *Pax-6* genes contain an alternative exon (5a) with flanking introns which are absent in the sea urchin *Pax-6* gene.

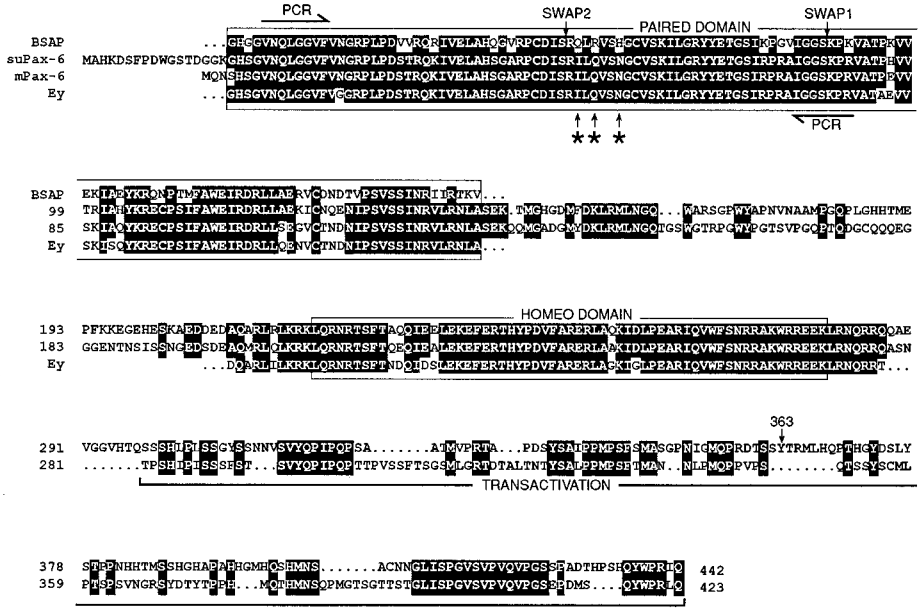


FIG. 2. Amino acid sequence comparison of the sea urchin, mouse, and *Drosophila* Pax-6 proteins. The deduced amino acid sequence encoded by the sea urchin *Pax-6* cDNA was aligned with the sequences of the entire mouse Pax-6 protein (mPax-6) (55), of the paired domain and homeodomain of the *Drosophila* Eyeless (Ey) protein (42), and the paired domain of BSAP (1). Amino acid residues that are identical with the suPax-6 sequence are highlighted by black overlay. Gaps introduced for optimal alignment of the protein sequences are denoted by dots. The paired domain and homeodomain are boxed, and the C-terminal region used for transactivation studies (Fig. 10) is underlined. The C-terminal break point of the mutant protein suPax-6(1–363) encoded by the cDNA clone λ1 is indicated. The amino acid sequences used for the design of degenerate PCR primers are indicated by arrows. SWAP1 and SWAP2 mark the positions where reciprocal parts of the suPax-6 and BSAP proteins were fused to generate chimeric proteins (Fig. 6). Asterisks indicate the positions of the three amino acid residues which were mutated in the protein suPax-6mut (Fig. 6).

tein of 442 amino acids contains both a paired domain and a paired-type homeodomain. Figure 2 shows a comparison of the deduced amino acid sequence with that of the mouse Pax-6 protein. The two proteins exhibit 70% identity along their entire sequence in comparison with the 96% identity observed between Pax-6 proteins of distantly related vertebrates (mouse and zebrafish) (41). Both the paired domains and the homeodomains are almost identical between the sea urchin and mouse Pax-6 proteins, indicating that the two domains have been subjected to extremely high selection pressure since the divergence of echinoderms and vertebrates. This high homology clearly identifies the cloned cDNA as the sea urchin homolog (*suPax-6*) of the vertebrate *Pax-6* genes. Recently, the *Drosophila eyeless* gene was shown to code for a Pax-6 protein (42) which is as highly conserved in the paired domain and homeodomain as the vertebrate and sea urchin Pax-6 proteins (Fig. 2). However, except for these two DNA-binding motifs, *Drosophila* Pax-6 shows little similarity to its sea urchin and vertebrate homologs because it is, with 838 amino acids, twice as long as other Pax-6 proteins (42).

Vertebrate *Pax-6* genes contain a small alternative exon (5a) which, upon differential mRNA splicing, gives rise to a 14-amino-acid insertion into the paired domain (13, 20, 41, 55). This exon and the corresponding alternative splice pattern have been highly conserved in all vertebrate *Pax-6* genes (13, 20, 41, 55) and result in a Pax-6 isoform with altered DNA-binding specificity (16, 30a). RT-PCR amplification of paired domain sequences from gastrula RNA with *suPax-6*-specific primers (Fig. 1B) generated only a single cDNA fragment, suggesting that the corresponding alternative splice product is absent from gastrula embryos. Moreover, a DNA fragment of the same length was obtained regardless of whether gastrula cDNA or genomic DNA was used as a template for PCR amplification (data not shown). Hence, the paired box of the sea urchin *Pax-6* gene does not contain an insertion at the position where vertebrate *Pax-6* genes contain the alternative exon 5a.

***Pax-6* gene expression during sea urchin ontogeny.** The expression of the *suPax-6* gene was analyzed at different stages of embryonic development and in adult tissues by RNase protection analysis (Fig. 3A). These experiments revealed that expression of the *Pax-6* gene is initiated at the gastrula stage of embryogenesis. Moreover, the *Pax-6* gene is expressed in a highly tissue-specific manner in the adult sea urchin, as its transcripts could be detected only in the tube foot of all the different tissues analyzed. As the tube foot is a heterogeneous tissue composed of distinct cell layers (28), we localized the domain of *Pax-6* expression by in situ hybridization on serial sections of the tube foot using RNA probes that correspond to C-terminal Pax-6 coding sequences (Fig. 1A). As shown in Fig. 3B, the antisense riboprobe, but not the control sense probe, labelled a single cell layer which is composed of the longitudinal muscle fibers of the tube foot (28). Furthermore, some Pax-6-specific labelling was also observed at the tip of the tube foot.

Indistinguishable DNA-binding properties of the sea urchin and mouse Pax-6 proteins. To characterize the full-length suPax-6 protein, we assembled the entire open reading frame from two overlapping cDNA clones (λ1 and λ3). In vitro translation of RNA derived from this construct resulted in a protein product of the expected size. However, the translation efficiency was very low (data not shown), suggesting that the translation start site of the *suPax-6* mRNA is not efficiently recognized by the mammalian translation machinery. We therefore replaced the initiation region of the *suPax-6* cDNA by the corresponding sequence of the mouse *Pax-6* gene, which

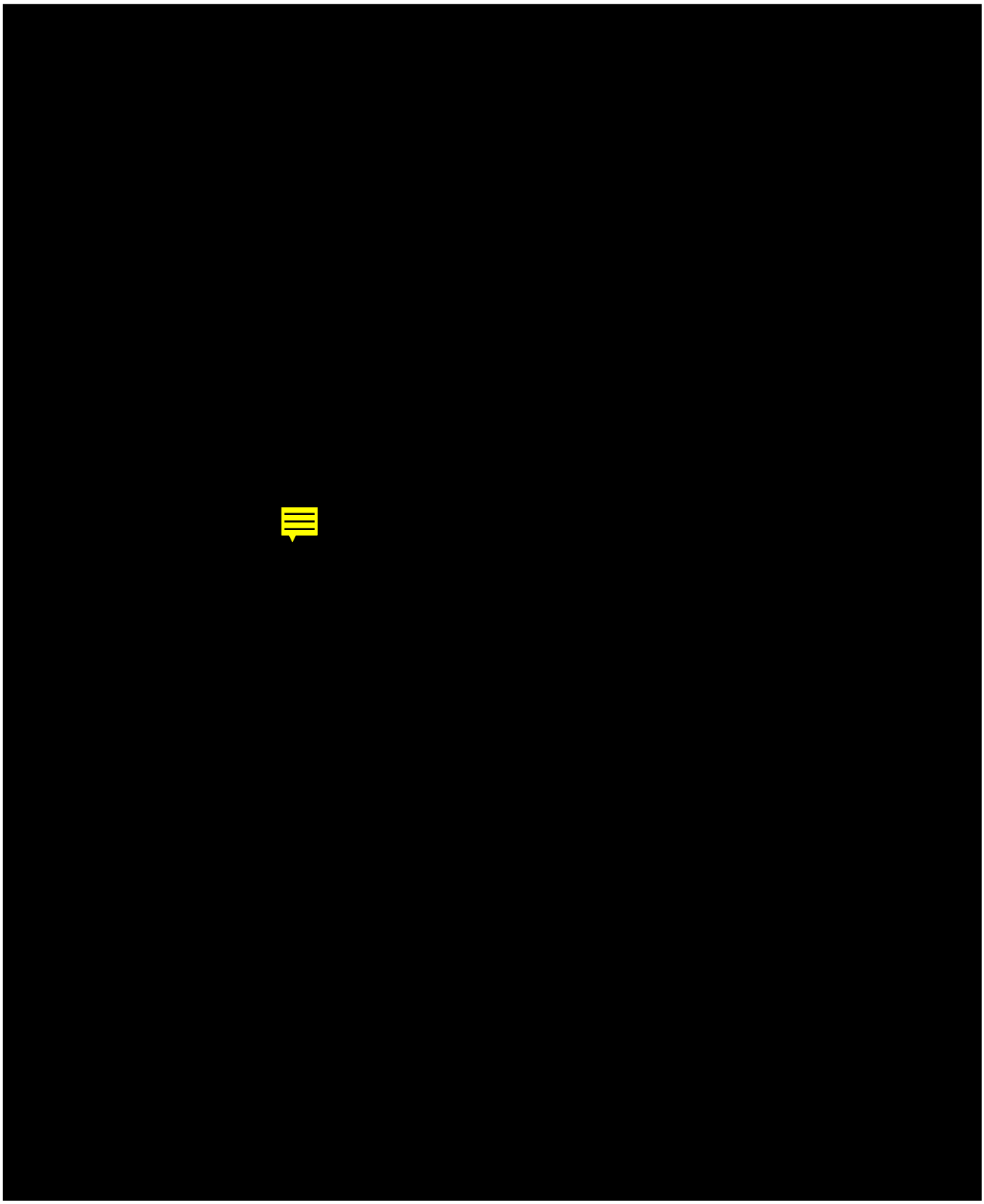


FIG. 3. Expression of the *Pax-6* gene during embryogenesis and in the tube foot of the adult sea urchin. (A) RNase protection analysis. Total RNA (10 μg) of the indicated developmental stages and adult tissues of the sea urchin *P. lividus* were analyzed for *suPax-6* transcripts as described in Materials and Methods. The sizes of two marker DNA fragments (M; in base pairs) are indicated to the left and right. The quality of each RNA preparation was checked by gel electrophoresis followed by Northern (RNA) blotting and methylene blue staining. (B) In situ hybridization analysis. Serial sections through a tube foot of *P. lividus* were hybridized with antisense and sense RNA probes derived from the C-terminal Pax-6 coding region (Fig. 1). Higher magnifications demonstrating specific hybridization to *Pax-6* mRNA in the longitudinal muscle fibers of the tube foot (marked by arrows) are shown on the right. The plane of sectioning is indicated on a schematic diagram of the tube foot.

