MOLECULAR AND CELLULAR BIOLOGY, Dec. 1996, p. 6677–6686 Vol. 16, No. 12
0270-7306/96/$04.00+0

An Alternative Splicing Event in the Pax-3 Paired Domain Identifies the Linker Region as a Key Determinant of Paired Domain DNA-Binding Activity

KYLE J. VOGAN,[1] D. ALAN UNDERHILL,[1] AND PHILIPPE GROS[1,2]*

Department of Biochemistry[1] and McGill Cancer Center,[2] McGill University, Montreal, Quebec, Canada

Received 24 April 1996/Returned for modification 14 June 1996/Accepted 11 September 1996

We have identified alternatively spliced isoforms of murine Pax-3 and Pax-7 which differ by the presence or absence of a single glutamine residue in a linker region which separates two distinct DNA-binding subdomains within the paired domain. By reverse transcription-PCR, these isoforms of Pax-3 and Pax-7 (Q+ and Q−) were detected at similar levels through multiple developmental stages in the early mouse embryo. DNA-binding studies using the Q+ and Q− isoforms of Pax-3 revealed that this alternative splicing event had no major effect on the ability of these isoforms to bind to an oligonucleotide specific for the Pax-3 homeodomain (P2) or to a paired domain recognition sequence (e5) that interacts primarily with the N-terminal subdomain of the paired domain. However, DNA-binding studies with sequences (P6CON and CD19-2/A) containing consensus elements for both the N-terminal and C-terminal subdomains revealed that the Q− isoform binds to these sequences with a two- to fivefold-higher affinity; further mutation of the GTCAC core N-terminal subdomain recognition motif of CD19-2/A generated binding sites with a high degree of specificity for the Q− isoform. These differences in DNA binding in vitro were also reflected in the enhanced ability of the Q− isoform to stimulate transcription of a reporter containing multiple copies of CD19-2/A upstream of the thymidine kinase basal promoter. In support of the observations made with these naturally occurring Pax-3 isoforms, introducing a glutamine residue at the analogous position in PAX6 caused a fivefold reduction in binding to P6CON and a complete loss of binding to CD19-2/A and to the C-terminal subdomain-specific probe 5aCON. These studies therefore provide direct evidence for a role for the paired-domain linker region in DNA target site selection, and they identify novel isoforms of Pax-3 and Pax-7 that have the potential to mediate distinct functions in the developing embryo.

Pax-3 belongs to a family of transcription factors which are essential for a variety of developmental processes during mammalian embryogenesis (reviewed in reference 30). Members of the Pax family are defined by the presence of a 128-amino-acid DNA-binding domain called the paired domain (34) which binds as a monomer to a degenerate consensus defined by aligning several naturally occurring and in vitro-derived Pax-binding sequences (9, 12). In addition to the paired domain, some Pax proteins, including Pax-3, Pax-6, and Pax-7, also contain a second sequence-specific DNA-binding domain, the paired-type homeodomain (30). The homeodomains in these proteins have been shown to bind preferentially as dimers to palindromic sequences containing two classical TAAT homeodomain core recognition motifs (8, 35, 37). In addition to these two DNA-binding domains, Pax-3 also contains domains for transcriptional activation and inhibition (6), as well as a conserved octapeptide motif of unknown function (30).

In the mouse embryo, *Pax-3* is expressed prior to the onset of neurogenesis and somitogenesis, and its expression becomes progressively restricted to the dorsal portion of the neural tube and to the ventrolateral portions of the developing somites (3, 18). Loss-of-function mutations in *Pax-3* produce the *Splotch* phenotype in mice (10, 11, 36), characterized by pigmentary disturbances in heterozygotes and by profound defects in neurogenesis and limb muscle development in homozygotes. In humans, loss-of-function *PAX3* mutations are associated with Waardenburg syndrome, a dominant disorder which includes pigmentary disturbances, craniofacial abnormalities, and sensorineuronal deafness (1, 31). Furthermore, a chromosomal translocation which fuses the DNA-binding domains of PAX3 with the transactivation domain of a forkhead-related transcription factor has been associated with the childhood solid tumor alveolar rhabdomyosarcoma, a malignancy in which muscle progenitors are blocked in an undifferentiated, transformed state (2, 17). The observation that both haplo-insufficiency (*Splotch* and Waardenburg syndrome) and misexpression (rhabdomyosarcoma) produce distinct aberrant phenotypes points at a critical dose dependence for Pax-3 at particular stages during embryonic development.

Recent studies (9, 13, 38) have demonstrated that the paired domain actually consists of two distinct subdomains, each of which makes a discrete contribution to DNA recognition. In the case of Pax-5, naturally occurring target sequences can be divided into two classes: class I sites, which contain an unusually long recognition sequence (~20 bp) and which require both the N-terminal and C-terminal subdomains for binding, and class II sites, which are shorter (~12 bp) and which require only the N-terminal subdomain for binding (9). In Pax-6, an alternative splicing event which introduces a 14-amino-acid insertion into the N-terminal subdomain generates an isoform that interacts with DNA exclusively through its C-terminal subdomain (13). Recently, the resolution of the crystal structure of the paired domain of *Drosophila* Prd bound to a 15-bp class II sequence has confirmed the presence of two structurally independent domains, each containing a DNA-binding helix-turn-helix motif, joined by an extended linker region (38). Taken together, these studies suggest that the two subdomains

* Corresponding author. Mailing address: Department of Biochemistry, Room 907, McGill University, 3655 Drummond St., Montreal, Quebec, Canada H3G 1Y6. Phone: (514) 398-7291. Fax: (514) 398-2603.

within the paired domain might act independently or may cooperate in the recognition of specific target genes in vivo.

In this paper, we report the discovery of an alternative splicing event within the linker region of the Pax-3 paired domain which generates isoforms with distinct DNA-binding and transactivation properties. The observed differences in the activity of these Pax-3 isoforms, together with the apparent sensitivity of this transcription factor to both loss- and gain-of-function mutations, suggest a functional relevance to this alternative splicing event in vivo.

MATERIALS AND METHODS

cDNA cloning and sequencing. The isolation of full-length *Pax-3* cDNA clones by reverse transcription (RT) and PCR from embryonic day 11.5 (E11.5) mouse embryos has been described previously (36). Additional *Pax-3* cDNA clones were amplified by using primers P_3I (5′ GTTGCGTCTTCTAAGATCCTG 3′) and P_3J (5′ GCGTCCTTGAGCAATTTGTC 3′). *Pax-7* cDNA clones were amplified using primers P_7A (5′ TGTCTCCAAGATTCTATGCC 3′) and P_7D (5′ GGA TTTCCCAGCTGAACATC 3′). Amplification products were resolved on 2% agarose gels, purified, and cloned into a dT-tailed pBluescript vector (22). Products were sequenced by the dideoxy-chain termination method (26).

Quantitation of Pax-3 and Pax-7 isoforms by RT-PCR. Staged embryos were generated from C3H/HeJ mice by using timed matings, with the morning of the vaginal plug designated E0.5. RNA was isolated from whole embryos and from frozen embryo sections by homogenization in 6 M guanidinium hydrochloride, followed by sequential ethanol precipitations and extraction with phenol and chloroform (7). First-strand cDNA was synthesized from 2 μg of total RNA as described previously (36). Primer pairs P_3I-P_3J and P_7A-P_7D, which flank the alternatively spliced codon at the 5′ end of exon 3, were then used to amplify *Pax-3* and *Pax-7* cDNA fragments by PCR from 50 ng of template, using the following reaction conditions: 94°C, 40 s; 60°C, 40 s; and 72°C, 40 s (20, 22, or 24 cycles). Products were radiolabeled by end labeling one primer in each reaction with T4 polynucleotide kinase and [γ-^{32}P]ATP (7,000 Ci/mmol; ICN Biochemicals). The 181- and 178-bp *Pax-3* cDNA fragments, and the 155- and 152-bp *Pax-7* cDNA fragments, were then resolved on 6% denaturing polyacrylamide gels. Following electrophoresis, gels were dried and exposed to X-ray film for autoradiography or to a phosphor imaging plate for quantitation.

