Development 124, 617-626 (1997)
Printed in Great Britain © The Company of Biologists Limited 1997
DEV4756

Positive and negative DNA sequence elements are required to establish the pattern of *Pax3* expression

Thomas A. Natoli, Mary Kay Ellsworth, Chuanzhen Wu, Kenneth W. Gross and Steven C. Pruitt*

Roswell Park Cancer Institute, Department of Molecular and Cellular Biology, Buffalo, NY 14263, USA

*Author for correspondence

SUMMARY

The transcription start site and DNA sequence elements required for the induction of *Pax3* expression in differentiating P19 embryonal carcinoma cells have been localized. These elements consist of a promoter and additional elements located within 1.6 kbp 5′ to the transcription start site. Sequence elements within this 1.6 kbp region are also sufficient to mediate the induction and dorsal restriction of *Pax3* in the neural tube and somites of transgenic mice throughout the hindbrain and trunk. Additional elements required for expression anterior to the hindbrain and in migrating myoblasts are located within 14 kbp 5′ to the transcription start site. This region also contains element(s) that repress *Pax3* expression in the ventral body wall mesoderm of the tail bud.

Key words: *Pax3*, P19 embryonal carcinoma, transgenic mouse, DNA sequence elements, enhancer, transcriptional regulatory elements, dorsoventral patterning

INTRODUCTION

Pax3 is one of nine members of a family of transcription factors that have homology to the *Drosophila paired* gene (reviewed in Stuart et al., 1993). Each of the Pax genes exhibits spatial and temporal regulation during development. In both the mouse and in humans, mutations in several of these genes have been shown to disrupt the normal development of the tissues in which they are expressed.

Pax3 is expressed in several different lineages during embryonic development. In early somite stage mouse embryos, at approximately embryonic day 8.5 (E8.5), expression is found in the dorsal margin of the neural plate (Goulding et al., 1991). Following neural tube closure, *Pax3* expression is maintained in the dorsal half of the neural tube (Goulding et al., 1991). Shortly after somitogenesis, *Pax3* expression is seen throughout the somite and this expression is subsequently limited to the dorsolateral surface (Goulding et al., 1991). Beginning at approximately E9.0, *Pax3* expression is enhanced in the lateral-most portion of the dermomyotome in a population of cells that migrate into the limb bud and body wall to form the limb and hypaxial musculature (Bober et al., 1994; Goulding et al., 1994; Williams and Ordahl, 1994).

Many of the lineages which express *Pax3* during embryonic development give rise to abnormal tissues in both mutant mouse lines and humans carrying mutations in the *Pax3* gene (reviewed in Stuart et al., 1993). Mutations of *Pax3* have been found in both the mouse *Splotch* mutant (e.g. Epstein et al., 1991) and Waardenburg's syndrome in humans (Tassabehji et al., 1992; Baldwin et al., 1992). In each case the phenotype is semi-dominant. *Splotch* heterozygotes show defects in neural crest derivatives which result in white spotting. In the homozygous condition the mutation is lethal by E13.5 and embryos display defects in neural tissues (exencephaly and spina bifida), neural crest derivatives (deficient melanocyte differentiation and spinal ganglia formation) and the limb musculature. In humans, Waardenburg's syndrome type I is characterized by facial malformations (broad nasal root and dystopia canthorum) and pigmentation defects (white forelock, partial albinism and heterochromia of the eye). Waardenberg's syndrome type III shows the same defects plus a deficiency in the limb musculature. The expression and function of *Pax3* during the development of the nervous system, neural crest and musculature have made it a useful marker for a variety of studies in each of these lineages.

P19 embryonal carcinoma (EC) cells can be induced to differentiate into neural and skeletal muscle lineages both in vivo and in vitro. Undifferentiated P19 stem cells can be maintained in monolayer cultures indefinitely. Growth as aggregates in the presence of 10^{-6} M retinoic acid (RA) results in differentiation towards neuroectodermal lineages, whereas growth in the presence of 1% dimethyl sulfoxide (DMSO) results in differentiation towards endodermal and mesodermal lineages (including skeletal myoblasts; McBurney et al., 1982). *Pax3* expression is also induced following either RA or DMSO treatment of P19 cells (Goulding et al., 1991; Pruitt, 1992, 1994). Since P19 cells express *Pax3* during differentiation towards the neural and skeletal muscle lineages, they may provide a useful model of normal *Pax3* regulation.

In the present study, P19 EC cells have been used to narrow DNA sequence elements required for *Pax3* expression to a region between −1.6 kbp and the promoter. When introduced into transgenic mice, this sequence is sufficient to mediate expression of a reporter gene in the dorsal neural tube and

somites. However, this analysis also reveals that additional elements are present in a larger construct carrying 14 kbp of DNA 5′ to the promoter that are required for reporter gene expression in migrating myoblasts and in neural derivatives anterior to the hindbrain. Additionally, this region contains one or more negative elements required for suppression of *Pax3* expression within the ventral portion of the tail bud.

MATERIALS AND METHODS

Cloning mouse *Pax3* promoter sequences

To clone the mouse *Pax3* promoter, oligonucleotide primers were designed to PCR amplify exon 1 of the mouse *Pax3* gene (5′-GTGAGCACCTTTGCCAGTAG-3′ and 5′-AGCTTCGCTCGCAAATTATA-3′). These primers were utilized to screen a mouse strain 129/SV P1 genomic library (Genome Systems, Inc.) and a single positive clone was obtained. Subfragments containing exon 1 and the 5′ flanking sequences were identified by Southern blot analysis and subcloned into pBluescript (Stratagene) using standard cloning techniques (Ausubel et al., 1994). Plasmids were sequenced using dideoxynucleotide incorporation and Sequenase v. 2.0 (USB) or dye termination using an ABI 373A automated sequencer, following the manufacturer's instructions. Sequence analysis was performed using the UWGCG package. The plasmid pPolyCAT was created by replacing the *Hin*dIII site of pCATBasic (Promega) with a *Not*I site and inserting the *Not*I-*Sal*I polylinker fragment of pBluescript KS+. The plasmid pPolyLac was created by replacing the *Hin*dIII site of pCH110 (Pharmacia) with a *Sal*I site, and inserting the *Sal*I-*Bam*HI *lacZ* fragment into pPolyCAT. Promoter fragments from the *Pax3* gene were inserted into these plasmids for functional analysis.

S1 nuclease protection and primer extension analyses

S1 nuclease protection analysis was performed essentially as described previously (Pruitt and Reeder, 1984a). Probes were generated by 5′ end labeling *Kpn*I (–971) to *Eco*RI (–494), *Eco*RI (–493) to *Bam*HI (–8) and *Hin*dIII (–46) to *Sma*I (+414) fragments and strand separating prior to annealing to 50 μg of total RNA per reaction.

Primer extension analysis was performed essentially as described previously (Pruitt and Reeder, 1984b). Probe was generated by 5′ end labeling an *Afl*III (+34) to *Tha*I (+75) fragment and strand separating prior to annealing to 50 μg of total RNA per reaction.

Cell culture and transfection

P19S18O1A1 cells (P19 cells; McBurney et al., 1982) were cultured and induced to differentiate as described previously (Pruitt and Natoli, 1992). Pools of mixed stable lines were derived as follows: test constructs were introduced into P19 cells in a 9:1 ratio with pSVNeoXh by calcium phosphate precipitation (Pruitt, 1994). Cells were selected in the presence of 400 μg/ml G418 for 10-14 days and colonies were counted. G418-resistant colonies were treated with 0.25% trypsin in phosphate-buffered saline (PBS) and pooled. Each pool represented 500-1500 independent transformants.

Northern and Southern blot analyses

Total RNA isolation and northern blot analysis was performed as described (Pruitt and Natoli, 1992). The probes used were: *Pax3*, a 200 bp *Cla*I to *Hin*dIII fragment derived from pBSPax3T7; CAT, a 1.4 kbp *Eco*RI fragment derived from pPolyCAT; *lacZ*, a 3.1 kbp *Sal*I to *Eco*RI fragment derived from pPolyLac; and TPI (used as an internal control for loading differences), a 1.7 kbp *Bam*HI fragment derived from pcDmTPI (Cheng et al., 1990). Genomic DNA was isolated (Ausubel et al., 1994), digested with *Eco*RI and electrophoresed on 1% agarose gels. Following electrophoresis, gels were denatured in 0.5 M NaOH, 1.5 M NaCl for 60 minutes and neutralized in 0.5 M Tris-acetate, pH 7.4, 3 M NaCl for 60 minutes before transfer and hybridization as for northern blot analysis.