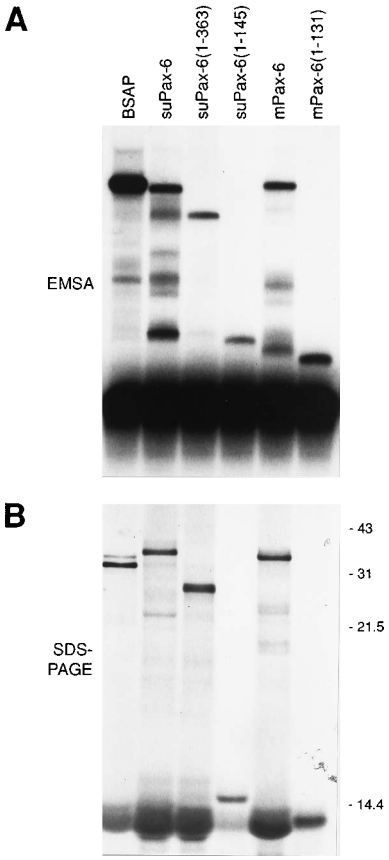


FIG. 4. Identical sequence recognition of sea urchin and mouse Pax-6 proteins. (A) EMSA analysis. The same molar amount of the indicated in vitro-translated proteins was analyzed by EMSA for binding to the CD19-2(A-ins) oligonucleotide (10). (B) Sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) analysis of in vitro-translated Pax-6 proteins. Only 0.2 μl of the in vitro-translation reaction mixtures of peptides suPax-6(1–145) and mPax-6(1–131) was analyzed, compared with 1 μl for all other proteins. Radioactive signals were quantitated on a PhosphorImager, and the relative protein amounts were calculated on the basis of the cysteine content of the different polypeptides. The positions of marker proteins (sizes in kilodaltons) are indicated on the right.

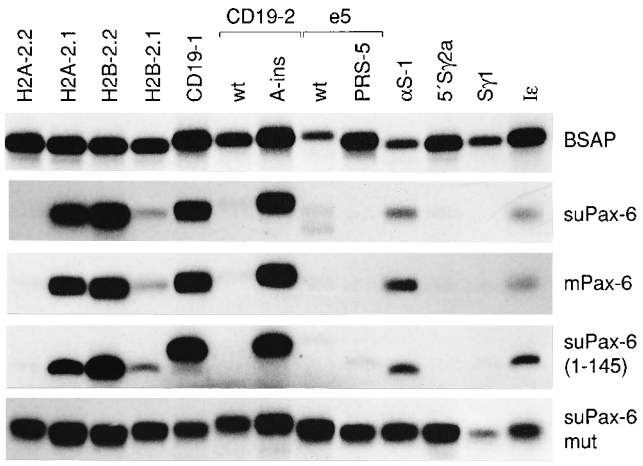


FIG. 5. Sequence specificities of Pax-6 proteins. The different Pax-6 proteins and BSAP were expressed by transient transfection in COP-8 cells, and whole-cell extracts were analyzed by EMSA for binding to a panel of BSAP recognition sequences. Only the relevant part of the autoradiograph containing the protein-DNA complexes is shown. The BSAP-binding sites originate from the sea urchin *H2A-2* and *H2B-2* genes (4), from the human *CD19* gene (32), from the *Drosophila even-skipped* promoter (e5 [52] and PRS-5 [6]), and from different regions of the immunoglobulin heavy-chain gene (αS-1 [57], 5′Sγ2a [35], Sγ1 [58], and Iε [43]; for sequences see Czerny et al. [10]). The protein suPax-6mut corresponds to the specificity mutant E shown in Fig. 6.

resulted in a significant improvement of the translation rate (see Materials and Methods).

The DNA-binding properties of the sea urchin and mouse Pax-6 proteins were compared by quantitative EMSA. For this purpose, full-length and truncated Pax-6 proteins were synthesized and quantitated by in vitro translation (Fig. 4B). Equivalent amounts of these proteins were then used for DNA-binding analysis with a high-affinity BSAP-binding site [CD19-2(A-ins)] which was previously shown to be efficiently recognized by the mouse Pax-6 protein (10). As indicated in Fig. 6A, the full-length Pax-6 proteins of sea urchin and mouse bound to this probe with the same affinity, as did a C-terminally truncated peptide, suPax-6(1–363), which still retained both DNA-binding domains. Interestingly, even the Pax-6 deletion mutants suPax-6(1–145) and mPax-6(1–131), which terminated immediately downstream of the paired domain, exhibited the same DNA-binding affinity as the full-length proteins of both species. The CD19-2(A-ins) site is therefore exclusively recognized by the Pax-6 paired domain with no contribution of the homeodomain.

The DNA-binding specificities of the sea urchin and mouse Pax-6 proteins were further compared by the use of a panel of different BSAP recognition sequences which originate from the sea urchin *H2A-2* and *H2B-2* histone genes, from the *CD19* gene, from the *Drosophila even-skipped* promoter (e5), and from immunoglobulin heavy-chain gene switch regions (for references, see the legend to Fig. 5). To this end, Pax-6 and BSAP proteins were expressed in COP-8 cells from transiently transfected expression plasmids (1), and whole-cell extracts were subsequently analyzed by EMSA. In contrast to BSAP, DNA binding of the suPax-6 protein to this panel of sequences varied considerably from site to site (Fig. 5). However, it is important to note that the binding pattern of sea urchin Pax-6 was strictly congruent with that of mouse Pax-6, thus indicating that both proteins possess the same DNA sequence specificity. Moreover, the paired domain peptide suPax-6(1–145) showed a DNA-binding behavior identical to that of the full-length Pax-6 proteins, and hence the paired domain rather than the homeodomain of Pax-6 interacts with the binding sites analyzed in Fig. 5.

The unique DNA-binding specificity of Pax-6 is determined by 3 residues within a 7-amino-acid stretch of the paired domain. Pax proteins have been divided into different subfamilies according to their sequence homology (56). This classification is also nicely reflected at the level of DNA-binding specificity (10). In particular, the Pax-6 paired domain binds to a panel of BSAP-binding sites with a unique pattern that is not observed with members of other Pax subfamilies (10). Interestingly, BSAP (Pax-5) and Pax-6 belong to two related subfamilies (39, 56), and yet the Pax-6 paired domain fails to interact with several BSAP-binding sites (Fig. 5). We therefore set out to map the corresponding specificity-determining region within the paired domain by creating chimeric suPax-6-BSAP proteins. Reciprocal parts of the two proteins were fused in the middle of the paired domain within a highly conserved region (Fig. 2; see Fig. 6), and the resulting proteins were analyzed by EMSA for binding to the H2A-2.2 site, which is recognized by wild-type BSAP only and not by Pax-6. The CD19-1 probe was used as a reference site which is bound by

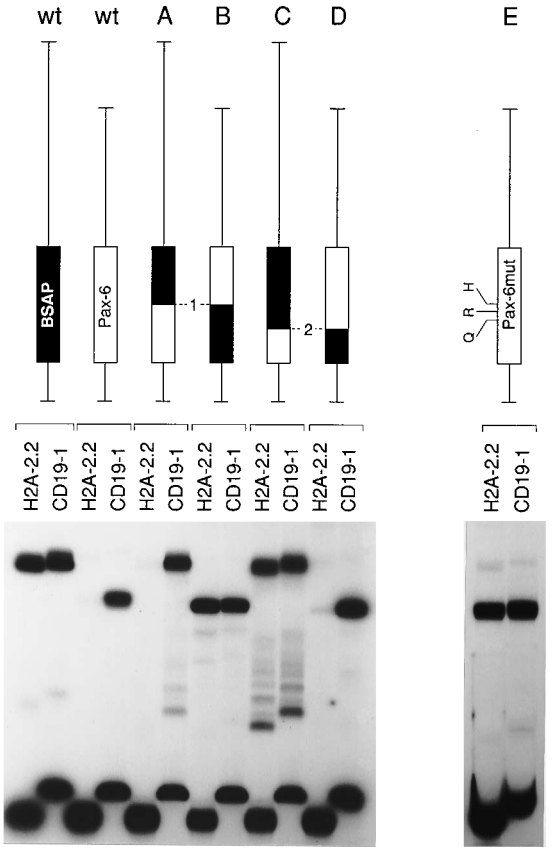


FIG. 6. The DNA-binding specificity of the Pax-6 paired domain is determined by 3 amino acid residues. Wild-type (wt) mouse BSAP and sea urchin Pax-6 proteins as well as chimeric proteins thereof were analyzed for binding to the H2A-2.2 and CD19-1 oligonucleotides (for sequences, see Fig. 7A). The different proteins were expressed by transient transfection in COP-8 cells, and whole-cell extracts were analyzed by EMSA. Positions 1 and 2, where reciprocal parts of BSAP and Pax-6 were fused in the chimeric proteins, are indicated as SWAP1 and SWAP2 in the sequence alignment of Fig. 2. The sea urchin protein Pax-6mut contained the following amino acid substitutions: I to Q at position 59 (I59Q), Q61R, and N64H.

both proteins (Fig. 5; see Fig. 6). The chimeric protein containing the N-terminal half of the Pax-6 paired domain (Fig. 6, construct A) failed to bind to the H2A-2.2 site and thus behaved as wild-type Pax-6 does. Hence, the specificity-determining region does not reside in the C-terminal half of the paired domain which is provided by BSAP. Consistent with this interpretation, the reciprocal fusion protein containing the N-terminal sequences of BSAP (Fig. 6, construct B) bound to both sites as BSAP does. These experiments therefore localized the specificity-determining sequences within the N-terminal half of the paired domain. In a next step, a second pair of fusion proteins was created by swapping Pax-6 and BSAP sequences in the middle of the N-terminal paired domain region (Fig. 2; see Fig. 6). The chimeric protein containing the very N-terminal suPax-6 sequences (Fig. 6, construct C) now exhibited the binding properties of BSAP, whereas the reciprocal protein (construct D) behaved as wild-type Pax-6 does. These data therefore allowed us to map the specificity-determining region to a stretch of 31 amino acids from position 58 to 88 of the suPax-6 protein.