Expression of Pax-3 isoforms in cell culture and immunodetection. A construct containing a full-length *Pax-3* cDNA fragment in the eukaryotic expression vector pMT2 has been described previously (35). The coding region of this clone is identical to the published Pax-3 sequence (18) and contains a glutamine residue at position 75 of its paired domain (Pax-3/Q+). A Pax-3 expression construct that lacks this glutamine residue (Pax-3/Q−) was engineered by replacing a 329-bp *Sma*I fragment (nucleotides 345 to 674) with the corresponding fragment from a Pax-3/Q− cDNA clone in pBluescript, isolated previously by RT-PCR (36). These constructs were then transfected into COS-7 cells by calcium phosphate coprecipitation, and cells were harvested 24 h later in phosphate-buffered saline (PBS)–citrate, pelleted, and lysed by sonicating for 30 s in a buffer containing 20 mM *N*-2-hydroxyethylpiperazine-*N*′-2-ethanesulfonic acid (HEPES; pH 7.6), 150 mM NaCl, 0.5 mM dithiothreitol, 0.2 mM EDTA, 0.2 mM EGTA, 25% glycerol, and protease inhibitors. To monitor protein expression, whole cell extracts were resolved on sodium dodecyl sulfate (SDS)–12% polyacrylamide gels and transferred to nitrocellulose membranes for immunodetection using an anti-Pax-3 polyclonal antibody as described previously (35).

Electrophoretic mobility shift assays (EMSAs). All binding assays were performed with whole cell extracts in a buffer containing 10 mM Tris (pH 7.5), 50 mM KCl, 1 mM dithiothreitol, 2 mM spermidine, 2 mg of bovine serum albumin per ml, and 10% glycerol. Double-stranded oligonucleotides P2 (37), e5 (34), P6CON (12), 5aCON (13), and CD19-2/A (9) and its derivatives (Fig. 8) were synthesized with 5′ overhangs for subsequent labeling with Klenow DNA polymerase and either [α-^{32}P]dATP or [α-^{32}P]dGTP (3,000 Ci/mmol; Dupont). Binding reactions were typically carried out with 5 fmol of radiolabeled probe and 1 μg of poly[(dI-dC) · (dI-dC)] as a nonspecific competitor, with the exception of P2, for which 2 μg of salmon sperm DNA was used as a nonspecific competitor. For quantitation, gels were dried, exposed to a phosphor imaging plate, and analyzed with a Fuji BAS 2000 phosphor imaging station.

Transactivation assays. A DNA fragment containing three copies of the CD19-2/A paired-domain recognition sequence was cloned into the vector pT109 (kindly provided by G. Shore, McGill University) upstream of the TATA element of the thymidine kinase minimal promoter fused to a luciferase reporter gene. This reporter plasmid (3 μg) was cotransfected with increasing amounts of the Pax-3/Q+ and Pax-3/Q− expression plasmids (0.01, 0.05, 0.1, 0.5, 1.0, and 2.0 μg) into NIH 3T3 cells by calcium phosphate coprecipitation, along with 2 μg of a reference plasmid carrying the *lacZ* gene under control of the Rous sarcoma virus promoter. Twenty-four hours after transfection, cells were harvested in PBS-citrate, pelleted, and lysed in a buffer containing 10 mM Tris (pH 7.4) and 0.5% Nonidet P-40. After pelleting insoluble material, supernatants were assayed for β-galactosidase and luciferase activities according to standard protocols (20).

Construction of *PAX6* expression plasmids. A full-length human *PAX6* cDNA cloned in the eukaryotic expression vector pCMV (13) was kindly provided by T. Glaser (University of Michigan). To introduce a glutamine residue between Arg-74 and Val-75 of the PAX6 paired domain, oligonucleotide primers were designed to amplify from the cloned *PAX6* cDNA template two overlapping *PAX6* cDNA fragments in which an additional CAG glutamine codon was introduced into the primers in the region of overlap between the two fragments. Primers used to amplify the 5′ and 3′ fragments by PCR were P6-5′ (sense, positions 271 to 280, 5′ GAAAGGATCCCTCATAAAGG 3′), P6-Q5 (antisense, positions 655 to 635, 5′ TCGCTACCTGTCTCGGTTTACTAC 3′), P6-Q3 (sense, positions 642 to 661, 5′ ACCGAGACAGGTAGCGACTCCAG 3′), and P6-3′ (antisense, positions 1178 to 1159, 5′ ATTCTTGCTTCAGGTA GATC 3′). These amplification products were then mixed, heated to 95°C, and slowly cooled to room temperature to allow the complementary ends of the two fragments to anneal. Fusion products were generated by extending the 3′ ends of the DNA, using T4 DNA polymerase. This *PAX6* cDNA fragment containing the additional glutamine codon was further amplified by PCR using primers P6-5′ and P6-3′ and was cloned into the pCMV-PAX6 expression vector by replacing a 634-bp *Apa*I/*Bgl*II fragment from the wild-type construct with the corresponding fragment from the mutagenized cDNA. To allow immunodetection of the expressed proteins, a double-stranded cDNA fragment encoding a 10-amino-acid epitope derived from the influenza virus hemagglutinin A (HA) protein (16) was introduced in-frame as a blunt-ended fragment into the unique *Msc*I site at the 3′ end of the *PAX6* cDNA (position 1669). Finally, these HA-tagged PAX6/Q− and PAX6/Q+ cDNAs were cloned as *Sac*II fragments (position 394 to vector) into the *Eco*RI site of the pMT2 expression vector to allow for high-level expression of the recombinant proteins in COS-7 cells.

All PCR amplifications were performed with Vent DNA polymerase (New England Biolabs) to improve the fidelity of the reactions. In addition, the *Apa*I/*Bgl*II fragment in the PAX6/Q+ construct was fully sequenced to verify that the CAG codon had been correctly introduced and to ensure that no additional mutations had been introduced during the amplification steps. The HA-tagged PAX6 isoforms were expressed by transient transfection in COS-7 cells as described for Pax-3, and protein levels were monitored by Western blotting (immunoblotting) using the anti-HA monoclonal antibody 12CA5 (Babco).

RESULTS

Alternative splicing generates novel Pax-3 and Pax-7 paired-domain isoforms. At the primary sequence level, Pax-3 and its most closely related family member, Pax-7, are distinguished from other mammalian Pax proteins by the presence of an additional glutamine residue between amino acids 74 and 75 of the paired domain (Fig. 1A). This glutamine lies within the extended linker region that joins the two DNA-binding subdomains and at the genomic level correlates with the presence of an intron just upstream of the glutamine codon (Fig. 1B). Cloning and sequencing of multiple *Pax-3* cDNA clones obtained from previous work (36) led to the identification of two distinct alternate splice forms that differed by the inclusion or removal of this single glutamine residue through alternate use of the CAG glutamine codon as a 3′ splice acceptor (Fig. 1C). Subsequently, sequence analysis of additional *Pax-3* cDNA clones obtained from various sources (C3H/HeJ and C57BL6/J E11.5 embryos and P19 embryonic carcinoma cells) confirmed that this alternative splicing event was not strain specific but was present in several independent mouse lines (data not shown).

Since the genomic structure and the codon usage in this portion of the gene has been conserved between *Pax-3* and *Pax-7*, we next wished to determine whether a similar alternative splicing event occurred in *Pax-7* transcripts. Indeed, sequence analysis of multiple *Pax-7* cDNA clones generated by RT-PCR confirmed the existence of two distinct Pax-7 isoforms which differ by the presence or absence of this glutamine residue (Fig. 1C). Moreover, the potential for this alternative splicing event to occur has been preserved in the human *PAX3* and *PAX7* genes (4), suggesting that this alternative splicing event may be a general feature among the vertebrate *Pax* genes belonging to the *Pax-3*/*Pax-7* subfamily.

Both Pax-3 and Pax-7 isoforms are expressed through multiple developmental stages. To further assess the relevance of this alternative splicing event, we used an RT-PCR-based strategy to determine the relative abundance of the two RNA isoforms of *Pax-3* and *Pax-7* at various stages during develop-

