

Two independent and interactive DNA-binding subdomains of the Pax6 paired domain are regulated by alternative splicing

Jonathan A. Epstein,^{1,2} Tom Glaser,² Jiexing Cai,² Lisa Jepeal,² David S. Walton,³ and Richard L. Maas^{2,4}

Divisions of Cardiology¹ and Genetics², Department of Medicine and Howard Hughes Medical Institute, Brigham and Women's Hospital, and Massachusetts Eye and Ear Infirmary,³ Harvard Medical School, Boston, Massachusetts 02115 USA

Vertebrate Pax proteins share a conserved 128-amino-acid DNA-binding motif, the paired domain. The *PAX6* gene, which is mutated in the murine *Small eye* and human aniridia developmental defects, also encodes a second protein with a 14-amino-acid insertion in the paired domain. This protein, which arises by alternative mRNA splicing, exhibits unique DNA-binding properties. Unlike other paired domains, which bind DNA predominantly by their amino termini, the extended Pax6 paired domain interacts with DNA exclusively through its carboxyl terminus. This property can be simulated by deletion of 30 amino-terminal residues from the Pax6 or Pax2 paired domains. Thus, the insertion acts as a molecular toggle to unmask the DNA-binding potential of the carboxyl terminus. The functional nonequivalence of the two Pax6 proteins is underscored by a T → C mutation at position -3 of the alternative splice acceptor site that changes the ratio of the two isoforms and causes a distinct human ocular syndrome.

[Key Words: Pax genes; transcription factor; alternative mRNA splicing; DNA binding; ocular development; aniridia]

Received May 17, 1994; revised version accepted July 18, 1994.

Transcription factors responsible for embryonic patterning and tissue differentiation regulate the expression of target genes by specific DNA sequence recognition. In large part, their unique functional and DNA-binding characteristics have evolved through the modular assembly of conserved domains (Kessel and Gruss 1990). These domains define classes of transcription factors, such as the basic leucine zipper (bZIP), basic helix-loop-helix (bHLH), zinc finger, and homeo box families. Mechanisms that refine DNA-binding specificity within these families are of particular interest, as they generate target gene diversity.

The Pax (paired box) family of developmentally regulated transcription factors share an amino-terminal DNA-binding motif known as the paired domain (for review, see Chalepakis et al. 1992). This domain consists of 128 amino acids and has been identified in a number of metazoan species, including *Drosophila* and humans (Noll 1993). The nine mammalian Pax genes are divided into four groups on the basis of their amino acid sequence, intron-exon boundaries, and the presence of

other conserved motifs (Walther et al. 1991; Wallin et al. 1993). The paired domain is predicted to contain at least three α -helical regions (Bopp et al. 1989) and circular dichroism spectral data suggest that a stable conformation with α -helical characteristics is adopted upon specific DNA binding (Epstein et al. 1994). Within the paired domain, an evolutionary discontinuity divides the relatively conserved 74-amino acid amino terminus from the more divergent 54-amino-acid carboxyl terminus and coincides with an exon boundary (Walther et al. 1991).

The *PAX6* gene is mutated in *Small eye* mice (Hill et al. 1991) and humans with the congenital eye disease aniridia (Ton et al. 1991; Glaser et al. 1994b). Because these phenotypes are semidominant and the mutations cause a loss of function, the Pax6 proteins appear to control rate-limiting steps in eye development. The *PAX6* gene spans 22 kb, is divided into 14 exons, and encodes a 422-amino-acid protein that is identical in humans and mice. In addition to the paired domain, Pax6 contains a homeo domain and a carboxy-terminal region capable of activating transcription (Glaser et al. 1994a). The paired domain (amino acids 3–130) is encoded by exons 5–7. An alternatively spliced exon (exon 5a) has been identified in several vertebrate *PAX6* genes and is included variably

⁴Corresponding author.

in the mature mRNA transcript, resulting in the insertion of a 14-amino-acid peptide in the paired domain after Gln47 (Walther and Gruss 1991; Glaser et al. 1992; Puschel et al. 1992). Both splice products are expressed in the developing eye, brain, spinal cord, and olfactory epithelium, with the exon 5a-containing transcripts at slightly lower levels (Ton et al. 1991; Walther and Gruss 1991; T. Glaser and R.L. Maas, unpubl.).

We have identified previously a high-affinity binding site for the Pax6 paired domain using an in vitro-binding site selection (Epstein et al. 1994). A similar sequence has been identified by stepwise replacement of positions in a potential genomic DNA-binding site for Pax5 (Czerny et al. 1993). Although Pax6 and Pax5 (Barberis et al. 1990; Czerny et al. 1993; Kozmik et al. 1993; Plaza et al. 1993) bind related DNA sequences such that up to 28 bp are protected from DNase I digestion, other paired domains exhibit different DNA-binding characteristics. The *Drosophila* Paired protein recognizes the e5 sequence located in the promoter region of the *even-skipped* gene. In this case, the paired domain footprints only 15 bp and deletion of the carboxyl terminus has no effect on e5 binding (Treisman et al. 1991). In contrast, the Pax1 paired domain contacts modified e5 sequences along a 24-bp distance, with the amino-terminal portion responsible for sequence specificity and the carboxy-terminal portion required for high-affinity binding or stability of the protein–DNA complex (Chalepakos et al. 1991). The Pax8 paired domain recognizes sequences within the promoter regions of two thyroid-specific genes that are not obviously related to e5 and footprints ~15 bp at each site (Zannini et al. 1992). Pax5 appears to contact DNA in a bipartite manner in which both regions of contact contribute to the overall affinity of binding (Czerny et al. 1993). Together, these results suggest that the amino terminus of the paired domain is required for specific DNA recognition. However, the role of the carboxyl terminus is unclear, and the function of the 5a insertion in Pax6 is unknown.

In this report we describe unique properties of the alternatively spliced form of Pax6, which have allowed us to examine the DNA-binding characteristics of the paired domain in detail. We demonstrate that (1) the paired domain functions as two discrete subdomains with distinct binding specificities, (2) insertion of the 5a peptide into the amino-terminal subdomain enables the carboxy-terminal subdomain to recognize its binding site, and (3) this binding is an important determinant of transcriptional activity. Finally, we describe a human PAX6 mutation that changes the ratio of the alternatively spliced forms and causes an unusual eye phenotype, indicating that the two Pax6 paired domains function differently during embryogenesis.

Results

A splice acceptor mutation in exon 5a

To investigate the biological function of the extended Pax6 paired domain, we tested DNA samples from pa-

tients with aniridia and related ocular disorders for point mutations in exon 5a and flanking genomic DNA. Heteroduplex and single-strand conformational analysis revealed a single-base substitution in one family (Bh) with a distinct clinical phenotype (Fig. 1). The disease affects a mother and daughter with bilateral juvenile cataracts, peripheral corneal opacification, glaucoma, and pendular nystagmus, an unsteady oscillating gaze attributed to a congenital sensory deficiency (see Materials and methods). However, in contrast to the typical manifestations of aniridia, such as iris aplasia or severe hypoplasia (Nelson et al. 1984), the irides in these two patients were relatively intact, the pupils were round, and the maculas, the central region of the retinas with the greatest density of photoreceptors, appeared normal. Abnormalities were noted in the anterior aspect of the iris derived from neural crest, including small radial defects, decreased vasculature and crypt density, and thinning of the iris stroma. The posterior pigmented iris layer derived from neural ectoderm was essentially normal.

The Bh family mutation is a T → C transition at position –3 in the exon 5a splice acceptor (Fig. 2A). The mutation appeared as a de novo change in the mother's DNA and creates an *MaeIII* restriction site, which was used to follow its transmission in the pedigree. The mutation is present only in the mother and daughter, correlating completely with the ocular phenotype, and was not observed in any of >50 unrelated DNA samples.

To determine the effect of the Bh mutation, we compared mutant and wild-type genomic fragments in an in vivo splicing assay (Fig. 2B). Transcripts containing ex-

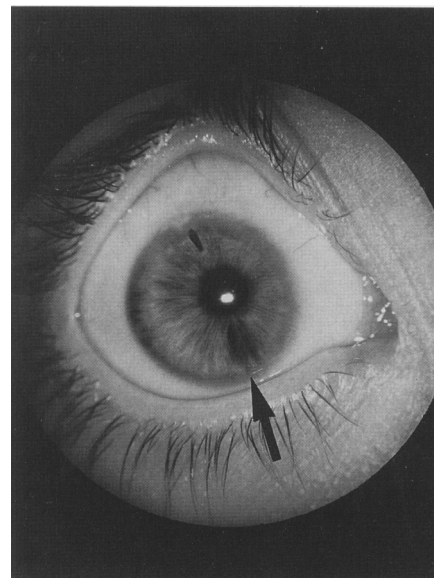


Figure 1. Ocular phenotype in family Bh. Anterior segment photograph of the daughter (right eye). The pupil is displaced nasally (corectopia) and an iris stromal defect is visible (arrow). The cornea is opacified by pannus formation at the superior margin. A surgical iridectomy is present superiorly. The left eye has similar findings.

