Development 117, 905-916 (1993)
Printed in Great Britain © The Company of Biologists Limited 1993

Two enhancer regions in the mouse *En-2* locus direct expression to the mid/hindbrain region and mandibular myoblasts

C. Logan[1,2], W. K. Khoo[2], D. Cado[2] and A. L. Joyner[1,2,*]

[1]Department of Molecular and Medical Genetics, University of Toronto, Canada
[2]Division of Molecular and Developmental Biology, Samuel Lunenfeld Research Institute, Mount Sinai Hospital, Toronto, Canada M5G 1X5

*Author for correspondence

SUMMARY

An *En-2/lacZ* gene fusion containing 9.5 kb of *En-2* genomic DNA was capable of directing *lacZ* expression in an *En-2*-specific manner both temporally and spatially during embryogenesis and in the adult. *lacZ* expression was confined in the embryo to cells within the mid/hindbrain and mandibular arch regions and in the adult to cells of the molecular and granular layers of the cerebellum, and within the pons and colliculi regions. Interestingly, in the adult, transgene expression patterns within the cerebellum in two lines appeared to mark distinct anterior-posterior compartments. Analysis of the expression pattern of this transgene, in fetal and adult mice lacking a functional *En-2* protein, provided evidence that the *En-2* gene in mouse is not autoregulated. Deletion analysis of the *En-2* genomic region and the use of a heterologous promoter identified two enhancer-containing regions of 1.5 and 1.0 kb in length, 5′ of the transcribed sequences, which independently directed expression in the embryo to either the mid/hindbrain region or mandibular myoblasts, respectively. The 1.5 kb fragment contains the most anterior neural enhancer and the 1.0 kb fragment, the earliest myogenic enhancer thus far characterized. These *En-2*-specific regulatory regions can now be used in a biochemical analysis to identify proteins important in anterior-posterior patterning of the vertebrate CNS and in the specification of muscle identity as well as in a mutational analysis to direct expression of other developmentally important genes to these regions.

Key words: transcriptional regulation, homeobox, *Engrailed*, transgenic mice, *lacZ* reporter, cerebellar compartments, myoblasts

INTRODUCTION

The vertebrate *Engrailed* (*En*) genes are expressed early in development in overlapping, spatially restricted domains in the presumptive mid/hindbrain region of the neural epithelium (reviewed in Joyner and Hanks, 1991), suggesting that they play a role in the regional specification of the central nervous system (CNS). By analogy to their *Drosophila* homologue, *engrailed* (*en*) (Morata and Lawrence, 1975; Kornberg, 1981; Jaynes and O'Farrell, 1988, 1991; Ohkuma et al., 1990), the vertebrate *En* genes are believed to function as transcription factors involved in specifying positional information along the anterior-posterior (A-P) axis. Consistent with this idea, recent experiments using chick/quail chimeric embryos have shown that induction and/or maintenance of *En* expression in neuroepithelial grafts correlates well with later morphological development into mid/hindbrain structures (Martinez and Alvardo-Mallart, 1990; Gardner and Barald, 1991; Itasaki et al., 1991; Martinez et al., 1991; Bally-Cuif et al., 1992). Furthermore, *En* gene expression not only precedes obvious morphological differentiation of the neural tube (Gardner et al., 1988; Patel et al., 1989; Davis et al., 1991; Gardner and Barald, 1992) but, at least in chick, also precedes determination of the corresponding region of the CNS (Alvarado-Mallart et al., 1990; Itasaki et al., 1991). A more direct mutational approach to determine the function of the vertebrate *En* genes has provided evidence for their functional redundancy during embryogenesis and shown that the mouse *En-2* gene is involved in development of the cerebellum (Joyner et al., 1991). When the *En-2* homeobox region was deleted by gene targeting in embryonic stem (ES) cells, mice homozygous for the mutation *En-2^{hd}*, were viable and had no obvious behaviourial defects but showed a distinct abnormality in patterning of the cerebellar folds.

In mouse, there are two *En* genes, *En-1* and *En-2* (Joyner et al., 1985; Joyner and Martin, 1987). For *En-2*, RNA in situ analysis has shown that endogenous transcripts are first detected at day 8 of embryogenesis at approximately the 5-somite stage in a triangular patch of cells across the anterior neural epithelium in a region that marks the presumptive mid/hindbrain junction (Davis et al., 1988; Davis and Joyner, 1988; McMahon et al., 1992). Expression continues in this region throughout development. Outside the

CNS, *En-2* is expressed in cells surrounding the developing pituitary (Davis et al., 1988, 1991). In addition, immunohistochemical analysis using a polyclonal antiserum (*aEnhb-1*) that detects both *En-1* and *En-2* protein has shown that one or both genes are expressed in presumptive myoblasts within the first branchial arch (Davis et al., 1991). In the adult, *En-2* is strongly expressed in cells of the internal granular layer of the cerebellum. No expression is detected in the neighbouring Purkinje cells, whereas weak expression is detectable in cells of the molecular layer (Davis et al., 1988; Davis and Joyner, 1988; K. Millen and A. L. J. unpublished data). Outside the cerebellum, expression has also been detected in specific groups of neurons in the pons region.

For *En-1*, RNA in situ analysis has shown that it is expressed in the anterior neural epithelium slightly earlier than *En-2*, at the 1-somite stage, in a domain which later overlaps with that of *En-2* (Davis and Joyner, 1988; McMahon et al., 1992). Unlike *En-2*, it is also expressed in specific domains within the spinal cord, somites and limb buds during embryogenesis and is not expressed in the adult cerebellum. Detailed studies of *En* expression in a number of vertebrate species have revealed similar expression patterns (reviewed in Joyner and Hanks, 1991).

Other mammalian genes have been identified that are also expressed in restricted domains within the developing CNS, including members of the *Hox* (reviewed in McGinnis and Krumlauf, 1992), *Wnt* (reviewed in Nusse and Varmus, 1992) and *Pax* (reviewed in Gruss and Walther, 1992) gene families. Recently, reporter gene constructs and their analysis in transgenic mice have been used to identify *cis*-acting DNA elements involved in the embryonic spatial- and temporal-specific expression of various members of the *Hox* gene family. A number of regulatory elements have been identified that are capable of directing reporter gene expression to specific regions within the CNS (Zakany et al., 1988, 1990; Kress et al., 1990; Tuggle et al., 1990; Schughart et al., 1991). However, multiple regulatory elements, some of which are shared between various members of a gene cluster, are required to reconstruct the spatial and temporal embryonic expression patterns of endogenous *Hox* genes in transgenic mice (Bieberich et al., 1990; Puschel et al., 1990, 1991; Whiting et al., 1991; Sham et al., 1992).

Although the expression patterns of the *En* genes have been analyzed in detail in several vertebrate species (reviewed in Joyner and Hanks, 1991), little is known about how their expression patterns are generated. In this study, we have used a transgenic approach to define the *cis*-acting DNA regulatory elements involved in the establishment and maintenance of the embryonic region-specific and adult cell-type-specific expression of the mouse *En-2* gene. Our analysis of the temporal and spatial expression patterns of various *En-2/lacZ* gene fusions in transgenic mice has defined a genomic region capable of correctly directing the expression of the *Escherichia coli lacZ* reporter gene in an *En-2*-specific manner. Subsequent analysis of this region has led to the identification of two separate enhancer regions capable of directing the expression of *lacZ* to cells within the embryonic mid/hindbrain or mandibular arch regions. Furthermore, the *En-2* gene in mouse was shown not to be autoregulated, by analyzing the expression of an *En-2/lacZ* transgene in fetal and adult mice lacking a functional *En-2* protein. Finally, analysis of reporter gene expression in the adult cerebellum supports previous mutant data suggesting that the cerebellum is divided into anterior and posterior compartments.

MATERIALS AND METHODS

Isolation of genomic clones

Five overlapping genomic DNA clones containing sequences extending further 5′ of the *En-2* coding region than had previously been isolated (Joyner and Martin, 1987; Logan et al., 1992) were obtained by screening a *Sau*3A partially digested, size selected, mouse genomic DNA library (Clontech) using a 350 bp *Bgl*II-*Hin*dIII *En-2* genomic DNA fragment 5′ of the homeobox. Approximately 10^6 clones were plated, transferred to nitrocellulose (Schleicher and Schuell) membrane and hybridized under conditions of high stringency (Joyner and Martin, 1987). The probe was radiolabelled to a high specific activity using the random priming procedure of Feinberg and Volgestein (1983). Final washes were done using 0.5× SSPE, 0.1% SDS at 60°C. Positive clones were plaque purified, restriction mapped and their inserts subcloned into pUC18/19 plasmid vectors. Restriction sites were confirmed by comparison with DNA fragment sizes visualized by Southern blot analysis of total mouse genomic DNA. A composite restriction map of part of the genomic region isolated as well as that contained within previously identified clones is shown in Fig. 1.