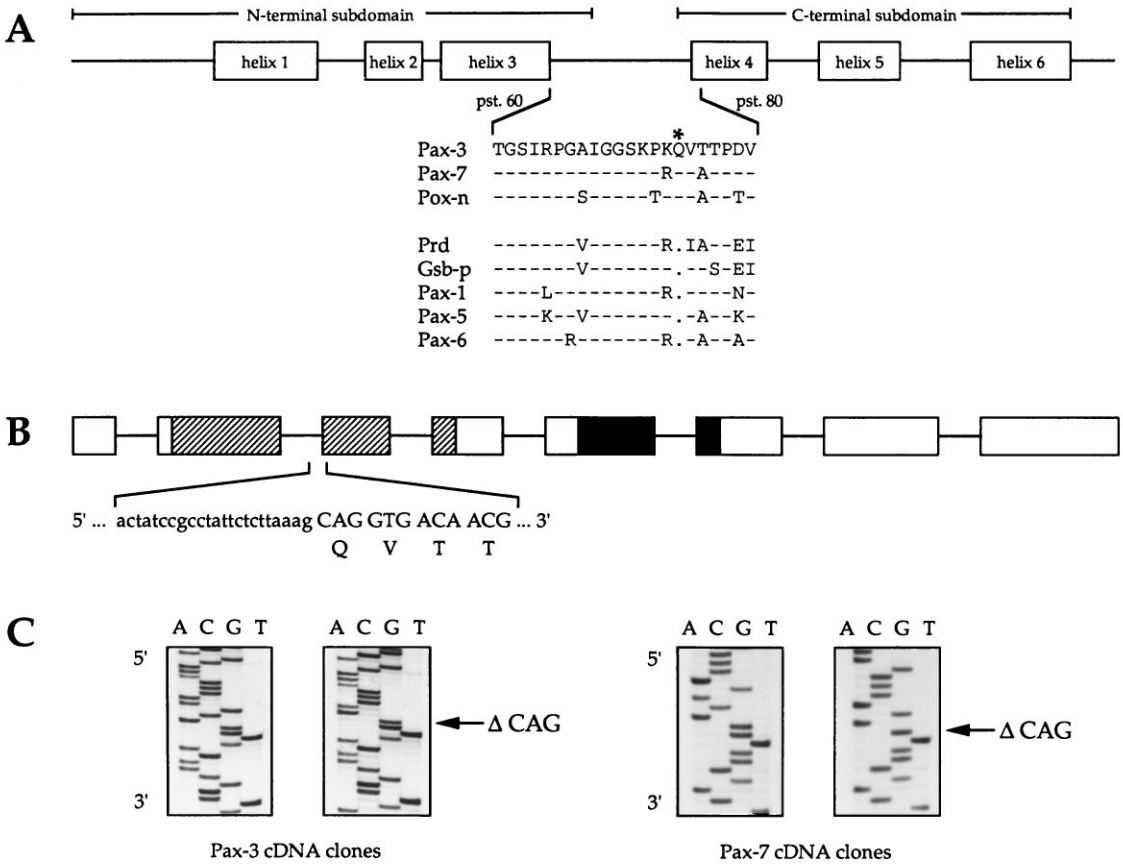

FIG. 1. Identification of novel alternate splice forms of Pax-3 and Pax-7. (A) Schematic representation of the paired domain. The N- and C-terminal subdomains are shown, with the six α-helices indicated with boxes. Shown below are the sequences of the linker regions (residues 60 to 80) of Pax-3, Pax-7, and other *Drosophila* and mammalian paired-domain-containing proteins (24, 30). Identical residues are indicated by dashes. Abbreviations for the *Drosophila* proteins: Pox-n, Pox-neuro; Prd, Paired; Gsb-p, Gooseberry-proximal. The additional glutamine in Pax-3, Pax-7, and *Drosophila* Pox-n is indicated by an asterisk. pst., position. (B) Genomic organization of mouse *Pax-3*. The nucleotide sequence at the end of intron 2 and the beginning of exon 3 is given below. Hatched box, paired domain; filled box, homeodomain. (C) Sequence analysis of multiple *Pax-3* (left) and *Pax-7* (right) cDNA clones. Arrows indicate positions of the three-nucleotide deletions identified in several *Pax-3* and *Pax-7* cDNA clones.

ment. This strategy makes use of primer pairs (P_3I-P_3J and P_7A-P_7D) that flank the alternatively spliced CAG codon and that generate two distinct amplification products corresponding to each isoform. The individual *Pax-3* (181- and 178-bp) and *Pax-7* (155- and 152-bp) amplification products differ from one another by three nucleotides and can therefore be resolved by electrophoresis through a 6% denaturing polyacrylamide gel. Both *Pax-3* and *Pax-7* are expressed during a fairly narrow window in the developing embryo, with peak levels at around E11.5 (18, 30). To assess potential temporal variation in the abundance of the two isoforms, RNA from different staged embryos (E9.5 through E14.5) was used to prepare first-strand cDNA, which was then used as a template for PCR amplification using the primer pairs described above (Fig. 2). Results for *Pax-3* indicated that the isoform containing the additional glutamine (hereafter Pax-3/Q+), corresponding to the slower migrating band, was approximately twice as abundant as the isoform lacking the glutamine (Pax-3/Q−), and moreover, that the ratio of the two isoforms did not differ at any of the stages examined (Fig. 2, upper left). A similar profile was obtained for *Pax-7* (Fig. 2, lower left). To ensure that PCR amplification was occurring in the linear range and that no misrepresentation in the abundance of the two isoforms was being introduced, amplification reactions were repeated for 20, 22, and 24 cycles (data not shown). Quantitation using a phosphor imaging screen yielded similar results at all time points, validating the use of this PCR-based strategy.

To assess potential differences in the spatial distribution of the two isoforms, we employed a similar RT-PCR approach

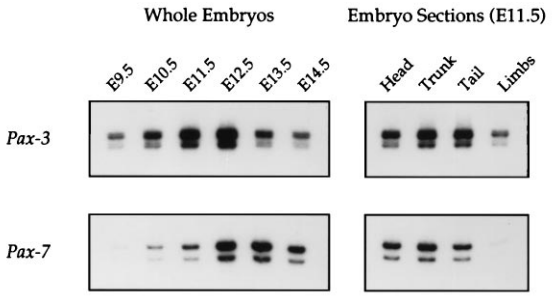

FIG. 2. Analysis of the temporal and spatial distribution of each *Pax-3* and *Pax-7* isoform during embryogenesis. RNA was prepared from whole embryos from day 9.5 to day 14.5 postcoitum (left panels) or from various anatomical sections isolated from day 11.5 embryos (right panels). Bands correspond to radiolabeled RT-PCR products that were generated by amplifying short *Pax-3* and *Pax-7* cDNA fragments by using primers that flank the alternatively spliced glutamine codon. Different-sized products corresponding to each isoform were separated by electrophoresis through a 6% denaturing polyacrylamide gel. In all cases, the upper band corresponds to the Q+ isoform, and the lower band corresponds to the Q− isoform.

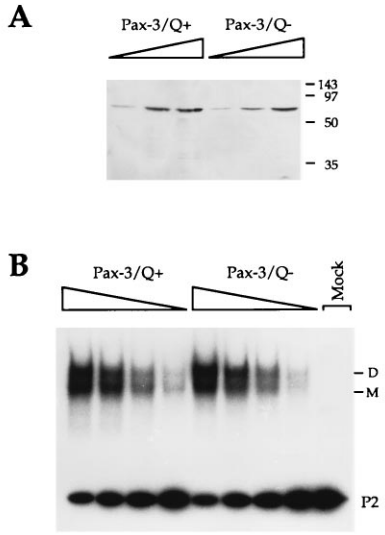

FIG. 3. Expression of full-length Pax-3/Q+ and Pax-3/Q− isoforms in COS-7 cells. (A) Western blot analysis of whole cell extracts prepared from COS-7 cells transiently transfected with full-length Pax-3/Q+ and Pax-3/Q− expression plasmids, using an anti-Pax-3 polyclonal antibody (35). Proteins were resolved by electrophoresis through an SDS–12% polyacrylamide gel and transferred to a nitrocellulose membrane. Adjacent lanes represent twofold increases in the amount of protein. The positions of the protein standards are shown to the right in kilodaltons. (B) EMSA using the Pax-3/Q+ and Pax-3/Q− whole cell extracts and the homeodomain-specific oligonucleotide P2 (37). Adjacent lanes represent twofold dilutions of protein extracts. A binding reaction using mock-transfected COS-7 cells is shown on the right (Mock). D, dimer; M, monomer; P2, free probe.

using RNA prepared from various anatomical sections (head, trunk, tail, and limbs) isolated from E11.5 mouse embryos. *Pax-3* and *Pax-7* show similar patterns of expression in the developing neural tube and in the adjacent somitic mesoderm (30); in addition, *Pax-3* is expressed in a population of myogenic cells that migrate into and populate the limb buds (3). As shown in Fig. 2 (upper right), both the Q+ and Q− isoforms of *Pax-3* were present in all the tissues analyzed, and as before, no variation was observed in the relative abundance of the two transcripts. A similar pattern was seen for *Pax-7* except that no expression was seen in the limbs (Fig. 2, lower right). These studies demonstrate that these novel isoforms of *Pax-3* and *Pax-7* are expressed together through multiple developmental stages in the developing mouse embryo.