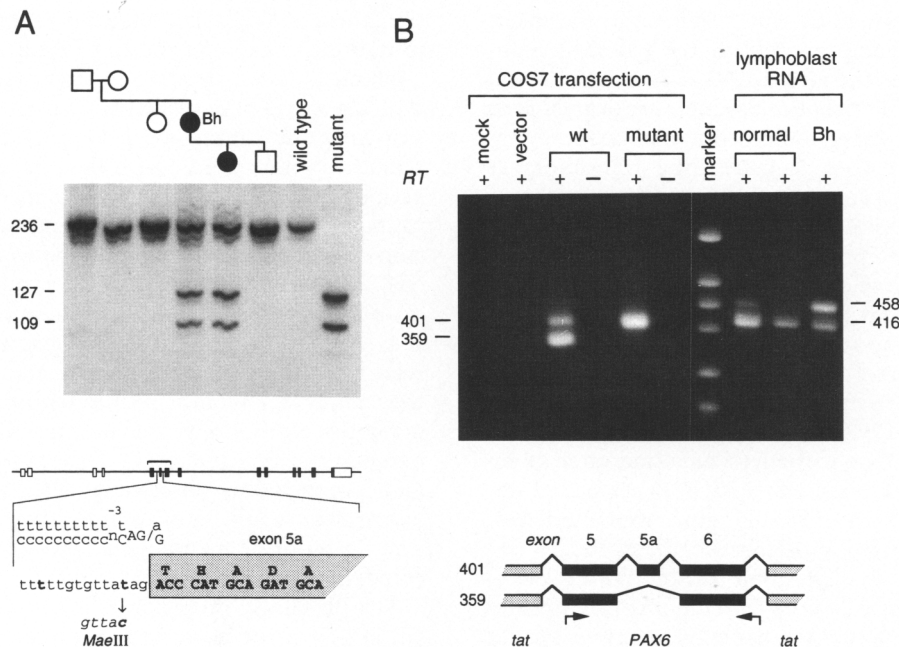


Figure 2. Mutational analysis of *PAX6* in family Bh. (A) Exon 5a splice acceptor mutation. (Top) *MaeIII* digest of radiolabeled PCR products spanning exon 5a. These products were amplified from genomic DNA of indicated family members and pSPL1 plasmid clones containing mutant or wild-type sequences. (□○) Normal; (■●) affected. (Bottom) Genomic map of *PAX6*, showing the 5' end of exon 5a and the T → C mutation at position -3. Coding regions are shaded. A consensus splice acceptor site (Senapathy et al. 1990) is aligned above for comparison. The mutation creates an *MaeIII* restriction site, which divides the 236-bp PCR product into 109- and 127-bp fragments. The sequence shown is identical to mouse (data not shown) and quail (Dozier et al. 1993) *PAX6* genomic DNA. The bracket indicates the genomic fragment used in the *in vivo* splicing assay. (B) The Bh mutation significantly increases the use of alternative exon 5a. (Top left) *In vivo* splicing assay. COS7 cells were transfected with expression plasmids containing mutant or wild-type (wt) *PAX6* genomic sequences, no DNA (mock), or pSPL1 DNA alone (vector). The relative expression of the two *PAX6* transcripts was assessed by RT PCR analysis. The addition or omission of RT enzyme from these reactions is indicated above the lanes. (Bottom) Schematic representation of PCR products detected in the *in vivo* splicing assay. (Solid bar) *PAX6*; (stippled bar) *HIVtat* sequences. The upstream primer is located in exon 5 and the downstream primer spans the *tat*-exon 6 junction. (Top right) RT PCR of *PAX6* transcripts in lymphoblasts derived from two healthy individuals and a heterozygous patient with the Bh mutation. The primers are located in exons 3 and 6. Fragment sizes are indicated in bp.

ons 5 and 6 were specifically amplified from transfected COS7 cells, and the relative abundance of exon 5a-containing products was assessed. The wild-type allele generates both expected splice forms, with an excess of the 359-bp product lacking exon 5a. The ratio is similar to that observed for *PAX6* transcripts in fetal brain and eye (data not shown). In comparison, the single-base-pair Bh mutation favors the use of the exon 5a splice acceptor, causing a marked increase in the abundance of the 401-bp product. Qualitatively similar results were obtained in other cell lines (data not shown). This increase was also reflected in the ratio of endogenous *PAX6* transcripts in lymphoblasts derived from the affected mother, who is heterozygous (Fig. 2B, lane Bh). Assuming the overall level of *PAX6* gene expression is not affected in these patients, the mutation is expected to increase the amount of the alternative Pax6 protein (Pax6-5a) and decrease the amount of the shorter form (Pax6).

Identification of a high-affinity DNA-binding sequence for Pd-5a

The two Pax6 paired domains (denoted Pd and Pd-5a)

were expressed in *Escherichia coli* as glutathione S-transferase (GST) fusion proteins, and their DNA-binding characteristics were examined in an electrophoretic mobility shift assay (EMSA). The extended paired domain (Pd-5a) does not bind P6CON (Fig. 3A), which is the consensus sequence identified previously using the shorter Pax6 paired domain, Pd (Epstein et al. 1994). Pd-5a also does not bind other paired domain recognition sites, including e5 and modified e5 (PRS4) sequences (data not shown), even under low stringency conditions; therefore, we performed a DNA-binding site selection with Pd-5a protein under the same conditions used to select P6CON. Similar results were obtained with oligonucleotide pools containing a core of 25 or 35 random positions (see Materials and methods). After nine rounds of selection using oligonucleotides containing 35 random positions, 75% (18/24) of the selected oligonucleotides contained a 22-bp consensus sequence (Fig. 3A) with no more than two mismatches. This sequence (denoted 5aCON) binds more tightly to Pd-5a by EMSA than any other sequence tested and was therefore chosen as an appropriate binding site for molecular analysis of the extended paired domain.

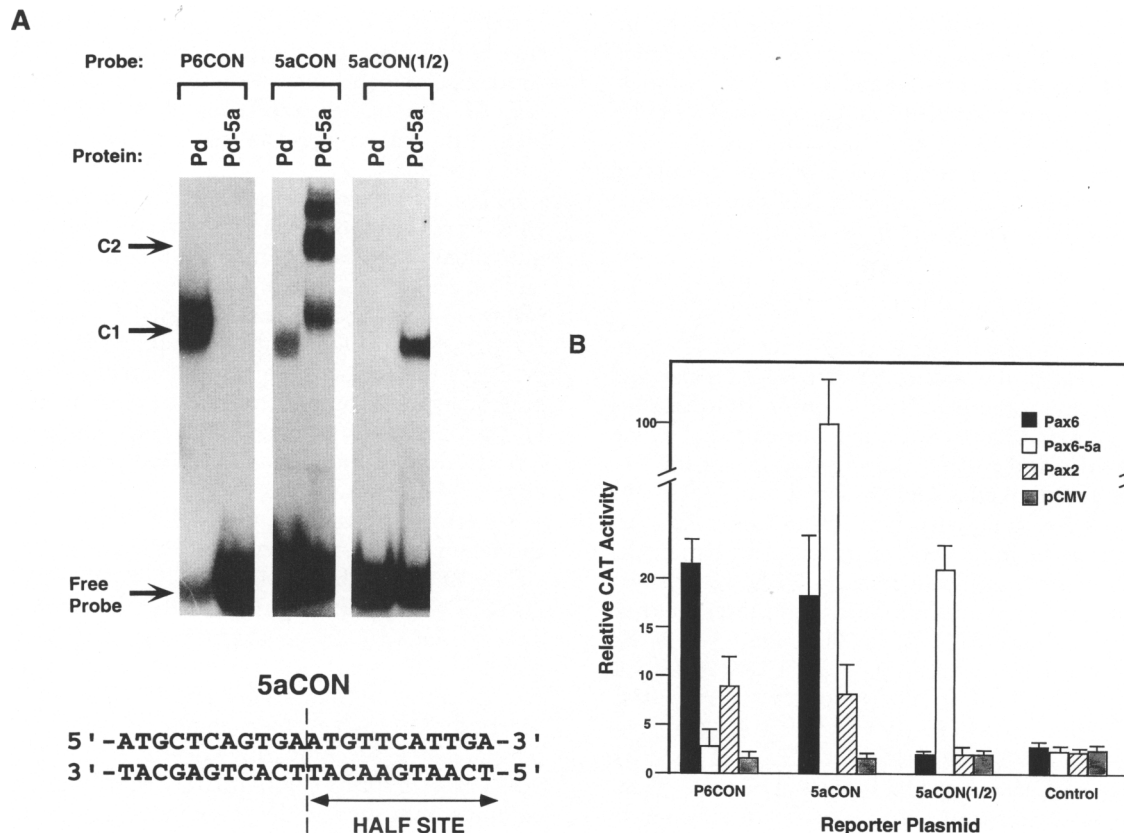


Figure 3. The 5a insertion alters DNA sequence recognition by the Pax6 paired domain. (A) EMSA experiment showing the binding of GST-Pd and GST-Pd-5a fusion proteins (0.5 μ M) to oligonucleotide probes identified by PCR selection. Pd binds its consensus sequence (P6CON), whereas Pd-5a does not. Both paired domains bind 5aCON, although the affinity and the ratio of faster (C1) to slower (C2) migrating complexes are different. 5aCON was selected with Pd-5a and contains two copies of an imperfect 11-bp direct repeat. Pd does not bind half-site probe 5aCON(1/2), whereas Pd-5a forms a single complex. (B) CAT assays. P19 cells were transfected with the indicated CAT reporter and pCMVPax expression plasmids. Pax6 stimulates CAT expression from a reporter containing six copies of P6CON, but Pax6-5a does not. All three Pax proteins *trans*-activate from a reporter containing two copies of the 5aCON-binding site. Pax6-5a, but not Pax6 or Pax2, *trans*-activates from a reporter containing six individual 5aCON half-sites. The control represents the pG₅ECAT vector with no Pax-binding sites. pCMV represents the expression vector alone with no Pax-coding sequences.