CAT assays

CAT assays were performed on mixed stable pools as described in Ausubel et al. (1994) and CAT activities were normalized for both transfection efficiency (based on the relative copy number of the construct carried in the mixed stable pool as determined by Southern blot analysis) and endogenous *Pax3* expression (based on *Pax3* transcript levels present in parallel cultures for each condition as determined by northern blot analysis).

Transgenic mouse production and analysis

Plasmids for pronuclear injection were digested with *Not*I and *Apa*I, electrophoresed on 1% low melting point agarose gels (SeaKem) and isolated using Gelase (Epicentre Technologies), following the manufacturers' protocols. DNA was introduced by pronuclear injection into fertilized eggs derived from a (C3Hf/HeRos×C57BL/10Ros-p^{d}) mating as described (Hogan et al., 1986). Southern blot analysis was used to screen tail DNAs for the presence of *lacZ* sequences. For analysis of embryos, males carrying the transgene were bred to BCF females. Noon on the day that the copulation plug was detected was defined as E0.5. Embryos were isolated and fixed in 2.5% glutaraldehyde for 30 (E7.5-E9.5), 60 (E10.5 and E11.5) or 90 minutes (E12.5 and older). Embryos were then rinsed three times in PBS containing 0.1% NP40, and β-galactosidase activity was detected as previously (Pruitt, 1992). Reactions were allowed to continue at room temperature from 30 minutes to several hours, or overnight at 4°C. Reactions were stopped by rinsing with PBS. Embryos were stored in PBS for photography. Some embryos were cleared as described by Marti et al. (1995). Embryos were staged according to Kaufman (1992).

RESULTS

Localization of the *Pax3* transcription start site and sequence analysis of the promoter region

Genomic DNA containing the *Pax3* promoter was cloned as described in Materials and Methods. Subclones encompassing exon 1 and the 5′ flanking region were isolated and sequenced (Fig. 1). Previous studies (Goulding et al., 1991) have demonstrated that *Pax3* transcripts of 3.3 and 3.6 kb are expressed in both differentiating P19 EC cells and E9 through E15 embryos. However, the longest previously isolated *Pax3* cDNA is only 2347 bp (Goulding et al., 1991). Since this clone contains a consensus polyadenylation signal, as much as 1 kbp of 5′ untranslated sequence could be missing from the cDNA clone. To attempt to localize the transcription start site, sequences from approximately 1 kbp 5′ to the cDNA sequence (position –971 in Fig. 1) to within the previously cloned cDNA sequence (position +414 in Fig. 1) were scanned using S1 nuclease protection and a series of three probes (Fig. 2A). Results from this study demonstrate that RNAs from RA treated, but not uninduced, P19 EC cell cultures give partial protection of only the 3′-most probe where the size of the protected fragment maps to approximately 70+/–5 nucleotides 5′ to the previously isolated cDNA (+5 to –5 in Fig. 1).

To confirm the presence of the 5′ end of the *Pax3* transcript at this location, primer extension analysis was performed (Fig. 2B). The primer used was complementary to sequences between positions +34 and +75 (Fig. 1) of the transcribed DNA strand and gives extension products of 74, 75 and 77 nt on

RNAs from either RA- or DMSO-induced P19 EC cell cultures. These extension products map to locations within the region identified by S1 nuclease protection analysis, where the position of the strongest primer extension product of 75 nt is indicated as +1 in Fig. 1.

The start site indicated by S1 nuclease protection and primer extension analyses is approximately 900 nt short of the location suggested by comparison of the lengths of the *Pax3* transcripts with that of the cDNA clone. To resolve this discrepancy, the transcript size of the *Pax3/lacZ* fusion carried by the EH3 cell line was determined by northern blot analysis using a *lacZ* probe (Fig. 2C). In this line *lacZ* sequences are fused to the *Pax3* gene by integration within the fourth intron. The size of the fusion gene transcript is approximately 4.1 kb, consistent with the transcription start site predicted by S1 nuclease protection and primer extension analyses. Together, these three lines of evidence demonstrate that the transcription start site of *Pax3* is located only 70 nt 5′ to the 5′ end of the previously identified cDNA and the additional, approximately 900 nt, discrepancy between the cDNA and transcript sizes must result from differences located 3′ to intron four.

Analyses of the sequences 5′ to the transcription start site identify two consensus CAAT boxes, a consensus TATA box and a consensus cap site in the flanking region (Bucher, 1990). The sequences of approximately 750 nt of 5′ flanking region have been reported for the human *Pax3* gene (Macina et al., 1995). Comparison of the mouse and human sequences demonstrates homology of 92% between position +15 and −125 of Fig. 1, which includes the consensus CAAT, TATA and cap sites. The level of homology decreases to 75% over the next 100 nt (−225 to −126) and is less than 50% for more distal sequences. These observations are consistent with the presence of a promoter, which initiates transcription at the site indicated as +1 in Fig. 1.

Elements necessary for *Pax3* regulation during P19 cell differentiation are located within 1.6 kbp of the transcription start site

As a starting point for deletion analysis of the *Pax3* regulatory sequences, 14 kbp of *Pax3* 5′ flanking sequence was fused to a CAT reporter gene, stably transfected into P19 EC cells, and a pool of approximately 1,000 transfectants was assayed for CAT activity at various times following chemically induced differentiation. CAT activity was not detected in uninduced stem cells but was detected within 4 days of RA treatment and 6 days of DMSO treatment (data not shown). Similar kinetics are seen for β-galactosidase induction using a P19 EC cell derivative (EH3) in which the *lacZ* gene is integrated into the *Pax3* locus (Pruitt, 1992). Therefore, elements sufficient for *Pax3* activation in P19 cells are located within the 14 kb 5′ of the *Pax3* gene.