Removal of a single glutamine from Pax-3 increases its affinity for full-length paired-domain consensus recognition sequences. To determine what effect this alternative splicing event had on the DNA-binding properties of the Pax-3 paired domain, we introduced full-length Pax-3/Q+ and Pax-3/Q− expression constructs into COS-7 cells by transient transfection. In addition to the paired domain, Pax-3, like several other Pax family members, also contains a second DNA-binding domain, a paired-type homeodomain (18), which is able to bind both monomerically and as a dimer to specific palindromic DNA sequences (37). Western blot analysis of whole cell extracts prepared from the Pax-3 transfectants by using a polyclonal anti-Pax-3 antibody (35) revealed that the recombinant proteins were intact and were expressed at similar levels in the two extracts (Fig. 3A). To further normalize Pax-3 levels in the two sets of extracts, serial dilutions of the Pax-3/Q+ and Pax-3/Q− transfectants were assayed for binding to P2, a homeodomain-specific oligonucleotide (37), in an EMSA. As shown in Fig. 3B, the two sets of extracts showed similar levels of monomeric and dimeric binding to P2; moreover, this binding was specific to the Pax-3-transfected cells, as no slower-mobility complexes were seen when P2 was incubated with whole cell extracts from mock-transfected COS-7 cells (Fig. 3B, right lane). These observations, together with the results of the Western blot analysis, indicated that both isoforms were present at similar levels in the extracts and suggested that the alternative splicing event in the linker region of the paired domain had no major effect on DNA binding through the homeodomain.

Having normalized protein levels in the two sets of extracts, we then tested the ability of the Pax-3/Q+ and Pax-3/Q− isoforms to bind to a series of paired-domain-specific oligonucleotides (Fig. 4). The first oligonucleotide tested was a typical class II sequence, e5, which has been shown to interact primarily with the highly conserved N-terminal subdomain found in all *Drosophila* and mammalian paired domains (9, 34). An EMSA with e5 using the Pax-3/Q+ and Pax-3/Q− transfectants at a single protein concentration revealed similar levels of binding by both isoforms (Fig. 4A, left lanes). To quantitatively determine the relative affinity of each isoform for e5, saturation analysis was performed with a fixed amount of protein and radiolabeled e5 and increasing amounts of unlabeled e5. As expected, the ratio of bound e5 to free e5 plotted against the amount of e5 bound (Scatchard plot) showed an inverse linear relationship whose slope reflected the relative affinity of each isoform for the oligonucleotide (Fig. 4B). The Pax-3/Q+ and Pax-3/Q− extracts yielded similar Scatchard plots with e5, indicating that the two isoforms bind to this representative class II sequence with the same affinity.

Since the alternatively spliced glutamine did not appear to have an effect on binding through the N-terminal subdomain, we went on to test the ability of each isoform to bind to oligonucleotides that include a recognition motif for the C-terminal subdomain. P6CON, which was derived in vitro as an optimal recognition sequence for the Pax-6 paired domain (12), contains a long consensus that can be aligned to the class I sequences that have been defined for Pax-5 (9). Using a set of binding assays similar to those described above for e5, we found that the Pax-3/Q− isoform bound to P6CON with twice the affinity of the Pax-3/Q+ isoform (Fig. 4A, middle lanes; Fig. 4C). Similarly, using CD19-2/A, a high-affinity binding site discovered for Pax-5 that contains strong consensus elements for both the N-terminal and C-terminal subdomains (9), we observed a fivefold-higher level of binding by the Pax-3/Q− isoform than by the Q+ isoform (Fig. 4A, right lanes; Fig. 4D). These results indicate that removal of this glutamine from the linker region of the Pax-3 paired domain increases its affinity for representative full-length paired-domain recognition sequences. This alternative splicing event therefore has the potential to generate greater diversity in the binding specificity of Pax-3 and Pax-7 in vivo.

Enhanced DNA binding by the Q− isoform of Pax-3 results in an increased transactivation potential. To extend these observations to the function of Pax-3 in vivo, we measured the ability of the two Pax-3 isoforms to activate transcription of a luciferase reporter plasmid containing three copies of the CD19-2/A binding site upstream of the TATA element of the thymidine kinase basal promoter (Fig. 5A). Pax-3, like other Pax proteins, has been shown to activate transcription over only a narrow range of protein concentrations (6, 8); at higher concentrations, Pax-3 has been found to inhibit transcription (6). In this experiment, NIH 3T3 cells were cotransfected with a fixed amount of reporter plasmid and increasing amounts of Pax-3/Q+ and Pax-3/Q− expression plasmid; control pMT2 vector with no insert was added to maintain the amount of expression vector constant in each transfection. In addition, a

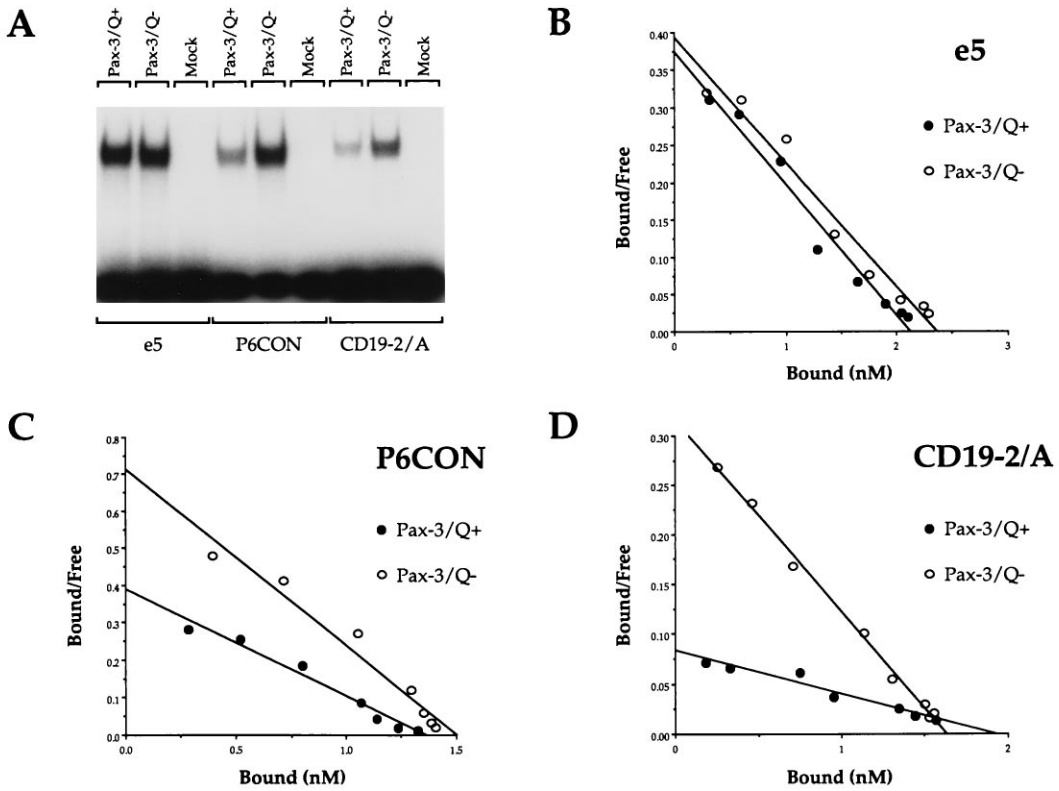

FIG. 4. Effect of the Q+/Q− alternative splicing event on the DNA-binding properties of the Pax-3 paired domain. (A) EMSAs using the N-terminal subdomain-specific oligonucleotide e5 and representative full-length paired-domain recognition sequences P6CON and CD19-2/A. The same amount of extract was used for each binding reaction, and a mock-transfected control extract is included for each probe. (B to D) Saturation analysis for e5, P6CON, and CD19-2/A, using a fixed amount of protein and radiolabeled probe and increasing amounts of unlabeled oligonucleotide. The total amount of oligonucleotide ranged from 25 to 2,000 fmol in a 20-μl reaction volume. Scatchard plots, generated by plotting the ratio of bound oligonucleotide to free oligonucleotide versus the amount of bound oligonucleotide, are shown. All plots represent averages of at least four experiments. The slope of the curve in panel B indicates that the two isoforms of Pax-3 bind to e5 with the same relative affinity. Likewise, the curves in panels C and D indicate that the Q− isoform of Pax-3 binds to P6CON and CD19-2/A with two- and fivefold-higher affinities, respectively.

fixed amount of RSV-β-gal expression plasmid was included in each transfection to control for transfection efficiency. Using the Pax-3/Q+ expression plasmid, we observed very weak stimulation of the luciferase reporter at low concentrations of expression plasmid (up to 1.3-fold over the levels obtained with the reporter plasmid alone) and weak inhibition at higher concentrations of expression plasmid (Fig. 5B). In contrast, the Pax-3/Q− expression vector produced up to a 2.2-fold stimulation in the activity of the promoter over the same range of concentrations and still showed weak activation even at the higher expression plasmid concentrations (Fig. 5B). We observed these differences consistently in several independent transfections; moreover, within a given assay, each data point was measured in triplicate. This experiment demonstrates that the enhanced DNA-binding activity of the Pax-3/Q− isoform increases its capacity to stimulate the transcription of specific target promoters, supporting the differences in DNA binding that we observe in vitro.