DNA constructs

Constructs 1 (MC4) and 2 (MC6) were made by first ligating a 2.5 kb *Bgl*II-*Pvu*I *En-2* genomic DNA fragment containing 1439 bp of 3′ untranslated sequence, the putative polyadenlyation signal and approximately 1.0 kb of 3′ flanking genomic DNA, into the *Bam*HI and *Xba*I (end-filled using Klenow) sites within the polylinker of a modified pBluescript (Stratagene) vector in which the *Sac*I site at position 657 had previously been changed to a *Sal*I site by insertion of a linker (M. Hanks, unpublished data). The bacterial *lacZ* gene originally from pMC1871 (Pharmacia) was then placed upstream by ligating an end-filled (Klenow) 3.0 kb *Xba*I fragment derived from pGT4.5A (derivative of pGT4.5 (Gossler et al., 1989) made by W. Skarnes) into the *Sma*I site within the polylinker. For construct 1, a 7.0 kb *Cla*I genomic DNA fragment containing the translational start site for *En-2* and 6.8 kb of 5′ flanking genomic DNA was then inserted into the *Cla*I site in the polylinker. For construct 2, a shorter 2.5 kb *Sal*I-*Cla*I fragment also containing the *En-2* translational start site but with only 2.3 kb of 5′ flanking genomic DNA was isolated from the 5′ end of the genomic phage clone λEn-2.1 (Joyner and Martin, 1987), end-filled with Klenow and ligated into the *Hin*dIII site (end filled with Klenow) within the polylinker. In both constructs 1 and 2, the *lacZ* coding region was fused in frame to the first 68 amino acids of the *En-2* protein and the *En-2* 3′ region provided the polyadenlyation signal. Sequence analysis confirmed that the correct reading frame had been maintained during construction of both constructs.

For construct 3 (MC3), a 1.8 kb *Sal*I-*Sma*I fragment from the 5′ end of λEn-2.1 (Joyner and Martin, 1987) in which the 3′ end is within the 5′ untranslated region of *En-2*, was cloned into the polylinker of pSP65 (Stratagene). A 3.6 kb *Hin*dIII-*Bam*HI fragment was then isolated from p*SDKlacZpA* (S. Darling, unpublished data), end-filled with Klenow and ligated into the *Sma*I site. This fragment contained a Kozak consensus sequence for eukaryotic translational initiation (Kozak, 1983) fused in frame to the

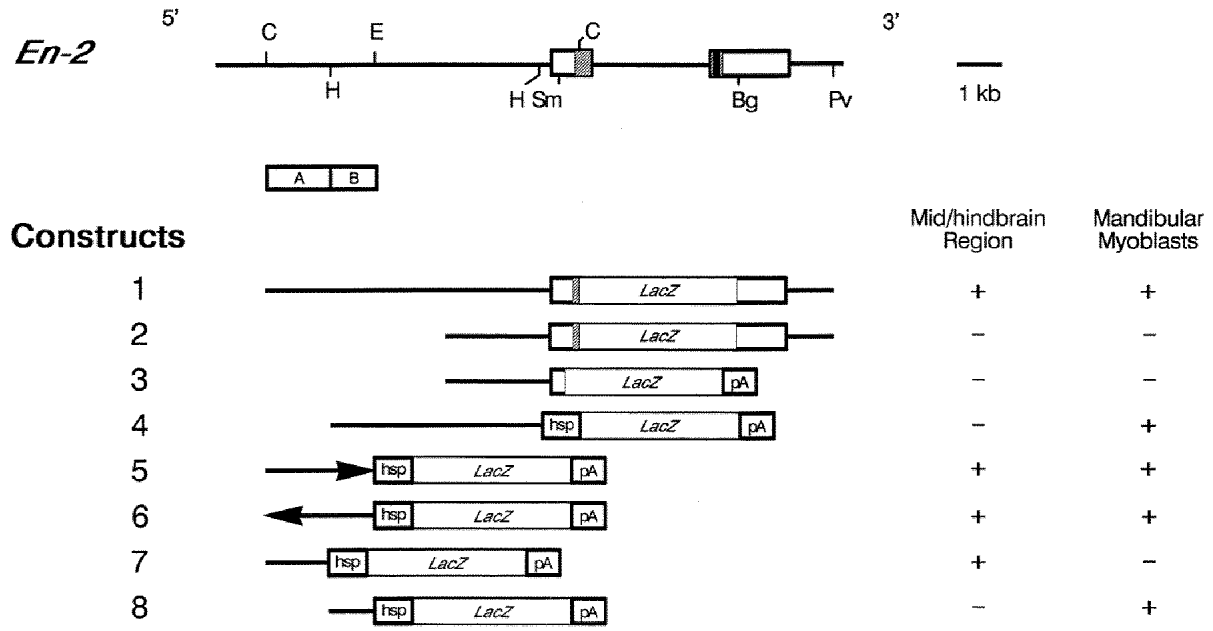

Fig. 1. Structure and *En-2*-specific expression of *En-2/lacZ* reporter constructs in transgenic mice. The top line represents a composite of the genomic fragments contained in overlapping phage clones isolated for the mouse *En-2* gene. Transcribed sequences are indicated by boxes. The black box represents the homeobox and the remaining coding sequences are represented by hatched boxes. Only the restriction sites used for making the *En-2/lacZ* constructs are indicated. Shown below are the *En-2/lacZ* constructs tested. The *En-2* sequences contained within each construct (see Materials and Methods) are aligned relative to the partial restriction map of the *En-2* locus. The arrows in constructs 5 and 6 indicate the 5′ to 3′ orientation with respect to *lacZ*, of the *En-2* genomic DNA fragment contained in each construct. The boxes labelled *lacZ*, hsp and pA represent the protein-coding region of the bacterial β-galactosidase gene, the promoter region of the *hsp68* gene and the polyadenlyation region from SV40 respectively. The *En-2*-specific expression patterns for each construct are summarized on the right. The enhancer regions (A and B) described in the text are indicated by the boxes labelled A and B, respectively. Bg, *Bgl*II; C, *Cla*I; E, *Eco*RI; H, *Hin*dIII; Pv, *Pvu*II; Sm, *Sma*I.

lacZ coding region from pMC1871 (Pharmacia) followed by an SV40 polyadenylation signal.

Constructs 4 to 8 (MC7-MC10) were made by inserting various *En-2* genomic fragments upstream of a 4.5 kb *Bam*HI fragment from p*hspPTlacZpA* (Kothary et al., 1989) which had previously been cloned into the *Bam*HI site within the polylinker of a modified pBluescript vector (see above). The *Bam*HI fragment from p*hspPTlacZpA* contains promoter sequences (−664 to +224 relative to the start of transcription; Kothary et al., 1989) from the mouse *hsp68* gene including the translational start site, fused in-frame to the *lacZ* gene from pMC1871, followed by an SV40 polyadenylation signal. Construct 4 (MC7) contains a 4.5 kb *Hin*dIII genomic fragment oriented 5′ to 3′ with respect to *lacZ*, ligated into the *Hin*dIII site within the polylinker. Construct 5 (MC9 ori1) and construct 6 (MC9 ori2) each contain a 2.5 kb *Cla*I-*Eco*RI genomic fragment (end-filled with Klenow) inserted in opposite orientations into the *Sma*I site within the polylinker. In construct 5, the 2.5 kb *Cla*I-*Eco*RI genomic fragment was oriented 5′ to 3′ with respect to *lacZ* and oriented in the opposite direction in construct 6. Construct 7 (MC8) contains a 1.5 kb *Cla*I-*Hin*dIII genomic fragment oriented 5′ to 3′ with respect to *lacZ*, that was end-filled with Klenow and ligated into the *Sma*I site within the polylinker. Construct 8 (MC10) was made by digesting construct 5 with *Hin*dIII and recircularizing to delete genomic sequences from *Cla*I to *Hin*dIII leaving a 1.0 kb genomic fragment extending from *Hin*dIII to *Eco*RI.

With the exception of construct 3, DNA fragments for injection were excised from vector sequences by digesting with *Sal*I and isolated either by equilibrium ultracentrifugation on a sucrose gradient or by electroelution following gel electrophoresis. Fragments isolated by gel electrophoresis were further purified prior to ethanol precipitation using a NACS (BRL) column. DNA from construct 3 was linearized prior to injection by digestion with *Sal*I, extracted with phenol-chloroform and ethanol precipitated.

Production and identification of transgenic mice

Outbred CD-1 mice were used to produce transgenic embryos and mouse lines as described by Hogan et al. (1986). Transgenic embryos and mice were identified by Southern blot analysis of DNA extracted from either yolk sacs or tail biopsies, respectively, using a *lacZ* or *En-2*-specific DNA probe. The temporal and spatial pattern of transgene expression in the various lines was analyzed either by interbreeding animals homozygous for the transgene or by breeding founder and/or F_1 males with CD-1 females. Pregnant females were killed at various days of gestation and their embryos analyzed for *lacZ* activity as described below. The day on which a vaginal plug was observed was designated day 0.5 of gestation.