To localize elements required for *Pax3* expression during P19 cell differentiation, 5′ deletions of the 14 kbp CAT construct were created, transfected into P19 cells and pools of stably transfected cells were tested for CAT activity in

```
      SstI
-1578 GAGCTCTAATGCTCCTCCCCAAATGTGGGTGGGATCGGGGGTGGGGTGTCCAGACATTTTCTGGAGTCCC

-1508 CAGAGGCCCGTCTAGCCTGCTCCCAGGGTCGATAAGCCCTTTTGACTTCAGGGGGAGCAAAGAGCTACCC
                                    HaeIII
-1438 AGCTTTCTCCACTGGGAGTGAACTCTTAAAGGGCCAGCTCAGGGACTTTGAACAACTACACGCACGCCCC

-1368 CCCCCCTTGAGTTTGCCTTAATTAGGGGGTGGGCAAGTAGGATTGGGGGCCTGAGTTCACATCAAAACTG

-1298 TCAAAGGAGATGGGGGGGGAGGACATGGAGAGAGGGGTGGAGGATTCAGGGGCAGGAAAAATTCAAATAT

-1228 CATAAAAAGAAAAGGGAACTTGAAGCCCAGAAGTCATTGTTTCCTTTATTTCAAAGGGGAAAAAACCTGC
                                  HaeIII
-1158 CTGATTCTCATCCTTTTGATTGATTAAGGCCCTGAATACCTCTCGGCCCCAGAGTGACAGTCCCTGAATA

-1088 GACTTGAGGGAATAAAGAGCAGCGCAGAACGTGGGGCTGGCACGGGAGGGGGGTGTAAGGGAGGAGTGTT
                                             KpnI
-1018 CGCTGGCACTTACCAAGTTATAAATAAAAGGCTAGGCACAATGGTACCTTCTCTAAGGACAGACAGTCTT

 -948 TACAACACTCCTGGCGTCATATCCTGCTGGGGACACTTCAGCTCCTAGCCAAGACGTTGCTTCTTTTATT
                                                                SmaI
 -878 TTTCCAGCAGTTTAGTCTGAATGCCATAATAAATTCCTGAGAACAAACGCTGCACCCGGGCAAAACCTCA

 -808 ACATATAGATGCAAGTGCATCGGGGATGAATGTGTACGTGGAGATTTAAAGTCCCCGCTTCTCAGAAGGG
                                                                     StuI
 -738 TTTTTAAACCAAGAAAAGGAGCTGGCTTCCCCATTTCAGCTTGCAACTTGGAGCCCAGGGGAGGCCTAAC

 -668 CTCTTCAGTCTCTGTCATGGCTATGCCACCTGTCTTTTCATTCCCAGCCTAAGAAAGACAAAAGGACCTT

 -598 CTGGGCACATCTTGGACTCTTCCTGAGAGATGCCTTTCTGGGAGGCTCACATAGGGAGCTCAGGCCCAGA
                                          EcoRI
 -528 GGCAGACCAGAGAAAGGAAAAGGAAACACATCCGGAATTCCTCTGTGCCCCTTTCCAAATTTGCTCCACC

 -458 GCCATCTCCGTATTAGTAATCCCGAAGCGAAATCCACAGGTGAAAAGCAGGGGCAGACAGAGGAAATAAG

 -388 GGGCGGATATAGCAAGGTTTCCAGAGAAAGGCGGAAAGAGATTAGGACAGAGTAATAGAAAACCAAGATG

 -318 ATCAAGGCTGAATCTCATGTAGAGGGACAGGGACACCGTGAGCCTTTTTCCATCCTCACTGGCTCTTGAC

 -248 ACCCAAGCTGCCTCTCCTAGTTTCAACCTGTCCACCCTTCTCTTGAGAAAGGGACACAGAGCTAGTCCCC
                 StuI
 -178 CTCTGGGTCGGCGGAGGCCTGGGGACCGTCAGGGATGGGAAGAGAAGTAGCTTCT**GTCGCCCAATCA**GCG
                                                                HindIII
 -108 CGTGTCTTTGCCACCCGGGACGGTCTCCTC**CTCGGCCAATCG**CAGCTCAGGGCTCCTGATCAAGCTTTGG
                                    -5  ↓ ↓↓ +5
  -38 GTGAAAGAAC**TAATAAATGCTCCCT**AGTCCGGATCCCTGCACTCGGTGTCACGACGGGAGGAGACTTGGG
      AflIII                                ThaI
  +33 ACGTGTTCCTGCCTCGTCACCACCCTCCCTCACCCGATCTTCGCGGTGCAGTGAGCACCTTTGCCAGTAG

 +103 CCCCAGCTGAGGGGCCGATCTGCTAGACTCGCACCAAACTCGTCCGCCCTGGGTTGGGATCCTGCACTCA

 +173 AGGGACTCCTCTGGAGCCTGTGGACTTGGATCTATTAGCGCCCCCCTCCCCGCTGCCCCGCCTCCCTCTC

 +243 TGGCCTTTTTTGGGGGAGGAGCTCTCCGAGATCCGGAGAGTTCCCGAGGGTCACCCGCCGCACTGTGCTC

 +313 GCTTTTTCGTCTCGCCTTCACCTGGATATAATTTGCGAGCGAAGCTGCCCCCAGGATGACCACGCTGGCC
                                    SmaI                          MetThrThrLeuAla
 +383 GGCGCTGTGCCCAGGATGATGCGGCCCGGCCCGGGGCAG +421
      GlyAlaValProArgMetMetArgProGlyProGlyGln
```

Fig. 1. *Pax3* transcription start site and promoter sequence elements. The nucleotide sequence of the *Pax3* promoter is shown (GenBank accession number U61230). Restriction endonuclease sites used to generate internal deletions, or in S1 nuclease protection and primer extension analyses, are indicated with a bar above the sequence. The putative CAAT boxes, TATA box and CAP site are indicated by the bold underlined, bold double-underlined and bold italicized sequences, respectively. Three transcription start sites detected by primer extension analysis are indicated by arrows. The region containing the transcription start site identified by the lower resolution S1 nuclease protection analysis is indicated by overlining between positions −5 and +5.

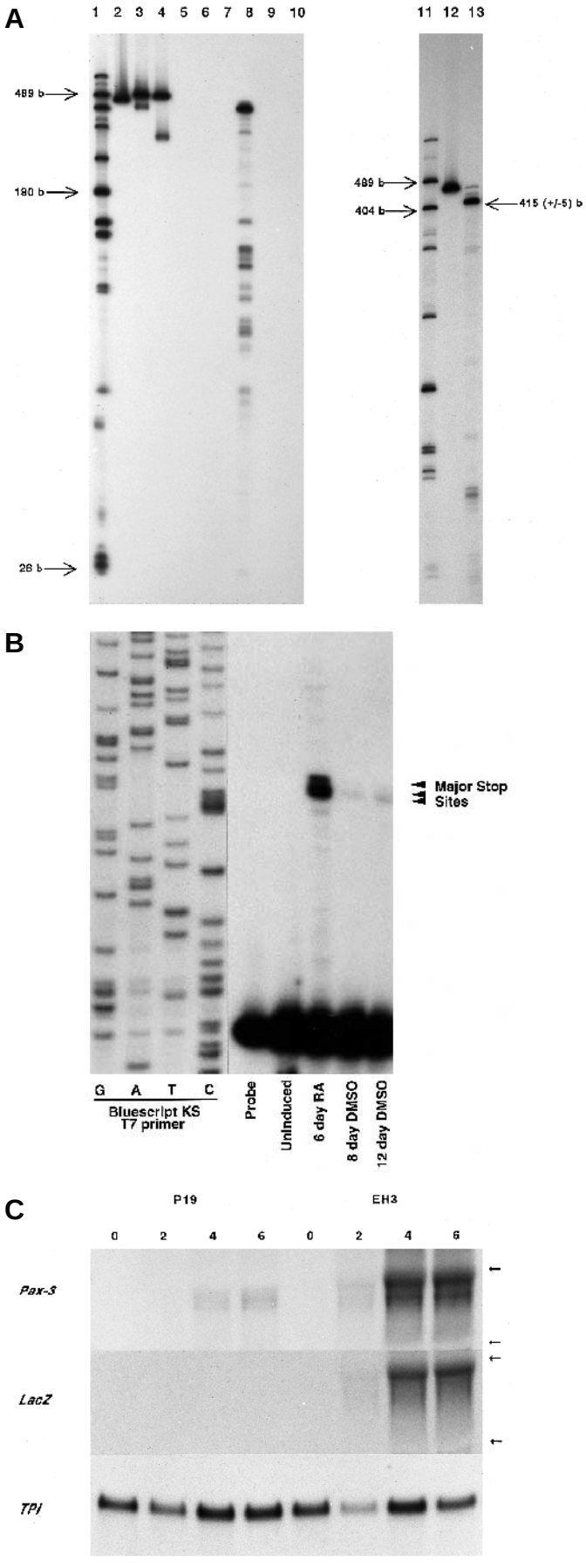

Fig. 2. Localization of the *Pax3* transcription start site. (A) S1 nuclease protection analysis. Three probe fragments spanning the regions between positions −46 to +414 (460 nt, lanes 2, 5, 8, 12 and 13), −493 to −8 (485 nt, lanes 3, 6 and 9) and −971 to −494 (477 nt, lanes 4, 7 and 10) electrophoresed on a 6% denaturing polyacrylamide gel following either no S1 nuclease treatment (lanes 2, 3, 4 and 12), hybridization with 50 μg of RNA derived from monolayer P19 EC cells plus S1 nuclease treatment (lanes 5, 6 and 7) or hybridization with 50 μg of RNA derived from RA-treated aggregate P19 EC cell cultures plus S1 nuclease treatment (lanes 8, 9, 10 and 13). Lanes 1 and 11 are *Hpa*II-treated pBluescript SK−, which were included as MW markers.(B) Primer extension analysis. A probe fragment from positions +34 to +75 of Fig. 1 was hybridized to 50 μg of RNA derived from uninduced, RA-treated, or DMSO-treated P19 EC cell cultures as indicated in the figure, extended using reverse transcriptase and electrophoresed on a 6% denaturing polyacrylamide gel. A sequencing ladder derived from extension of the T7 primer on pBluescript KS+ is included as a molecular mass marker. Arrows indicate extension products of 74, 75 and 77 nucleotides. (C) Northern blot analysis of *Pax3/lacZ* fusion transcripts in EH3 cells. RNAs derived from either monolayer (0), or 2-, 4- or 6-day RA-treated P19 and EH3 cell cultures were probed for the presence of *Pax3* or *lacZ* transcripts as indicated. The *lacZ* probe detects a single major transcript of approximately 4.1 kb, of which 3.1 kb is expected to be derived from the *lacZ* portion and the remainder from *Pax3* sequences 5′ to intron 4. Positions of the 4.7 and 1.9 nt rRNAs are indicated by arrows. Blots were probed for TPI (1.4 kb) to control for loading differences.

monolayer cultures (Fig. 3A), or 8 days after RA (Fig. 3B) or DMSO (data not shown) treatment. Two controls were performed for each assay. First, the average copy number of the reporter construct was compared between pools of stable transfectants carrying the different constructs, by assaying for the presence of the CAT gene (Fig. 3C), and found to vary by a factor of approximately 3-5. Second, the efficiency of differentiation towards lineages that express *Pax3* was assayed using cultures seeded in parallel to those used for CAT assays, 8 days following treatment with either RA or DMSO, by assaying for the level of endogenous *Pax3* expression using northern blot analysis (data not shown). Endogenous *Pax3* expression was found to vary by a factor of approximately 3-10 in most experiments although, in the case of DMSO-treated cultures, differences of as great as a factor of 20 were observed in one experiment. Reassays of several of these mixed stable populations suggested that the variability was a result of the specific induction conditions and not a property of the transfected cells since at higher cell densities, endogenous *Pax3* expression and CAT activities increased in parallel. Further, these results are consistent with previous observations using the EH3 cell line where it has been found that differences in seeding density of as little as twofold can affect the kinetics of *Pax3* induction for both RA- and DMSO-treated cultures and, in the case of DMSO, the proportion of cells in which *Pax3* is expressed.