5aCON contains two copies of an imperfect 11-bp direct repeat and has no significant homology to P6CON. It forms several complexes when incubated with Pd-5a protein (Fig. 3A), indicating that one, two, or more protein molecules can bind the DNA probe. Bands representing higher order complexes predominate even under conditions where DNA is in excess. The Pd protein also binds the 5aCON probe, although the pattern of shifted bands is different (Fig. 3A), and there is no such preference for higher order complexes, although a low intensity second band is sometimes observed at high protein concentrations (see Fig. 6B, below). The formation of DNA-protein complexes is not affected by the GST moiety, as Pd and Pd-5a proteins containing an amino-terminal polyhistidine tail instead of GST give identical patterns of shifted complexes in EMSA experiments (data not shown).

To compare the protein-DNA ratios directly within the complexes formed upon Pd-5a binding to 5aCON, we

performed double-label experiments in which Pd-5a was expressed as a GST fusion protein with a single cAMP-dependent protein kinase phosphorylation site. This protein, in EMSA with 5aCON, gave an identical pattern of shifted complexes to the fusion protein lacking the phosphorylation site (data not shown). After phosphorylation with [γ -³²P]ATP, the protein was incubated with 5aCON end-labeled with [γ -³³P]ATP. The relative protein to DNA content of shifted complex C2 compared with C1 was 1.95 ± 0.14 (mean \pm s.e.), suggesting that C2 is a dimer of Pd-5a binding to 5aCON. Previous studies of Pd binding to P6CON (Epstein et al. 1994) suggest that C1 likely represents monomeric binding. For technical reasons (see Materials and methods) the protein/DNA ratio for the most slowly migrating Pd-5a/5aCON complex seen in Figure 3A could not be determined.

An equilibrium dissociation constant (K_d) was estimated for binding of both paired domain forms to the 5aCON probe by performing EMSA reactions with a

fixed, limiting amount of DNA in the absence of non-specific DNA competitor (data not shown). We obtained K_d values of 5.5×10^{-9} M for Pd and 3×10^{-9} M for Pd-5a. These values are comparable to the K_d observed for binding of Pd to P6CON (2.5×10^{-9} M) (Epstein et al. 1994), and represent an upper limit, as they were calculated assuming bimolecular, monomeric binding reactions with 5aCON (Fried 1989).

Pax6 activates transcription upon binding 5aCON

To assess the functional significance of the 5aCON consensus, we tested the ability of full-length Pax6 and Pax6-5a proteins to activate transcription from this sequence in vivo, in comparison with murine Pax2, which has been shown to activate transcription in murine P19 embryo carcinoma cells (Fickenscher et al. 1993). We cotransfected P19 cells with Pax6, Pax6-5a, or Pax2 expression plasmids and a reporter plasmid in which the chloramphenicol acetyltransferase (CAT) gene was placed downstream of a minimal E1b promoter and two to six copies of the P6CON- or 5aCON-binding sites. We also tested a reporter containing six 5aCON half-sites, each separated by 6–14 bp [see Materials and methods]. The Pax6-5a protein efficiently stimulates CAT expression when 5aCON is present, whereas Pax6 is effective when either P6CON or 5aCON is present (Fig. 3B). Transcriptional activation correlates with the in vitro ability of the Pax6 paired domain, but not Pd-5a, to bind P6CON (Fig. 3A). Pax2 also stimulates CAT expression from P6CON, but to a lesser extent than Pax6, in accord with its lower affinity for this site (Epstein et al. 1994). Thus, this activation is independent of the presence of a homeo domain, as Pax2 lacks a homeo domain. In accord with DNA-binding data, Pax6-5a can activate transcription from separated 5aCON half-sites [5aCON(1/2)] in Fig. 3B] whereas Pax6 and Pax2 cannot.

Pd-5a binds to adjacent half-sites

The tendency of Pd-5a to form higher order complexes upon binding 5aCON suggests that individual protein molecules bind to each of the two predicted half-sites of the direct repeat composing 5aCON. Pd-5a binds a probe containing one copy of this 11-bp half-site [5aCON(1/2)], yielding a single band in EMSA experiments (Fig. 3A). This probe is not bound by the Pd protein, indicating that the two forms of the Pax6 paired domain bind 5aCON by different mechanisms.

Methylation interference assays were performed to identify specific nucleotides contacted by Pd and Pd-5a proteins in each complex. The residues contacted by Pd-5a in the faster migrating complex (C1) cluster predominantly around the right-hand half-site, whereas residues contacted in the more slowly migrating complex (C2) span the entire length of the direct repeat (Fig. 4E). The two half-sites are not identical; therefore, differences in sequence may account for the preference for the right-hand half-site. The Pd-5a protein, therefore, binds 5aCON by recognizing one of the two half-sites prefer-

entially and subsequently occupying the second half-site. In contrast, the interference pattern for Pd reveals contacts along the entire length of the direct repeat in 5aCON, although the migration of the Pd/5aCON complex in the EMSA suggests that only one protein molecule is bound.

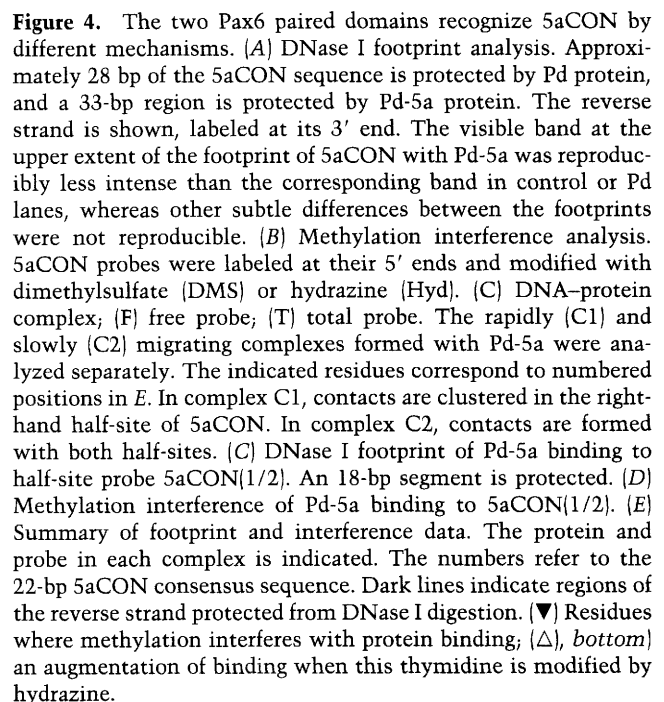
DNase I footprinting experiments are also consistent with Pd-5a binding to 5aCON by consecutively filling individual half-sites, whereas Pd recognizes 5aCON as one long binding site (Fig. 4A,E). The Pd-5a domain protects ~33 bp of 5aCON from DNase I digestion but only 18 bp of the 5aCON(1/2) half-site probe (Fig. 4C). When Pd binds 5aCON, ~28 bp is protected, as was shown previously for Pd binding to P6CON (Epstein et al. 1994). Pd-5a consistently protected a slightly longer sequence from DNase I digestion, as compared with Pd, when incubated with 5aCON (Fig. 4A).

We tested the importance of the relative position of the two half-sites with respect to each other by synthesizing DNA probes with two (5aCON+2) or four (5aCON+4) A-T base pairs inserted between the half-sites of 5aCON. Pd-5a binding and the formation of higher order complexes is reduced as the distance between half-sites is increased (Fig. 5). The effect of the separation on higher order complex formation can best be appreciated by comparing lanes 2, 7 and 12, in which Pd-5a protein forms approximately equal amounts of C2 complex with each of the three probes, respectively. As the half-sites are separated, the relative amount of probe in the C1 complex, and the ratio of C1 to C2, increases. The 2- and 4- nucleotide insertions not only physically separate the two half-sites from each other but change progressively their rotational relationship in the DNA helix. Pd also binds very poorly to the 5aCON+4 probe (Fig. 6C).