In one of the transgenic lines, Tg4.35, analysis of DNA extracted from the founder revealed that the transgene had inserted at two different sites within the genome. These two integrations (a and b) have subsequently been separated from one another by breeding. Analysis of the expression pattern in this line was done prior to the segregation of the two integrations. Subsequent analysis has shown that the *lacZ* expression pattern in Tg4.35a is identical to that seen when both insertions are present, whereas *lacZ* expression in Tg4.35b although similar is at a much lower level in both the embryo and adult brain (data not shown).

lacZ staining of whole-mount embryos and adult brain sections

Embryos to be stained were dissected in PBS, fixed in 0.2% glutaraldehyde (Sigma), 2 mM $MgCl_2$, 5 mM EGTA in PBS at room temperature for 15 to 90 minutes, depending on size, and then washed in three changes of PBS plus 0.02% NP40, 2 mM $MgCl_2$ at room temperature for 30 minutes each. They were then stained in the dark in 1 mg/ml X-gal (BRL), 5mM $K_3Fe(CN)_6$, 5 mM $K_4Fe(CN)_6$, 2 mM $MgCl_2$ and 0.02% NP40 in PBS at 37°C overnight. Whole-mount 15.5 day embryos were fixed in 4% formaldehyde, 0.2% glutaraldehyde prior to staining. Beyond 12.5 days however, this staining procedure was relatively ineffective due to poor penetration of both fixative and stain. Following *lacZ* staining, embryos to be sectioned were dehydrated and embedded in paraffin wax. Sections were cut at 6-10 μm, mounted on glass slides, dewaxed and counterstained with either eosin or fast nuclear red.

Founder embryos were analyzed between 10.5 to 12.5 days of gestation (Table 1) whereas embryos from each of the five transgenic lines that express the large *En-2/lacZ* construct (see results) were analyzed at 8.5, 9.5, 10.5, 11.5 and 12.5 days of gestation. One line (Tg5) was also analyzed at 15.5 days.

In the adult brain, β-galactosidase activity was analyzed by *lacZ* staining of cryostat sections. Dissected tissues were first rapidly frozen in O.C.T. (Tissue-Tek). Sections were then cut at 10 μm, mounted on slides and subsequently fixed in 0.2% glutaraldehyde, 2% formaldehyde, 2 mM $MgCl_2$, 5 mM EGTA in PBS for 5 minutes, washed three times for 5 minutes each and stained as above. Adult brain tissue from each of the ten founder animals obtained carrying the large *En-2/lacZ* transgene was analyzed as well as adult brain tissue from four of the five established lines. After several generations, embryonic expression in the fifth line (Tg7) was no longer detectable and this line was not continued. Adult tissues from this line were therefore not analyzed.

Mutant mouse strains

Mice carrying a mutation in the *En-2* gene, *En-2^{hd}* (Joyner et al., 1991) on an outbred background, were bred with mice from the transgenic line Tg5, carrying DNA from construct 1. *lacZ* expression in 8.5, 9.5, 10.5, 11.5 and 12.5 day embryos homozygous for the *En-2^{hd}* mutation and hemizygous for the transgene were analyzed via whole-mount staining. In addition, 11.5 day embryos obtained from females after the following cross were genotyped by Southern blot analysis of yolk sac DNA and stained whole mount for *lacZ* activity: *En-2^{hd}*/+;+/+ × *En-2^{hd}*/+;Tg5/Tg5. Adult brain tissue from littermates of an identical cross, which were hemizygous for Tg5 and were either wild type, heterozygous or homozygous for the *En-2^{hd}* mutation, was sectioned and also stained for *lacZ* activity.

Meander tail mutant mice (*mea^{2J}/mea^{2J}*) on a C57BL/KSJ background were obtained from Jackson Laboratories (Bar Harbour, ME) and bred with mice from the transgenic line, Tg5. Adult brain tissues from *mea* homozygous littermates either lacking or hemizygous for the transgene were sectioned and stained as above for *lacZ* activity.

Table 1. Transgene expression in founder embryos

			Transgene expression patterns compared to endogenous *En-2*		
			Characteristic of *En-2*		
Construct*	Embryo age (days)	No. exp Tg†/ Total no. Tg	Mid/ hindbrain	Mandibular myoblasts	Ectopic‡
1	11.5-12.5	4/4	4	4	1
2	12.5	5/8	−	−	5
3	11.5-12.5	3/6	−	−	3
4	10.5-11.5	2/6	−	2	2
5	10.5-11.5	3/3	3	3	2
6	10.5	5/8	5	5	4
7	10.5-11.5	5/6	5	−	4
8	10.5-11.5	6/8	−	6	5

*Construct numbers correspond to those used in Fig. 1.

†No. exp Tg refers to the number of founder embryos that expressed detectable levels of *lacZ*.

‡Expression patterns inappropriate for *En-2* or patterns that varied from line to line were defined as ectopic.

Whole-mount RNA in situ analysis

Whole-mount RNA in situ analysis of the *En-2* expression pattern in 8.5 day CD-1 mouse embryos was done as in Conlon and Rossant (1992).

RESULTS

An *En-2/lacZ* transgene is correctly expressed embryonically in transgenic mice

As a first step in looking at the *cis*-acting DNA regulatory elements involved in the expression of *En-2*, we have used reporter gene constructs and their analysis in transgenic mice to define a large genomic region capable of correctly directing the expression of the *Escherichia coli lacZ* reporter gene in an *En-2*-specific manner during embryogenesis. An *En-2/lacZ* fusion construct containing a total of 9.5 kb of *En-2* genomic DNA (1 in Fig. 1) was used to generate transgenic mice. Embryos derived either directly from injected zygotes (founder embryos) or by mating of established lines were analyzed via histochemical staining for β-galactosidase activity (see Materials and Methods).

In a preliminary experiment, twelve founder embryos were analyzed at 11.5 or 12.5 days of development. Four were found to be transgenic by Southern blot analysis of yolk sac DNA and all four expressed the transgene in a similar pattern within the CNS which, in whole-mount staining, resembled that of the endogenous *En-2* gene (Table 1). *lacZ* expression was detected in a ring of cells spanning the mid/hindbrain junction. In addition, expression was detected in the mandibular region in presumptive myoblasts. Given that recent experiments have shown that *En-2*, but not *En-1*, transcripts are present in certain jaw muscle precursors in zebrafish (Ekker et al., 1992) and that *En-2* protein is present in presumptive myoblasts within the first branchial arch in chick (Gardner and Barald, 1992), zebrafish (Hatta et al., 1990) and *Xenopus* (Hemmati-Brivanlou et al., 1991), it seems likely that the *lacZ* expression in this region in mouse reflects expression of the endogenous *En-2* gene. Expression was not detected in cells surrounding the developing pituitary.

To examine the spatial and temporal pattern of the transgene expression in more detail, transgenic lines were established using this construct. Ten transgenic founder animals were identified by Southern blot analysis of DNA extracted from tail biopsies and were bred further. Seven of these founder animals transmitted the transgene through to their offspring, and mid-gestation embryos from five of these seven lines expressed detectable levels of *lacZ*. The copy number of the transgene varied between lines and ranged

from 2 copies in line 5 (Tg5) to in excess of 50 copies in line 7 (Tg7) (data not shown).

Mid-gestation (10.5 to 12.5 days) transgenic embryos from all five expressing lines exhibited moderate to strong expression of the *lacZ* reporter gene in the mid/hindbrain region as well as in the mandibular region in presumptive myoblasts (Fig. 2). This pattern matched that seen in the founder embryos and was consistent with the normal expression pattern of the endogenous *En-2* gene (Davis et al., 1988; Davis and Joyner, 1988; Davis et al., 1991). As in the founder embryos, expression was not detected in the region of the developing pituitary.

All five transgenic lines were subsequently analyzed at different developmental stages (see Materials and Methods) to determine the precise temporal pattern of reporter gene activity in the mid/hindbrain and mandibular regions and to examine the spatial pattern of transgene expression in more detail. All lines displayed similar patterns of β-galactosidase activity in these two regions at each developmental stage examined. An example of one such line (Tg5) is shown in Fig. 2.

Expression of the transgene was first detected in 8.5 day embryos at approximately the 5-somite stage in a few scattered cells within the anterior neuroepithelium (data not shown). Slightly later, at approximately 10-12 somites, two triangular patches of expressing cells were readily detectable in this region (Fig. 2B). At 9.5 days, expression spanned the mid/hindbrain junction (Fig. 2D, E). As shown in Fig. 2F,G (12.5 and 15.5 days, respectively), expression was maintained in this region until at least 15.5 days of gestation, the latest embryonic stage examined. Differences in the width of the band of expressing cells across the mid/hindbrain junction were seen from line to line (eg. Fig. 2D,E). No correlation, however, was found between the transgene copy number and the domain of *lacZ* expression. When the embryonic expression pattern within one line (Tg5) was analyzed in littermates either hemizygous or homozygous for the transgene, the width of the band of expressing cells was greatest in homozygotes. This result suggests that the band width is correlated with the level of expression of the transgene.

Outside the CNS, bilateral populations of scattered *lacZ*-expressing cells were detected in the mandibular region beginning at 9.5 days. In slightly older embryos (10.5 days), *lacZ*-expressing cells formed a densely associated cell mass which extended into the center of the first branchial arch and surrounded the mandibular branch of the trigeminal nerve (Fig. 2C). Expression was also maintained in this region until at least 15.5 days (Fig. 2G) at which time *lacZ*-expressing cells appear to mark differentiating jaw muscles (D. Sassoon, personal communication). Based on lineage analysis of craniofacial muscles in the chick embryo (Noden, 1991), these results are consistent with the hypothesis that *En*-expressing cells mark paraxial mesoderm-derived cells, which differentiate into a specific set of jaw muscles. Support for this idea comes from analysis of *En-2* protein and mRNA expression patterns in chick (Gardner and Barald, 1992), zebrafish (Hatta et al., 1990, Ekker et al., 1992) and *Xenopus* (Hemmati-Brivanlou et al., 1991).