The corresponding paired domain region differs only at 6 amino acid positions between Pax-6 and BSAP (Fig. 2). The first three of these amino acids (I-59, Q-61, and N-64 in suPax-6) are characteristic of all Pax-6 proteins from *Drosophila melanogaster* to humans (Fig. 2), while amino acids R-81, R-83, and A-84 are also found at equivalent positions in the paired domains of other Pax proteins (56). To test the hypothesis that the three Pax-6-specific amino acids are responsible for the unique binding specificity of Pax-6, we replaced these three residues in the Pax-6 protein with the corresponding amino acids of BSAP. Interestingly, the resulting protein (Fig. 6, Pax-6mut) bound to the H2A-2.2 site similarly to wild-type BSAP. Moreover, this mutant Pax-6 protein also recognized the whole panel of BSAP-binding sites (Fig. 5) similarly to BSAP. Hence, substitution of only three amino acids is sufficient to switch the DNA-binding specificity of the paired domain from Pax-6 to BSAP. These 3 amino acid residues, which are located at positions 42, 44, and 47 of the paired domain, therefore determine the unique sequence recognition of Pax-6.

Amino acids 42, 44, and 47 of the paired domain discriminate between nucleotides at the primary divergent position of the Pax-6 and BSAP consensus recognition sequences. As the differential sequence recognition by the Pax-6 and BSAP paired domains could be attributed to a difference of only 3 amino acids, we next investigated which corresponding nucleotide position(s) in Pax-6 and BSAP-binding sites might be discriminated by these residues. To do so, we had to take into account that there is a considerable overlap in Pax-6- and BSAP-binding sites (Fig. 5) and that the affinity of a given binding site can be influenced by compensatory base changes in both half-sites of the recognition sequence (10). To bypass these potential problems, we compared the consensus recognition sequences of the Pax-6 and BSAP paired domains rather than individual binding sites with each other. First, we updated both consensus recognition sequences (10, 15) by including recently published binding sites in the sequence compilation (for detailed description, see the legend to Fig. 7A). For instance, the four high-affinity binding sites identified within the panel of BSAP recognition sequences (Fig. 5) and the binding sites present in the genes coding for αA-crystallin (9) and the neural cell adhesion molecule L1 (7) were added to the list of Pax-6 recognition sequences (Fig. 7A). Moreover, as Pax-2, Pax-5 (BSAP), and Pax-8 are known to bind DNA in an indistinguishable manner (31), we also used Pax-2- and Pax-8-binding sites to further optimize the consensus sequence for this subfamily of Pax proteins. As shown by the comparison in Fig. 7A, the two improved consensus recognition sequences deviate from each other primarily at nucleotide position 19, where an A residue is present in 76% of all Pax-6-binding sites as opposed to a C residue in 68% of the Pax-2-, Pax-5-, and Pax-8-binding sequences.

To study the influence of nucleotide 19 on binding of the Pax-6 and BSAP paired domains, we decided to mutate this position in a high-affinity binding site that closely resembled the two consensus recognition sequences and that was consequently bound by both proteins. For this purpose we chose a mutant H2A-2.2 site (H2A-C) with a G-to-C substitution at position 14, which resulted in additional binding of Pax-6 (10). The relevant C-to-A mutation was subsequently introduced at position 19 to generate oligonucleotide H2A-CA (Fig. 7A). As shown by the EMSA analysis of in vitro-translated BSAP, Pax-6, and Pax-6mut proteins (Fig. 7B), this C-to-A substitution led to increased binding of Pax-6 and decreased binding of BSAP, as predicted by the two consensus recognition sequences. More interestingly, the same mutation also had opposite effects on binding of Pax-6 (twofold increase) and Pax-6mut (eightfold decrease), further demonstrating that the latter protein behaves similarly to BSAP. Replacement of the three Pax-6-specific amino acids at positions 42, 44, and 47 of

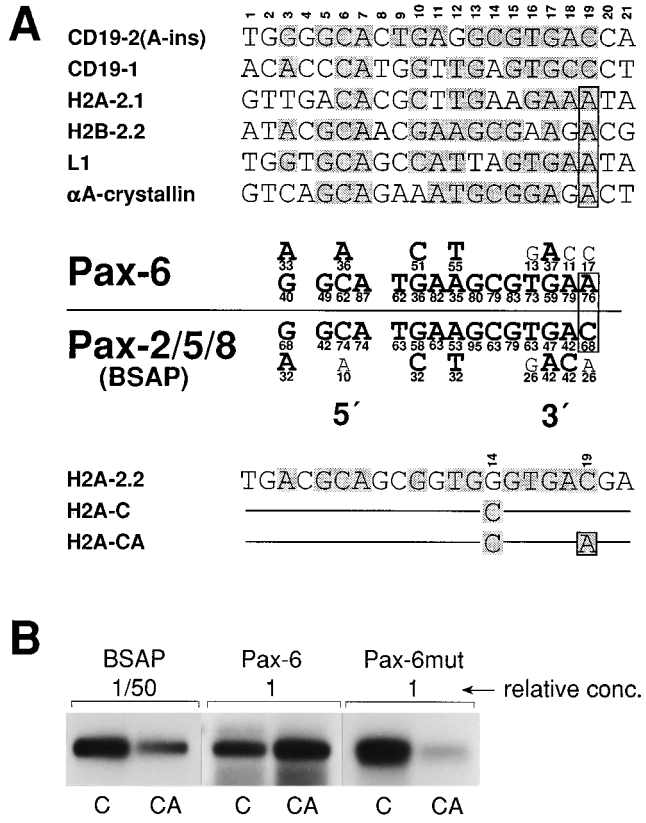


FIG. 7. Three amino acids of the paired domain discriminate between Pax-6- and BSAP-specific nucleotides at a divergent position of the two recognition sequences. (A) Comparison of the consensus recognition sequences for the Pax-6 and BSAP (Pax-5) paired domains. The sequences of the four high-affinity Pax-6-binding sites identified within a panel of BSAP-binding sites (Fig. 5) are shown together with Pax-6 recognition sequences present in the genes coding for αA-crystallin (9) and the neural cell adhesion molecule L1 (7). These binding sites and a consensus sequence previously identified by PCR-based binding-site selection assay (15) were used in a ratio of 1:2 to derive the indicated consensus recognition sequence of the Pax-6 paired domain. The number below each nucleotide denotes the percentage at which this nucleotide is represented at the corresponding position of Pax-6 binding sites. The consensus recognition sequence for the paired domain of the Pax-2, Pax-5, and Pax-8 subfamily results from a compilation of the 13 binding sites analyzed by Czerny et al. (10), the sequence of oligonucleotide Pax2Con (15), and the recognition sites present in the immunoglobulin heavy chain 3′α enhancer (47) and in the *blk* (62), thyroglobulin (61), thyroperoxidase (61), and N-CAM (27) genes. The two half sites of the recognition sequence are symbolized by 5′ and 3′. The derivation of the H2A-C and H2A-CA oligonucleotides (with a total length of 42 bp) from the binding site of the H2A-2.2 gene (4) is indicated below. (B) Binding of the proteins BSAP, suPax-6, and suPax-6mut to the oligonucleotides H2A-C (C) and H2A-CA (CA). In vitro-translated proteins were quantitated by fluorography and used for EMSA analysis at the relative concentrations indicated. Only the radioactive signals corresponding to the protein-DNA complexes are shown. Quantitation on a PhosphoImager revealed that the C-to-A mutation at position 19 reduced DNA binding of BSAP and Pax-6mut by factors of 2.5 and 8, respectively, while Pax-6 binding was increased twofold.

the paired domain by those of BSAP reduced the affinity of the Pax-6 protein for its preferred recognition sequence (H2A-CA), therefore, by a factor of 16. Hence, we conclude that amino acids 42, 44, and 47 of the paired domain are critical for discrimination between a Pax-6- and a BSAP-specific nucleotide (A versus C) at the primary divergent position of the two consensus recognition sequences. The experiment shown in Fig. 2B furthermore demonstrated that only the sequence specificity but not the overall affinity for DNA is determined by amino acids 42, 44, and 47, as the protein Pax-6mut had to be used at a 50-fold higher concentration than BSAP to obtain equivalent DNA binding.