Quantitation of the CAT activities resulting from either RA or DMSO induction of transfected P19 EC cell pools carrying a series of 5′ deletion constructs is shown in Fig. 4, following normalization for both the average copy number of the reporter construct and the level of endogenous *Pax3* expression in cultures seeded in parallel. Since a CAT signal could not be detected in monolayer cultures, activities are expressed relative to the −4.1 kbp construct. (The percentage CAT conversion to

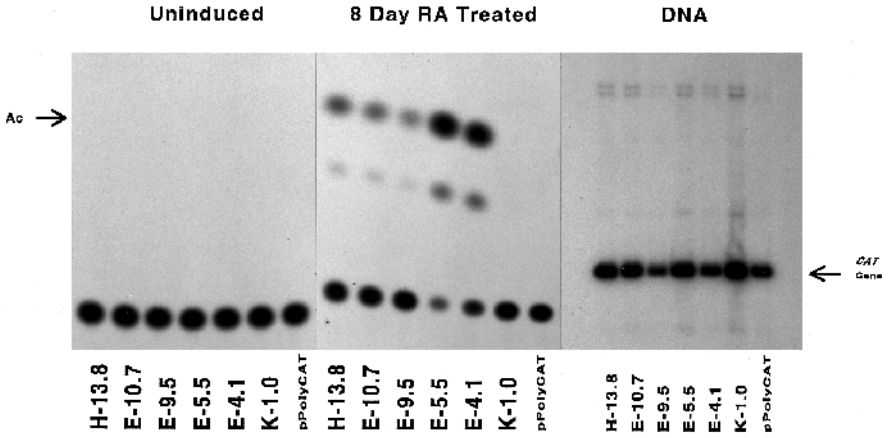

Fig. 3. Analysis of 5′ *Pax3* promoter deletions in the P19 EC cells. CAT activity was measured for pools of transfected P19 EC cells carrying several 5′ deletion constructs as indicated, or the reporter construct in the absence of *Pax3* sequences (pPolyCAT). Assays for CAT activities present in uninduced monolayer cultures (left) or RA-treated cultures 8 days following treatment (center), are shown. The average level of reporter construct present in DNAs derived from each of the mixed stable pools was determined following *Eco*RI digestion and probing for the presence of the CAT-containing portion of each construct (right panel).

the acetylated form was 64% for the −4.1 reference construct following RA treatment. This is over 500× the minimal detectable conversion of 0.1% in these assays.) All constructs containing more than 1.6 kbp of *Pax3* 5′ flanking sequence showed expression following RA or DMSO treatment, whereas constructs containing 1.0 kbp or less of *Pax3* 5′ flanking sequence showed no expression. The level of induction following RA treatment was routinely approximately fivefold higher than the corresponding DMSO induction, consistent with the higher levels of *Pax3* mRNA expression following RA treatment. This difference is not reflected in the relative activities shown in Fig. 4, since the results have been normalized for endogenous *Pax3* expression. These results demonstrate that, although DNA sequences located 5′ to position −1.6 may modulate the level of expression, and these effects may differ between the different cell lineages formed following RA or DMSO treatment, the minimum sequence elements necessary for *Pax3* expression under either induction condition are located between −1.6 kbp and the promoter.

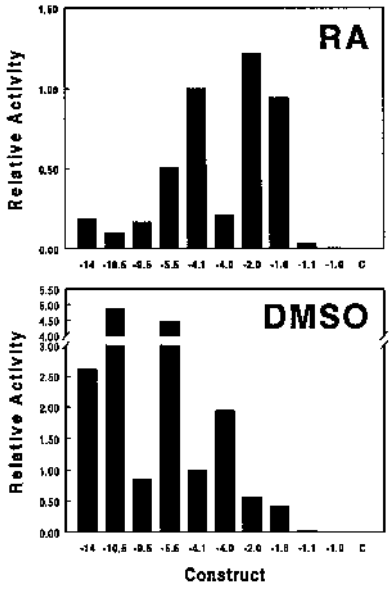

Fig. 4. Quantitation of expression from 5′ *Pax3* promoter deletions in RA or DMSO treated P19 EC cells. Parallel pools of stably transformed cells carrying 5′ deletion constructs indicated by the position of the 5′ truncation were aggregated in the presence of RA (upper panel) or DMSO (lower panel) for 8 days and harvested for CAT assays and northern blot analysis. CAT activity was normalized for both endogenous *Pax3* expression and transfection efficiency. Relative activity was determined using the −4.1 construct as an internal reference where expression from this construct is defined as 1.0.

Deletion analysis within the 1.6 kbp domain demonstrates the presence of at least two discrete elements required for *Pax3* expression in P19 EC cells

To more precisely define the sequence elements within the −1.6 kbp construct that are required for *Pax3* expression, internal deletions were created in this sequence and assayed for expression in P19 EC cells (Fig. 5). Expression was normalized as previously except that the −1.6 kbp 5′ deletion was used as the reference. Two regions were found to be required for

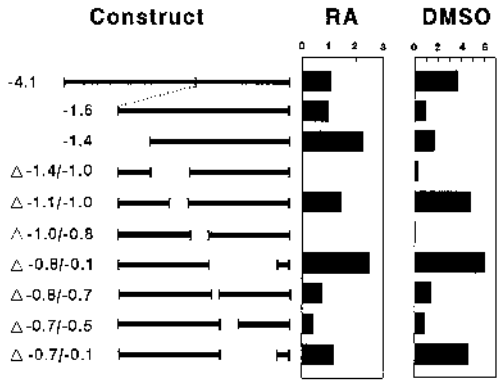

Fig. 5. Quantitation of expression from internal deletions within the −1.6 kbp *Pax3* promoter construct in RA- or DMSO-treated P19 EC cells. P19 EC cells carrying the indicated *Pax3* promoter/CAT fusion constructs were assayed as described in Fig. 4, except that the data are expressed relative to the −1.6 construct, which was included as a reference in all experiments. Expression from this construct is defined as a value of 1.0 for both RA and DMSO treatments. For several key constructs results represent averaged data from either three (−1.6, −1.4/−1.0, −1.1/−1.0) or two (−1.0/−0.8, −0.8/−0.1) mixed stable pools resulting from independent transfections. Locations of each deletion are shown schematically (left).

expression following either RA or DMSO treatment: one located between nucleotides −1405 and −1130, the other between nucleotides −971 and −822.

The 1.6 kbp region contains sequence elements sufficient for correct induction of *Pax3* in vivo and dorsal restriction between the hindbrain and node

To determine if the region identified as sufficient for *Pax-3* expression in P19 EC cells is active in vivo, transgenic mice were created using constructs in which either the 1.6 kbp or 14 kbp promoter fragments drive *lacZ* expression. Progeny from four independently derived founder animals carrying the 1.6 kbp construct [lines TgN(1.6P3Z)II1, TgN(1.6P3Z)II2, TgN(1.6P3Z)II3 and TgN(1.6P3Z)II5] and a single founder animal carrying the 14 kbp construct [line TgN(14P3Z)I1] were examined. Three of the four lines carrying the 1.6 kbp construct gave similar expression patterns when assayed on days E9.5 or E12.5. Line TgN(1.6P3Z)II3 showed the same expression pattern, but in addition expressed β-galactosidase in a lateral trunk domain in E12.5 embryos. Lines TgN(14P3Z)I1 and TgN(1.6P3Z)II1 were chosen for more extensive analysis.