With the exception of the cells surrounding the developing pituitary, the transgene expression patterns correlated well with the previously described spatial and temporal pattern of expression of the endogenous *En-2* gene during embryogenesis (eg. Fig. 2A,B). We conclude that *En-2*-specific regulatory element(s), which are capable not only of correctly directing the expression of *lacZ* to the mid/hindbrain region and mandibular myoblasts but also of correctly regulating such expression temporally, are contained within the transgene.

Transgene expression that was neither consistent between lines nor with the endogenous *En-2* expression pattern was also detected in each of the transgenic lines and in one of the founder embryos. We interpret this as ectopic expression due to the influence of regulatory elements at the different sites of integration. Similar observations have been made for other *lacZ* transgenes (eg. Puschel et al., 1990; Kress et al., 1990). The ectopic expression patterns observed in our transgenic lines were not analyzed in detail.

Identification of two separate *En-2*-specific enhancer regions

To begin to determine which *En-2* sequences were responsible for correctly directing expression of the transgene to the mid/hindbrain region and mandibular myoblasts during embryogenesis, two constructs (2 and 3, Fig. 1) were tested by analyzing β-galactosidase activity in mid-gestation founder embryos, in which 4.5 kb of genomic sequence from the 5′ end of construct 1 was deleted. In construct 3, the *En-2* 3′ untranslated region and flanking genomic DNA sequences were also removed and replaced with an SV40 polyadenylation signal. Transgenic founder embryos carrying either construct revealed no consistent pattern of *lacZ* expression. As shown in Table 1, 5 out of 8 embryos and 3 out of 6 embryos carrying construct 2 or 3 respectively, expressed detectable levels of β-galactosidase. Their patterns, however, were unrelated to each other and to the patterns described above for the large *En-2/lacZ* transgene (construct 1). Such ectopic expression was presumably due to the presence of a minimal *En-2* promoter and its sensitivity to regulatory elements at the site of integration. No expression was detected in any of the embryos in either the mid/hindbrain or mandibular region. These results demonstrate that the 5′ genomic sequences contained in these fusion constructs have promoter activity but are incapable of directing expression of the transgene in an *En-2*-specific manner. Furthermore, these results suggest that the *En-2*-specific expression seen using the large *En-2/lacZ* transgene (construct 1), was due to regulatory element(s) contained within the 4.5 kb 5′ genomic fragment deleted from constructs 2 and 3.

To test this hypothesis, various genomic fragments 5′ of the *En-2* gene (Fig. 1) were placed upstream of a minimal promoter from the mouse heat-inducible gene, *hsp68* fused in frame to the *lacZ* gene (Kothary et al., 1989). These constructs were also analyzed in mid-gestation transgenic founder embryos and the results obtained are shown in Table 1. Previous studies had shown that this promoter fragment had no detectable basal activity in transgenic mouse embryos (Kothary et al., 1989), but could be activated in distinct patterns by defined heterologous enhancer elements (eg. Tuggle et al., 1990; Whiting et al., 1991).

Our analysis defined a 2.5 kb genomic fragment approx-

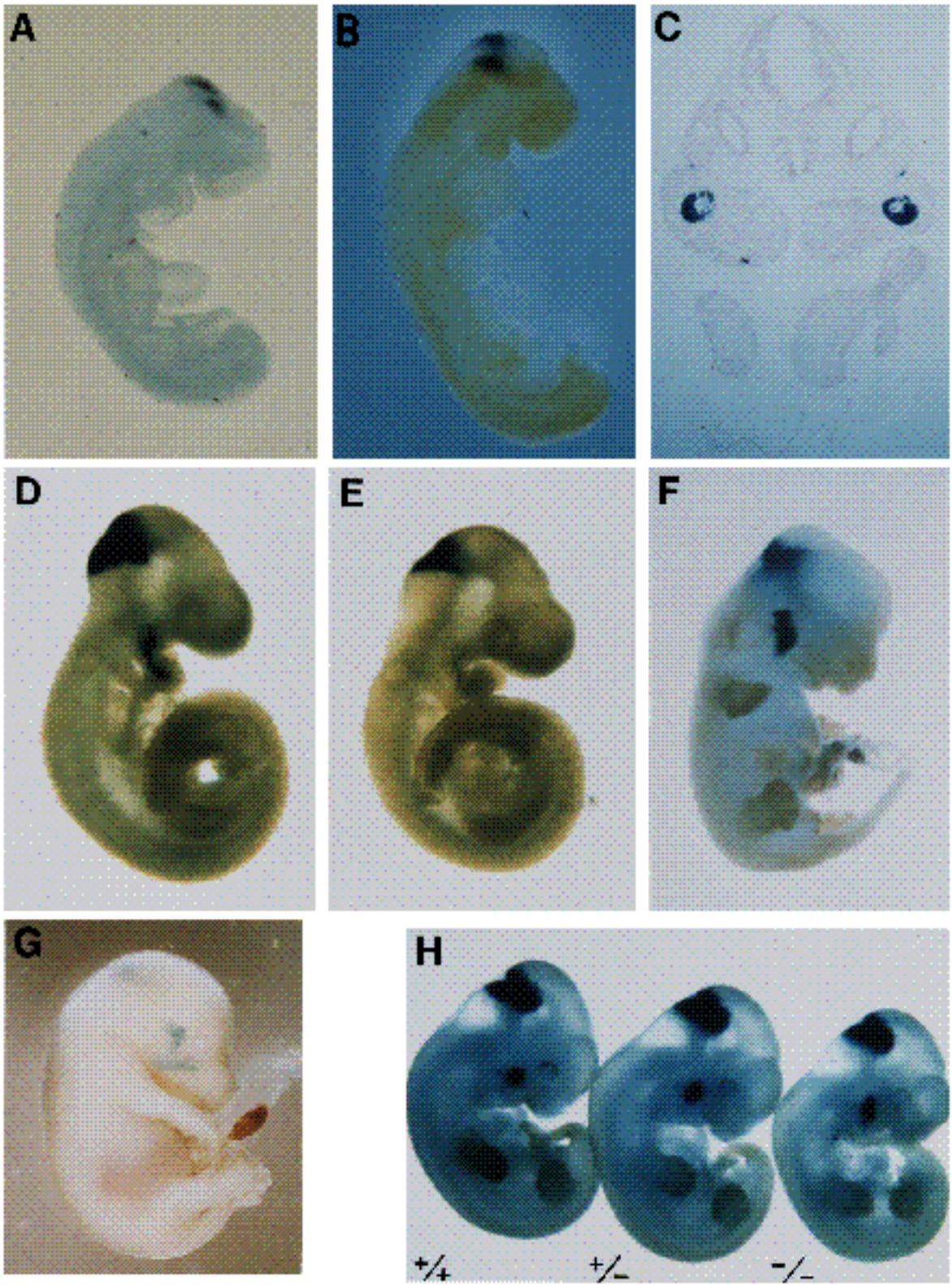

Fig. 2. Spatial and temporal pattern of β-galactosidase activity in transgenic embryos carrying the large *En-2/lacZ* transgene (construct 1; Fig. 1). (A) Whole-mount RNA in situ analysis of the endogenous *En-2* expression pattern in an 8.5 day (10- to 12-somite stage) embryo (kindly provided by Drs S. L. Ang and R. Conlon). Histochemical analysis of β-galactosidase activity in transgenic embryos from line Tg5 at 8.5 days (B), 9.5 days (D), 12.5 days (F) and 15.5 days (G). (C) Cross section through the mandibular arch region of a 10.5 day transgenic embryo from line Tg5 after X-gal staining. (E) A 9.5 day transgenic embryo from line Tg4.32 stained with X-gal. (H) *lacZ* expression in 11.5 day embryos that were hemizygous for Tg5 and either wild-type (+/+), heterozygous (+/−), or homozygous (−/−) for the *En-2*^{hd} mutation (Joyner et al., 1991).