Cooperative dimerization of the Pax-6 homeodomain on palindromic binding sites. To investigate the DNA-binding properties of the Pax-6 homeodomain, we performed EMSA experiments with a series of homeodomain recognition sequences containing a single TAAT core motif in different sequence contexts. As binding of Pax-6 to all of these sites was very inefficient, we decided to use a PCR-based binding-site selection assay to define an optimal recognition sequence for the Pax-6 homeodomain (see Materials and Methods). Sequences that were specifically bound by a bacterially expressed Pax-6 homeodomain polypeptide (Pax-6HD) were selected from a random mixture of degenerate oligonucleotides, amplified by PCR, and subjected to three further rounds of selection and PCR amplification prior to cloning and sequencing analysis. All sequenced oligonucleotides contained multiple TAAT motifs. Moreover, inverted repeats of the TAAT motif separated by a 3-bp spacer (so-called P3 sites [59]) were present in 45% of all cases (Fig. 8A), while our collection of binding sites did not contain any palindromic sequence with two central base pairs (P2 sites). In agreement with this, cooperative dimerization of the Pax-6HD protein was more efficient on an optimal P3 site than on a P2 site, as shown by the EMSA analysis of Fig. 8B. Moreover, the importance of the central base pairs is emphasized by the fact that replacement of a CpG dinucleotide (P2) by GpC (P2*) almost completely interfered with cooperative DNA binding of Pax-6HD (Fig. 8B). In summary, we conclude that the Pax-6 homeodomain preferentially binds to P3 sites through cooperative dimerization. Interestingly, the homeodomain of the *Drosophila* Prd protein was previously shown to preferentially dimerize on P2 sites (59).

We next compared the two DNA-binding activities of the paired domain and homeodomain in the context of the full-length Pax-6 protein. For this purpose, optimal paired domain (H2A-C) and homeodomain (P3) recognition sequences were used for EMSA analysis with increasing amounts of full-length Pax-6 protein which was synthesized in transiently transfected COP-8 cells (Fig. 8C). Under the in vitro binding conditions used, the paired domain proved to be more effective, by ~2 orders of magnitude, in DNA binding than the homeodomain. These data therefore suggest that the paired domain is the primary DNA-binding motif of Pax-6.

The C-terminal sequences of Pax-6 code for a potent transactivation function. The transactivation function of Pax-6 was analyzed by transient transfection of J558L plasmacytoma cells, which are devoid of endogenous Pax proteins, with Pax-6 expression vectors, and with Pax-6-responsive genes. The reporter gene lucCD19 was generated by insertion of three multimerized Pax-6-binding sites [CD19-2(A-ins)] upstream of a TATA box and a luciferase gene (see Fig. 10A). As shown in Fig. 9A, the sea urchin and mouse Pax-6 proteins reproducibly stimulated luciferase gene transcription four- and sixfold, respectively, in comparison with the basal expression level obtained with the empty expression vector pKW10 (1). Hence, the sea urchin Pax-6 protein is a transcriptional activator like mouse Pax-6 despite the fact that heterologous mammalian cells were used for the transactivation assay. During the course of these experiments, we realized that transactivation by Pax-6 was critically dependent on the concentration of the Pax-6 expression vector. A strong increase in transcriptional stimulation was seen at low concentrations (0.2 to 1 μg) of the transactivator plasmid, while higher amounts (2 to 8 μg) of expression vector resulted in a sharp decrease of transcriptional activity (Fig. 9B). To investigate the basis of this narrow concentration dependence, we have taken advantage of the

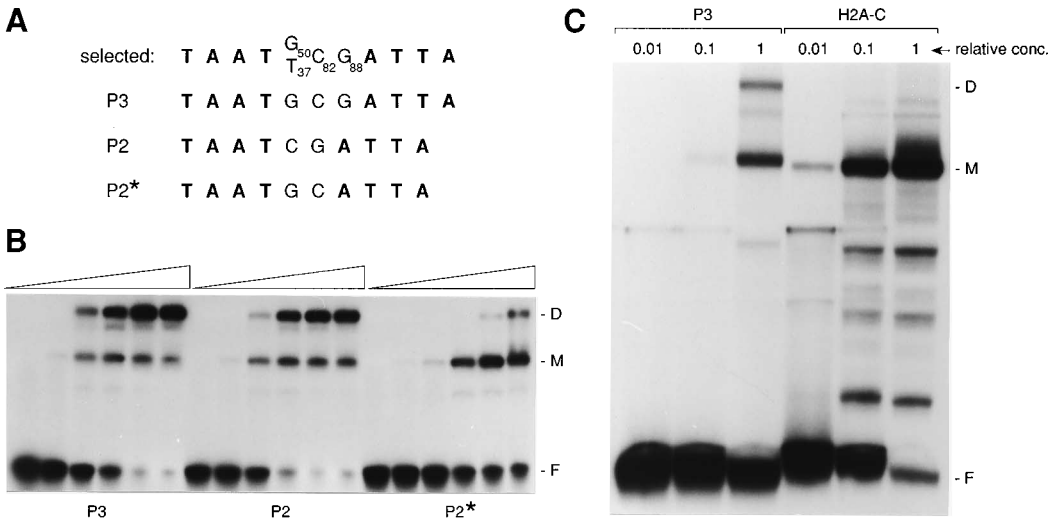


FIG. 8. Cooperative dimerization of the Pax-6 homeodomain on palindromic binding sites. (A) Binding-site selection assay. The homeodomain polypeptide Pax-6HD was synthesized in *E. coli*, purified on Ni-NTA agarose, and used for binding-site selection assay as described in Materials and Methods. The compilation of 17 selected perfect P3 sites is shown with the corresponding numbers indicating the percentage at which each nucleotide is represented in the center of the recognition sequence. The inverted repeat sequences of the P3, P2, and P2* oligonucleotides used for EMSA analysis are shown below. (B) Cooperative dimerization of the Pax-6HD polypeptide on DNA. Samples (5 fmol) of the indicated end-labeled DNA probes were used for EMSA analysis with increasing amounts of purified Pax-6HD polypeptide. The differences in protein amounts were threefold between each lane, with the first lane (to the left) containing 50 fmol of protein. The positions of monomeric (M) and dimeric (D) protein-DNA complexes and of the free probe (F) are indicated. (C) Different contributions of the paired domain and homeodomain to the DNA-binding activity of full-length Pax-6. A whole-cell extract of COP-8 cells transiently transfected with the expression vector pKW10-mPax6 was used at the indicated relative concentrations for EMSA analysis with DNA probes specific for the Pax-6 homeodomain (P3) or paired domain (H2A-C). Quantitation on a PhosphorImager indicated that the paired domain bound DNA by 2 orders of magnitude more efficiently than the homeodomain. The difference in binding of the two domains was similar even at different concentrations of monovalent and divalent cations (unpublished data). For transcriptional stimulation via the two DNA-binding domains of Pax-6, see Addendum in Proof.

mouse Pax-6 isoform containing a 14-amino-acid insertion in the paired domain which abolishes binding to the CD19-2(A-ins) site present in the reporter gene (30a). Figure 9C shows the result of a titration experiment, in which a constant amount (1 μg of DNA) of wild-type Pax-6 expression vector was co-transfected with increasing amounts of an expression plasmid coding for the non-DNA-binding isoform of Pax-6. Increasing concentrations of this protein nevertheless resulted in a strong decrease of the transcriptional activity of Pax-6, indicating that the observed effect is independent of DNA binding and may thus reflect self-squelching of the Pax-6 transactivation domain.

The C-terminally truncated protein suPax-6(1–363) exhibited low, if any, transcriptional activity in the same transient transfection assay (Fig. 9A), suggesting that the deleted C-terminal 79 amino acids constitute at least part of the Pax-6 transactivation domain. To verify this hypothesis, we separated the C-terminal region from other Pax-6 sequences by fusing it to the DNA-binding domain of the yeast GAL4 transcription factor. The resulting proteins, GAL-suPax6 and GAL-mPax6, stimulated transcription of a GAL4-responsive luciferase gene 40- and 75-fold in transiently transfected J558L cells (Fig. 9D). For comparison, the GAL-VP16 protein (45), consisting of the strong VP16 transactivation domain fused to the GAL4 DNA-binding region, was only 5 to 10 times more active than the GAL-Pax6 proteins (Fig. 9D). We conclude, therefore, that a potent transactivation function resides in the C-terminal region of Pax-6. Similar results with GAL-Pax6 proteins were obtained in transactivation assays with the RAC65 cell line (40), a differentiation-deficient subclone of P19 embryonal carcinoma cells (data not shown). Moreover, Glaser et al. (18) have recently reported similar findings by studying fusion proteins between GAL4 and the human Pax-6 protein in transiently transfected glioblastoma cells. Together, these data suggest therefore that the C-terminal transactivation function of Pax-6 is neither cell type nor species specific.

Pax proteins including Pax-6 are transcriptionally active from a distant position. The activities of transcription factors can be strongly influenced by the positions and sequence contexts of their binding sites (45). To see whether Pax-6 can act over a large distance, we introduced multimerized GAL4- and Pax-6-binding sites 2 kb downstream of the transcription start site of a TK-luciferase gene (Fig. 10C). For control purposes, the same arrangement of binding sites was also inserted directly upstream of the TK promoter (Fig. 10B). As shown by the transient transfection experiments illustrated in Fig. 10, the transcriptional activities of Pax-6 were similar regardless of whether its binding sites were present in the proximal promoter (Fig. 10A and B) or the distal enhancer (Fig. 10C). In addition, representative members of three other Pax subfamilies, Pax-1, Pax-3, and BSAP (Pax-5), also displayed similar position-independent transactivation of the different reporter genes (Fig. 10), indicating that Pax proteins can act as enhancer-binding transcription factors. Moreover, the chimeric protein GAL-Pax6 proved to be highly active not only from a promoter location but also from an enhancer position (Fig. 10B and C). Hence, the C-terminal transactivation domain of Pax-6 is capable of stimulating transcription initiation over a long distance.