A mutant form of PAX6 with a glutamine inserted at position 75 of its paired domain shows a drastic reduction in DNA-binding activity. To extend these observations to other Pax proteins, we introduced an additional glutamine residue at the analogous position in the PAX6 paired domain and tested the ability of this mutant to bind to representative paired-domain recognition sequences (Fig. 6). In general, the PAX6 paired domain shows a greater dependence on the C-terminal subdomain than other Pax proteins, as evidenced by its propensity to select a relatively long recognition sequence in vitro (12) and its failure to interact with many class II sequences that do not contain a recognition motif for the C-terminal subdomain (9). Thus, if an additional glutamine at position 75 of the paired domain interferes specifically with the function of the C-terminal subdomain, as suggested by our observations with the naturally occurring isoforms of Pax-3, its effect on PAX6 binding would be expected to be quite pronounced. Wild-type (PAX6/Q−) and mutant (PAX6/Q+) isoforms were expressed by transient transfection in COS-7 cells, and levels of recombinant protein in the extracts were assessed by means of an HA epitope tag inserted at the C-terminal end of each polypeptide (Fig. 6A). Western blot analysis using the anti-HA monoclonal antibody 12CA5 revealed a major band of approximately 60 kDa in each extract and a number of higher-molecular-mass species that may be the result of different states of posttranslational modification; no immunoreactive bands of this size were seen in the mock-transfected COS-7 whole cell extracts (Fig. 6B). Moreover, similar levels of immunoreactivity were seen in both extracts over a series of different protein concentrations, indicating that the recombinant proteins were present in equal amounts in the two extracts.

EMSAs using P6CON, the consensus recognition sequence derived in vitro for PAX6, revealed that the addition of a glutamine residue at position 75 of the PAX6 paired domain caused approximately a fivefold reduction in binding to this probe (Fig. 6C), as assessed by quantitating the amount of specific complex formed at each protein level by phosphor imaging. An even more dramatic reduction in binding was

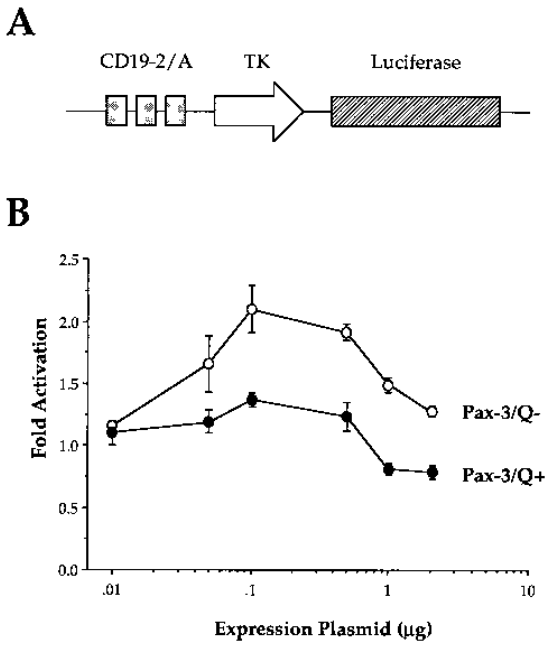

FIG. 5. Transactivation properties of the Pax-3 isoforms. (A) Schematic representation of the luciferase reporter used in these experiments. Three copies of the CD19-2/A paired-domain-binding sequence were placed upstream of the thymidine kinase (TK) basal promoter driving expression of a luciferase reporter gene. (B) Transactivation assays. NIH 3T3 cells were transiently transfected with a fixed amount of reporter plasmid (3 μg) and increasing amounts of the Pax-3/Q+ and Pax-3/Q− pMT2 expression plasmids. A reference plasmid (2 μg) carrying the *lacZ* gene under the control of the Rous sarcoma virus promoter was also included in each transfection to control for transfection efficiency. Plots show fold increase in the activity of the luciferase reporter in response to different amounts of Pax-3/Q+ and Pax-3/Q− expression plasmid. Fold activation is expressed relative to the activity measured in cells transfected with the reporter plasmid alone. Each data point represents an average of at least three separate transfections. Error bars represent standard errors of the mean.

observed with CD19-2/A; with this probe, we were no longer able to detect specific protein-DNA complexes with the PAX6/Q+ isoform, while a complex between wild-type PAX6 and the radiolabeled CD19-2/A probe was seen even at the lowest dilution of protein (Fig. 6D). Prolonged exposure of this EMSA showed no specific binding in the PAX6/Q+-transfected extract compared to the mock-transfected COS-7 cell extract (data not shown). Based on the dilutions shown, this represents at least a 32-fold reduction in the level of binding by the Q+ mutant isoform. These findings extend the observations made with the naturally occurring isoforms of Pax-3 and reveal for PAX6 a high degree of dependence on residues within the linker region during the recognition of specific sequences as targets for regulation.

To further examine the effect of this glutamine residue on C-terminal subdomain function, we tested the abilities of these wild-type and mutant forms of PAX6 to bind to the C-terminal subdomain-specific probe 5aCON (Fig. 7). 5aCON was selected as an optimal recognition sequence in vitro by an alternatively spliced isoform of PAX6 defective for N-terminal subdomain binding (13) and therefore serves as a useful probe for isolating the DNA-binding properties of the C-terminal subdomain. As shown in Fig. 7, only the wild-type PAX6/Q− isoform formed a specific complex with 5aCON (lane 1); the presence of the additional glutamine residue eliminated formation of a complex with this probe (lane 2). In addition, neither isoform of Pax-3 showed any specific binding to 5aCON over the background level (Fig. 7, lanes 3 and 4); this finding is consistent with the high degree of sequence divergence between the C-terminal subdomains of Pax-3 and PAX6, as noted previously (13). Nevertheless, this experiment demonstrates that a glutamine residue at position 75 of the PAX6 paired domain linker region is sufficient to abrogate binding to 5aCON, highlighting the importance of this linker region in mediating C-terminal subdomain DNA-binding activity.

Point mutations within the N-terminal subdomain recognition motif of CD19-2/A generate oligonucleotides with a high specificity for the Pax-3/Q− isoform. The inability of the Q+ mutant isoform of PAX6 to recognize some sequences effectively bound by the wild-type form of the protein suggested that the Q− isoform of Pax-3 might also bind to some DNA sequences not recognized by its Q+ counterpart. To investigate this possibility, we introduced single point mutations into the 5′ end of the CD19-2/A recognition sequence and tested the ability of the Q+ and Q− isoforms of Pax-3 to bind to these oligonucleotides in EMSAs (Fig. 8). The mutated bases fall within the core GTCAC pentanucleotide motif which forms base-specific contacts with helix 3 of the N-terminal subdomain in the paired-domain crystal structure (38). Since the additional glutamine residue in Pax-3/Q+ appears to specifically disrupt C-terminal subdomain function, this isoform is highly dependent on N-terminal subdomain-mediated major groove contacts during binding to the wild-type CD19-2/A sequence, and mutations in the GTCAC core should have a profound effect on the ability of the Q+ isoform to bind stably to this sequence. Conversely, the enhanced DNA-binding properties of the C-terminal subdomain of the Q− isoform should serve to stabilize the interaction of Pax-3/Q− with these sequences and compensate for the loss of N-terminal subdomain-mediated DNA contacts.

Initial DNA-binding experiments with these mutant derivatives of CD19-2/A indicated that the G-to-T substitution at position 4 resulted in approximately a fivefold reduction in binding by the Q− isoform relative to the parent CD19-2/A oligonucleotide, while the C-to-A mutation at position 8 reduced complex formation with Pax-3/Q− by a factor of 20 (data not shown). Consequently, for the EMSA presented in Fig. 8B, we adjusted protein levels in the binding reactions to produce equivalent levels of complex formation by Pax-3/Q− with all probes. Using CD19-2/A as a control (Fig. 8B, left), we observed approximately a four- to fivefold difference in binding between the Q+ and Q− isoforms, as seen previously (Fig. 4). We then assayed the first of the mutant derivatives, CD19-2/A (4T), and found that the Pax-3/Q+ isoform had a highly reduced affinity for this probe (Fig. 8B, middle). Using a control binding reaction with an equivalent amount of mock-transfected COS-7 whole cell extract to estimate background (not shown), we measured approximately a 30-fold difference in complex formation between the two isoforms. Similarly, the Q+ isoform of Pax-3 also showed a markedly reduced affinity for the second mutant derivative, CD19-2/A (8A) (Fig. 8B, right); prolonged exposure of the gel showed no specific complex formation by Pax-3/Q+ to this probe compared to a mock-transfected control extract (not shown). This experiment therefore demonstrates that the enhanced DNA-binding properties of the C-terminal subdomain of the Q− isoform of Pax-3 can compensate for the loss of N-terminal subdomain-mediated DNA interactions, allowing this isoform to recognize target sequences not effectively bound by the Q+ form of the protein. Altogether, these studies provide direct evidence for a critical role for the paired-domain linker region in DNA target site selection and identify novel isoforms of Pax-3 and Pax-7 that have the potential to mediate distinct functions in the developing embryo.