Different protein subdomains determine unique DNA-binding characteristics of Pd and Pd-5a

To define the parts of each paired domain that are required for binding DNA, we prepared a set of Pd and Pd-5a proteins with amino- or carboxy-terminal deletions and tested their ability to bind P6CON and 5aCON by EMSA (Fig. 6). Deletion of the amino-terminal 23 amino acids from Pd abolishes binding to P6CON, whereas deletion of the carboxy-terminal 33 amino acids decreases binding to P6CON only slightly and reduces the DNase I footprint from 28 to 16 bp (Fig. 6A). This 16-bp footprint of the amino-terminal subdomain contains the region of homology among several paired domain binding sites (Czerny et al. 1993; Epstein et al. 1994). We conclude that the amino subdomain determines the majority of Pd-DNA-binding affinity, whereas the carboxyl subdomain makes less important contacts with adjacent nucleotides.

Binding of Pd and Pd-5a to 5aCON requires the carboxy-terminal subdomain. Deletion of the carboxy-terminal 33 amino acids from Pd or Pd-5a abolishes 5aCON binding (Fig. 6B). In contrast, amino-terminal deletions of Pd confer the ability to bind 5aCON, form higher or-



In accord with these observations, we found that amino-terminal deletions of the Pax2 and Pax3 paired domains also abolish binding to P6CON and enhance

Our findings provide evidence for two autonomous but interacting subdomains that reside in the amino- and carboxy-terminal portions of the paired domain. This

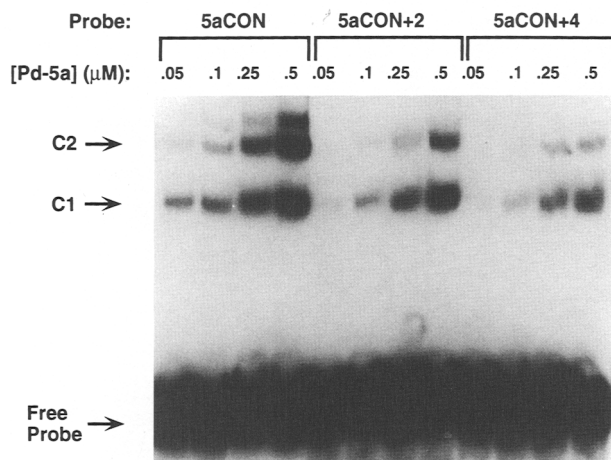


Figure 5. Relative spacing of adjacent half-sites in 5aCON affects Pd-5a binding. As the concentration of Pd-5a is increased, the formation of slowly migrating DNA–protein complexes (C2) is enhanced. If two (5aCON+2) or four (5aCON+4) A–T base pairs are inserted between the half-sites, formation of the more slowly migrating complex is diminished. The formation of monomeric complexes is also reduced, but to a lesser extent.

model extends previous DNA-binding studies that suggested a bipartite nature of Pax5 binding to some B-cell lineage-specific activator protein-binding sites (Czerny et al. 1993). It is also consistent with the evolutionary discontinuity among paired domains at residue 74 (Walther et al. 1991) and the presence of DNA-binding elements on both sides, including a potential helix–turn–helix (HTH) motif in the relatively divergent carboxyl terminus, at positions 80–106 (Bopp et al. 1989).

Dual mechanisms for DNA binding by alternative Pax6 paired domains

On the basis of our results, we propose a model in which the two subdomains of the Pax6 paired domain can bind independently to unique recognition sequences, or interact to bind a single extended site (Fig. 7). The amino subdomain of Pd is shown binding to the P6CON sequence, whereas the carboxyl subdomain contacts adjacent nucleotides and is unable to bind to P6CON independently. The amino subdomains of Pax2 and Pax3 recognize this site in a similar fashion. Insertion of the 5a peptide in Pax6 disrupts or refolds the amino subdomain, preventing Pd-5a from binding P6CON.

Pd does not recognize 5aCON(1/2) either because the DNA-binding surface of the carboxyl subdomain is shielded by the amino subdomain, as shown in the model, or because the carboxyl subdomain is folded into a different structure. Disruption of the amino subdomain in Pd-5a or deletion of the amino terminus from Pd allows the carboxyl subdomain to recognize an isolated half-site.

When the two half-sites are juxtaposed in 5aCON, Pd contacts DNA with both subdomains; the carboxyl subdomain is absolutely required for this binding and a sin-

gle Pd molecule can footprint both half-sites. There is no extended homology between P6CON and 5aCON. Although both P6CON and 5aCON probes used in our study contain a GTTC, a different subdomain recognizes this motif in each case. Pd-5a recognizes 5aCON with the carboxyl subdomain only and occupies the direct repeats as two protein molecules. Because each half-site is composed of 11 bp, or about one turn of the double helix, both molecules should dock on the same side of the double helix, perhaps contacting adjacent major grooves. This arrangement, with the potential for protein–protein interactions, is consistent with the observed cooperativity of binding, although it is also possible that an induced change in the DNA conformation favors binding of additional Pd-5a molecules.

Our model suggests an overall similarity between the paired domain and the POU bipartite DNA-binding domain, which consists of amino-terminal POU_s and carboxy-terminal POU_{HD} subdomains connected by a flexible tether of 15–27 amino acids (Klemm et al. 1994). Unlike the Pax6 paired domain, neither POU subdomain is capable of independent high affinity binding (Verrijzer et al. 1992).

Alternative splicing changes the population of sequences recognized by the paired domain, potentially allowing Pax6-5a to regulate an expanded or restricted set of genes in comparison to Pax6. Although we do not know whether half-sites or tandem repeat 5aCON-like sequences are used in nature, it is clear that some DNA sequences are recognized by both forms of Pax6, whereas others are recognized specifically by one form only. The use of variously oriented half and full DNA-binding sites has been demonstrated in nature among members of the nuclear hormone receptor superfamily, which can bind DNA as monomers or dimers (Näär et al. 1991; Umesono et al. 1991; Kato et al. 1992; Parker 1993).

An alternative splicing mutation in PAX6 causes a distinct ocular phenotype

The importance of the disparate DNA-binding properties of the alternative paired domains is underscored by our finding of a PAX6 mutation that disturbs the ratio of the two splice products and causes an ocular phenotype. The T → C mutation at position –3 of the exon 5a splice acceptor is unusual in that it increases splicing efficiency significantly. Although not absolutely essential (Fu et al. 1988), the presence of a pyrimidine at position –3 is thought to be important for effective interaction between the splice acceptor site of the pre-mRNA and small nuclear ribonucleoproteins (snRNPs) U1 and U5 of the spliceosome (Green 1991; Guthrie 1991). In a survey of 3724 splice acceptor sites, cytosine was found 3.5 times more frequently than thymidine at this position (75% vs. 21%; Senapathy et al. 1990), suggesting that cytosine may be enzymatically preferred. In the model proposed by Steitz (1992), the conserved AG dinucleotide of the splice acceptor site pairs with the C₉U₁₀ dinucleotide in the 5' tail of U1 snRNA during assembly of the mammalian spliceosome. In this configuration, the cy-