In early somite stage embryos (approximately E8.5), *Pax3* is expressed in the dorsal neural tube and dermomyotome compartment of the somites (Goulding et al., 1991). *LacZ* expression is seen in these regions in mice carrying either the 1.6 kbp construct (Fig. 6G-I,L) or the 14 kbp construct (Fig. 6J,K), although the expression seen from the 1.6 kbp construct was stronger. This observation demonstrates that sequence elements sufficient for both tissue-specific expression and dorsal restriction are present within the −1.6 kbp construct.

Expression from this line was also characterized at earlier stages of development (Fig. 6A-F). The earliest detectable expression occurs in a patch of cells near the base of the allantois in neural plate stage embryos (Fig. 6A). By mid- to late-headfold stages, this expression extends along the neural plate-surface ectoderm boundaries in two continuous lines, which begin at the hindbrain and meet at the base of the allantois (Fig. 6B-F). Expression along this line appears to

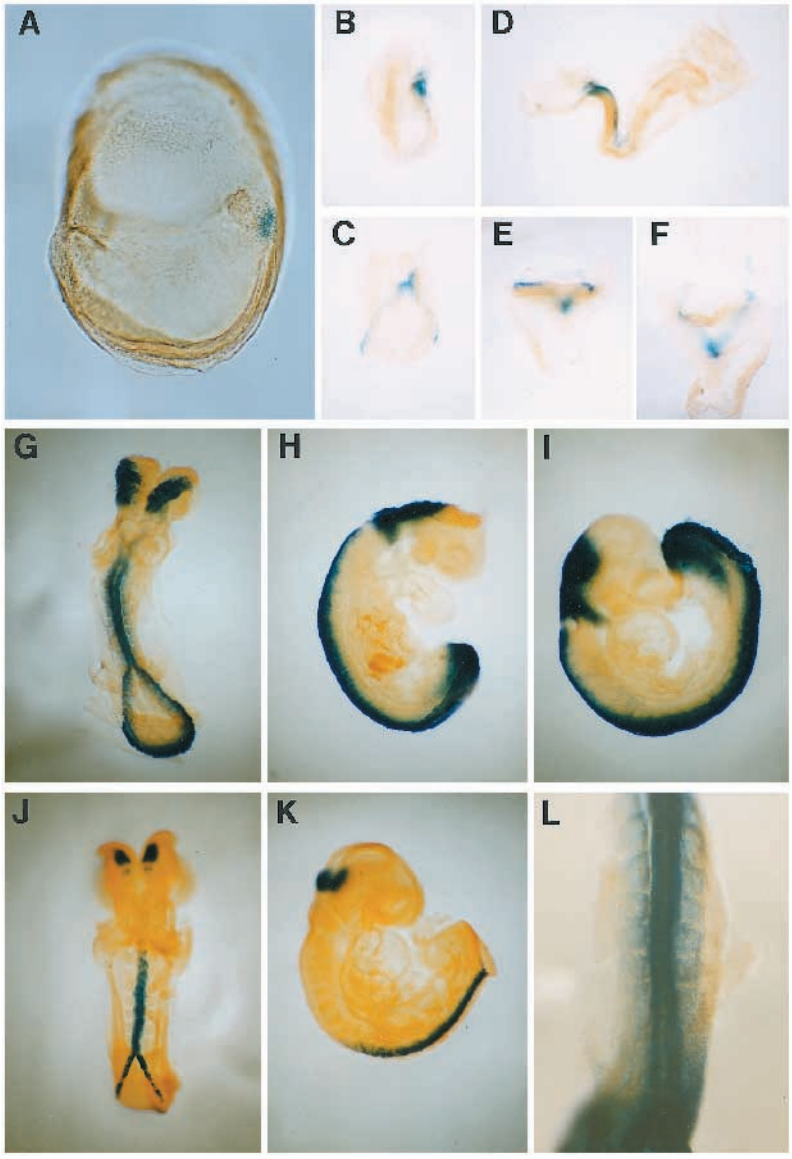

Fig. 6. Initial activation of *Pax3* expression in transgenic mice. Transgenic mice carrying *Pax3* promoter/*lacZ* fusion constructs were isolated at E7.5 (A-F) or E8.5 (G-L) and stained for β-galactosidase activity. (A-I and L) Embryos carrying the 1.6 kbp *Pax3* promoter/*lacZ* fusion construct [line TgN(1.6P3Z)II1]. (A) Neural plate stage embryo (side view, anterior to the left). Staining in this embryo is limited to the base of the allantois. (B) (Side view, anterior to left) and (C) (posterior view), early headfold stage embryo. Staining in this embryo is localized to the neural plate/surface ectoderm boundary in two continuous lines, which begin at the lateral margins of the prospective hindbrain and meet at the base of the allantois. (D) Late headfold/early somite stage embryo (side view, anterior to right). (E,F) Posterior half of the embryo shown in D following dissection (E, posterior view, dorsal to top; F, dorsal view, anterior to top). Lines of expression can be seen to parallel the neural plate/surface ectoderm boundaries, to the neural plate side, in E and to extend beyond the most posterior position of the morphologically discrete neural plate in F. (G) An 8-10 somite stage embryo (dorsal view); staining is seen in the hindbrain, dorsal neural tube and somites. An expanded region of expression at the caudal tip of the embryo is starting to become detectable. (H,I) Progressively more mature somite stage embryos (side views). Expression is present in the hindbrain, dorsal neural tube and somites. Expression in neural crest migrating into the second branchial arch can be seen, as well as expression in the expanded caudal domain. (J,K) Embryos carrying the 14 kbp *Pax3* promoter/*lacZ* fusion construct [line TgN(14P3Z)I1 at stages similar to those in G and H. In J, expression is present in the hindbrain, dorsal neural tube and somites but is reduced relative to that seen for the 1.6 kbp construct in G. The difference in caudal expression between the two constructs is readily observed in the slightly later stage embryos shown in H and K. (L) A dorsal view of an embryo at approximately the stage as shown in I, to highlight expression in the neural tube and somites. Magnification, 200× (A), 40× (G-K) and 100× (L).

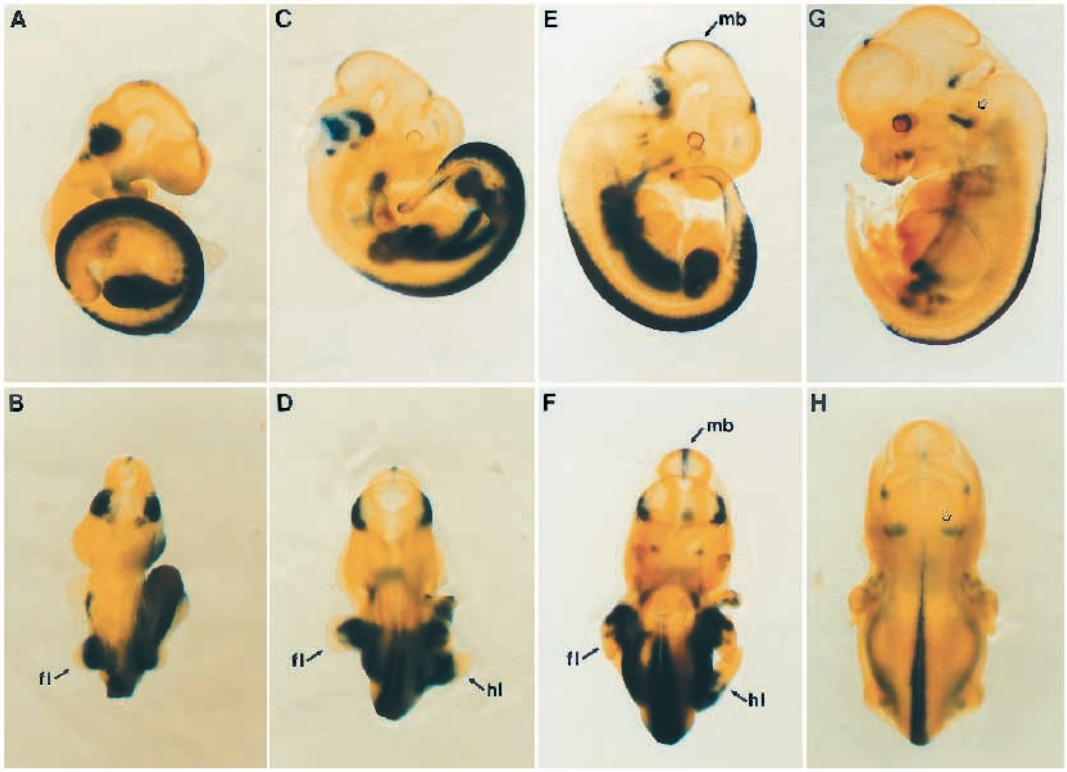

Fig. 7. In vivo expression from a *Pax3* construct containing 14 kbp of 5′ sequence. Transgenic mice carrying the 14 kbp *Pax3* promoter/*lacZ* fusion construct were isolated and stained for β-galactosidase activity. Embryos were isolated at E9.5 (A and B), E10.5 (C and D), E11.5 (E and F) or E12.5 (G and H). A, C, E and G: side views. B, D, F and H: dorsal views. Abbreviations: fl, forelimb; hl, hindlimb; mb, midbrain. The open arrowhead indicates staining in the cochlea. Magnification, 50× (A,B), 30× (C,D), 28× (E,F) and 20× (G,H).

precede the morphological differentiation of the neural plate in the most posterior regions of the embryo.