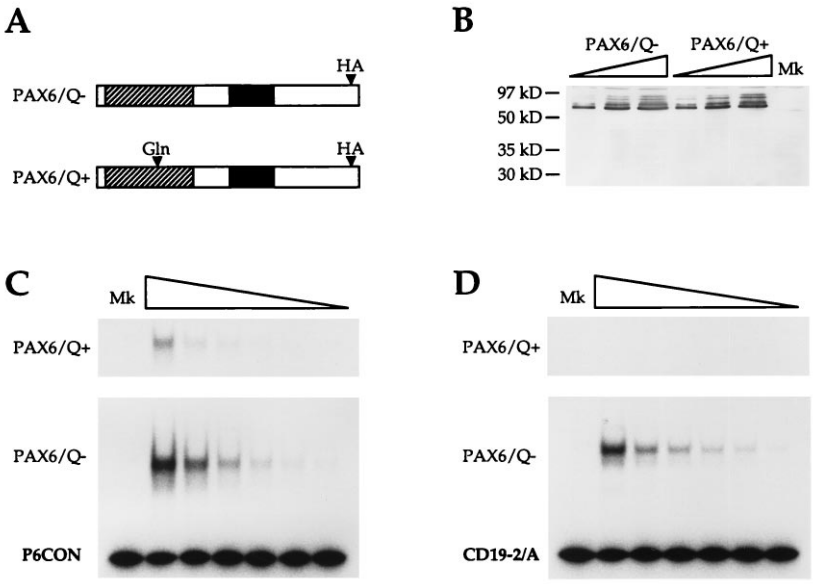

FIG. 6. DNA-binding properties of wild-type (Q−) and mutant (Q+) forms of PAX6. (A) Schematic representation of the proteins generated from the PAX6/Q− and PAX6/Q+ expression constructs. In the PAX6/Q+ mutant, a single glutamine (Q) was introduced between residues 74 and 75 of the PAX6 paired domain, and in both constructs, a 10-amino-acid HA epitope was inserted near the C-terminal end of each polypeptide to allow for immunodetection. (B) Western blot of PAX6/Q−- and PAX6/Q+-transfected COS-7 cells. Expression plasmids were introduced into COS-7 cells by transient transfection, and levels of recombinant protein in the resulting whole cell extracts were analyzed by Western blotting using the anti-HA monoclonal antibody 12CA5. For each extract (Q− and Q+), three different protein concentrations, representing twofold increases in protein, are shown. Lane Mk shows an extract of mock-transfected COS-7 cells at the highest protein concentration. (C) EMSAs showing binding of PAX6/Q− (top) and PAX6/Q+ (bottom) to P6CON. In both panels, lane Mk shows a binding reaction performed with an extract prepared from mock-transfected COS-7 cells. The adjacent lane shows the binding activity of whole cell extracts expressing an equivalent amount of PAX6/Q− or PAX6/Q+ protein. Subsequent lanes show twofold dilutions of protein extracts. Free probe (bottom) is shown only for the lower panel. (D) EMSAs showing binding of PAX6/Q− (top) and PAX6/Q+ (bottom) to CD19-2/A. Details are as for panel C.

DISCUSSION

Paired-domain-containing proteins occupy key positions in a variety of different developmental cascades, both in invertebrates and in a number of higher vertebrate organisms (24, 30). Within these proteins, the 128-amino-acid paired domain itself is the most highly conserved structure, with over 50 residues found to be invariant in all paired domains described to date. The majority of these residues reside in the more highly conserved N-terminal subdomain, which folds into a highly ordered structure containing a classical helix-turn-helix motif (38). In agreement with this high degree of conservation in the N-terminal subdomain, all known paired-domain-containing proteins are able to recognize a similar core motif found in e5, a paired-domain recognition sequence identified in the *Drosophila even-skipped* promoter, and other typical class II recognition sequences (9). However, naturally occurring binding sites for Pax-5 (9) and Pax-8 (39), and in vitro-derived optimal recognition sequences for Pax-2 and Pax-6 (12), contain bases outside this core which are highly conserved and which are required for optimal binding (9). Furthermore, deletion analyses have clearly demonstrated that the C-terminal subdomain of the paired domain makes an important contribution to binding to these class I sequences, suggesting a form of cooperativity between the two subdomains in the recognition of many naturally occurring target sequences (9). The published three-dimensional structure of the Prd paired domain bound to a 15-bp class II recognition sequence identified within the C-terminal subdomain an additional helix-turn-helix motif that is predicted to interact with the major groove of DNA (38). However, because the oligonucleotide used for cocrystallization was too short to interact with the C-terminal subdomain, it is not clear how the two subdomains and the linker region interact to recognize full-length class I sequences or how each subdomain contributes to the recognition of target sites in vivo.

In this study, we have identified an alternative splicing event in *Pax-3* and *Pax-7* transcripts which generates isoforms that differ by the presence or absence of a single glutamine residue within the paired-domain linker region. While the two isoforms of Pax-3 show similar levels of homeodomain and N-terminal subdomain-mediated paired domain DNA-binding activity, we find that the Q− isoform binds with a higher affinity to conventional full-length paired-domain recognition sequences (P6CON and CD19-2/A). This finding suggests that the presence of this additional glutamine interferes with the ability of the C-terminal subdomain to interact with a typical

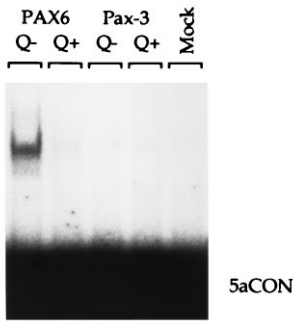

FIG. 7. EMSA showing binding of the PAX6 and Pax-3 Q+ and Q− isoforms to the C-terminal subdomain-specific probe 5aCON. Binding reactions were performed with equivalent amounts of whole cell extracts prepared from COS-7 cells transiently transfected with the PAX6/Q−, PAX6/Q+, Pax-3/Q−, and Pax-3/Q+ expression plasmids. A control reaction using whole cell extract from mock-transfected COS-7 cells is shown to the right.

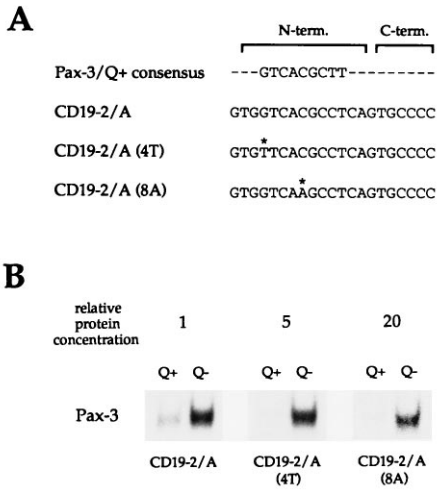

FIG. 8. Binding of the Q+ and Q− isoforms of Pax-3 to derivatives of CD19-2/A carrying single base substitutions in the GTCAC core N-terminal subdomain recognition motif. (A) Alignment of CD19-2/A and mutant derivatives to a consensus derived in vitro for the Q+ isoform of Pax-3 (5, 14). As shown, the N-terminal subdomain recognition motif of CD19-2/A differs from the Pax-3/Q+ consensus by a single T-to-C substitution at position 11. Additional mutations made to the GTCAC core of CD19-2/A are indicated by asterisks. The positions of the mutated bases are identified according to the numbering system used in the published paired-domain crystal structure (38). (B) EMSAs showing binding of Pax-3/Q+ and Pax-3/Q− to CD19-2/A and derivatives. Binding reactions with CD19-2/A were performed as before, using equivalent amounts of whole cell extracts from Pax-3/Q+- and Pax-3/Q−-transfected COS-7 cells. To obtain a similar level of binding by the Q− isoform with each probe, CD19-2/A (4T) and CD19-2/A (8A) were shifted when 5- and 20-fold-higher concentrations of extract, respectively, were used. For each probe, only the portion of the gel showing the shifted protein-DNA complexes is shown.

consensus defined for this subdomain by comparing naturally occurring and in vitro-derived paired-domain recognition sequences (9, 12, 13). In this regard, it is interesting that selection of optimal binding sites for the Q+ isoform of Pax-3 generated a relatively short consensus that did not extend into the region normally recognized by the C-terminal subdomain (5, 14). A possible interpretation of these results is that removal of this additional glutamine residue by alternative splicing is required to unmask the latent DNA-binding properties of the C-terminal subdomain of the Pax-3 paired domain. Again, it is significant that the Q+ isoform of Pax-3 cannot bind to any of the known oligonucleotides which show an absolute dependence on both subdomains (so-called class I sequences) for binding by other Pax proteins (9). The demonstration that mutations in the N-terminal subdomain recognition motif of CD19-2/A can generate binding sites specific for the Q− isoform raises the possibility that this isoform has retained the ability to bind additional class I sequences not recognized by the Q+ form of the protein and could eventually lead to the identification of specific cellular promoters differentially regulated by the two isoforms.