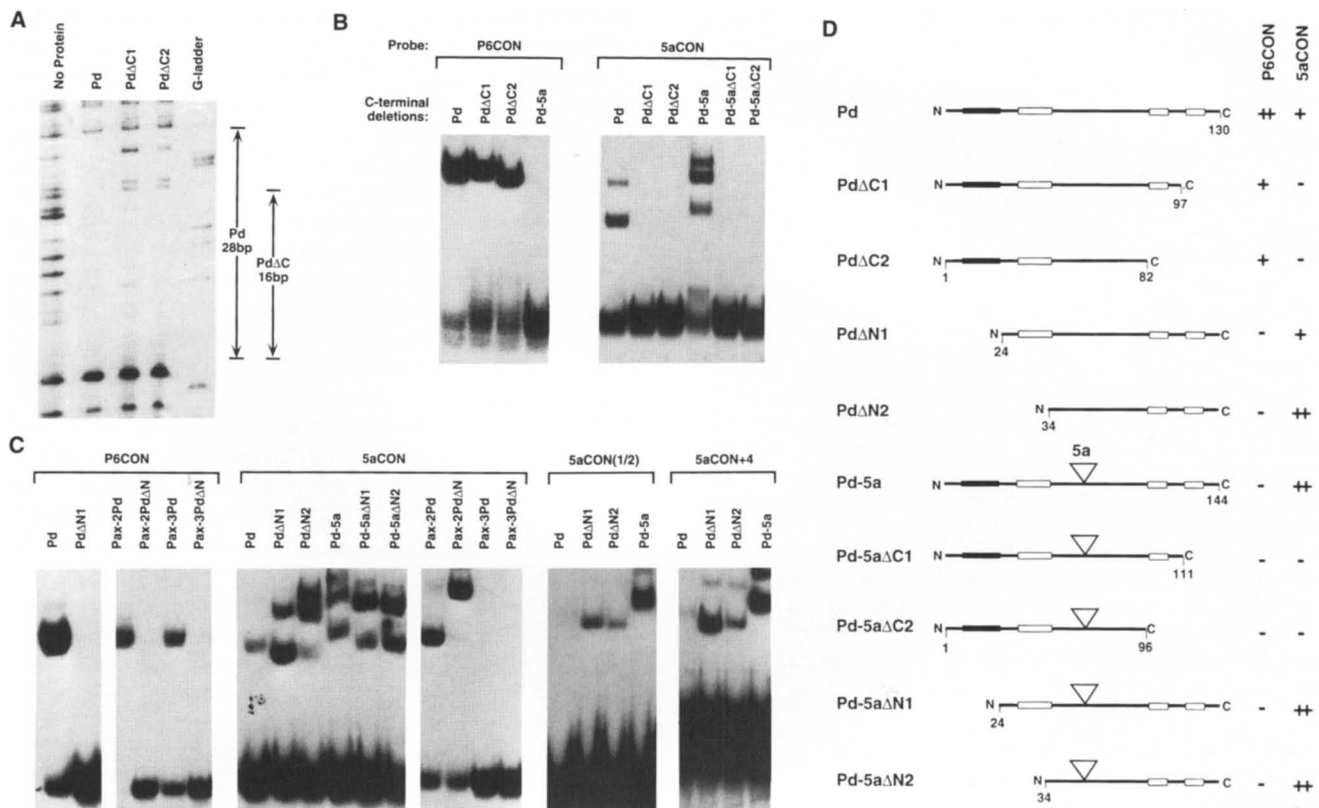


Figure 6. Paired domain DNA binding is mediated by two distinct subdomains. (A) DNase I footprint of a Pd-selected sequence (related to P6CON; see Epstein et al. 1994). The 5' end of the forward strand was radiolabeled. The Pd protein protects a 28-bp segment. Deletion of 33 (PdΔC1) or 48 (PdΔC2) amino acids from the carboxyl terminus of Pd reduces the footprint to 16 bp. The 5' boundary of the footprint is unchanged. (B) EMSA analysis of paired domains with carboxy-terminal deletions. Both carboxy-terminal deletions prevent binding of Pd and Pd-5a to 5aCON but do not prevent binding of Pd to P6CON. Pd-5a does not bind P6CON. (C) EMSA analysis of paired domains with amino-terminal deletions. Deletion of 24 (PdΔN1) or 34 (PdΔN2) amino acids from the amino terminus of Pd prevents binding to P6CON. The Pax2 and Pax3 paired domains also bind P6CON and deletion of 30 amino acids from the amino terminus of each prevents binding. Amino-terminal deletions of Pd promote the formation of higher order complexes with 5aCON, similar to the pattern observed for Pd-5a binding to 5aCON. Likewise, Pax-2Pd binds 5aCON and Pax-2PdΔN binds 5aCON forming high order complexes, whereas Pax-3Pd and Pax-3PdΔN do not bind this sequence. Amino-terminal deletions enable Pd to bind a half-site probe 5aCON(1/2) and probe 5aCON + 4, in which the two half-sites are separated by 4 bp. (D) Summary showing Pax6 paired domain deletion constructs and DNA-binding data. Relative binding to P6CON and 5aCON is indicated at right. Amino acids are numbered starting from the initiation methionine. The paired domain extends from amino acid 3 to 130. The solid bar indicates a 16-amino-acid segment that is highly conserved among paired domains. Open bars indicate α -helical regions predicted from primary sequence data (Bopp et al. 1989). The 5a peptide is inserted after amino acid 47.

tosine at position -3 is aligned with G₁₁ of the U1 snRNA. If pairing of these complementary bases also occurs, it could increase splicing efficiency and explain the effect of the Bh mutation. The few human mutations reported in splice acceptors at position -3 introduce purines at this position, and all appear to decrease splicing efficiency (Krawczak et al. 1992). A 195-bp fragment encompassing exon 5a is 97.0% identical in human and mouse genomic DNA (data not shown), and the T at position -3 is invariant among quail (Dozier et al. 1993), human, and mouse *PAX6* genes. The exact sequence of the acceptor site, in particular the pyrimidine at position -3, therefore plays an important role in modulating the relative expression of the two paired domain forms.

Mammalian ocular and central nervous system devel-

opment is exquisitely sensitive to *PAX6* gene dosage. Mutant phenotypes range from anophthalmia with profound olfactory and brain malformations in *PAX6* homozygotes (Schmahl et al. 1993; Glaser et al. 1994a) to relatively mild findings in the present case. Aniridia is caused by heterozygous null mutations, either cytogenetic deletions of 11p13 in individuals with the WAGR syndrome (Wilms tumor, aniridia, genitourinary anomalies, mental retardation) (Ton et al. 1991) or intragenic nonsense or frameshift mutations (Glaser et al. 1994b). In both groups of patients, the phenotype is characterized by severe hypoplasia of the iris, with anterior segment and retinal defects (Nelson et al. 1984). All mutations described thus far are predicted to affect both Pax6 and Pax6-5a transcripts (Glaser et al. 1994b). The Bh mu-

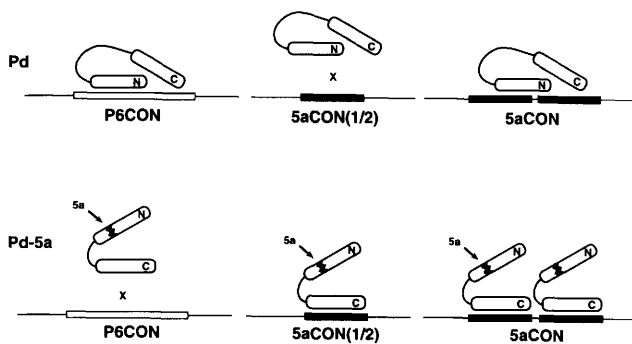


Figure 7. A model for DNA binding by the two alternative Pax6 paired domains. The paired domain is shown consisting of amino and carboxyl subdomains. The Pd protein is folded so that the amino subdomain binds its preferred sequence, P6CON. The carboxyl subdomain makes limited contacts with P6CON but is shielded partially from or conformationally incapable of binding its preferred sequence, 5aCON(1/2). Insertion of the 5a peptide deforms the amino subdomain, preventing it from binding DNA and exposing the independent binding capability of the carboxyl subdomain. This conformational change allows the carboxyl subdomain to bind 5aCON(1/2), and promotes binding of Pd-5a molecules to adjacent half-sites in 5aCON.

tation causes a distinct ocular phenotype, with relative sparing of the fovea and iris tissue. This could arise from a deficiency of the shorter Pax6 protein, an excess of the Pax6-5a protein, or an altered ratio of the two isoforms. The latter mechanism has been implicated in the etiology of Denys-Drash syndrome, which is caused by missense or alternative splicing mutations in the *WT1* locus (Bruening et al. 1992). The Bh phenotype suggests that the two isoforms are not functionally interchangeable. In view of their radically different binding specificities, Pax6-5a and Pax6 are likely to regulate different target genes and play distinct functional roles in eye development.

The 5a insertion functions as a molecular toggle

The exon 5a insertion has been conserved during vertebrate evolution, with only a single amino acid substitution distinguishing human, mouse, quail, axolotl, and zebrafish 5a peptides (Glaser et al. 1994b). This implies that the 5a insertion itself has an important functional role. As indicated by amino-terminal deletion analysis, the 5a peptide is not required for dimerization upon DNA binding or recognition of the 5aCON site. Instead, we suggest it functions as a molecular toggle, preventing the amino subdomain from binding DNA and allowing the carboxyl terminus to bind.

Structural studies of the *Drosophila* Paired protein reveal the existence of a HTH motif in the amino subdomain that is not predicted from the amino acid sequence (W. Xu, M. Rould, S. Jun, C. Desplan, and C. Pabo, unpubl.). The 5a insertion site is located at the end of the first α -helix of the HTH motif. There are several prece-

dents for variable length insertions into the turn region of the HTH motif (for review, see Brennan 1993; Klemm et al. 1994). Among these, the LFB1 homeo domain contains a 21-amino-acid insertion in the HTH turn region (Finney 1990), which extends the first helix by 8 residues and converts the turn into a flexible disordered loop (Ceska et al. 1993; Leiting et al. 1993). Because insertions in this region do not generally disturb DNA binding or the relative configuration of the α -helices, the turn region has been viewed as relatively tolerant to the insertion of additional residues (Brennen 1993; Klemm et al. 1994). The 5a insertion into the Pax6 paired domain is the first example of an insertion into the turn region of a HTH motif that destabilizes its ability to interact with DNA. Like the 21-residue insertion in LFB1 (Bach et al. 1992; Zapp et al. 1993), the 5a insertion occurs at an intron-exon boundary in genomic DNA, indicating that modification of the turn region may be an evolutionary strategy for modulating the DNA-binding properties of the HTH motif.

A distinct evolutionary origin for the two paired subdomains is supported by the structure of the *Drosophila* gene *Lune*, which encodes a partial paired domain (Wilson et al. 1993). *Lune* is homologous to other paired domains in the carboxy-terminal region but diverges within 5 amino acids of the 5a insertion site (S. Jun, B. Kalionis, and C. Desplan, unpubl.). *Lune* may represent the prototype of a class of DNA-binding proteins with homology to the carboxyl subdomain.

