DEVELOPMENTAL BIOLOGY **180,** 227–241 (1996)
ARTICLE NO. 0297

# Restricted Expression of a Novel Murine *atonal*-Related bHLH Protein in Undifferentiated Neural Precursors

**Gérard Gradwohl, Carol Fode, and François Guillemot[1]**

*Institut de Génétique et de Biologie Moléculaire et Cellulaire, CNRS/INSERM/Université Louis Pasteur, B.P.163, 67404 Illkirch Cédex, C.U. de Strasbourg, France*

**Tissue-specific bHLH proteins play important roles in the specification and differentiation of neural cell lineages in invertebrate and vertebrate organisms. Two groups of bHLH proteins, atonal and achaete-scute, have proneural activities in *Drosophila*, and the mouse achaete-scute homolog MASH1 is required for the differentiation of several neural lineages. In a screen for proteins interacting with MASH1, we have isolated a novel bHLH protein related to atonal, named MATH4A, which is broadly expressed in neural precursor cells in the mouse embryonic CNS and PNS. Interaction assays in yeast and *in vitro* demonstrate that MATH4A interacts efficiently with both MASH1 and the ubiquitous bHLH protein E12. MATH4A-E12 heterodimers, but not MATH4A-MASH1, bind to a consensus E-box sequence. *Math4A* expression is restricted to undifferentiated neural precursors and is complementary to that of *Mash1* in most regions of the nervous system. In particular, *Math4A* is transcribed at high levels in the cerebral cortex, dorsal thalamus, and epibranchial placodes, which present little or no *Mash1* expression. However, expression of the two genes shows limited overlap in certain CNS regions (retina, preoptic area of the hypothalamus, midbrain, hindbrain). Its structure and expression pattern suggest that MATH4A may regulate an early step of neural development, either as a partner of ubiquitous bHLH proteins or associated with other neural-specific bHLH proteins.** 

## INTRODUCTION

Transcription factors of the basic helix-loop-helix (bHLH) class (Murre *et al.*, 1989) play important roles in the determination and differentiation of cell lineages in vertebrate and invertebrate species (Jan and Jan, 1993; Olson and Klein, 1994). The functions of bHLH proteins in the development of neural lineages have been extensively examined, particularly in *Drosophila*. In this organism, several families of proneural genes regulate the production of neural precursors in the peripheral nervous system (PNS) and central nervous system (CNS). Genes of the *achaete-scute* family encode bHLH proteins required for the development of external sensory organs and subsets of precursors in the CNS (Campuzano and Modolell, 1992). The proneural gene *atonal* also encodes a bHLH factor required for development of chordotonal organs and photoreceptors (Jarman *et al.*, 1993, 1994).

*Mash1* is the only known mouse gene homologous to *achaete-scute* genes which is expressed in the developing nervous system. *Mash1* transcripts and protein are present in proliferating precursors in the CNS and the autonomic nervous system (Lo *et al.*, 1991; Guillemot and Joyner, 1993). A null mutation of *Mash1* in mice leads to a specific loss of olfactory and sympathetic neurons (Guillemot *et al.*, 1993). In *Mash1* mutant embryos, neural precursors are being generated but their development is blocked before the stage of terminal differentiation (Sommer *et al.*, 1995). Recently, several mammalian genes related to *Drosophila atonal* have also been isolated. The genes *MATH1*, *MATH2/Nex1*, *BETA3*, and *kw8/NDRF* were identified by sequence conservation within their bHLH domains (Bartholomä and Nave, 1994; Akazawa *et al.*, 1995; Shimizu *et al.*, 1995; Kume *et al.*, 1996; Peyton *et al.*, 1996; Yasunami *et al.*, 1996), and *NeuroD/BETA2* by its ability to dimerize with ubiquitous bHLH factors and bind to a specific DNA sequence in yeast (Lee *et al.*, 1995; Naya *et al.*, 1995). *NeuroD* was also shown to activate neuronal differentiation in ven-

Sequence data from this article have been deposited with the EMBL/GenBank and DDBJ Data Libraries under Accession No. Y07621.

[1] To whom correspondence should be addressed. E-mail: francois@igbmc.u-strasbg.fr.