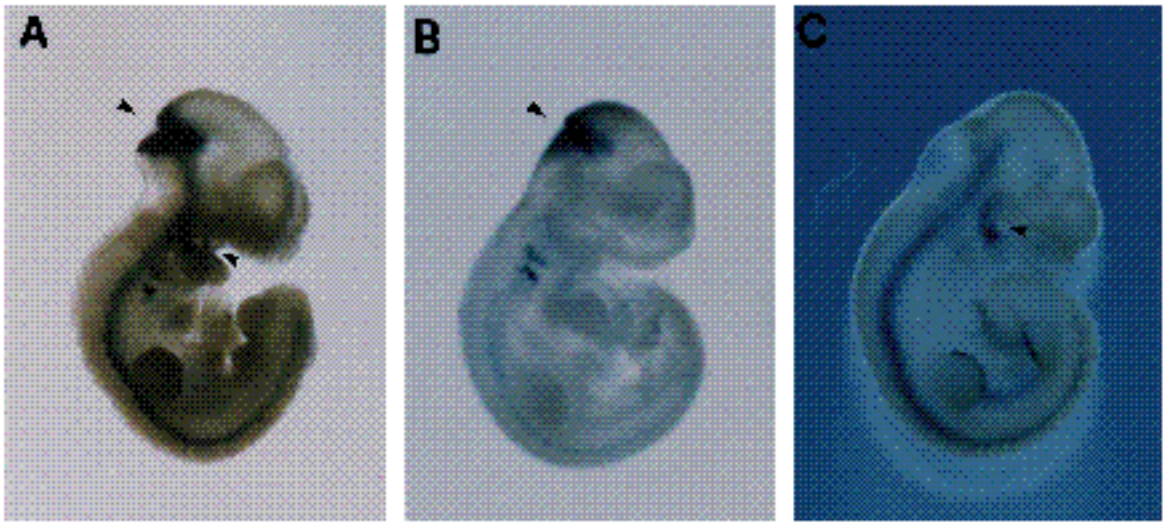

Fig. 3. Enhancer regions A and B direct expression of *lacZ* to the mid/hindbrain region and mandibular myoblasts. Histochemical analysis of the β-galactosidase activity in 10.5 day embryos carrying DNA from construct 6 (A), construct 7 (B) or construct 8 (C). Arrowheads indicate *En-2*-specific *lacZ* expression.

imately 3.7 kb 5′ of the putative *En-2* transcriptional start site (Logan et al., 1992), which correctly directed the expression of the transgene in an *En-2*-specific manner to the mid/hindbrain region and mandibular myoblasts in mid-gestation embryos (Table 1). This fragment conferred spatial specificity on a heterologous promoter and functioned in either orientation (Fig. 3A). Further analysis (summarized in Table 1) defined two regions, A and B (Fig. 1), which independently directed expression of the transgene to either the mid/hindbrain region (Fig. 3B) or mandibular myoblasts (Fig. 3C), respectively. Fragment A was a 1.5 kb genomic fragment whereas fragment B was a 1.0 kb genomic fragment. These results identify two regions (A and B; Fig. 1) 5′ of the *En-2* gene, which function as enhancers since they operate in an orientation-independent manner and confer spatial specificity onto a heterologous promoter.

As shown in Table 1, ectopic expression patterns were observed in a number of founder embryos analyzed. Most patterns were unrelated to each other and were presumably due to the sensitivity of the transgenes to regulatory elements at the site of integration. For constructs 4 to 8, consistent expression however was observed in the spinal cord, which was independent of the enhancer elements used (data not included in Table 1). When the *hsp68* promoter in construct 5 was replaced with the *En-2* promoter, the spinal cord expression was lost and the *En-2*-specific expression pattern was retained (D. L. Song and A. L. J., unpublished observations). We therefore suggest that the *hsp68* promoter fragment used in these studies contains an element that is capable of directing expression to the spinal cord and that such expression is only detectable when the promoter is flanked by a strong enhancer element. Similar expression has also been seen in lines carrying a transgene in which several copies of a retinoic acid response element were placed upstream of the p*hspPTlacZpA* vector (J. Rossant, personal communication).

For constructs 5, 6 and 7, a consistent pattern of *lacZ* expression (data not included in Table 1) was also seen in three distinct groups of cells posterior to the first branchial arch (eg. Fig. 3A,B). Similar expression was seen transiently between 9.5 and 10.5 days in four out of the five founder lines carrying DNA from construct 1. Thus, the 1.5 kb genomic fragment (region A) appears to contain elements capable of directing expression to these sites. These may represent sites of expression of the endogenous *En-2* gene that were undetected in our previous RNA in situ and immunohistochemical analysis.

En-2/lacZ transgene expression in the adult brain

We have analyzed the expression pattern of the largest *En-2/lacZ* transgene (construct 1) in the adult brain to determine if it is appropriately expressed in comparison to the endogenous *En-2* gene. Cryostat sections of adult brain tissue from four of the five transgenic lines that expressed *lacZ* embryonically were analyzed for β-galactosidase activity (see Materials and Methods). As summarized in Table 2, all four lines expressed the transgene in cells of the external molecular and internal granular layers of the cerebellum and not in the Purkinje cells. Outside the cerebellum, a consistent pattern of transgene expression was seen in the pons region in all lines examined. As shown in Fig. 4, expression extended from the junction of the cerebellum and pons rostrally into the midbrain in a pattern that is similar to that of the endogenous *En-2* gene (Davis et al., 1988; Davis and Joyner, 1988). Interestingly, rostral to the cerebellum and dorsal to the pons, *lacZ* expression was also seen in scattered cells within the colliculi in all of the lines analyzed. Endogenous *En-1* and *En-2* transcripts were not detected in this region in previous RNA in situ analyses of adult brain tissue although they were present at high levels during embryogenesis and postnatally (Davis et al., 1988; Davis and Joyner, 1988). Recent immunohistochemical analysis, using an antibody that detects both *En-1* and

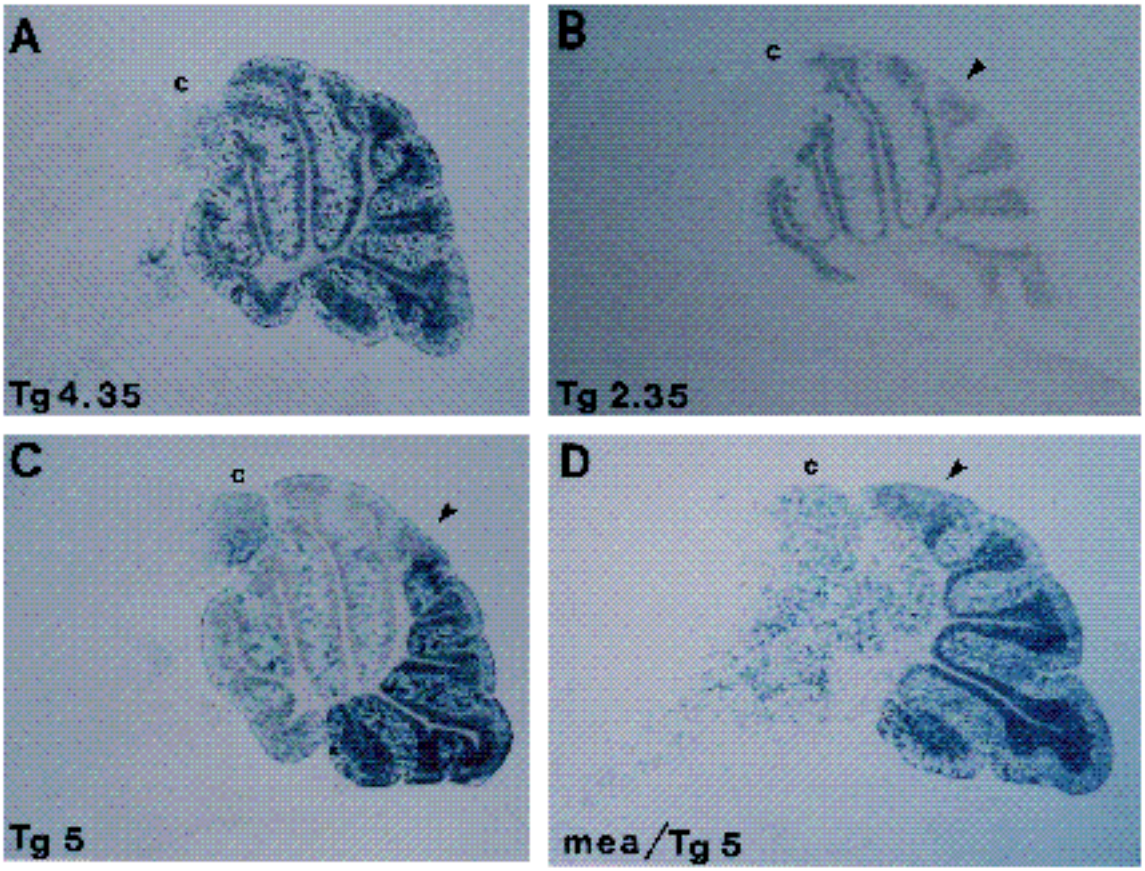

Fig. 4. *En-2/lacZ* transgene expression marks compartments in the adult cerebellum. Mid-sagittal sections of adult brain tissue (cerebellum and pons region shown) from lines Tg4.35 (A), Tg2.35 (B) and Tg5 (C) carrying DNA from construct 1 stained for β-galactosidase activity. Analysis of line Tg4.35 was done prior to segregation of the two integrations, a and b (see Materials and Methods). (D) Transgene expression pattern in the adult brain of mice homozygous for *mea*^{2J} and hemizygous for Tg5 (mea/Tg5). The arrowheads indicate the transition in intensity of *lacZ* staining between the anterior and posterior halves of the cerebellum. c, colliculi.