On the basis of sequence comparisons, the mammalian Pax family is divided into four groups, consisting of Pax1/9, Pax4/6, Pax3/7, and Pax2/5/8 (Walther et al. 1991; Noll 1993). Although the intron-exon boundary at which the 5a insertion occurs is restricted to the Pax4/6 group, the interactive nature of the amino and carboxyl subdomains of the paired domain appears to be independent of the presence of the homeo domain as Pax2, which lacks a homeo domain, displays similar DNA-binding and functional characteristics as Pax6 in our experiments. Most Pax proteins are able to bind sequences that are recognized by the highly conserved amino subdomain (Czerny et al. 1993; Epstein et al. 1994). Carboxyl subdomain-binding specificity is less homogeneous. Pax6 and Pax2, which share carboxy-terminal homology (64.8% identity), bind 5aCON, whereas the more divergent Pax3 (55.5% identity with Pax6) does not bind 5aCON. The carboxyl subdomain modifies the binding specificity of the paired domain and significantly affects functional activity. This may account for differences in target gene regulation among Pax gene products.

Materials and methods

Patients

Clinical records of both affected Bh family members were reviewed and their eyes were examined carefully. The mother was first noted to have poor vision, pendular nystagmus, and cortical cataracts at age 2, at which time bilateral iridectomies and cataract surgery were performed. On examination at age 33, visual acuity was 20/200 in each eye with aphakic correction. The

gaze was unsteady and neither eye could fixate on objects for an extended period. The corneas were small, measuring 10 mm in horizontal diameter and were opacified at the superior and inferior margins by corneal pannus. The intraocular pressures were slightly elevated while receiving topical glaucoma medication. A moderately broad iris leaf was present bilaterally with a surgical coloboma above in each eye. The optic discs showed evidence of glaucomatous cupping bilaterally. There was no macular abnormality.

The daughter had poor vision through early childhood. Cataracts were first reported at age 12, consisting of cortical equatorial opacities and diffuse posterior pole opacities that progressed during 3 years of observation. At age 13, the visual acuity was 20/80 (right) and 20/400 (left). After bilateral cataract extractions at age 16, unsteady pendular fixation persisted, suggesting a congenital sensory nystagmus. The corneas measured 9.5 mm in horizontal diameter. Pannus defects extended 2 mm inward from the superior and inferior corneal margins and stained well with fluorescein. The pupils were round and dilated normally, but were displaced nasally. The irides were blue and smooth with minimal circumferential folds and no collarette. Radial stromal defects were present inferiorly in the right eye, allowing visualization of the anterior surface of the iris pigment epithelium. Intraocular pressures were normal. On gonioscopy, the anterior chamber angles were found open, with unobstructed trabecular meshwork and normal ciliary body bands connected by frequent uveal processes. The optic discs were small but well vascularized, and the macular regions appeared normal.

PAX6 mutational analysis

Genomic DNA was extracted from venous blood samples obtained with informed consent. Fragments spanning PAX6 exons were amplified by the polymerase chain reaction (PCR) using previously described primers and conditions (Glaser et al. 1992) and screened for mutations by heteroduplex (White et al. 1992) and single-strand conformational (Orita et al. 1989) analysis. The Bh mutation was detected by altered migration of exon 5a heteroduplex molecules. Mutant and wild-type PCR products were cloned into pGEM3Z and sequenced. To confirm the exon 5a mutation, radiolabeled PCR products were digested with *Mae*III. The presumed genetic relationships within the family were confirmed using seven highly polymorphic DNA markers on chromosome 11.

Reverse transcriptase PCR analysis

RNA was extracted from Epstein-Barr virus-transformed lymphoblast cell lines by the phenol-guanidinium thiocyanate method (Chomczynski and Sacchi 1987). Twenty microliter reverse transcriptase (RT) reactions were performed at 37°C as described (Kawasaki 1990) using 2 µg RNA as template and 0.2 µg pd(N)₆ as primer. Fragments spanning the exon 5a insertion site were amplified from the resulting cDNA by PCR, using upstream primer 5'-GGAAGACTTTAACTAGGGGC-3' (exon 3) and downstream primer 5'-ATGGACGGGCACTCCCGC-TT-3' (exon 6) as follows: 1 min at 95°C, 2 min at 60°C, and 2 min at 72°C, for 45 cycles. The expected products are 416 and 458 bp long and extend between nucleotides 292 and 707 of PAX6 clone λHPX-2 (Glaser et al. 1992).

In vivo splicing assay

A 1.3-kb genomic fragment encompassing PAX6 exons 5, 5a, and 6 was amplified by PCR from λFix-3 (Glaser et al. 1992) and

cloned in the sense orientation into the *Bam*HI site of exon amplification vector pSPL1 (Buckler et al. 1991). This fragment begins 64 bp upstream of exon 5 and ends 54 bp downstream of exon 6. The cloning site is located within a 2.3-kb intron derived from the human immunodeficiency virus (HIV-1) *tat* gene. The intron forms part of a transcription unit that begins with the SV40 *ori* promoter and is flanked by *tat* splice donor and acceptor sites. An isogenic derivative containing the Bh point mutation was created by site-directed mutagenesis (Deng and Nickoloff 1992). The DNA sequence surrounding exon 5a was confirmed in both plasmid clones. COS7 cells (10⁶) were plated in 60-mm Petri dishes and transfected with 1.5 µg of plasmid DNA coated with a cationic lipopolyamine (transfectam, Promega, Madison, WI). Total cellular RNA was extracted after 72 hr and PAX6-*tat* chimeric transcripts were amplified by RT-PCR using an upstream primer in exon 5 (5'-GTACACGCGAGTGAATCAGCTCGG-3') and a downstream primer that crosses the *tat*-exon 6 splice junction (5'-TTGGGAGGTGGGTGCTTGGTATGTTATCG-3') (Fig. 2B). The PCR conditions were 1 min at 95°C, 2 min at 63°C, and 2 min at 72°C, for 45 cycles. The expected 359- and 401-bp products were verified by Southern analysis with internal oligonucleotide probes.

Fusion protein preparations

The construction of pGEX-Pax6 and pGEX-Pax2 GST-paired domain expression plasmids has been described (Epstein et al. 1994). Pax6-5a (codons 1–144) and Pax3 (codons 34–162) sequences were amplified similarly by PCR from human fetal eye and mouse embryonic brain RNA, respectively, and cloned into pGEX2T (Smith and Johnson 1988). Paired domain deletion constructs were made by amplifying partial coding sequences from Pax6, Pax6-5a, Pax2, and Pax3 plasmid templates and cloning the resulting PCR products into pGEX2T. GST fusion proteins were expressed in *E. coli* strain BL21 (DE3), purified from bacterial extracts using a glutathione-agarose resin (Epstein et al. 1994), and assessed by polyacrylamide gel electrophoresis. Aliquots were frozen immediately in liquid nitrogen and stored at –80°C. Comparable plasmids expressing Pax6 Pd and Pd-5a domains with an amino-terminal polyhistidine tail (Studier et al. 1990) were constructed in pET14b (Novagen, Madison, WI) and the resulting proteins were purified from *E. coli* strain BL21 (DE3) by Ni affinity chromatography as described (Epstein et al. 1994). The coding region of each construct was verified by DNA sequencing.

DNA-binding site selection

A modified binding-site selection (Wilson et al. 1993) was performed using a degenerate oligonucleotide with a core of 25 random positions (see Epstein et al. 1994). However, because the region of Pd-5a binding defined by DNase I protection analysis extended into the flanking sequences, a longer oligonucleotide (93 bp) with a core of 35 random positions was used for all further experiments. Approximately 100 ng of purified GST-Pd-5a protein was attached to glutathione-agarose beads and incubated with 5 µg of oligonucleotide in a buffer containing 20 mM HEPES (pH 7.4), 150 mM KCl, 250 µg/ml of poly[d(I-C)], 20% glycerol, 0.5 mg/ml of bovine serum albumin (BSA), and 1 µM DTT. Samples were incubated at room temperature for 45 min with gentle agitation and washed extensively with cold binding buffer. Pd-5a/DNA complexes were released from the GST moiety by thrombin cleavage and used for subsequent PCR amplification. After nine rounds of selection, bound oligonucleotides were cloned into pGEM3Z and sequenced. The results of the selection performed with 35 random positions are presented.

The consensus obtained from the experiment using 25 random positions (24 selected sequences analyzed) contained the same 22-bp imperfect direct repeat except that position 4 (as numbered in Fig. 4E) was T instead of C, and position 14 was A instead of G.

EMSA

EMSA reactions were performed essentially as described (Epstein et al. 1994), in a buffer consisting of 10 mM Tris (pH 7.4), 50 mM KCl, 100 μ g/ml of poly[d(I-C)], 20% glycerol, 2.5 mg/ml of BSA, 1 mM β -mercaptoethanol, and 1 μ M protein, unless otherwise noted. For measurement of K_d , the poly[d(I-C)] competitor was omitted. Double-stranded probes were made by annealing an oligonucleotide template with a short complementary primer and filling in the remaining sequence using Klenow DNA polymerase. Probes were radiolabeled by kinasing the oligonucleotide with [γ - 32 P]ATP or including [α - 32 P]dCTP in the polymerase reaction, and were purified on polyacrylamide gels. Binding reactions were performed at 4°C for 30 min followed by electrophoresis through 6% polyacrylamide gels in 0.5 \times TBE [1 \times TBE = 45 mM Tris-borate, 1 mM EDTA, (pH 8.0)] for 90 min at 17 v/cm. Gels were dried and exposed to X-ray film, or quantitated using a PhosphorImager (Molecular Dynamics). The probes are P6CON, 5'-TGGAATTCAGGAAAAATTTTCACGCTTGAGTTCACAGCTCGAGTA-3'; 5aCON, 5'-AATAAATCTGAACATGCTCAGTGAATGTTTCATTGACTCTCGAGTCA-3'; 5aCON(1/2), 5'-GATCCAATGTTTCATTGACTCTCGAG-3'.