0012-1606/96 $18.00

tral ectoderm cells, by forced expression in *Xenopus* embryos (Lee *et al.*, 1995). The expression patterns of the vertebrate *atonal*-related genes suggest that these genes are required in different regions and at different stages of development of the nervous system. *NeuroD* transcripts are present in cranial and dorsal root ganglia, in differentiating neurons in several regions of the CNS, and in pancreatic endocrine cell lines (Lee *et al.*, 1995; Naya *et al.*, 1995). *MATH1* is expressed in the dorsal neural tube and the external granular layer of the cerebellum during embryonic and early postnatal development and in adult intestinal epithelial cells (Akazawa *et al.*, 1995). *MATH2/NEX1* is transcribed in postmitotic neurons in the dorsal telencephalon and in cerebellar granular cells, in both embryos and adults (Bartholomä and Nave, 1994; Shimizu *et al.*, 1995). kw8/NDRF is broadly expressed in adult brain (Kume *et al.*, 1996; Yasunami *et al.*, 1996). Thus, the mouse *atonal*-related genes described so far are mostly expressed at later stages in neural lineages than *Mash1*.

The activity of bHLH transcription factors requires their dimerization and binding to sequences that contain the consensus motif CANNTG, referred to as an E-box (Murre *et al.*, 1989). Tissue-specific bHLH proteins dimerize preferentially with ubiquitous bHLH proteins such as E12. They usually homodimerize or heterodimerize poorly with other tissue-specific bHLH proteins (Cabrera and Alonso, 1991; Sun and Baltimore, 1991; Brown and Baer, 1994; Hsu *et al.*, 1994), although the neural bHLH protein HEN1 has been shown to self-associate efficiently (Brown and Baer, 1994). The DNA binding properties of bHLH proteins are acquired upon dimerization (Murre *et al.*, 1989). Binding site selection experiments (Blackwell and Weintraub, 1990) and crystallographic studies (Ferré-D'Amaré *et al.*, 1993; Ellenberg *et al.*, 1994; Ma *et al.*, 1994) have shown that each bHLH partner of a dimer binds predominantly to one half of the E-box, suggesting that the specific composition of a bHLH dimer could influence its DNA-binding specificity. Indeed, distinct DNA-binding properties have been reported for homodimers of the tissue-specific protein HEN1, compared to heterodimers of HEN1 and the ubiquitous protein E12, or to E12 homodimers (Brown and Baer, 1994). In addition, dimerization with the Id HLH proteins, which lack a basic region, abolishes the DNA-binding activity of bHLH proteins (Benezra *et al.*, 1990). The activities of bHLH proteins have also been shown to be modulated by interactions with other classes of nuclear factors, such as MADS domain proteins (Molkentin *et al.*, 1995), leucine zipper proteins (Bengal *et al.*, 1992), and LIM domain proteins (Larson *et al.*, 1996).

In order to identify factors that could regulate MASH1 activity, we have performed a two-hybrid screen in yeast using the bHLH domain of MASH1 as an interaction partner. In this article, we report the cloning of a cDNA encoding a novel bHLH protein, referred to as MATH4A, which belongs to a subfamily of *atonal*-related proteins. MATH4A could dimerize with both the ubiquitous bHLH protein E12 and the neural-specific protein MASH1, but bound to an E-box consensus sequence only in association with E12. Similar to *Mash1*, *Math4A* expression was confined to undifferentiated neural precursors. In many regions of the embryonic CNS and PNS, expression of the two genes was mutually exclusive. However, there were domains of overlap, particularly in posterior regions of the brain. Thus, MATH4A may function during early stages of development of neural lineages, as a partner of ubiquitous bHLH proteins, or associated with MASH1 or other neural-specific bHLH proteins.

## MATERIALS AND METHODS

### *Plasmids*

In the yeast two-hybrid assays, LexA DNA-binding domain (DBD) fusion proteins were expressed from the pBTM116 plasmid (*TRP* marker) (Vojtek *et al.*, 1993) and VP16 transcription activation domain (TAD) fusion proteins from the pASV3 plasmid (*LEU2* marker) (Le Douarin *et al.*, 1995). The following sequences were used to produce LexA and/or VP16 fusion proteins: mouse MASH1, G71 to A181 (Guillemot and Joyner, 1993); rat E12, A507 to L649 (Nelson *et al.*, 1990); mouse Twist, S103 to H206 (Wolf *et al.*, 1991); mouse MATH4A, G81 to S200 (Fig. 1).

For *in vitro* transcription/translation experiments, an ATG start codon and a consensus Kozak sequence were introduced by PCR at the 5′ end of the cDNAs to be translated and the corresponding fragments were subcloned in pBluescript II(KS$^-$). RNAs were synthesized using the T7 RNA polymerase. The *Math4A* clone encodes the full-length MATH4A protein preceded by: MAILAGRRARFRAAGTFPDTHRSSSCAATHLEPRR, resulting from the translation of the pASV3 adaptor region and *Math4A* 5′ untranslated sequence. The *Mash1* and *NeuroD* clones encode the full-length mouse MASH1 and NeuroD proteins (Guillemot and Joyner, 1993; Lee *et al.*, 1995). The *E12* clone encodes mouse E12 starting at residue D469 of the corresponding rat sequence (Nelson *et al.*, 1990). *Twist* encodes mouse Twist from G102 to Q203 (Wolf *et al.*, 1991).