In support of the observations made with the naturally occurring Pax-3 isoforms, introducing an additional glutamine residue into the PAX6 paired domain at the analogous position causes a significant reduction in binding to an optimized PAX6 recognition sequence, P6CON, and completely abolishes binding to CD19-2/A, a recognition sequence that is bound with high affinity by wild-type PAX6 and all paired domains tested to date (9). Compared to other Pax proteins, PAX6 contains a number of amino acid substitutions in helix 3 of its N-terminal subdomain, and as a result, it is not able to bind to most class II paired-domain recognition sequences (8, 9). However, in the case of the wild-type CD19-2 sequence, a typical class II sequence which is bound very poorly by PAX6 (9), the introduction of a single A residue at a specific position within this oligonucleotide generates a recognition sequence (CD19-2/A) that contains consensus binding sites for both the N-terminal and C-terminal subdomains and is bound by PAX6 with a very high affinity (9). Thus, it appears that either the absence of a C-terminal subdomain consensus (CD19-2) or disruption of C-terminal subdomain structure (PAX6/Q+) is sufficient to prevent PAX6 from binding to this sequence. The dramatic effect of adding a glutamine within the linker region of the PAX6 paired domain serves to underscore the high degree of interdependence between the N- and C-terminal subdomains of PAX6 in DNA sequence recognition and illustrates the importance of residues in the linker region in mediating the DNA-binding properties of this protein.

From a structural point of view, this study raises some interesting questions with regards to the role of the linker region in modulating C-terminal subdomain-DNA interactions. In the published crystal structure (38), the C-terminal tail of the N-terminal subdomain (residues 68 to 72) remains extended and sits in the minor groove of the DNA, making base-specific contacts. The residues immediately following, which include the additional glutamine in the Q+ isoforms of Pax-3 and Pax-7, form the N-terminal arm of the C-terminal subdomain. In the *Hin* recombinase, a prokaryotic DNA-binding protein that has significant structural and sequence homology with the C-terminal subdomain of the paired domain (38), this N-terminal arm sits in the minor groove of the DNA and makes critical base-specific contacts (15). Indeed, mutations in the N-terminal arm of this protein destroy its ability to bind to its cognate sequence elements (15). Moreover, in the *Hin* recombinase and other related prokaryotic DNA-binding proteins, a proline and an invariant basic residue which interact with DNA are always followed by a hydrophobic residue which sits at the interface of a hydrophobic pocket formed by helices 1 and 3 (15). Pax proteins also contain a proline followed by a basic residue (K or R) and a hydrophobic residue (I, V, or L) in this region of the N-terminal arm; strikingly, the additional glutamine in the Q+ isoforms of Pax-3 and Pax-7 is located between these basic and hydrophobic residues (Fig. 1A). Thus, the presence of a glutamine at this position may disrupt the normal folding of the C-terminal subdomain or may direct the C-terminal subdomain to adopt a novel conformation. Alternatively, this glutamine may sit in the minor groove and disrupt normal DNA contacts. Additional mutagenesis of the linker region may help to unravel the structural constraints placed on this region during binding to conventional full-length paired-domain recognition sequences.

Mechanistically, the alternative splicing of the CAG glutamine codon in *Pax-3* and *Pax-7* transcripts provides an interesting example of competition between adjacent 3′ splice acceptor sites during pre-mRNA processing in vivo. Canonical 3′ splice acceptor sequences are characterized by an invariant AG dinucleotide at the −2 and −1 positions of the intron, usually preceded by C or U at the −3 position (72 and 23% of naturally occurring human introns, respectively), and a polypyrimidine tract which begins at the −5 position and extends 12 to 18 nucleotides upstream (29). The AG dinucleotide at the 3′ end of the intron has been shown to interact with the splicing machinery through base pairing with a CU dinucleotide in the 5′ tail of the U1 small nuclear RNA (25); a guanine immediately downstream of this CU dinucleotide is also in a favorable position for base pairing, which may explain the selection of a cytosine at the intron −3 position. In addition, U5 small nuclear RNA has been shown to interact with nucleotides at

positions +1 and +2 of the downstream exon (23); while this interaction does not involve traditional Watson-Crick base pairing, naturally occurring intron/exon boundaries show a weak preference for G and U at the +1 and +2 positions, respectively (29). As well, the polypyrimidine tract has been shown to interact with a variety of nuclear factors, including the essential splicing factor U2AF, and plays an important role in defining 3′ splice site utilization (28).

At the *Pax-3* intron 2/exon 3 border, the 3′ splice acceptor defined by the downstream glutamine codon (. . .cagGU. . .) matches the optimal consensus for interaction with the U1 and U5 small nuclear RNAs (Fig. 1B). However, the position of this downstream splice acceptor with respect to the polypyrimidine tract may interfere with efficient use of this acceptor site by the splicing machinery. The upstream splice acceptor, on the other hand, is favorably positioned with respect to the polypyrimidine tract but deviates from the optimal consensus typically found at 3′ splice acceptor sites (. . .aagCA. . .), including an A instead of a C at the −3 position, resulting in a less efficient interaction of this acceptor site with the U1 and U5 components of the spliceosome. These competing factors (efficiency of U1/U5-acceptor site interactions versus proximity to the U2AF-bound polypyrimidine tract) may therefore be responsible for the roughly equal selection of the two alternate 3′ splice acceptor sites during transcript processing to produce mature *Pax-3* and *Pax-7* mRNAs in vivo.

In addition to the novel isoforms of Pax-3 and Pax-7 identified in this study, other *Pax* genes have also made use of alternative splicing to increase functional diversity within the family. In Pax-8, multiple isoforms differing in their transactivation potential are generated by alternative splicing of exons downstream of the DNA-binding domains (21). An interesting parallel also exists with PAX6, where an alternative splicing event introducing an additional 14 amino acids between residues 45 and 46 of the paired domain destroys the ability of this PAX6-5a isoform to bind to the normal PAX6 recognition sequence, P6CON (13). Instead, this 5a isoform selects a novel recognition sequence, 5aCON, that interacts exclusively with the C-terminal subdomain (13). Thus, this alternative splicing event appears to act as a molecular switch which toggles between (i) an isoform (wild-type PAX6) in which both paired-domain subdomains are functional and contribute to sequence recognition and (ii) an isoform (PAX6-5a) in which only the C-terminal subdomain retains the ability to interact functionally with DNA. Importantly, we have found that a mutant isoform of PAX6 carrying a glutamine at position 75 can no longer bind to 5aCON, demonstrating that this additional glutamine has a specific effect on the C-terminal subdomain-mediated DNA-binding activity of this protein. Thus, the removal of the glutamine from the Pax-3 linker region by alternative splicing may also serve as a type of molecular switch, toggling between (i) an isoform (Pax-3/Q+) which exhibits primarily N-terminal subdomain DNA-binding activity and (ii) an isoform (Pax-3/Q−) in which both subdomains make an important contribution to sequence recognition.

The developmental importance of the PAX6-5a alternative splicing event is supported by the discovery of an intronic mutation which alters the ratio of the two PAX6 isoforms, producing a distinct ocular phenotype in humans (13). However, as with the alternative splicing events described herein for Pax-3 and Pax-7, the inclusion of this 5a-exon does not appear to be regulated during embryogenesis, the relative abundance of the two isoforms remaining constant throughout multiple developmental stages (13). This apparent lack of regulation of a developmentally relevant alternative splicing event is also seen with the *Drosophila* gene transcripts *I-POU* and *twin of I-POU*, which make use of an alternate 3′ splice acceptor to generate proteins that differ by only two amino acids within the POU-type homeodomain (33). As with PAX6, these isoforms, though displaying very distinct activities, are expressed at similar levels throughout development (33). This finding suggests that many developmentally important alternative splicing events may serve simply to increase the functional diversity of particular genes during embryogenesis.