Double-label binding experiments

To calculate the relative ratio of protein to DNA in the complexes formed upon Pd-5a binding to 5aCON, we cloned the coding region of Pd-5a into pGEX2TK (Pharmacia) that encodes a phosphorylation site for cAMP-dependent protein kinase between the GST moiety and the polylinker. The resulting fusion protein was expressed as described above and labeled with [γ - 32 P]ATP according to the manufacturer's instructions to a specific activity of $\sim 2 \times 10^6$ cpm/ μ g before elution from the agarose beads. 5aCON was labeled with 33 P to a specific activity of $\sim 4 \times 10^6$ cpm/ μ g using polynucleotide kinase and [γ - 33 P]ATP (Dupont NEN) before gel purification. EMSA was performed in the absence of competitor DNA at various molar ratios of total protein to DNA ranging from 0.5 to 2. Four percent polyacrylamide gels were used to minimize comigration of free protein with specific complexes. Complexes were excised from dried gels and rehydrated in 0.5 ml of water. Comparable slices were excised from lanes in which no DNA was added to the binding reaction to calculate the amount of comigrating free protein. Complexes were eluted from the gel slices by adding 0.5 ml of Solvable (Dupont NEN) and incubating for 3 hr at 50°C before neutralization with 10 ml of fluor (Formula 989, Dupont NEN). Dual-channel scintillation counting was used to calculate the relative ratios of 32 P to 33 P in complexes within a given binding reaction, that is, from the same lane of the gel. We were unable to determine the ratio of protein to DNA in complexes migrating more slowly than C2 because of their low abundance and comigrating free protein.

CAT assays

P19 murine embryo carcinoma cells (5×10^4) were plated in 60-mm petri dishes and transfected with plasmid DNA coated with cationic lipopolyamine (transfectam). The pG₅ECAT reporter plasmid contains the adenovirus E1bTATA promoter upstream

of the bacterial CAT gene (Lillie and Green 1989). Derivative plasmids with six copies of the P6CON sequence, two copies of a 5aCON sequence, or six copies of the 5aCON(1/2) sequence upstream of the promoter were prepared by inserting two copies of PRDCONx3 (5'-TCGACAAAATTTTCACGCTTGAGTTCACAGCTCGAAAAATTTTCACGCTTGAGTTCACAGCTCGAAAAATTTTCACGCTTGAGTTCACAGCTC-3'), two copies of 5aCON-ava (5'-CTAGTGAACATGCTCAGTGAATGTTTCATTGACTCTC-3'), or three copies of 5aCON(1/2)-6A (5'-CTAGTGAACATGTTTCATTGAAAAAATGTTTCATTGACTCTC-3') into the *SalI* site of pG₅ECAT after the addition of *SalI* linkers. Expression plasmids were derived by replacing the *NotI* β -gal cassette of pCMV β (MacGregor et al. 1987) with full-length coding sequences for Pax6, Pax6-5a, or Pax2, and expression of Pax proteins was confirmed by Western analysis. The 1.6-kb human Pax6 DNA fragment was amplified by PCR from pHPX-2 (Glaser et al. 1992) using an upstream primer in the 5'-untranslated region (5'-AAACCATTGCGGCCGCAAGCAACAACAGCAGCAGCACA-3') and a downstream primer in flanking vector DNA (5'-AATACAAGCGGCCGCATACGACTACTATAGGCG-3'). Mouse Pax2 sequences were recombined as a 2.6-kb *NotI* fragment from two partial cDNA clones (Dressler et al. 1990). The Pax6-5a plasmid was derived by replacing the 371-bp *NheI* fragment of pHPX-2 with the corresponding 359-bp *NheI* fragment containing exon 5a. Each dish was transfected with 2 μ g of reporter and 1 μ g of expression plasmid DNA. Cell extracts were prepared after 48 hr and assayed for CAT activity as described by Gorman et al. (1982). Samples were diluted as necessary to measure activity in the linear range and equalize the amount of protein added. The percent acetylation was quantitated from thin-layer chromatography plates using a PhosphorImager.

DNase I footprinting and methylation interference

Footprinting and interference experiments were performed as described (Epstein et al. 1994). For DNase I footprint analysis, double-stranded oligonucleotides P6CON, 5aCON, and 5aCON(1/2) were cloned in the polylinker of plasmid pGEM3Z or pGEM4Z. Restriction fragments containing these sequences were end-labeled by kinasing the oligonucleotide with [γ - 32 P]ATP or including [α - 32 P]dCTP in the polymerase reaction, gel purified, and mixed with 1 μ M GST-Pd or GST-Pd-5a protein in 20 μ l of EMSA buffer including 5 μ g/ml of poly[d(I-C)], 10 mM CaCl₂, and 2 mM MgCl₂. After DNase I digestion, samples were precipitated with ethanol and run on denaturing 6% or 10% polyacrylamide gels. For interference assays, complementary oligonucleotides were end-labeled by kinasing with [γ - 32 P]ATP, annealed, and extended with Klenow DNA polymerase. These probes were gel purified, modified with dimethylsulfate or hydrazine, precipitated with ethanol, and resuspended in 20 μ l of EMSA buffer with 0.1 μ M GST-Pd or GST-Pd-5a. After 30 min at 4°C, samples were electrophoresed through a denaturing 6% polyacrylamide gel in 0.5 \times TBE and exposed to X-ray film. DNA-protein complexes and free probes were excised separately, electroeluted in 0.2 \times TBE, and precipitated with ethanol. Samples were then resuspended in 1 M piperidine, cleaved for 30 min at 95°C, lyophilized three times, and run on a denaturing 10% polyacrylamide gel.

Acknowledgments

J.A.E. and T.G. both contributed substantially to this work. We are grateful to the aniridia patients for participating in this study, and to Drs. Carl Pabo and Claude Desplan and members

of their laboratories for critical reading of the manuscript and helpful discussions. R.L.M. is an Assistant Investigator of the Howard Hughes Medical Institute. This work was supported in part by a grant from the National Institutes of Health (EY10123).

The publication costs of this article were defrayed in part by payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 USC section 1734 solely to indicate this fact.