For bacterial expression of the glutathione *S*-transferase (GST) fusions proteins, constructs were made in the pGEX-4T1 vector (Pharmacia) and contained the following sequences: MASH1 (Q60–F227), MATH4A (full-length protein preceded by the peptide RARFRAAGTFDTHRSSSCAATHLEPRR), E12 (D469 to the end), Twist (G102 to Q203).

### *Yeast Interaction Screen*

Two hundred milligrams of a mouse embryonic (9.5–12.5 d.p.c.) cDNA library cloned in the yeast *LEU2* VP16 TAD expression vector (vom Baur *et al.*, 1996) was introduced by LiAc transformation into the yeast L40 reporter strain [*MATa trp1-901 leu2-3, 112 his3-D200 ade2 LYS:: (LexAop)4-HIS$_3$ URA3:: (LexAop)$_8$-LacZ;* 3] (Vojtek *et al.*, 1993) expressing the fusion protein LexA-MASH1. Yeast transformants ($4 \times 10^6$) were selected on Trp$^-$ Leu$^-$ plates and replated at a multiplicity of 10 onto Trp$^-$ Leu$^-$ His$^-$ plates containing 25 m*M* 3-amino-1, 2, 4-triazole (3AT). After 6 days of incubation, 626 clones were isolated. Two hundred clones were analyzed further by purification, conversion to the plasmid form, and retransformation with LexA-MASH1, LexA-E12, LexA-Twist, LexA-Lamin, (Vojtek *et al.*, 1993) or unfused LexA to test for specificity of interactions. The clones showing a specific interaction with LexA-MASH1 were sequenced.

### β-Galactosidase Assay

Expression plasmids encoding LexA and VP16 fusion proteins were introduced sequentially in the yeast strain L40 by electroporation. Three independent transformants were grown in selective liquid media lacking Leu and Trp. β-Galactosidase activity was measured as described (Ausubel *et al.*, 1987). Efficient translation of the fusion proteins was verified by Western blot using specific antibodies against VP16 and LexA (data not shown). The ability of the different bHLH fusion proteins to dimerize was monitored by interaction with either LexA-E12 or VP16-E12.

### In Vitro Protein Interaction Assays

*Escherichia coli* lysates containing GST fusion proteins were incubated with glutathione–Sepharose beads (Pharmacia) in NETN buffer (20 m*M* Tris, pH 8, 100 m*M* NaCl, 1 m*M* EDTA, 1 m*M* DTT, 0.1% NP-40) for 30 min at room temperature and washed three times in NETN buffer. Five microliters of lysate containing the *in vitro* $^{35}$S-labeled translated protein was added to 30 μl of a 30% suspension of GST fusion protein-coated beads in 200 μl of NETN, for 1 hr at 4°C. The resin was washed two times in NETN buffer at room temperature, and bound proteins were eluted in SDS sample buffer and analyzed by SDS–PAGE and autoradiography (Patrick *et al.*, 1993).

### DNA Binding Assay

A double-stranded oligonucleotide with the sense strand sequence CTGCAGCTTCAGCCCCTCTGGCCATCTGCTGATCCG containing the E-box of the rat insulin II enhancer RIPE3 (Shieh and Tsai, 1991) was labeled with [$^{32}$P]ATP using the T4 polynucleotide kinase. *In vitro* transcription and translation were performed using the TnT7 kit (Promega) according to manufacturer's instructions. Two to six microliters of lysate was used in electrophoretic mobility shift assays. Binding reactions contained 5 m*M* Hepes, pH 7.9, 2.5 m*M* $MgCl_2$, 5 m*M* EDTA, 5 m*M* NaCl, 2.5 m*M* DTT, 1 mg/ml BSA, 2 m*M* spermidine, 0.2 ng of end-labeled probe ($10^5$ cpm), and 1 μg of poly(dI-dC) in a total volume of 20 μl. After 15 min at room temperature, binding reactions were loaded onto a 5% polyacrylamide gel (acrylamide:bisacrylamide ratio, 29:1) and run in 0.5× TBE.