To date, no *Splotch* or Waardenburg syndrome mutations have been discovered in the vicinity of the alternatively spliced glutamine residue. Thus, in the absence of a means to easily modulate levels of the two isoforms of Pax-3 in vivo, it remains difficult to assess the functional relevance of this alternative splicing event during the process of development. However, Pax proteins have been shown to display a critical dose dependence in a variety of experimental systems. In transactivation assays in cultured cells, Pax-3 (6) and other Pax proteins (8) activate expression of reporter genes over only a very narrow concentration range, a phenomenon which we also observe in this current study (Fig. 5). Similarly, in vivo, losses of single alleles from the mouse *Pax-3* and *Pax-6* genes (10, 19), or from the human *PAX2*, *PAX3*, and *PAX6* genes (1, 27, 31, 32), all produce readily observable phenotypes. These observations suggest that regulatory events such as alternative splicing that modulate the activity of Pax proteins will have important functional consequences. The differences that we observe in the abilities of the Pax-3 isoforms to transactivate a common reporter plasmid in cotransfection assays serve to underscore the functional nonequivalence of these proteins. The possibility that the two isoforms of Pax-3 and Pax-7 recognize distinct target genes in vivo therefore remains an interesting avenue for further study.

ACKNOWLEDGMENTS

We thank C. Desplan, R. Maas, and J. Epstein for communicating results prior to publication, and we thank P. Lepage and M. Dehbi for help with the transactivation experiments.

This work was supported by a grant to P.G. from the Howard Hughes Medical Institute. P.G. is a recipient of a scientist award from the Medical Research Council of Canada. K.J.V. is a Howard Hughes Medical Institute predoctoral fellow.

REFERENCES

1. **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.
2. **Barr, F. G., N. Galili, J. Holick, J. A. Biegel, G. Rovera, and B. S. Emanuel.** 1993. Rearrangement of the *PAX3* paired box gene in the paediatric solid tumour alveolar rhabdomyosarcoma. Nat. Genet. **3:**113–117.
3. **Bober, E., T. Franz, H.-H. Arnold, P. Gruss, and P. Tremblay.** 1994. *Pax-3* is required for the development of limb muscles: a possible role for the migration of dermomyotomal muscle progenitor cells. Development **120:**603–612.
4. **Burri, M., Y. Tromvoukis, D. Bopp, G. Frigerio, and M. Noll.** 1989. Conservation of the paired domain in metazoans and its structure in three isolated human genes. EMBO J. **8:**1183–1190.
5. **Chalepakis, G., and P. Gruss.** 1995. Identification of DNA recognition sequences for the Pax3 paired domain. Gene **162:**267–270.
6. **Chalepakis, G., F. S. Jones, G. M. Edelman, and P. Gruss.** 1994. Pax-3 contains domains for transcription activation and transcription inhibition. Proc. Natl. Acad. Sci. USA **91:**12745–12749.
7. **Chirgwin, J. M., A. A. Przybyla, R. J. MacDonald, and W. J. Rutter.** 1979. Isolation of biologically active ribonucleic acid from sources enriched in ribonuclease. Biochemistry **18:**5294–5306.
8. **Czerny, T., and M. Busslinger.** 1995. 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). Mol. Cell. Biol. **15:**2858–2871.
9. **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.
10. **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.
11. **Epstein, D. J., K. J. Vogan, D. G. Trasler, and P. Gros.** 1993. A mutation within intron 3 of the *Pax-3* gene produces aberrantly spliced mRNA transcripts in the splotch (*Sp*) mouse mutant. Proc. Natl. Acad. Sci. USA **90:** 532–536.
12. **Epstein, J. A., J. Cai, T. Glaser, L. Jepeal, D. S. Walton, and R. L. Maas.** 1994. Identification of a Pax paired domain recognition sequence and evidence for DNA-dependent conformational changes. J. Biol. Chem. **269:** 8355–8361.
13. **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. **17:** 2022–2034.
14. **Epstein, J. A., D. N. Shapiro, J. Cheng, P. Y. P. Lam, and R. L. Maas.** 1996. Pax3 modulates expression of the c-Met receptor during limb muscle development. Proc. Natl. Acad. Sci. USA **93:**4213–4218.
15. **Feng, J., R. C. Johnson, and R. E. Dickerson.** 1994. Hin recombinase bound to DNA: the origin of specificity in major and minor groove interactions. Science **263:**348–355.
16. **Field, J., J. Nikawa, D. Broek, B. MacDonald, L. Rodgers, I. A. Wilson, R. A. Lerner, and M. Wigler.** 1988. Purification of a RAS-responsive adenylyl cyclase from *Saccharomyces cerevisiae* by use of an epitope addition method. Mol. Cell. Biol. **8:**2159–2165.
17. **Galili, N., R. J. Davis, W. J. Fredericks, S. Mukhopadhyay, F. J. Rauscher, 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. Nat. Genet. **5:**230–235.
18. **Goulding, M. D., G. Chalepakis, U. Deutsch, J. Erselius, and P. Gruss.** 1991. Pax-3, a novel murine DNA binding protein expressed during early neurogenesis. EMBO J. **10:**1135–1147.
19. **Hill, R. E., J. Favor, B. L. Hogan, C. C. T. 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.
20. **Lepage, P., D. A. Underhill, and P. Gros.** 1995. Activation of a MMTV/*mdr3* fusion transcript from a cryptic viral promoter is stimulated by *mdr*-derived sequences located in intron 1. Virology **210:**244–253.
21. **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 transactivation properties. Mol. Cell. Biol. **13:**6024–6035.
22. **Marchuk, D., M. Drumm, A. Saulino, and F. S. Collins.** 1991. Construction of T-vectors, a rapid and general system for direct cloning of unmodified PCR products. Nucleic Acids Res. **19:**1154.
23. **Newman, A. J., and C. Norman.** 1992. U5 snRNA interacts with exon sequences at 5′ and 3′ splice sites. Cell **68:**743–754.
24. **Noll, M.** 1993. Evolution and role of Pax genes. Curr. Opin. Genet. Dev. **3:**595–605.
25. **Reich, C. I., R. W. Van Hoy, G. L. Porter, and J. A. Wise.** 1992. Mutations at the 3′ splice site can be suppressed by compensatory base changes in U1 snRNA in fission yeast. Cell **69:**1159–1169.
26. **Sanger, F., S. Nicklen, and A. R. Coulson.** 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA **74:**5463–5467.
27. **Sanyanusin, P., L. A. Schimmenti, L. A. McNoe, T. A. Ward, M. E. M. Pierpont, M. J. Sullivan, W. B. Dobyns, and M. R. Eccles.** 1995. Mutation of the PAX2 gene in a family with optic nerve colobomas, renal anomalies and vesicoureteral reflux. Nat. Genet. **9:**358–364.
28. **Singh, R., J. Valcarcel, and M. R. Green.** 1995. Distinct binding specificities and functions of higher eukaryotic polypyrimidine tract-binding proteins. Science **268:**1173–1176.
29. **Stephens, R. M., and T. D. Schneider.** 1992. Features of spliceosome evolution and function inferred from an analysis of the information at human splice sites. J. Mol. Biol. **228:**1124–1136.
30. **Stuart, E. T., C. Kioussi, and P. Gruss.** 1994. Mammalian *Pax* genes. Annu. Rev. Genet. **28:**219–236.
31. **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.
32. **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 of a paired box- and homeobox-containing gene from the aniridia region. Cell **67:**1059–1074.
33. **Treacy, M. N., L. I. Neilson, E. E. Turner, X. He, and M. G. Rosenfeld.** 1992. Twin of I-POU: a two amino acid difference in the I-POU homeodomain distinguishes an activator from an inhibitor of transcription. Cell **68:**491–505.
34. **Treisman, J., E. Harris, and C. Desplan.** The paired box encodes a second DNA-binding domain in the paired homeo domain protein. 1991. Genes Dev. **5:**594–604.
35. **Underhill, D. A., K. J. Vogan, and P. Gros.** 1995. Analysis of the mouse *Splotch-delayed* mutation indicates that the Pax-3 paired domain can influence homeodomain DNA-binding activity. Proc. Natl. Acad. Sci. USA **92:** 3692–3696.
36. **Vogan, K. J., D. J. Epstein, D. G. Trasler, and P. Gros.** 1993. The *Splotch-delayed* (*Sp^d*) mouse mutant carries a point mutation within the paired box of the *Pax-3* gene. Genomics **17:**364–369.
37. **Wilson, D., G. Sheng, T. Lecuit, N. Dostatni, and C. Desplan.** 1993. Cooperative dimerization of Paired class homeodomains on DNA. Genes Dev. **7:**2120–2134.
38. **Xu, W., M. A. Rould, S. Jun, C. Desplan, and C. O. Pabo.** 1995. Crystal structure of a paired domain-DNA complex at 2.5 A resolution reveals structural basis for Pax developmental mutants. Cell **80:**639–650.
39. **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.