Table 2. Transgene expression in adult brain

Tg line	Cerebellum			Colliculi	Pons/ midbrain	Embryonic[4] expression
	Exp. region	Granular layer	Molecular layer			
Tg 5	vermus	+ a +++ p	+ a +++ p	+++	+++	++++
Tg 4.32	vermus	+ a[1] + p[1]	+ a + p	+	+	++
Tg 2.35	vermus and hemispheres	++ a + p	++ a + p	+	+	+
Tg 4.35	vermus and hemispheres	++++ a ++++ p	++++ a[2] ++++ p[2]	+++	+++	++++
Tg 7[3]	N/A	N/A	N/A	N/A	N/A	++

Exp., expressing
+, the number of +s indicates the intensity of *lacZ* staining
a, anterior region of cerebellum
p, posterior region of cerebellum

[1]a stronger level of *lacZ* expression was seen in the tuber and pyramis lobules
[2]*lacZ* staining in molecular layer decreased in the hemispheres
[3]adult expression pattern was not analyzed (N/A) in line Tg 7
[4]the level was judged by the width of the mid/hindbrain band

En-2 protein (αEnhb-1; Davis et al., 1991), has shown that *En* protein is indeed present within the colliculi of the adult mouse brain (K. J. Millen and A. L. J., unpublished data). Overall, *lacZ* expression in the adult brain was more readily detectable in lines that expressed the transgene at higher levels during embryogenesis as judged by a broader band across the mid/hindbrain junction.

Although in the cerebellum the cell types that expressed *lacZ* were those expected for *En-2*, the spatial pattern of *lacZ* expression within the cerebellar folia varied consider-

ably from line to line. In two lines, Tg5 and Tg4.32, expression was only detected in the medial lobes (vermus) of the cerebellum whereas, in Tg2.35 and Tg4.35, expression was detected in both the medial and lateral lobes (hemispheres). In addition, reproducible differences were seen in the intensity of *lacZ* staining between various folia. In Tg5, expression was much stronger in folia in the posterior half of the cerebellum (Fig. 4A) whereas in Tg2.35 expression was stronger in the anterior half (Fig. 4B). The tuber and pyramis lobules stained more intensely in Tg4.32. Only one line, Tg4.35, expressed the transgene at similar levels within all the various folia, consistent with the endogenous *En-2* expression pattern in the adult cerebellum.

Transgene expression patterns were also analyzed in the adult brain of each of the founder animals. The patterns described above for each of the four lines were identical to those seen in the adult brains of their respective founders. The remaining six founders had little or no detectable *lacZ* activity within the adult cerebellum, pons or colliculi. Ectopic *lacZ* expression was occasionally seen in some of the lines and/or founders within the frontal cortex.

En-2/lacZ transgene expression patterns mark domains within the adult cerebellum

In two of our transgenic lines, Tg2.35 and Tg5, the intensity of *lacZ* staining was reproducibly stronger within the anterior or posterior halves of the cerebellum, respectively. For both lines, the transition in intensity occurred within the declive, dividing the cerebellum into either a strongly stained anterior half and a less intensely stained posterior half (Tg2.35; Fig. 4B) or vice versa (Tg5; Fig. 4C).

The phenotype of a number of naturally occurring mouse mutants had previously suggested that the cerebellum was divided into anterior and posterior compartments. One of these mouse mutants, *meander tail* (*mea*) (Hollander and Waggie, 1977), displays normal foliation and cytoarchitecture in the posterior folia, whereas the anterior folia are completely disorganized (Ross et al., 1990). As in our two transgenic lines, a boundary exists within the declive separating the affected anterior and the non-affected posterior areas. To determine more precisely whether the boundary defined morphologically in *mea* homozygous mutants was the same as that seen in our lines, we generated mice that were homozygous for *mea* and hemizygous for one of the transgenes (Tg5). As shown in Fig. 4D, the expression pattern of the transgene on the *mea* background matches that seen in the original transgenic line (Fig. 4C) and the boundary in *lacZ* expression is similar in position to the boundary defined morphologically in *mea* (Ross et al., 1990).

En-2 does not appear to be autoregulated

Since *En-2* is itself a putative transcription factor and, in *Drosophila*, *en* is required during a certain developmental period (3 to 7 hours) to autoregulate positively its own transcription (Heemsterk et al., 1991), it was of interest to address the question of autoregulation at the *En-2* locus in mouse. Our *En-2/lacZ* transgenes provide a tool with which to address this question. We generated mice (see Materials and Methods) that were hemizygous for the largest *En-2/lacZ* transgene (construct 1) and homozygous for a targeted mutation, *En-2^{hd}*, which deletes the homeobox of the *En-2* gene (Joyner et al., 1991), and analyzed the expression pattern of the transgene at various developmental stages (see Material and Methods). In the *En-2* mutant, a truncated protein containing the first 219 amino acids may still be produced from the mutant locus; however, by analogy to *Drosophila* homeodomain proteins, such a protein would lack DNA-binding activity (eg. Desplan et al., 1985, 1988; Hoey et al., 1988), presumably resulting in loss of *En* gene function.

Transgenic embryos homozygous for the *En-2^{hd}* mutation continued to express *lacZ* at each of the developmental stages analyzed between 8.5 and 12.5 days (see Materials and Methods for details). Furthermore, littermates from mid-gestation embryos (11.5 days) hemizygous for the transgene and either wild type, heterozygous or homozygous for the *En-2^{hd}* mutation showed no obvious difference in the pattern or intensity of *lacZ* staining when analyzed via whole-mount staining for β-galactosidase activity (Fig. 2H). Analysis of cryostat sections of adult brain tissue from similar littermates produced identical results suggesting that the *En-2* gene in mouse is not positively autoregulated. Slight changes in β-galactosidase activity reflecting small changes in transcription would not have been detected in our analysis.

DISCUSSION

Reporter gene constructs were analyzed in transgenic mice and a large genomic region was defined that was capable of directing the expression of *lacZ* in an *En-2*-specific manner both temporally and spatially during embryogenesis as well as to the correct cell types within the adult cerebellum, pons and colliculi. Interestingly, in the adult, transgene expression patterns within the cerebellum in two of four lines marked distinct anterior-posterior compartments. Deletion analysis of the large *En-2/lacZ* transgene and the use of a heterologous promoter identified two enhancer-containing regions 5′ of the *En-2* transcribed region, which directed expression in the embryo to either the mid/hindbrain region or mandibular myoblasts. The large *En-2/lacZ* transgene (construct 1) was also used as a tool with which to analyze *En-2* expression in fetal and adult mice lacking a functional *En-2* protein and no evidence was found for autoregulation of the mouse *En-2* gene.

The two *En-2*-specific enhancer regions identified in this study lie adjacent to one another approximately 3.7 kb 5′ of the putative *En-2* transcriptional start site (Logan et al., 1992). One (region A, Fig. 3), located furthest upstream, is a 1.5 kb DNA fragment, which directs expression to a restricted domain within the CNS spanning the mid/hindbrain junction, whereas the second, more proximal region (B, Fig. 3), is a 1.0 kb DNA fragment directing expression to presumptive myoblasts within the first branchial arch. Recent analysis of the temporal and spatial pattern of transgene expression in two lines carrying construct 5, which contains both enhancer regions, showed that it is correctly expressed in the CNS and mandibular regions during embryonic development (D.-L. Song and A. L. J., unpublished data). Taken together, these results suggest a simple

regulatory model in which separate enhancer elements near the *En-2* gene function to direct its expression in a spatially and temporally restricted manner to two different regions within the embryo and suggest that a third as yet undefined embryonic enhancer region exists that directs expression to the region surrounding the developing pituitary. This element is presumably missing from construct 1, which contains the largest genomic region tested.

In the adult brain, the large *En-2/lacZ* transgene (construct 1) was expressed, as expected, in cells of the molecular and granular layers of the cerebellum as well as in specific groups of cells in the pons and colliculi regions. In accordance with the above model for the embryonic regulation of *En-2* gene expression, these results suggest that the large transgene contains *En-2*-specific enhancer element(s) that can correctly direct expression of *lacZ* to particular cell types within the adult brain. Further analysis of the expression patterns of constructs 2 to 8 in the adult brain is necessary to determine the location and nature of such element(s). Since fate mapping studies in birds show that the cerebellum is derived from the mes-metencephalic region of the neural tube (Hallonet et al., 1990), it is possible that the enhancer region directing expression of *En-2* to this region during embryogenesis also directs expression to specific cell types within the adult cerebellum. Alternatively, the embryonic region-specific enhancer element(s) may well be separate from the adult cell-type-specific enhancer element(s).

Although the transgene was expressed in the correct cell types within the adult cerebellum, interesting variations were seen in the spatial pattern of transgene expression between various lines. Transgene expression patterns within the adult cerebellum in two of the four lines analyzed marked distinct anterior or posterior compartments that matched those defined morphologically in *mea* mutant mice (Ross et al., 1990). Since it is unlikely that the variation would be due to such cerebellar enhancers present at each of the integration sites, these results suggest that, in addition to cell-type-specific enhancers, multiple spatial- or compartment-specific enhancers important in defining expression patterns within the adult cerebellum are contained within the transgene. If this is the case, then the variations observed in the spatial pattern of transgene expression may reflect the sensitivity of such elements to position effects. Alternatively, additional elements required for a reproducibly high level of *En-2* expression throughout the adult cerebellum may not be present on the genomic region tested.