Transgene expression in midgestation embryos reveals additional sequence elements required for correct *Pax3* expression

Although the domains in which both the 1.6 and 14 kbp *Pax3* promoter constructs are expressed are initially similar, differences in the expression from these two constructs are seen beginning at E9.5. The 14 kbp construct efficiently mimics the expected pattern of expression for the endogenous *Pax3* gene (Goulding et al., 1991 and Fig. 7). Expression is limited to the roof and alar plate of the closed neural tube. Down-regulation at the most rostral level of the hindbrain and within rhombomere 5 is also observed. Expression within the midbrain is first observed at E9.5. In the tail region, expression occurs around the dorsal margin of the open neural plate. Expression from the 14 kbp *Pax3* promoter construct also correctly mimics expression from the endogenous *Pax3* gene in migratory muscle precursors of both the developing limbs and body wall. Expression here is first detected at E9.5 in the ventrolateral margin of each somite (Fig. 7A,B). The signal increases between E9.5 through E11.5, and is found in progressively more distal locations within the limb (Fig. 7A-F), but decreases at later times (Fig. 7G,H). Segregation of cells expressing the −14 kbp construct into dorsal and ventral domains within the limb is detected by E10.5-E11.5 (Fig. 7C-F). Expression from this construct is also detected in the cochlea and craniofacial neural crest derivatives between E11.5 and E12.5.

A similar survey was performed using line TgN(1.6P3Z)II1, which carries the 1.6 kbp construct (Fig. 8). As for the longer construct, expression within neural tissues is confined to the dorsal region of the closed neural tube and mimics the expression seen for the endogenous *Pax3* gene, with three exceptions. First, this construct fails to support expression within the midbrain (compare Figs 8E and 7E). Second, the domain of expression within the open neural plate region of the tail is expanded both within the neural plate itself and into the surrounding ventral mesoderm (compare Figs 8A and 7A). Third, expression in the cochlea is not detected (although expression in the pinnae is detected in all of the transgenic lines carrying this construct).

Expression of the short construct within the dorsal somite appears normal (see Fig. 6L). However, no expression is detected in the population of migrating myoblasts present in the ventrolateral region of the somites (compare Figs 7 and 8 and 9A,B). Additionally, no signal is detectable at any time in the forelimbs of these embryos. Signal is detected in the posterior portion of the hindlimbs by day E10.5 (Fig. 8C,D). However, this expression appears to be a continuation of the expanded expression in the ventral mesoderm of the tailbud that is seen in younger embryos. Within the hindlimb, expression from this construct appears to occur in a different population of cells from those that express the 14 kbp construct (compare Figs 8C-H and 7C-H).

DISCUSSION

Elements sufficient for the initial induction of *Pax3* both in vitro and in vivo localize to within 1.6 kbp of the promoter

The initial localization of DNA sequence elements required for *Pax3* expression was performed using RA- or DMSO-induced P19 EC cells. It has been well established that, under these conditions, P19 cells differentiate into neurons and skeletal

muscle, two lineages known to express *Pax3* in vivo (e.g. McBurney et al., 1982) and that P19 cells induced towards these lineages express *Pax3* (Goulding et al., 1991; Pruitt, 1992). These observations suggested that analysis of P19 EC cells could provide an accurate in vitro model for at least some aspects of *Pax3* induction. However, a DNA sequence element present in the *wnt1* gene that is apparently not required in vivo has been identified on the basis of function in P19 EC cells (St-Arnaud and Moir, 1993; Echelard et al., 1994). Hence, as for any in vitro model, extrapolations from the P19 EC cell model to the developing embryo must be made cautiously. Nonetheless, in vitro models have been an essential component in the analysis of most eukaryotic promoters, including those expressed in specific cell lineages during development (e.g. Calzonetti et al., 1995; Goldhamer et al., 1995). Further, the P19 EC cell system has allowed identification of elements in the cardiac actin, *HoxA1* and *Oct3/4* promoters, which are utilized in vivo (Skerjanc and McBurney, 1994; Pratt et al., 1993; Yeom et al., 1996). The *Oct3/4* promoter elements utilized for expression in P19 EC cells are those required for expression in the epiblast, not those required for germ-line expression, which supports the use of these cells as a model for early gastrulation events (Yeom et al., 1996). The observation that the minimal 5′ deletion of the *Pax3* promoter sufficient to allow induction in P19 EC cells is also sufficient for the initial expression in both the neuroectodermal and somitic compartments of the developing embryo suggests that neural and skeletal muscle differentiation of P19 EC cells in vitro is dependent on many of the same mechanisms that are utilized in vivo. Further, the in vitro analysis suggests that, in addition to the promoter, two separate elements located between positions −1.4 to −1.1 and −1.0 and −0.8 are each required for *Pax3* expression.

Two different mechanisms mediate dorsoventral restriction of *Pax3* expression

Mechanisms responsible for the dorsal localization of *Pax3* expression in both the neural tube and somites have been studied extensively. In the chick, notochord ablation results in expansion of the *Pax3*-expressing domain ventrally within the neural tube. Conversely, ectopic notochord or floor plate grafts cause the localized suppression of *Pax3* expression (Goulding et al., 1993). In each of these cases, the shift in the domain of *Pax3* expression is likely to reflect the respecification of the neural tube to dorsal or ventral fates, respectively. The role of sonic hedgehog in mediating the ventralizing activity of the notochord and floorplate is well established (Echelard et al., 1993; Rashbash et al., 1994). It is not known whether the inhibitory effect of sonic hedgehog on *Pax3* expression is direct or indirect. The present study demonstrates that the 1.6 kbp region, which is sufficient to activate *Pax3* expression, also contains the elements sufficient to restrict its expression dorsally throughout the majority of the body axis.

While sequence elements contained within 1.6 kbp of the promoter allow correct dorsal localization throughout the majority of the trunk, additional elements are required to suppress *Pax3* expression in the tailbud. In each of four transgenic lines carrying the 1.6 kbp construct, inappropriate expression was observed in a ventrally expanded domain within the tailbud of early- to mid-somite stage embryos. This continues to be present through E12.5, at which point it extends from the hindlimb through the tip of the tailbud. The expanded domain of expression occurs in both the neuroectoderm and ventral mesodermal lineages. It is unlikely that this expression reflects an integration site-dependent position effect, since all four transgenic lines show the same pattern. It is also unlikely that this expanded expression domain is a consequence of deleting an element necessary for repression in response to sonic hedgehog signalling in this region. First, the dorsally restricted expression from this construct in anterior regions is a strong argument that elements sufficient to allow suppression by sonic hedgehog are present. Second, sonic hedgehog is not detected at positions posterior to the regressing node in early to mid-somite stage embryos (Marti et al., 1995).

There are two possible explanations for the ventrally expanded expression from the 1.6 kbp construct in the caudal domain. First, positive sequence elements unrelated to the mechanisms responsible for regulating *Pax3* expression may have been introduced into the reporter construct. However, the observation that the construct containing 14 kbp of 5′ sequence shows appropriate restriction of *Pax3* expression in the caudal domain of E9.5-12.5 embryos suggests that this is not the case. An alternative explanation is that the 1.6 kbp construct has lost elements that normally act to suppress *Pax3* expression in this caudal domain. This result would be consistent with the ability of sequences within the 14 kbp construct to restore dorsally restricted expression in the tail bud. It would also be consistent with the presence of signals necessary to activate *Pax3* through elements in the 1.6 kbp region in this domain (S.C.Pruitt, unpublished observations).