DISCUSSION

The *Pax* gene family, which codes for important regulators of early development, is known to be highly conserved during evolution (39). In particular, the *Pax-6* gene of vertebrates has been subject to extremely high selection pressure, as evidenced by the fact that the human and mouse Pax-6 proteins differ over their entire length of 423 amino acids by only 1 amino acid

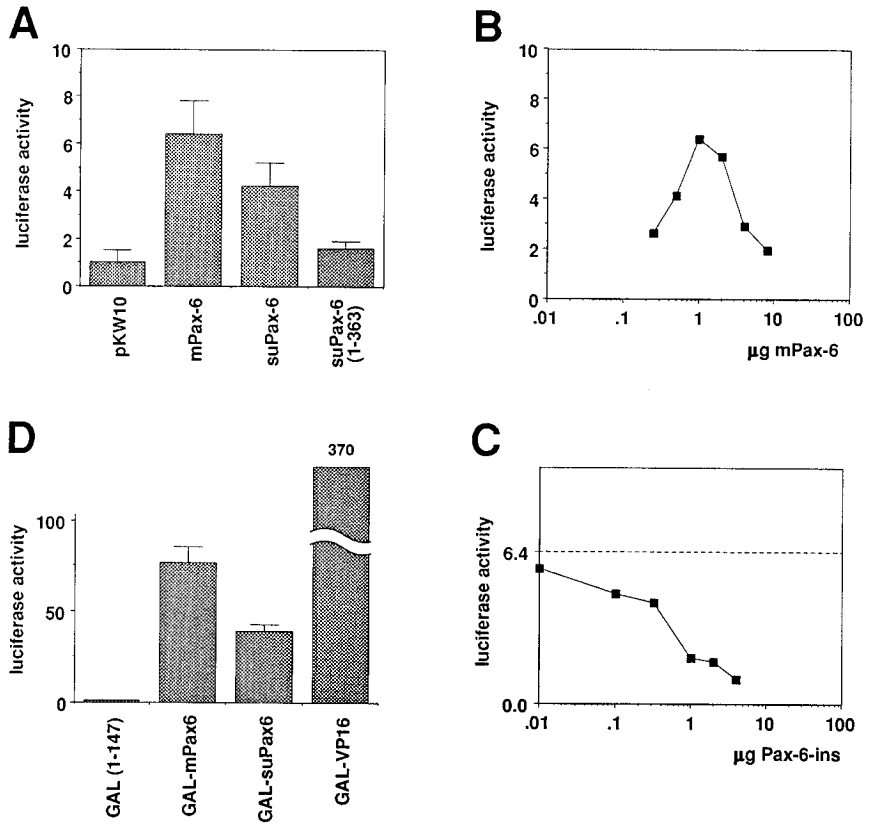


FIG. 9. Transactivation potential of Pax-6 proteins. (A) Transactivation by full-length and C-terminally truncated Pax-6 proteins. Expression plasmids (1 μg of DNA) directing the synthesis of mouse and sea urchin Pax-6 proteins were electroporated into J558L cells together with the reporter plasmid lucCD19 (schematically shown in Fig. 10A). The translation initiation sequence of the mouse *Pax-6* cDNA was engineered into the sea urchin Pax-6 expression vectors to ensure equal translation efficiencies. The average values of four independent transfection experiments are shown as relative levels of luciferase activity compared with the basal expression level obtained with the empty expression vector (pKW10). (B) Concentration dependence of the transactivation function of the mouse Pax-6 protein. Increasing amounts of the Pax-6 expression vector were cotransfected with the lucCD19 reporter gene. The amount of expression plasmid was equalized by the addition of the empty expression vector pKW10 to normalize for promoter interference effects. Average values for three independent experiments are shown. (C) Self-squelching of the Pax-6 transactivation domain. A fixed amount of the mPax-6 expression plasmid (1 μg) was cotransfected with the lucCD19 reporter gene and increasing amounts of an expression vector directing the synthesis of the mouse Pax-6 isoform, which contains a 14-amino-acid insertion in the paired domain (mPax-6-ins). A dashed line indicates the level of transactivation observed by Pax-6 in the absence of the Pax-6-ins protein. The amounts of expression plasmids were equalized as described for panel B. (D) Transactivation by chimeric GAL4–Pax-6 fusion proteins. The GAL4-responsive reporter plGC-luc was cotransfected with the indicated expression plasmids (1 μg of DNA). The fusion proteins GAL-suPax6 and GAL-mPax6 were obtained by fusing the DNA-binding domain of GAL4 (amino acids 1 to 147) to the C-terminal region of suPax-6 (amino acids 298 to 442) or mPax-6 (amino acids 281 to 423), respectively. The average values for four independent transfection experiments are shown. Note that the transcriptional activity of GAL-Pax6 is ~10-fold higher than that of the full-length Pax-6 protein (panel A). Contrary to expectation, deletion of part or all of the intervening sequences between the paired domain and transactivation region did not increase the transcriptional activity of Pax-6 (data not shown). The data from all transactivation experiments (panels A to D) were normalized for equal transfection efficiencies by coelectroporation of a reference CMV-CAT gene and by standardizing the luciferase activities to the CAT values. The error bars in panels A and D indicate standard errors of the mean.

substitution (51, 55). cDNA cloning and characterization of the sea urchin Pax-6 protein have now extended this finding by demonstrating high evolutionary conservation of the Pax-6 protein even between the distantly related sea urchins and vertebrates. The functionally important sequences of the paired domain and homeodomain are almost identical between the sea urchin and vertebrate Pax-6 proteins, while regions of extensive homology are additionally found in the C-terminal transactivation domain. Consequently, the sea urchin and mouse Pax-6 proteins are virtually indistinguishable in their DNA-binding and transactivation properties. However, the *Pax-6* gene of the sea urchin differs at least in one aspect of its genomic organization from the vertebrate genes. The paired box of vertebrate *Pax-6* genes is interrupted after codon 44 by an intron and alternative exon 5a which, upon differential mRNA splicing, gives rise to a Pax-6 isoform with a 14-amino-acid insertion in the paired domain (20) resulting in a dramatically altered DNA-binding specificity (16, 30a). The corresponding exon 5a and flanking introns are missing in the sea urchin *Pax-6* gene, suggesting that this additional level of *Pax-6* gene regulation has been acquired during vertebrate evolution. While we were characterizing the sea urchin *Pax-6* cDNA, Quiring et al. (42) identified the *Drosophila eyeless* gene as another invertebrate homolog of the *Pax-6* gene family. The *Drosophila* and sea urchin *Pax-6* genes have two features in common. Both genes lack the alternative exon 5a of vertebrate *Pax-6* genes and code for proteins that are highly homologous in the paired domain and homeodomain. However, the *Drosophila* Pax-6 protein has twice the size of the sea urchin and vertebrate Pax-6 proteins and consequently shows little homology in the putative transactivation domain.

The phenotypic analysis of homozygous *Small eye* mice and rats has defined an important role for Pax-6 in the morphogenesis of the forebrain and formation of the eye and nose (18, 26, 37). Consistent with these functions, the *Pax-6* expression pattern in the developing central nervous system, eye, and

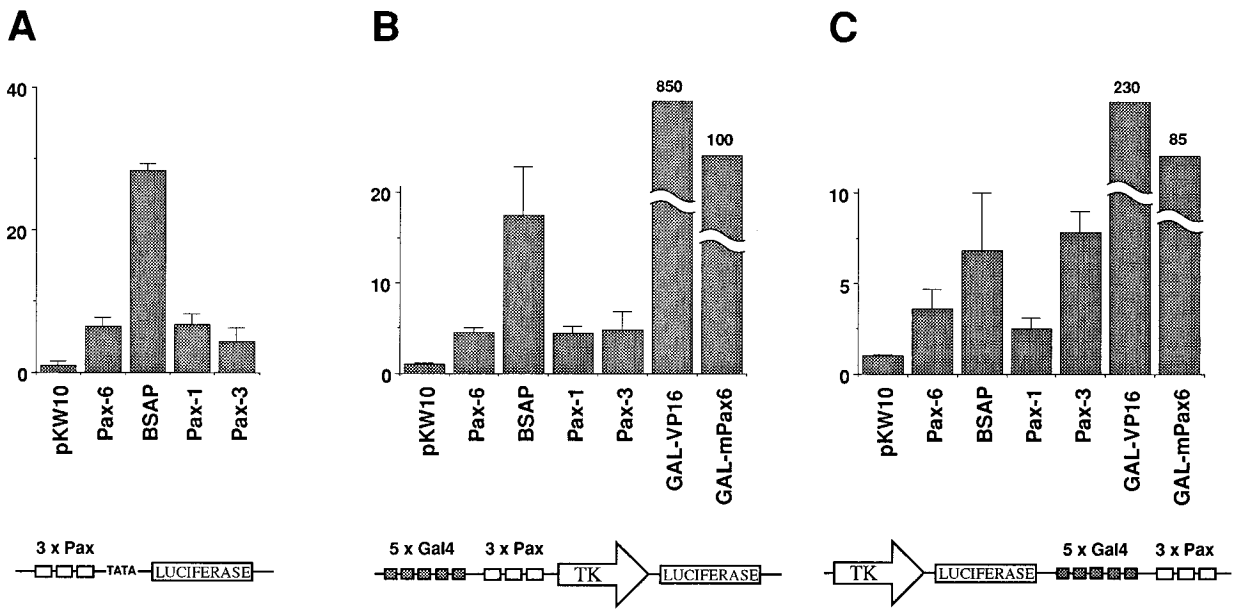


FIG. 10. Pax proteins are transcriptionally active from enhancer positions. The indicated expression plasmids (1 μg) were cotransfected with three different reporter constructs into J558L cells. The average levels of luciferase activity relative to the basal expression level observed with pKW10 alone are shown for three independent experiments. All Pax cDNAs used were of mouse origin with the exception of human BSAP. Schematic diagrams of the reporter constructs lucCD19 (A), lucTK-GAL/CD19 (B), and lucE-GAL/CD19 (C) are shown in the bottom part of the figure. The CD19-2(A-ins) sequence (10) was used as a Pax-6-binding site. The GAL4- and Pax-6-binding sites in the circular plasmid lucE-GAL/CD19 are 2 kb downstream and 3 kb upstream of the transcription initiation site of the TK gene. A reporter gene containing the GAL4- and Pax-6-binding sites in the opposite orientation to the lucTK-GAL/CD19 gene was transactived to the same degree by the different proteins (data not shown).

nasal epithelia has been highly conserved among vertebrates (34, 41, 55), which may explain the high constraint on vertebrate Pax-6 evolution. Even the *Drosophila Pax-6* (*eyeless*) gene is expressed in the primordia of the eye imaginal disks, indicating that eye morphogenesis is under similar genetic control in both vertebrates and insects (42). It is, however, difficult to fit the expression pattern of the sea urchin *Pax-6* gene into this general scheme, because of the lack of any analogy between the rudimentary nervous system of the sea urchin and the eyes and central nervous system of vertebrates and insects (28). In the sea urchin, expression of the *Pax-6* gene is first detected during gastrulation of the embryo and is later restricted to a muscle cell layer in the highly innervated tube foot of the adult organism (Fig. 3).