The 1.5 kb DNA fragment (region A), which directs expression to the mid/hindbrain region, represents the most anterior, region-specific neural enhancer identified to date within the developing CNS. As discussed above, it likely contains all of the element(s) required for the spatial and temporal control of *En-2* gene expression within the embryonic CNS. In contrast, multiple positive as well as negative regulatory regions, which are shared between neighbouring genes, are required to similarly reconstruct endogenous *Hox* gene expression patterns (Bieberich et al., 1990; Puschel et al., 1990, 1991; Whiting et al., 1991; Sham et al., 1992). This requirement provides a regulatory basis for their highly conserved clustered organization. Unlike the *Hox* genes, the vertebrate *En* genes are not organized in clusters but map to separate chromosomes (Joyner and Martin, 1987; Logan et al., 1989). Further analysis is necessary to determine whether multiple *En-2* regulatory elements reside on the 1.5 kb fragment which independently regulate the spatial and temporal pattern seen.

The 1.0 kb DNA fragment (region B), which directs expression to presumptive myoblasts within the first branchial arch, represents a novel muscle-specific enhancer. Recent RNA in situ analysis has shown that both the mouse *En-1* and *En-2* genes are expressed at low levels in 9.5 day embryos in presumptive myoblasts within the first branchial arch (C. C. Hui and A. L. J., unpublished data). In contrast to other regulatory genes known to be involved in skeletal muscle formation in mouse, such as members of the *myoD* gene family (reviewed in Buckingham et al., 1992), *En-2*, based on the transgene expression pattern presented here, appears to mark a specific set of craniofacial muscles and their precursors as has been demonstrated in zebrafish (Hatta et al., 1990). Furthermore, it is expressed at an earlier developmental stage than *myf-5*, the earliest known molecular marker of myogenesis in the mouse embryo, which is first detectable in the visceral arches beginning at 10 days (Ott et al., 1991). Taken together, these results suggest that one or both *En* genes are involved in the specification of muscle identity in mammals.

By comparative sequence analysis, Renucci et al. (1992) recently identified highly conserved regions upstream of the mouse and human *Hox 4.4* homologues and subsequently used reporter gene constructs to show that these regions function as spatially restricted enhancers in transgenic mouse embryos. We have similarly identified highly conserved regions located immediately upstream and within the untranslated regions of the mouse and human *En-2* genes (Logan et al., 1992). However, these regions do not function as embryonic enhancers as no *En-2*-specific expression was detected when they were tested using transgenic mouse embryos (constructs 2 and 3). The conserved sequences identified 5′ of the *En-2* gene are presumably involved in promoter activity and not in directing the *En-2*-specific expression pattern, whereas the conserved sequences within the untranslated regions may be involved in fine tuning the regulation of *En-2* gene expression, perhaps at the level of mRNA stability and/or efficiency of polyadenylation. Such regulatory effects would not be easily detected in our transgenic studies in which only the steady state level of β-galactosidase activity was analyzed. It will be of interest to compare the mouse genomic sequence of the two enhancer regions (A and B) identified in this study with the corresponding regions from the human *EN2* gene.

In *Drosophila*, the regulation of *en* gene expression during embryogenesis is complex involving at least four distinct modes of control (Heemskerk et al., 1991). Following activation by pair rule genes, both an extracellular signal provided by the *wingless* (*wg*) protein and the *en* protein itself are required for the early maintenance of *en* expression. Autoregulation of *en* then becomes independent of *wg*. Following this autoregulatory phase, new regulators, both positive and negative, have been identified that control the late expression of *en*. Interestingly, recent analysis of *En* expression in mice homozygous for a null allele of

Wnt-1, a mouse homologue of the *Drosophila* gene *wg*, demonstrated that, as in *Drosophila*, *wnt-1* is not required for the activation of *En* expression but may be required for its maintenance (McMahon et al., 1992), suggesting that the *En* signalling pathway has been conserved during evolution. It was of interest therefore, to explore the possibility of autoregulation at the *En-2* locus in mouse using the large transgene in *En-2*^{hd} homozygous mutant mice. Our results provided no functional evidence for autoregulation. It is possible that, in the embryo, the presence of an intact potentially redundant *En-1* protein accounts for the continued expression of the transgene. However, in the adult cerebellum, no *En-1* protein is present to regulate expression of the transgene. In contrast, recent experimental evidence suggesting conservation of an autoregulatory feed-back loop between *Drosophila* and mouse has been obtained for a homeobox-containing gene of the *Hox/HOM-C* gene complex (Awgulewitsch and Jacobs, 1992; Malicki et al., 1992).

The identification of *cis*-acting DNA regulatory regions capable of correctly directing the expression of a reporter gene in an *En-2*-specific manner during embryogenesis represents the first step in defining the molecular mechanisms that establish and maintain its expression. Further deletion mapping of these regions should uncover shorter DNA sequence(s) involved in the regulation of *En-2* that can subsequently be used in a biochemical analysis to identify interacting proteins. Since *En* is one of the few early markers of regionalization in the anterior neural epithelium, such proteins may play a more global role in establishment of the A-P pattern in the vertebrate CNS. In addition, the *En-2*-specific enhancer regions identified in this study provide valuable tools with which to express selectively other developmentally important genes in the embryonic mid/hindbrain region or in mandibular myoblasts to gain valuable insight into their biological functions.

We thank the members of the laboratory and Dr J. Rossant for many helpful discussions and D. Nallainathan, K. Harpal and M. Chiu for excellent technical help. This work was funded by grants to A. L. J. from the Medical Research Council (MRC) and National Cancer Institute of Canada and Bristol-Myers Squibb Ltd. C. L. was supported by an MRC Studentship and A. L. J. is an MRC Scientist and a Howard Hughes International Scholar.

REFERENCES

Alvarado-Mallart, R. M., Martinez, S. and Lance-Jones, C. C. (1990). Pluripotentiality of the 2-day old avian germinative neuroepithelium. *Dev. Biol.* **139**, 75-88.

Awgulewitsch, A. and Jacobs, D. (1992). Deformed autoregulatory element from *Drosophila* functions in a conserved manner in transgenic mice. *Nature* **358**, 341-344.

Bally-Cuif, L., Alvarado-Mallart, R. M., Darnell, D. K. and Wassef, M. (1992). Relationship between *Wnt-1* and *En-2* expression domains during early development of normal and ectopic met-mesencephalon. *Development* **115**, 999-1009.

Bieberich, C. J., Utset, M. F., Awgulewitsch, A. and Ruddle, F. H. (1990). Evidence for positive and negative regulation of the *Hox-3.1* gene. *Proc. Natl. Acad. Sci. USA* **87**, 8462-8466.

Buckingham, M. E., Lyons, G. E., Ott, M.-O. and Sassoon, D. A. (1992). Myogenesis in the mouse. *Ciba Foundation Symposium* **165**, 111-131. Chichester: Wiley.

Conlon, R. A. and Rossant J. (1992). Exogenous retinoic acid rapidly induces anterior ectopic expression of murine *Hox-2* genes in vivo. *Development* (in press)

Davis, C. A., Holmyard, D. P., Millen, K. J. and Joyner, A. L. (1991). Examining pattern formation in mouse, chicken and frog embryos with an *En*-specific antiserum. *Development* **111**, 287-298.

Davis, C. A., Noble-Topham, S. E., Rossant, J. and Joyner, A. L. (1988). Expression of the homeo box-containing gene *En-2* delineates a specific region of the developing mousebrain. *Genes Dev.* **2**, 361-371.

Davis, C. A. and Joyner, A. L. (1988). Expression patterns of the homeo box-containing genes *En-1* and *En-2* and the proto-oncogene *int-1* diverge during mouse development. *Genes Dev.* **2**, 1736-1744.

Desplan, C., Theis, J. and O'Farrell, P. H. (1985). The *Drosophila* developmental gene, *engrailed*, encodes a sequence-specific DNA binding activity. *Nature* **318**, 630-635.

Desplan, C., Theis, J. and O'Farrell, P. H. (1988). The sequence specificity of homeodomain-DNA interaction. *Cell* **54**, 1081-1090.

Ekker, M., Wegner, J., Akimenko, M. A. and Westerfield, M. (1992). Coordinate embryonic expression of three zebrafish *engrailed* genes. *Development* (in press).

Feinberg, A. and Vogelstein, B. (1983). A technique for radiolabeling DNA restriction endonuclease fragments to high specific activity. *Analyt. Biochem.* **132**, 6-13.

Gardner, C. A. and Barald, K. F. (1991). The cellular environment controls the expression of *engrailed*-like protein in the cranial neuroepithelium of quail-chick chimeric embryos. *Development* **113**, 1037-1048.

Gardner, C. A. and Barald, K. F. (1992). Expression patterns of *engrailed*-like proteins in the chick embryo. *Dev. Dynam.* **193**, 370-388.

Gardner, C. A., Darnell, D. K., Poole, S. J., Ordahl, C. P. and Barald, K. F. (1988). Expression of an *engrailed*-like gene during development of the early embryonic chick nervous system. *J. Neurosci. Res.* **21**, 426-437.

Gossler, A., Joyner, A. L., Rossant, J. and Skarnes, W. (1989). Mouse embryonic stem cells and reporter constructs detect developmentally regulated genes. *Science* **244**, 463-465.

Gruss, P. and Walther, C. (1992) Pax in development. *Cell* **69**, 719-722.

Hallonet, M. E. R., Teillet, M.-A. and Le Douarin, N. M. (1990). A new approach to the development of the cerebellum provided by the quail-chick marker system. *Development* **108**, 19-31.