DNA sequence element requirements for *Pax3* expression differ in medial and ventrolateral myotomal populations

Pax3 is initially up-regulated uniformly throughout the dorsolateral half of newly formed somites (Goulding et al., 1993). This expression is observed for both the longer and shorter *Pax3* promoter constructs at E9.5. In contrast, up-regulation of *Pax3* expression in the ventrolateral third of each somite, which is observed for only the longer construct, begins at approximately E9.5, coincident with maturation of the somite.

During maturation, two myogenic precursor populations are established in the somite. Myocytes present in the medial somite give rise to epaxial muscles of the vertebral column, whereas laterally located myocytes migrate and give rise to the hypaxial muscles of the limb and body wall (Watchtler and Christ, 1992). Establishment of different myogenic lineages is correlated with differential expression of a variety of genes, including *Pax3* and the myogenic factors *MyoD1* and *Myf5* (Bober et al., 1994; Goulding et al., 1994; Williams and Ordahl, 1994; Pownall and Emerson, 1992). Down-regulation of *Pax3* expression in the medial region of the somite begins at approximately the time that expression of the myogenic factors *MyoD1* and *Myf5* initiates. Both the longer and shorter *Pax3* promoter constructs show down-regulation in these cells. In contrast, both the down-regulation of *Pax3* and expression of the myogenic factors is delayed in the migrating myoblasts that give rise to the hypaxial musculature. In these cells *Pax3* continues to be expressed for an additional 2 days during their migration to their final locations in the limb or body wall.

It has been suggested that the signals responsible for the differential expression of *Pax3* in these two muscle types are

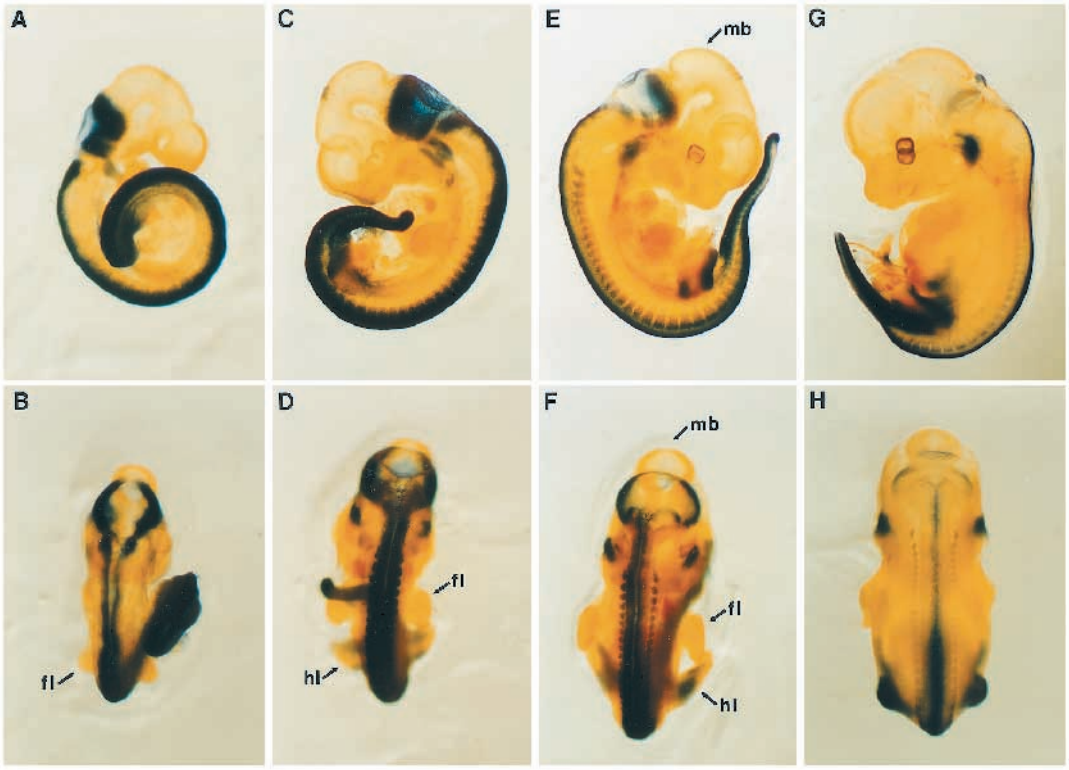

Fig. 8. In vivo expression from a *Pax3* construct containing 1.6 kbp of 5′ sequence. Transgenic mice carrying the 1.6 kbp *Pax3* promoter/*lacZ* fusion construct were isolated and stained for β-galactosidase activity. Embryos were isolated at E9.5 (A and B), E10.5 (C and D), E11.5 (E and F) or E12.5 (G and H). A, C, E and G: side views. B, D, F and H: dorsal views. Abbreviations: fl, forelimb; hl, hindlimb; mb, midbrain. Magnification, 50× (A,B), 30× (C,D), 28× (E,F) and 20× (G,H).

different, since *Pax3* expression in the ventrolateral myotome is dependent on a signal from the lateral plate mesoderm, while expression in medially located myoblasts does not require this signal (Pourquie et al., 1995). The present study provides strong support for this hypothesis by demonstrating that the signalling mechanisms responsible for *Pax3* expression in the hypaxial myocytes require DNA sequence elements in addition to those sufficient for *Pax3* regulation in the epaxial muscles. This study does not, however, resolve whether elements mediating *Pax3* expression in the epaxial muscles are also required for expression in the migrating myoblast.

Pax3 expression in the mesencephalon and neural crest derivatives

Several additional sites of endogenous *Pax3* expression also show differences in expression between the longer and shorter *Pax3* promoter constructs. The longer, but not the shorter, construct contains elements sufficient for expression in the developing midbrain and craniofacial neural crest derivatives. These results suggest that *Pax3* expression in these regions is dependent on different signaling mechanisms from those mediating expression in the hindbrain or trunk, similar to the situation in the migrating myoblasts. Expression from the longer construct occurs in the cochlea at E12.5. Although expression of the endogenous *Pax3* gene has not been reported for this site, and only a single founder animal was obtained carrying this construct, this expression may reflect a normal site of *Pax3* function since *splotch* homozygotes and many humans with Waardenburg's syndrome show malformations of the inner ear, including cochlear defects (reviewed by Steel and Smith, 1992). Expression from the 1.6 kbp construct was not detected in the cochlea; however, expression in all four transgenic lines containing this construct was detected in the pinnae. It is unclear whether this expression is related to the cochlear expression observed for the longer construct.

Conclusions

This study demonstrates that sequence elements necessary for the initial activation of *Pax3* in the dorsal neural tube and

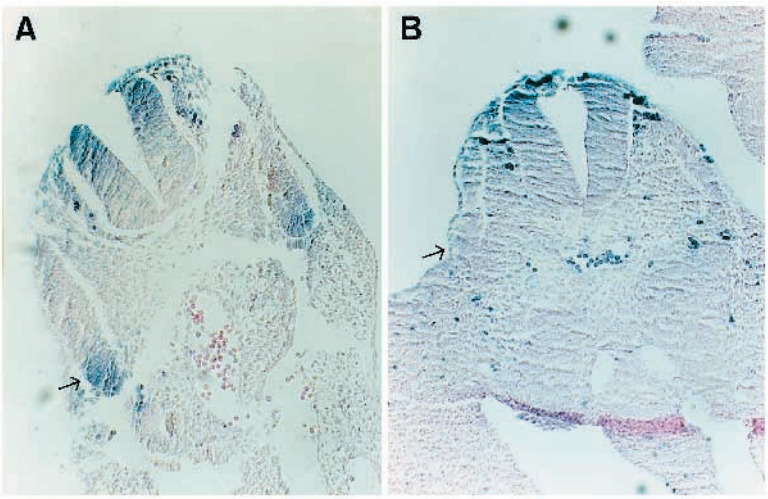

Fig. 9. Comparison of expression from the 14 and 1.6 kbp *Pax3* constructs in E9.5 cross sections. Cross sections from the mid-trunk region of E9.5 embryos carrying either the 14 kbp (A) or 1.6 kbp (B) constructs are shown where the arrows indicate the ventrolateral portion of the somite. Magnification, 100×.

somites are present within 1.6 kbp of the *Pax3* promoter and that this region is sufficient for expression in the P19 EC cell model. However, the region between −14 and −1.6 contains additional elements that act both positively and negatively at later stages of development to modulate *Pax3* expression. Essentially all of the elements required for the endogenous pattern of *Pax3* expression are contained within the region between −14 kbp and the promoter.