The paired domains of Pax-6 and BSAP (Pax-5) recognize similar DNA sequences, although both proteins belong to different subfamilies of Pax proteins (10, 15). Analysis of the consensus binding sites for both proteins revealed only one primary divergent nucleotide position (position 19 [Fig. 7A]). However, it is important to note that naturally occurring binding sites present in BSAP and Pax-6 target genes deviate at several positions from their consensus sequence because of compensatory base changes in both half sites (reference 10 and Fig. 7A). It may, therefore, not be surprising that the binding specificities of BSAP and Pax-6 differ vastly for individual binding sites, indicating that deviations from the consensus sequence are differently tolerated by the two proteins. This difference in binding specificity could be attributed to 3 amino acid residues within a heptapeptide sequence in the N-terminal part of the paired domain (amino acids 42, 44, and 47). Replacement of these three residues in Pax-6 by the corresponding amino acids of BSAP completely switched the sequence specificity of the paired domain from Pax-6 to BSAP (Fig. 6). A corollary of this is that the C-terminal halves of the Pax-6 and BSAP paired domains make similar contributions to the binding specificity of the two proteins despite their relatively high sequence divergence. The three specificity-determining amino acids are also capable of discriminating between the Pax-6-specific A residue and the BSAP-specific C residue at position 19 of the two consensus binding sites. However, it is worth noting that a C-to-A mutation at position 19 of the H2A-2.2 site did not convert this BSAP recognition sequence into a Pax-6-binding site (unpublished data). Position 14 additionally had to be mutated to its consensus nucleotide in order to facilitate Pax-6 binding (H2A-C [Fig. 7B]). This evidence clearly indicates that other nucleotide positions along the recognition sequence also have a significant influence on the binding specificity of Pax-6.

Mutational analysis of the paired domain and its recognition sequence revealed a bipartite structure for this DNA-binding motif (10). According to this model, the paired domain is composed of two subdomains that bind to two distinct half sites in adjacent major grooves of the DNA helix. Xu et al. (60) have recently confirmed our model by X-ray crystallographic analysis of the paired domain-DNA complex (see Fig. 11). These structural studies revealed that each of the N-terminal and C-terminal subdomains accomodates a helical region resembling the homeodomain. The N-terminal subdomain, which consists of two antiparallel β-sheets and three α-helices, binds to the 3′ half site of the recognition sequence (Fig. 7A) by extensive interaction with the phosphate backbone and bases in the major and minor groove of the DNA (Fig. 11). The recognition helix 3 is responsible for all base contacts in the major groove. Interestingly, the specificity-determining amino acids 42, 44, and 47 are located at the end of helix 2, in the linker, and at the beginning of helix 3 (Fig. 11). Histidine 47 of the *Drosophila* Prd protein was shown to interact with guanine 19 (in the complementary strand of the consensus sequence

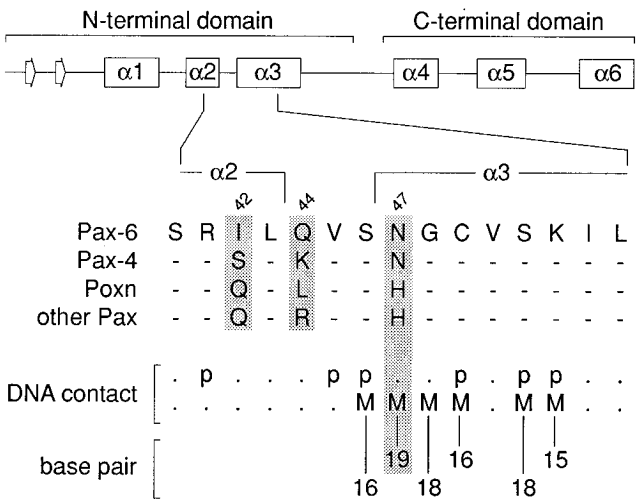


FIG. 11. Locations of amino acids 42, 44, and 47 in the structure of the paired domain-DNA complex. The secondary structure of the paired domain, which has recently been determined by X-ray crystallographic analysis (60), is schematically shown together with the amino acid sequences of different Pax proteins from position 40 to position 54 of the paired domain. Antiparallel β-sheets and α-helices are represented by arrows and boxes, respectively. Contacts of the paired domain with the phosphate backbone (p) and bases in the major groove (M) of the DNA are indicated. The nucleotide positions of the DNA-binding site are numbered as in Fig. 7A. For details, see Xu et al. (60).

shown in Fig. 7A) by hydrogen bonding. These structural data therefore support our observation that amino acid 47 is involved in discrimination between nucleotides at position 19 in the Pax-6- and BSAP-binding sites. We consider it, however, likely that this discrimination also depends on amino acids 42 and 44. First, these 2 amino acids may influence local protein conformation, thereby facilitating or interfering with phosphate backbone and major groove contacts in the region. Second, the same two positions are mutated not only in Pax-6 but also in Pax-4 and Pox neuro (Fig. 11), suggesting that the latter two proteins may exhibit yet other sequence preferences. All other Pax proteins from *D. melanogaster* to humans contain the same amino acids at positions 42 (Q), 44 (R), and 47 (H) as BSAP. However, the majority of these proteins also differ in their DNA-binding specificities from BSAP (10), and hence additional determinants in other regions of the paired domain that are responsible for differential sequence recognition by these proteins must exist. In this context, it is also important to emphasize that amino acids 42, 44, and 47 determine only the sequence specificities and not the overall DNA affinities of the Pax-6 and BSAP paired domains (Fig. 7B).

The homeodomain is the second DNA-binding motif of Pax-6, which, as demonstrated in this report, interacts by cooperative dimerization with palindromic binding sites consisting of inverted TAAT repeats. Previously, the homeodomains of the *Drosophila* Prd (59) and mammalian Pax-3 and Pax-7 proteins (44) have been shown to cooperatively dimerize on DNA. The Pax-6 homeodomain preferentially binds to inverted TAAT repeats with a 3-nucleotide spacer (P3 sites), compared with the preference of the Prd homeodomain for P2 sites (59). Moreover, optimal DNA binding of the Pax-6 and Prd homeodomains appears to depend on different nucleotide sequences in the center of the palindrome (59) which strongly influence cooperativity (Fig. 8). Interestingly, the dimerization property of the homeodomain allows different Pax proteins to interact with each other, as was demonstrated for *Drosophila* Pax proteins (59) and mammalian Pax-3 and Pax-7 (44). Heterodimerization via the homeodomain may therefore be an important mechanism for combinatorial control of gene expression by different Pax proteins. In the context of the full-length Pax-6 protein, the paired domain proved to have a higher affinity for DNA than the homeodomain (Fig. 8), suggesting that the paired domain may be the primary DNA-binding motif of Pax proteins. This hypothesis is supported by genetic evidence, as no missense mutation within the homeodomain has yet been identified in mouse developmental mutants and human disease syndromes, in marked contrast to the paired domain (reviewed in reference 49). The paired domain and homeodomain of Pax-6 are able to bind simultaneously to composite DNA elements consisting of recognition seqences for each of the two DNA-binding domains. This is best illustrated by the interaction of the specificity mutant Pax-6mut with the e5 site (Fig. 5). In contrast to wild-type Pax-6, this protein was able to bind to the e5 site by virtue of the BSAP-like binding specificity of its paired domain. Because of the additional presence of the homeodomain, the protein Pax-6mut bound to this site in comparison with simple paired domain recognition sequences with higher affinity than the homeodomainless BSAP protein (Fig. 5). In summary, we conclude that the existence of two functional DNA-binding domains in one and the same molecule renders a subclass of Pax proteins highly versatile in DNA sequence recognition.

Pax-6 has been proposed to be a transcription factor (18, 22). Here, we have provided evidence that full-length Pax-6 is indeed capable of stimulating transcription of a gene containing high-affinity Pax-6-binding sites in its control region. Pax-6, like other members of the Pax protein family, is able to activate transcription over a long distance and in a position-independent manner and, hence, has all the properties of an enhancer-binding protein. Even though the *Pax-6* gene is expressed in a tissue-restricted fashion (33, 55), transactivation by the Pax-6 protein is apparently independent of the cell type-specific background. A potent transactivation domain has been identified within the C-terminal serine-threonine-proline-rich region of Pax-6 by fusion to the heterologous DNA-binding motif of the yeast GAL4 protein (Fig. 9D). The importance of this domain for transactivation is, furthermore, emphasized by the discovery of several nonsense and frameshift mutations in the corresponding region of the *Pax-6* gene of aniridia patients (11, 18, 25, 29) and of one *Small eye* mutant (26). Most remarkably, the C-terminal transactivation function has been conserved between the sea urchin and mouse Pax-6 proteins. This functional conservation is also reflected at the sequence level, as stretches of up to 16 identical amino acids are present in the C-terminal domains of the echinoderm and vertebrate Pax-6 proteins. These homology regions almost certainly contain the critical amino acid residues involved in transactivation and thus provide an interesting starting point for detailed analysis of the sequence requirement of this Pax-6 function.

ACKNOWLEDGMENTS

We are grateful to C. Gache for help with sea urchin embryo cultures, to W. Schaffner and Z. Kozmik for providing the GAL-VP16 and Pax-6-ins expression plasmids, to I. Fetka for in situ hybridization analysis, to G. Schaffner for oligonucleotide synthesis, to R. Kurzbauer for DNA sequencing, to H. Tkadletz for graphical work, and to M. Cotten for critical reading of the manuscript.

This work was supported in part by a grant from the Austrian Industrial Research Promotion Fund.