Hatta, K., Schilling, T. F., BreMiller, R. A. and Kimmel, C. B. (1990). Specification of jaw muscle identity in zebrafish: correlation with *engrailed*-homeoprotein expression. *Science* **250**, 802-805.

Heemskerk, J., DiNardo, S., Kostriken, R. and O'Farrell, P. H. (1991). Multiple modes of *engrailed* regulation in the progression towards cell fate determination. *Nature* **352**, 404-410.

Hemmati-Brivanlou, A., de la Torre, J. R., Holt, C. and Harland, R. M. (1991). Cephalic expression and molecular characterization of *Xenopus En-2*. *Development* **111**, 715-724.

Hoey, T., Warrior, R., Manak, J. and Levine, M. (1988). DNA-binding activities of the *Drosophila melanogaster even-skipped* protein are mediated by its homeo domain and influenced by protein context. *Mol. Cell Biol.* **8**, 4598-4607.

Hogan, B., Costantini, F. and Lacy, E. (1986). *Manipulating the mouse embryo. A laboratory manual*. Cold Spring Harbor, New York: Cold Spring Harbor Laboratory.

Hollander, W. F. and Waggie, K. S. (1977). Meander tail, a recessive mutant located on chromosome 4 of the mouse. *J. Hered.* **68**, 403-406.

Itasaki, N., Ichijo, H., Hama, C., Matsuno, T. and Nakamura, H. (1991). Establishment of rostrocaudal polarity in tectal primordium: *engrailed* expression and subsequent tectal polarity. *Development* **113**, 1133-1144.

Jaynes, J. B. and O'Farrell, P. (1988). Activation and repression of transcription by homoeodomain-containing proteins that bind a common site. *Nature* **336**, 744-748.

Jaynes, J. B. and O'Farrell, P. (1991). Active repression of transcription by the *engrailed* homeodomain protein. *EMBO J.* **10**, 1427-1433.

Joyner, A. L. and Hanks, M. (1991). The engrailed genes: evolution of function. *Sem. Dev. Biol.* **2**, 435-445.

Joyner, A. L. and Martin, G. R. (1987). *En-1* and *En-2*, two mouse genes with sequence homology to the *Drosophila engrailed* gene: expression during embryogenesis. *Genes Dev.* **1**, 29-38.

Joyner, A. L., Kornberg, T., Coleman, K. G., Cox, D. R. and Martin, G. R. (1985). Expression during embryogenesis of a mouse gene with sequence homology to the *Drosophilaengrailed* gene. *Cell* **43**, 29-37.

Joyner, A. L., Herrup, K., Auerbach, B. A., Davis, C. A. and Rossant, J.

(1991). Subtle cerebellar phenotype in mice homozygous for a targeted deletion of the *En-2* homeobox. *Science* **251**, 1239-1243.

Kornberg, T. (1981). *Engrailed*; a gene controlling compartment and segment formation in *Drosophila*. *Proc. Natl. Acad. Sci. USA* **78**, 1095-1099.

Kothary, R., Clapoff, S., Darling, S., Perry, M. D., Moran, L. A. and Rossant, J. (1989). Inducible expression of an *hsp68-lacZ* hybrid gene in transgenic mice. *Development* **105**, 707-714.

Kozak, M. (1983). Comparison of initiation of protein synthesis in procaryotes, eukaryotes and organelles. *Micro Rev.* **47**, 1-45.

Kress, C., Vogels, R., De Graaff, W., Bonnerot, C., Meijlink, F., Nicolas, J. F. and Deschamps, J. (1990). *Hox-2.3* upstream sequences mediate *lacZ* expression in intermediate mesoderm derivatives of transgenic mice. *Development* **109**, 775-786.

Logan, C., Willard, H. F., Rommens, J. M. and Joyner, A. L. (1989). Chromosomal localization of the human homeo box-containing genes, *EN1* and *EN2*. *Genomics* **4**, 206-209.

Logan, C., Hanks, M. C, Noble-Topham, S., Nallainathan, D., Provart, N. J. and Joyner, A. L. (1992). Cloning and sequence comparison of the mouse, human and chick *Engrailed* genes reveal potential functional domains and regulatory regions. *Dev. Gen.* (in press).

Malicki, J., Cianetti, L. C., Peschle, C. and McGinnis, W. (1992). A human *HOX4B* regulatory element provides head-specific expression in *Drosophila* embryos. *Nature* **358**, 345-347.

Martinez, S. and Alvarado-Mallart, R. M. (1989). Rostral cerebellum originates from the caudal portion of the so-called 'mesencephalic' vesicle: a study using chick/quail chimeras. *Eur. J. Neurosci.* **1**, 549-560.

Martinez, S. and Alvarado-Mallart, R. M. (1990). Expression of the homeobox *chick-en* gene in chick/quail chimeras with inverted mesmetencephalic grafts. *Dev. Biol.* **139**, 432-436.

Martinez, S., Wassef, M. and Alvarado-Mallart, R. M. (1991). Induction of a mesencephalic phenotype in the 2-day old chick prosencephalon is preceded by the early expression of the homeobox gene *en*. *Neuron* **6**, 971-981.

McGinnis, W. and Krumlauf, R. (1992). Homeobox genes and axial patterning. *Cell* **68**, 283-302.

McMahon, A. P., Joyner, A. L., Bradley, A. and McMahon, J. A. (1992). The midbrain-hindbrain phenotype of *Wnt-1⁻/Wnt-1⁻* mice results from stepwise deletion of *engrailed*-expressing cells by 9.5 days postcoitum. *Cell* **69**, 581-595.

Morata, G. and Lawrence, P. A. (1975). Control of compartment development by the *engrailed* gene of *Drosophila*. *Nature* **255**, 614-617.

Noden, D. (1991). Vertebrate craniofacial development: the relation between ontogenetic process and morphological outcome. *Brain Behav. Evol.* **38**, 190-225.

Nusse, R. and Varmus, H. E. (1992). *Wnt* genes. *Cell* **69**, 1073-1088.

Ohkuma, Y., Horikoshi, M., Roeder, R. G. and Desplan, C. (1990). Binding site-dependent direct activation and repression of in vitro transcription by *Drosophila* homeodomain proteins. *Cell* **61**, 475-484.

Ott, M.-O., Bober, E., Lyons, G., Arnold, H. and Buckingham, M. (1991). Early expression of the myogenic regulatory gene, *myf-5*, in precursor cells of skeletal muscle in the mouse embryo. *Development* **111**, 1097-1107.

Patel, N. H., Martin-Blanco, E., Coleman, K. G., Poole, S. J., Ellis, M. C., Kornberg, T. B. and Goodman, C. S. (1989). Expression of *engrailed* proteins in arthropods, annelids, and chordates. *Cell* **58**, 955-968.

Puschel, A. W., Balling, R. and Gruss, P. (1990). Position-specific activity of the *Hox1.1* promoter in transgenic mice. *Development* **108**, 435-442.

Puschel, A. W., Balling, R. and Gruss, P. (1991). Separate elements cause lineage restriction and specify boundaries of *Hox-1.1* expression. *Development* **112**, 279-287.

Renucci, A., Zappavigna, V., Zakany, J., Izpisua-Belmonte, J. C., Burki, K. and Duboule, D. (1992). Comparison of mouse and human *HOX-4* complexes defines conserved sequences involved in the regulation of *Hox-4.4*. *EMBO J.* **11**, 1459-1468.

Ross, M. E., Fletcher, C., Mason, C. A., Hatten, M. E. and Heintz, N. (1990). Meander tail reveals a discrete developmental unit in the mouse cerebellum. *Proc. Natl. Acad. Sci. USA* **87**, 4189-4192.

Schughart, K., Bieberich, C. J., Eid, R. and Ruddle, F. H. (1991). A regulatory region from the mouse *Hox-2.2* promoter directs gene expression into developing limbs. *Development* **112**, 807-811.

Sham, M. H., Hunt, P., Nonchev, S., Papalopulu, N., Graham, A., Boncinelli, E. and Krumlauf, R. (1992). Analysis of the murine *Hox-2.7* gene: conserved alternative transcripts with differential distributions in the nervous system and the potential for shared regulatory regions. *EMBO J.* **11**, 1825-1836.

Tuggle, C. K., Zakany, J., Cianetti, L., Peschle, C. and Nguyen-Huu, M. C. (1990). Region-specific enhancers near two mammalian homeo box genes define adjacent rostrocaudal domains in the central nervous system. *Genes Dev.* **4**, 180-189.

Whiting, J., Marshall, H., Cook, M., Krumlauf, R., Rigby, P. W. J., Stott, D. and Allenmann, R. K. (1991). Multiple spatially specific enhancers are required to reconstruct the pattern of *Hox-2.6* gene expression. *Genes Dev.* **5**, 2048-2059.

Zakany, J., Tuggle, C. K., Patel, M. D. and Nguyen-Huu, M. C. (1988). Spatial regulation of homeobox gene fusions in the embryonic central nervous system of transgenic mice. *Neuron* **1**, 679-691.

Zakany, J., Tuggle, C. K. and Nguyen-Huu, C. M. (1990). The use of *lacZ* gene fusions in the studies of mammalian development. developmental regulation of mammalian homeobox genes in the CNS. *J. Physiol.* **84**, 21-26.

(Accepted 17 November 1992)