The authors thank L. Mielnicki for construction of pPolyCAT and pPolylacZ, R. Heinaman for plasmid preparations and C. Jones for advice on preparation of DNA for pronuclear injections. The initial stages of this work were supported by the ACS and production of transgenic animals was supported by grants from the NIH.

REFERENCES

Ausubel, F. M., Brent, R., Kingston, R. E., Moore, D. D., Seidman, J. G., Smith, J. A. and Struhl, K. (1994). *Current Protocols in Molecular Biology.* Current Protocols, USA.

Baldwin, C. T., Hoth, C. F., Amos, J. A., da-Silva, E. O. and Milunsky, A. (1992). An exonic mutation in the HuP2 paired domain gene causes Waardenburg's syndrome. *Nature* **355**, 637-638.

Bober, E., Franz, T., Arnold, H.-H., Gruss, P. and Tremblay, P. (1994). *Pax-3* is required for the development of limb muscles: a possible role for the migration of dermomyotomal muscle progenitor cells. *Development* **120**, 603-616.

Bucher, P. (1990). Weight matrix description of four eukaryotic RNA polymerase II promoter elements derived from 502 unrelated promoter sequences. *J. Mol. Biol.* **212**, 563-578.

Calzonetti, T., Stevenson, L. and Rossant, J. (1995). A novel regulatory region is required for trophoblast-specific transcription in transgenic mice. *Dev. Biol.* **171**, 615-626.

Cheng, J., Mielnicki, L. M., Pruitt, S. C. and Maquat, L. E. (1990). Nucleotide sequence of murine triosephosphate isomerase cDNA. *Nucleic Acids Res.* **18**,4261.

Echelard, Y., Epstein, D. J., Stjacques, B., Shen, L., Nohler, J., McMahon, J. A. and McMahon, A. P. (1993). *Sonic-hedgehog*, a member of a family of putative signaling molecules, is implicated in the regulation of CNS polarity. *Cell* **75**, 1417-1430.

Echelard, Y., Vassileva, G. and McMahon, A. P. (1994). *Cis*-acting regulatory sequences governing *Wnt-1* expression in the developing mouse CNS. *Development* **120**, 2213-2224.

Epstein, D. J., Vekemans, M. and Gros, P. (1991). Splotch (Sp^{2H}), a mutation affecting development of the mouse neural tube, shows a deletion within a paired homeodomain of *Pax-3*. *Cell* **67**, 767-774.

Goldhamer, D. J., Brunk, B. P., Faerman, A., King, A., Shani, M. and Emerson Jr, C. P. (1995). Embryonic activation of the myoD gene is regulated by a highly conserved distal control element. *Development* **121**, 637-649.

Goulding, M. D., Chalepakis, G., Deutsch, U., Erselius, J. R. and Gruss, P. (1991). *Pax-3*, a novel murine DNA binding protein expressed during early neurogenesis. *EMBO J.* **10**, 1135-1147.

Goulding, M. D., Lumsden, A. and Gruss, P. (1993). Signals from the notochord and floor plate regulate the region-specific expression of two Pax genes in the developing spinal cord. *Development* **117**, 1001-1016.

Goulding, M., Lumsden, A. and Paquette, A. J. (1994). Regulation of *Pax-3* expression in the dermomyotome and its role in muscle development. *Development* **120**,957-71.

Hogan, B., Constantini, F. and Lacy, E. (1986). *Manipulating the Mouse Embryo*. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.

Kaufman, M. H. (1992). *The Atlas of Mouse Development*. San Diego: Academic Press Limited.

Macina, R. A., Barr, F. G., Galili, N. and Reithman, H. C. (1995). Genomic organization of the human *Pax3* gene: DNA sequence analysis of the region disrupted in alveolar rhabdomyosarcoma. *Genomics* **26**, 1-8.

Marti, E., Takada, R., Bumcrot, D. A., Sassaki, H. and McMahon, A. P. (1995). Distribution of Sonic hedgehog peptides in the developing chick and mouse embryo. *Development* **121**, 2537-2547.

McBurney, M. W., Jones-Villeneuve, E. M. V., Edwards, M. K. S. and Anderson, P. J. (1982). Control of muscle and neuronal differentiation in a cultured embryonal carcinoma cell line. *Nature* **299**, 165-167.

Pourquie, O., Coltey, M., Breant, C. and Le Douarin, N. M. (1995). Control of somite patterning by signals from the lateral plate. *Proc. Natl. Acad. Sci. USA* **92**, 3219-3223.

Pownall, M. E. and Emerson Jr., C. P. (1992). Sequential activation of three myogenic regulatory genes during somite morphogenesis in quail embryos. *Dev. Biol.* **151**, 67-79.

Pratt, M. A., Langston, A. W., Gudas, L. J. and McBurney, M. W. (1993). Retinoic acid fails to induce expression of Hox genes in differentiation-defective murine embryonal carcinoma cells carrying a mutant gene for alpha retinoic acid receptor. *Differentiation* **53**, 105-113.

Pruitt, S. C. and Reeder, R. H. (1984a). Effect of topological constraint on transcription of ribosomal DNA in *Xenopus* oocytes. *J. Mol. Biol.* **174**, 121-139.

Pruitt, S.C. and Reeder, R.H. (1984b). Effect of intercalating agents on RNA polymerase I promoter selection in *Xenopus laevis*. *Mol. Cell. Biol.* **4**, 2851-2857.

Pruitt, S. C. (1992). Expression of *Pax3*- and neuroectoderm-inducing activities during differentiation of P19 embryonal carcinoma cells. *Development* **116**, 573-583.

Pruitt, S. C. (1994). Primitive streak mesoderm-like cell lines expressing *Pax3* and Hox gene autoinducing activities. *Development* **120**, 37-47.

Pruitt, S. C. and Natoli, T. A. (1992). Inhibition of differentiation by leukemia inhibitory factor distinguishes two induction pathways in P19 embryonal carcinoma cells. *Differentiation* **50**, 57-65.

Rashbash, P., Wilson, V., Rosen, B. and Roelink, H., Augsburger, A., Heemskerk, J., Korzh, V., Norlin, S., Altaba, A. R. I., Tanabe, Y., Placzed, M., Edlund, T., Jessell, T. M. and Dodd, J. (1994). Floor plate and motor neuron induction by VHH-1, a vertebrate homolog of *hedgehog* expressed by the notochord. *Cell* **76**, 761-775.

Skerjanc, I. S. and McBurney, M. W. (1994). The E box is essential for activity of the cardiac actin promoter in skeletal but not in cardiac muscle. *Devel. Biol.* **163**,125-32.

St-Arnaud, R. and Moir, J. M. (1993). *Wnt-1*-inducing factor-1: a novel G/C box-binding transcription factor regulating the expression of *Wnt-1* during neuroectodermal differentiation. *Mol. Cell. Biol.* **13**, 1590-1598.

Steel, K. P. and Smith, R. J. (1992). Normal hearing in *Splotch* (Sp/+), the mouse homologue of Waardenburg syndrome type 1. *Nature Genetics* **2**, 75-9.

Stuart, E. T., Kioussi, C. and Gruss, P. (1993). Mammalian PAX genes. *Annu. Rev. Genet.* **27**, 219-236.

Tassabehji, M., Read, A. P., Newton, V. E., Harris, R., Balling, R., Gruss, P. and Strachan, T. (1992). Waardenburg's syndrome patients have mutations in the human homologue of the Pax-3 paired box gene. *Nature* **355**, 635-636.

Wachtler, F. and Christ, B. (1992). The basic embryology of skeletal muscle formation in vertebrates: the avian model. *Sem. Dev. Biol.* **3**, 217-227.

Williams, B. A. and Ordahl, C. P. (1994). Pax-3 expression in segmental mesoderm marks early stages in myogenic cell specification. *Development* **120**, 785-796.

Yeom, Y. I., Fuhrmann, G., Ovitt, C. E., Brehm, A., Ohbo, K., Gross, M., Hubner, K. and Scholer, H. R. (1996). Germline regulatory element of Oct-4 specific for the totipotent cycle of embryonal cells. *Development* **122**, 881-894.

(Accepted 31 October 1996)