ADDENDUM IN PROOF

Recent transient-transfection experiments have indicated that full-length Pax-6 can stimulate transcription from minimal promoters containing either a single high-affinity paired domain [CD19-2(A-ins)]- or homeodomain (P3)-binding site with equal efficiency in RAC65 cells (unpublished data). Hence, both DNA-binding domains of Pax-6 are able to mediate similar levels of transcriptional activation of target genes despite the extreme differences observed in in vitro DNA-binding assays (Fig. 8C).

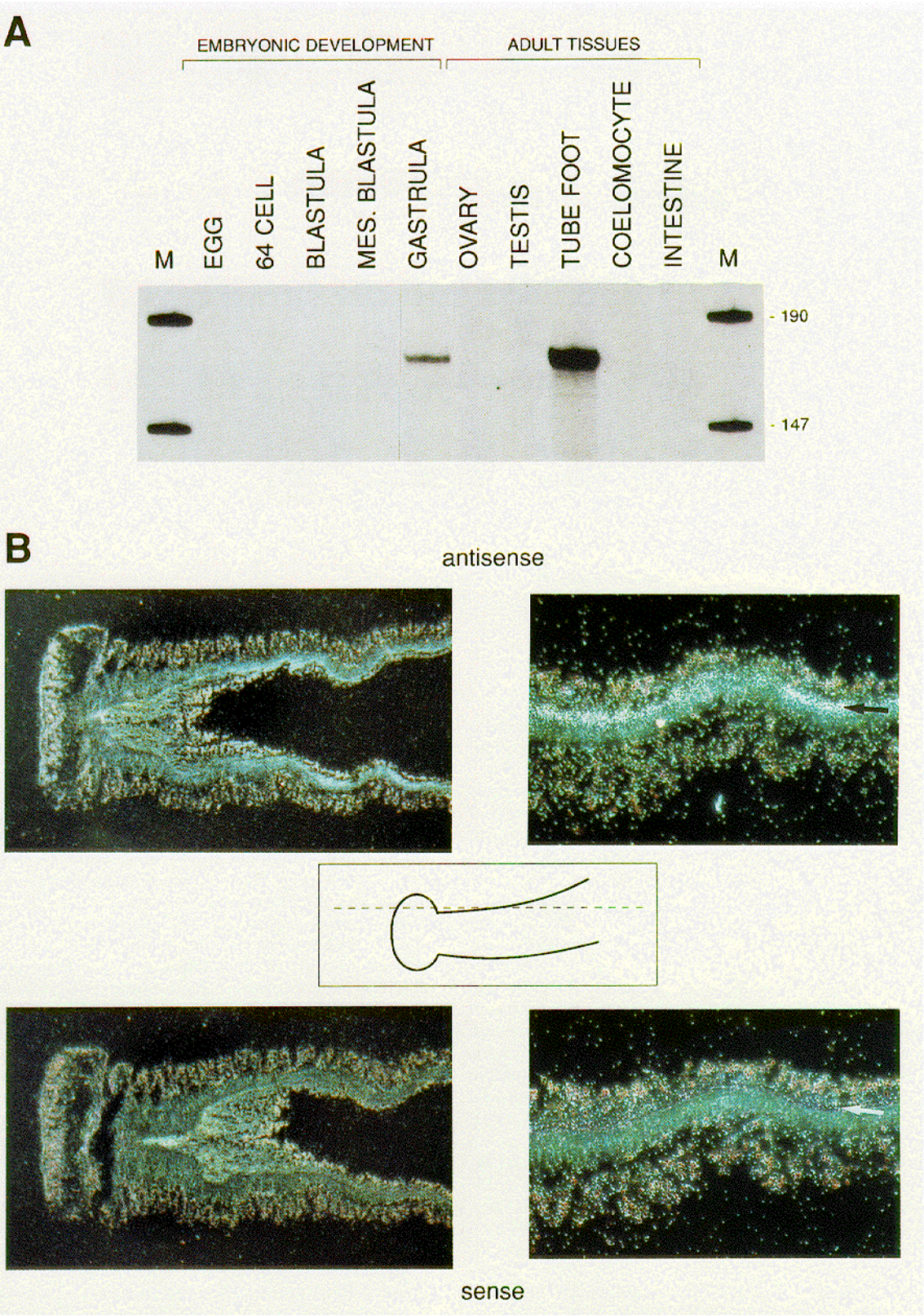
REFERENCES

1. **Adams, B., P. Dörfler, A. Aguzzi, Z. Kozmik, P. Urbánek, I. Maurer-Fogy, and M. Busslinger.** 1992. *Pax-5* encodes the transcription factor BSAP and is expressed in B lymphocytes, the developing CNS, and adult testis. Genes Dev. **6:**1589–1607.
2. **Baldwin, C. T., C. F. Hoth, J. A. Amos, E. O. da-Silva, and A. Milunsky.** 1992. An exonic mutation in the HuP2 paired domain gene causes Waardenburg's syndrome. Nature (London) **355:**637–638.
3. **Balling, R., U. Deutsch, and P. Gruss.** 1988. *undulated*, a mutation affecting the development of the mouse skeleton, has a point mutation in the paired box of *Pax 1*. Cell **55:**531–535.
4. **Barberis, A., G. Superti-Furga, L. Vitelli, I. Kemler, and M. Busslinger.** 1989. Developmental and tissue-specific regulation of a novel transcription factor of the sea urchin. Genes Dev. **3:**663–675.
5. **Braselmann, S., P. Graninger, and M. Busslinger.** 1993. A selective transcriptional induction system for mammalian cells based on Gal4-estrogen receptor fusion proteins. Proc. Natl. Acad. Sci. USA **90:**1657–1661.
6. **Chalepakis, G., R. Fritsch, H. Fickenscher, U. Deutsch, M. Goulding, and P. Gruss.** 1991. The molecular basis of the *undulated*/*Pax-1* mutation. Cell **66:**873–884.
7. **Chalepakis, G., J. Wijnholds, P. Giese, M. Schachner, and P. Gruss.** 1994. Characterization of Pax-6 and Hoxa-1 binding to the promoter region of the neural cell adhesion molecule L1. DNA Cell Biol. **13:**891–900.
8. **Chirgwin, J. M., A. E. Przybyla, R. J. MacDonald, and W. J. Rutter.** 1979. Isolation of biologically active ribonucleic acid from sources enriched in ribonuclease. Biochemistry **18:**5294–5299.
9. **Cvekl, A. C., C. M. Sax, E. H. Bresnick, and J. Piatigorsky.** 1994. A complex array of positive and negative elements regulates the chicken αA-crystallin gene: involvement of Pax-6, USF, CREB and/or CREM, and AP-1 proteins. Mol. Cell. Biol. **14:**7363–7376.
10. **Czerny, T., G. Schaffner, and M. Busslinger.** 1993. DNA sequence recognition by Pax proteins: bipartite structure of the paired domain and its binding site. Genes Dev. **7:**2048–2061.
11. **Davis, A., and J. K. Cowell.** 1993. Mutations in the Pax6 gene in patients with hereditary aniridia. Hum. Mol. Genet. **2:**2093–2097.
12. **Davis, R. J., C. M. D'Cruz, M. A. Lovell, J. A. Biegel, and F. G. Barr.** 1994. Fusion of *PAX7* to *FKHR* by the variant t(1;13)(p36;q14) translocation in alveolar rhabdomyosarcoma. Cancer Res. **54:**2869–2872.
13. **Dozier, C., C. Carrière, D. Grévin, P. Martin, B. Quatannens, D. Stéhelin, and S. Saule.** 1993. Structure and DNA-binding properties of *Pax-QNR*, a paired box- and homeobox-containing gene. Cell Growth Diff. **4:**281–289.
14. **Epstein, D. J., M. Vekemans, and P. Gros.** 1991. *Splotch* (*Sp^{2H}*), a mutation affecting development of the mouse neural tube, shows a deletion within the paired homeodomain of *Pax-3*. Cell **67:**767–774.
15. **Epstein, J., J. Cai, T. Glaser, L. Jepeal, and R. Maas.** 1994. Identification of a *Pax* paired domain recognition sequence and evidence for DNA-dependent conformational changes. J. Biol. Chem. **269:**8355–8361.
16. **Epstein, J. A., T. Glaser, J. Cai, L. Jepeal, D. S. Walton, and R. L. Maas.** 1994. Two independent and interactive DNA-binding subdomains of the Pax6 paired domain are regulated by alternative splicing. Genes Dev. **8:**2022–2034.
17. **Galili, N., R. J. Davis, W. J. Fredericks, S. Mukhopadhyay, F. J. Rauscher III, B. S. Emanuel, G. Rovera, and F. G. Barr.** 1993. Fusion of a fork head domain gene to *PAX3* in the solid tumour alveolar rhabdomyosarcoma. Nature Genet. **5:**230–235.
18. **Glaser, T., L. Jepeal, J. G. Edwards, S. R. Young, J. Favor, and R. L. Maas.** 1994. *PAX6* gene dosage effect in a family with congenital cataracts, aniridia, anophthalmia and central nervous defects. Nature Genet. **7:**463–471.
19. **Glaser, T., J. Lane, and D. Housman.** 1990. A mouse model of the aniridia-Wilms tumor deletion syndrome. Science **250:**823–827.
20. **Glaser, T., D. S. Walton, and R. L. Maas.** 1992. Genomic structure, evolutionary conservation and aniridia mutations in the human *PAX6* gene. Nature Genet. **2:**232–239.
21. **Goulding, M. D., A. Lumsden, and P. Gruss.** 1993. Signals from the notochord and floor plate regulate the region-specific expression of two Pax genes in the developing spinal cord. Development **117:**1001–1016.
22. **Gruss, P., and C. Walther.** 1992. Pax in development. Cell **69:**719–722.
23. **Gubler, U., and B. J. Hoffman.** 1983. A simple and very efficient method for generating cDNA libraries. Gene **25:**263–269.
24. **Hanson, I. M., J. M. Fletcher, T. Jordan, A. Brown, D. Taylor, R. J. Adams, H. H. Punnett, and V. van Heyningen.** 1994. Mutations at the *PAX6* locus are found in heterogeneous anterior segment malformations including Peters' anomaly. Nature Genet. **6:**168–173.
25. **Hanson, I. M., A. Seawright, K. Hardman, S. Hodgson, D. Zaletayev, G. Fekete, and V. van Heyningen.** 1993. *PAX6* mutations in aniridia. Hum. Mol. Genet. **2:**915–920.
26. **Hill, R. E., J. Favor, B. L. Hogan, C. C. Ton, G. F. Saunders, I. M. Hanson, J. Prosser, T. Jordan, N. D. Hastie, and V. van Heyningen.** 1991. Mouse *Small eye* results from mutations in a paired-like homeobox-containing gene. Nature (London) **354:**522–525.
27. **Holst, B. D., R. S. Goomer, I. C. Wood, G. M. Edelman, and F. S. Jones.** 1994. Binding and activation of the promoter of the neural cell adhesion molecule by Pax-8. J. Biol. Chem. **269:**22245–22252.
28. **Hyman, L.** 1955. The invertebrates: Echinodermata, vol. 4. McGraw-Hill, New York.
29. **Jordan, T., I. Hanson, D. Zaletayev, S. Hodgson, J. Prosser, A. Seawright, N. Hastie, and V. van Heyningen.** 1992. The human PAX6 gene is mutated in two patients with aniridia. Nature Genet. **1:**328–332.
30. **Kozak, M.** 1991. Structural features in eukaryotic mRNAs that modulate the initiation of translation. J. Biol. Chem. **266:**19867–19870.