References

- Bach, I., M. Pontoglio, and M. Yaniv. 1992. Structure of the gene encoding hepatocyte nuclear factor 1 (HNF1). *Nucleic Acid Res.* **20**: 4199–4204.
- Barberis, A., K. Widenhorn, L. Vitelli, and M. Busslinger. 1990. A novel B-cell lineage specific transcription factor present at early but not late stages of differentiation. *Genes & Dev.* **4**: 849–859.
- Bopp, D., E. Jamet, S. Baumgartner, M. Burri, and M. Noll. 1989. Isolation of two tissue-specific Drosophila paired box genes, Pox meso and Pox neuro. *EMBO J.* **8**: 3447–3457.
- Brennan, R.G. 1993. The winged-helix DNA-binding motif: Another helix-turn-helix takeoff. *Cell* **74**: 773–776.
- Bruening, W., N. Bardeesy, B.L. Silverman, R.A. Cohn, G.A. Machin, A.J. Aronson, D. Housman, and J. Pelletier. 1992. Germline intronic and exonic mutations in the Wilms' tumour gene (*WT1*) affecting urogenital development. *Nature Genet.* **1**: 144–148.
- Buckler, A.J., D.D. Chang, S.L. Graw, D. Brook, D.A. Haber, P.A. Sharp, and D.E. Housman. 1991. Exon amplification: A strategy to isolate mammalian genes based on RNA splicing. *Proc. Natl. Acad. Sci.* **88**: 4005–5009.
- Ceska, T.A., M. Lamers, P. Monaci, A. Nicosia, R. Cortese, and D. Suck. 1993. The X-ray structure of an atypical homeodomain present in the rat liver transcription factor LFB1/HNF1 and implications for DNA binding. *EMBO J.* **12**: 1805–1810.
- Chalepakakis, 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.
- Chalepakakis, G., P. Tremblay, and P. Gruss. 1992. Pax genes, mutants and molecular function. *J. Cell Sci. (Suppl.)* **16**: 61–67.
- Chomczynski, P. and N. Sacchi. 1987. Single-step method of RNA isolation by guanidinium thiocyanate-phenol-chloroform extraction. *Anal. Biochem.* **162**: 156–159.
- 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.
- Deng, W.P. and J.A. Nickoloff. 1992. Site-directed mutagenesis of virtually any plasmid by eliminating a unique site. *Anal. Biochem.* **200**: 81–88.
- Dozier, C., C. Carriere, D. Grevin, P. Martin, B. Quatannens, D. Stehelin, and S. Saule. 1993. Structure and DNA-binding properties of Pax-QNR, a paired box- and homeobox-containing gene. *Cell Growth Differ.* **4**: 281–289.
- Dressler, G.R., U. Deutsch, K. Chowdhury, H.O. Nornes, and P. Gruss. 1990. Pax2, a new murine paired-box-containing gene and its expression in the developing excretory system. *Development* **109**: 787–795.
- 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.
- Fickenscher, H.R., G. Chalepakakis, and P. Gruss. 1993. Murine Pax-2 protein is a sequence-specific trans-activator with expression in the genital system. *DNA Cell Biol.* **12**: 381–391.
- Finney, M. 1990. The homeodomain of the transcription factor LF-B1 has a 21 amino acid loop between helix 2 and helix 3. *Cell* **60**: 5–6.
- Fried, M.G. 1989. Measurement of protein–DNA interaction parameters by electrophoretic mobility shift assay. *Electrophoresis* **10**: 366–376.
- Fu, X.-Y., H. Ge, and J.L. Manley. 1988. The role of the polypyrimidine stretch at the SV40 early pre-mRNA 3' splice site in alternative splicing. *EMBO J.* **7**: 809–817.
- 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.
- Glaser, T., L. Jepeal, J.G. Edwards, S.R. Young, J. Favor, and R.L. Maas. 1994a. PAX6 gene dosage effect in a family with congenital cataracts, aniridia, anophthalmia and central nervous system defects. *Nature Genet.* **7**: 463–471.
- Glaser, T., D.S. Walton, J. Cai, J.A. Epstein, L. Jepeal, and R.L. Maas. 1994b. PAX6 gene mutations in aniridia. In *Molecular genetics of human ocular disorders* (ed. J. Wiggs), pp. 51–80. Wiley, New York.
- Gorman, C.M., L.F. Moffat, and B.H. Howard. 1982. Recombinant genomes which express chloramphenicol acetyltransferase in mammalian cells. *Mol. Cell Biol.* **2**: 1044–1051.
- Green, M.R. 1991. Biochemical mechanisms of constitutive and regulated pre-mRNA splicing. *Annu. Rev. Cell Biol.* **7**: 559–599.
- Guthrie, C. 1991. Messenger RNA splicing in yeast: Clues to why the spliceosome is a ribonucleoprotein. *Science* **253**: 157–163.
- 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* **354**: 522–525 (erratum. 1992. *Nature* **355**: 750).
- Kato, S., L. Tora, J. Yamauchi, S. Masushige, M. Bellard, and P. Chambon. 1992. A far upstream estrogen response element of the ovalbumin gene contains several half-palindromic 5'-TGACC-3' motifs acting synergistically. *Cell* **68**: 731–742.
- Kawasaki, E.S. 1990. Amplification of RNA. In *PCR protocols: A guide to methods and applications* (ed. M.A. Innis, D.H. Gelfand, J.J. Sninsky, and T.J. White), pp. 21–27. Academic Press, San Diego, CA.
- Kessel, M. and P. Gruss. 1990. Murine developmental control genes. *Science* **249**: 374–379.
- Klemm, J.D., M.A. Rould, R. Aurora, W. Herr, and C. Pabo. 1994. Crystal structure of the Oct-1 POU domain bound to an octamer site: DNA recognition with tethered DNA-binding modules. *Cell* **77**: 1–20.
- Kozmik, Z., R. Kurzbauer, P. Dorfler, and M. Busslinger. 1993. Alternative splicing of Pax-8 gene transcripts is developmentally regulated and generates isoforms with different trans-activation properties. *Mol. Cell Biol.* **13**: 6024–6035.
- Krawczak, M., J. Reiss, and D.N. Cooper. 1992. The mutational spectrum of single base-pair substitutions in mRNA splice junctions of human genes: Causes and consequences. *Hum. Genet.* **90**: 41–54.
- Leiting, B., R. De Francesco, L. Tomei, R. Cortese, G. Otting, and K. Wüthrich. 1993. The three-dimensional NMR-solution structure of the polypeptide fragment 195–286 of the LFB1/HNF1 transcription factor from rat liver comprises a non-classical homeodomain. *EMBO J.* **12**: 1797–1803.
- Lillie, J.W. and M.R. Green. 1989. Transcription activation by the adenovirus E1a protein. *Nature* **33**: 39–44.

- MacGregor, G.R., A.E. Mogg, J.F. Burke, and C.T. Caskey. 1987. Histochemical staining of clonal mammalian cell lines expressing *E. coli* β -galactosidase indicates heterogeneous expression of the bacterial gene. *Somat. Cell. Mol. Genet.* **13**: 253–265.
- Näär, A.M., J.-M. Boutin, S.M. Lipkin, V.C. Yu, J.M. Holloway, C.K. Glass, and M.G. Rosenfeld. 1991. The orientation and spacing of core DNA-binding motifs dictate selective transcriptional responses to three nuclear receptors. *Cell* **65**: 1267–1279.
- Nelson, L.B., G.L. Spaeth, T.S. Nowinski, C.E. Margo, and L. Jackson. 1984. Aniridia. A review. *Surv. Ophthalmol.* **28**: 621–642.
- Noll, M. 1993. Evolution and role of Pax genes. *Curr. Opin. Genet. Dev.* **3**: 595–605.
- Orita, M., Y. Suzuki, T. Sekiya, and K. Hayashi. 1989. Rapid and sensitive detection of point mutations and DNA polymorphisms using the polymerase chain reaction. *Genomics* **5**: 874–879.
- Parker, M.G. 1993. Steroid and related receptors. *Curr. Opin. Cell Biol.* **5**: 499–504.
- Plaza, S., C. Dozier, and S. Saule. 1993. Quail PAX-6 (PAX-QNR) encodes a transcription factor able to bind and transactivate its own promoter. *Cell Growth Differ.* **4**: 1041–1050.
- Puschel, 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.
- Schmahl, W., M. Knoedlseder, J. Favor, and D. Davidson. 1993. Defects of neuronal migration and the pathogenesis of cortical malformations are associated with small eye (Sey) in the mouse, a point mutation at the Pax-6 locus. *Acta Neuropathol.* **86**: 126–135.
- Senapathy, P., M.B. Shapiro, and N.L. Harris. 1990. Splice junctions, branch point sites, and exons: Sequence statistics, identification and application to the genome project. *Methods. Enzymol.* **183**: 252–278.
- Smith, D.B. and K.S. Johnson. 1988. Single-step purification of polypeptides expressed in *Escherichia coli* as fusions with glutathione S-transferase. *Gene* **67**: 31–40.
- Steitz, J.A. 1992. Splicing takes a holiday. *Science* **257**: 888–889.
- Studier, F.W., A.H. Rosenberg, I.J. Dunn, and J.W. Dubendorf. 1990. Use of T7 RNA polymerase to direct expression of cloned genes. *Methods Enzymol.* **185**: 60–89.
- 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.
- 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.
- Umesono, K., K.K. Murakami, C.C. Thompson, and R.M. Evans. 1991. Direct repeats as selective response elements for the thyroid hormone, retinoic acid, and vitamin D₃ receptors. *Cell* **65**: 1255–1266.
- Verrijzer, C.P., M.J. Alkema, W. van Weperen, H.C. Van Leeuwen, M.J.J. Strating, and P.C. van der Vliet. 1992. The DNA binding specificity of the bipartite POU domain and its subdomains. *EMBO J.* **11**: 4993–5003.
- Wallin, J., Y. Mizutani, K. Imai, N. Miyashita, K. Moriwaki, M. Taniguchi, H. Koseki, and R. Balling. 1993. A new Pax gene, Pax-9, maps to mouse chromosome 12. *Mamm. Genome* **4**: 354–358.
- Walther, C. and P. Gruss. 1991. Pax-6, a murine paired box gene, is expressed in the developing CNS. *Development* **113**: 1435–1449.
- 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.
- White, M.B., M. Carvalho, D. Derse, S.J. O'Brien, and M. Dean. 1992. Detecting single base substitutions as heteroduplex polymorphisms. *Genomics* **12**: 301–306.
- 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.
- 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.
- Zapp, D., S. Bartkowski, C. Zoidl, L. Klein-Hitpass, and G.U. Ryffel. 1993. Genomic structure of the *Xenopus laevis* liver transcription factor LFB1. *Gene* **134**: 251–256.



Two independent and interactive DNA-binding subdomains of the Pax6 paired domain are regulated by alternative splicing.

J A Epstein, T Glaser, J Cai, et al.

Genes Dev. 1994, **8**:

Access the most recent version at doi:[10.1101/gad.8.17.2022](https://doi.org/10.1101/gad.8.17.2022)

References

This article cites 56 articles, 17 of which can be accessed free at:
<http://genesdev.cshlp.org/content/8/17/2022.full.html#ref-list-1>

License

Email Alerting Service

Receive free email alerts when new articles cite this article - sign up in the box at the top right corner of the article or [click here](#).