30a. **Kozmik, Z. et al.** Unpublished data.

31. **Kozmik, Z., R. Kurzbauer, P. Dörfler, and M. Busslinger.** 1993. Alternative splicing of *Pax-8* gene transcripts is developmentally regulated and generates isoforms with different transactivation properties. Mol. Cell. Biol. **13:**6024–6035.
32. **Kozmik, Z., S. Wang, P. Dörfler, B. Adams, and M. Busslinger.** 1992. The promoter of the CD19 gene is a target for the B-cell-specific transcription factor BSAP. Mol. Cell. Biol. **12:**2662–2672.
33. **Krauss, S., T. Johansen, V. Korzh, U. Moens, J. U. Ericson, and A. Fjose.** 1991. Zebrafish *pax[zf-a]*: a paired box-containing gene expressed in the neural tube. EMBO J. **10:**3609–3619.
34. **Li, H.-S., J.-M. Yang, R. D. Jacobson, D. Pasko, and O. Sundin.** 1994. Pax-6 is first expressed in a region of ectoderm anterior to the early neural plate: implications for stepwise determination of the lens. Dev. Biol. **162:**181–194.
35. **Liao, F., S. L. Giannini, and B. K. Birshtein.** 1992. A nuclear DNA-binding protein expressed during early stages of B-cell differentiation interacts with diverse segments within and 3′ of the IgH chain gene cluster. J. Immunol. **148:**2909–2917.
36. **Martin, P., C. Carrière, C. Dozier, B. Quatannens, M. A. Mirabel, B. Vandenbunder, D. Stéhelin, and S. Saule.** 1992. Characterization of a paired box- and homeobox-containing quail gene (*Pax-QNR*) expressed in the neuroretina. Oncogene **7:**1721–1728.
37. **Matsuo, F., N. Osumi-Yamashita, S. Noji, H. Ohuchi, E. Koyama, F. Myokai, N. Matsuo, S. Taniguchi, H. Doi, S. Iseki, Y. Ninomiya, M. Fujiwara, T. Watanabe, and K. Eto.** 1993. A mutation in the *Pax-6* gene in rat *small eye* is associated with impaired migration of midbrain crest cells. Nature Genet. **3:**299–304.
38. **Maulbecker, C. C., and P. Gruss.** 1993. The oncogenic potential of Pax genes. EMBO J. **12:**2361–2367.
39. **Noll, M.** 1993. Evolution and role of *Pax* genes. Curr. Opin. Genet. Dev. **3:**595–605.
40. **Pratt, M. A. C., J. Kralova, and M. W. McBurney.** 1990. A dominant negative mutation of the alpha retinoic acid receptor gene in a retinoic acid-nonresponsive embryonal carcinoma cell. Mol. Cell. Biol. **10:**6445–6453.
41. **Püschel, A. W., P. Gruss, and M. Westerfield.** 1992. Sequence and expression pattern of *pax-6* are highly conserved between zebrafish and mice. Development **114:**643–651.
42. **Quiring, R., U. Walldorf, U. Kloter, and W. J. Gehring.** 1994. Homology of the *eyeless* gene of *Drosophila* to the *Small eye* gene in mice and *aniridia* in humans. Science **265:**785–789.
43. **Rothman, P., S. C. Li, B. Gorham, L. Glimcher, F. Alt, and M. Boothby.** 1991. Identification of a conserved lipopolysaccharide-plus-interleukin-4-responsive element located at the promoter of germ line ε transcripts. Mol. Cell. Biol. **11:**5551–5561.
44. **Schäfer, B. W., T. Czerny, M. Bernasconi, M. Genini, and M. Busslinger.** 1994. Molecular cloning and characterization of a human PAX-7 cDNA expressed in normal and neoplastic myocytes. Nucleic Acids Res. **22:**4574–4582.
45. **Seipel, K., O. Georgiev, and W. Schaffner.** 1992. Different activation domains stimulate transcription from remote ('enhancer') and proximal ('promoter') positions. EMBO J. **11:**4961–4968.
46. **Shapiro, D. N., J. E. Sublett, B. Li, J. R. Downing, and C. W. Naeve.** 1993. Fusion of *PAX3* to a member of the forkhead family of transcription factors in human alveolar rhabdomyosarcoma. Cancer Res. **53:**5108–5112.
47. **Singh, M., and B. K. Birshtein.** 1993. NF-HB (BSAP) is a repressor of the murine immunoglobulin heavy-chain 3′α enhancer at early stages of B-cell differentiation. Mol. Cell. Biol. **13:**3611–3622.
48. **Stapleton, P., A. Weith, P. Urbánek, Z. Kozmik, and M. Busslinger.** 1993. Chromosomal localization of seven *PAX* genes and cloning of a novel family

member, *PAX-9*. Nature Genet. **3:**292–298.
49. **Strachan, T., and A. P. Read.** 1994. PAX genes. Curr. Opin. Genet. Dev. **4:**427–438.
50. **Tassabehji, M., A. P. Read, V. E. Newton, R. Harris, R. Balling, P. Gruss, and T. Strachan.** 1992. Waardenburg's syndrome patients have mutations in the human homologue of the *Pax-3* paired box gene. Nature (London) **355:**635–636.
51. **Ton, C. C. T., H. Hirvonen, H. Miwa, M. M. Weil, P. Monaghan, T. Jordan, V. van Heyningen, N. D. Hastie, H. Meijers-Heijboer, M. Drechsler, B. Royer-Pokora, F. Collins, A. Swaroop, L. C. Strong, and G. F. Saunders.** 1991. Positional cloning and characterization of a paired box- and homeobox-containing gene from the aniridia region. Cell **67:**1059–1074.
52. **Treisman, J., E. Harris, and C. Desplan.** 1991. The paired box encodes a second DNA-binding domain in the Paired homeo domain protein. Genes Dev. **5:**594–604.
53. **Urbánek, P., Z.-Q. Wang, I. Fetka, E. F. Wagner, and M. Busslinger.** 1994. Complete block of early B cell differentiation and altered patterning of the posterior midbrain in mice lacking Pax5/BSAP. Cell **79:**901–912.
54. **Vitelli, L., I. Kemler, B. Lauber, M. L. Birnstiel, and M. Busslinger.** 1988. Developmental regulation of micro-injected histone genes in sea urchin embryos. Dev. Biol. **127:**54–63.
55. **Walther, C., and P. Gruss.** 1991. *Pax-6*, a murine paired box gene, is expressed in the developing CNS. Development **113:**1435–1449.
56. **Walther, C., J. L. Guenet, D. Simon, U. Deutsch, B. Jostes, M. D. Goulding, D. Plachov, R. Balling, and P. Gruss.** 1991. Pax: a murine multigene family of paired box-containing genes. Genomics **11:**424–434.
57. **Waters, S. H., K. U. Saikh, and J. Stavnezer.** 1989. A B-cell-specific nuclear protein that binds to DNA sites 5′ to immunoglobulin Sα tandem repeats is regulated during differentiation. Mol. Cell. Biol. **9:**5594–5601.
58. **Williams, M., and N. Maizels.** 1991. LR1, a lipopolysaccharide-responsive factor with binding sites in the immunoglobulin switch regions and heavy-chain enhancer. Genes Dev. **5:**2353–2361.
59. **Wilson, D., G. Sheng, T. Lecuit, N. Dostatni, and C. Desplan.** 1993. Cooperative dimerization of Paired class homeo domains on DNA. Genes Dev. **7:**2120–2134.
60. **Xu, W., M. A. Rould, S. Jun, C. Desplan, and C. O. Pabo.** 1995. Crystal structure of the paired domain-DNA complex at 2.5 Å resolution reveals structural basis for Pax developmental mutations. Cell **80:**639–650.
61. **Zannini, M., H. Francis-Lang, D. Plachov, and R. Di Lauro.** 1992. Pax-8, a paired domain-containing protein, binds to a sequence overlapping the recognition site of a homeodomain and activates transcription from two thyroid-specific promoters. Mol. Cell. Biol. **12:**4230–4241.
62. **Zwollo, P., and S. Desiderio.** 1994. Specific recognition of the *blk* promoter by the B-lymphoid transcription factor B-cell-specific activator protein. J. Biol. Chem. **269:**15310–15317.


A
EMBRYONIC DEVELOPMENT
ADULT TISSUES
M
EGG
64 CELL
BLASTULA
MES. BLASTULA
GASTRULA
OVARY
TESTIS
TUBE FOOT
COELOMOCYTE
INTESTINE
M
- 190
- 147
B
antisense
sense