### RNA in Situ Hybridization

Whole mount *in situ* hybridizations were performed as described in Décimo *et al.* (1995). *In situ* hybridizations of $^{35}$S-labeled probes were performed on frozen sections of mouse embryos as described in Décimo *et al.* (1995). Briefly, 12.5-d.p.c. and younger embryos were frozen in OCT embedding medium (Miles) placed on dry ice, and older embryos were immersed in isopentane chilled in liquid nitrogen. Frozen embryos were cut in 10-μm-thick sections using a cryostat. Frozen sections were dipped in ice-cold acetone for 3 min, air-dried, fixed in 4% formaldehyde in phosphate-buffered saline (PBS) at 4°C for 15 min, rinsed in PBS, acetylated in 0.25% acetic anhydride in 0.1 *M* triethanolamine buffer at room temperature for 10 min, and incubated in 50% formamide, 150 m*M* sodium chloride, 15 m*M* sodium citrate, pH 7.0, at 60°C for 10 min and then in 50, 70, and 100% ethanol at −20°C. Sections were hybridized overnight at 55°C in a buffer containing 50% formamide, 300 m*M* sodium chloride, 20 m*M* Tris, pH 8.0, 10 m*M* sodium phosphate, and 10% dextran sulfate and washed at 65°C in a solution containing 50% formamide, 300 m*M* sodium chloride, and 30 m*M* sodium citrate.

*In situ* hybridization of digoxygenin-labeled probes on tissue sections was performed essentially as described by Myat *et al.* (1996). Briefly, embryos were fixed in 4% paraformaldehyde in PBS at 4°C for 1–2 hr, equilibrated in 20% sucrose in PBS at 4°C for 4–6 hr, mounted in 7.5% gelatin, 15% sucrose in PBS at 37°C, frozen on dry ice, and cut in 15-μm-thick sections. Thawed sections were hybridized in a buffer containing 50% formamide, 200 m*M* sodium chloride, 20 m*M* sodium acetate, and 10% dextran sulfate at 65°C overnight and washed in 50% formamide, 150 m*M* sodium chloride, 15 m*M* sodium citrate, 0.1% Tween 20 at 65°C. Sections were then incubated with an alkaline phosphatase (AP)-coupled anti-digoxygenin antibody (Boehringer Mannheim), washed, and processed for AP activity with NBT/BCIP.

The *Math4A* antisense RNA probe was synthesized from the 1.4-kb full-length cDNA using T7 RNA polymerase. Similar results were obtained using a probe corresponding to the last 29 amino acids of the protein (excluding the bHLH domain) and 543 bp of 3′ untranslated sequence. The *Mash1* probe was synthesized from the 2.0-kb full-length c2 cDNA clone (Guillemot and Joyner, 1993).

For BrdU incorporation experiments, 2 mg of BrdU was injected intraperitoneally into pregnant mice 2 hr before they were sacrificed, and embryos were harvested and fixed as described above. Sections were processed for RNA *in situ* hybridization, washed in PBS, incubated in 2 *M* HCl at room temperature for 45 min, rinsed three times in PBS, and processed for immunocytochemistry as described in Guillemot *et al.* (1993). Anti-class III β tubulin monoclonal antibody and anti-BrdU antibody were from Sigma. Biotinylated anti-mouse immunoglobulin antibody and avidin–biotin complex reagents were from the Vectastain kit (Vector).

## RESULTS

### Screening for Proteins Interacting with MASH1 Using the Yeast Two-Hybrid System

To identify proteins interacting with MASH1, we performed a screen in yeast using the two-hybrid system (Fields and Sternglanz, 1994). In this experiment, a plasmid encoding a portion of MASH1 containing the bHLH domain (aa 71–181; Guillemot and Joyner, 1993) in fusion with the LexA DNA-binding domain was introduced in the L40 yeast strain which carried the two reporter genes *LacZ* and *HIS3* under the control of upstream LexA-binding sites (Vojtek *et al.*, 1993). In control tests, the L40/LexA-MASH1 strain showed no *LacZ* expression and did not grow in synthetic medium lacking His and containing 25 m*M* inhibitor 3AT, which eliminated background *HIS3* gene activity (Vojtek *et al.*, 1993). When LexA-MASH1 was coexpressed with a plasmid encoding the bHLH domain of E12 in fusion with the VP16 transcription activation domain (TAD), both reporter genes were activated as a result of MASH1-E12 heterodimerization (data not shown). Thus, the L40/LexA-MASH1 strain was suitable for the identification of novel interacting partners of MASH1.

A mouse embryonic (9.5–12.5 d.p.c.) cDNA expression library fused to VP16 TAD (vom Baur *et al.*, 1996) was transformed into the L40/LexA-MASH1 strain. From a total of

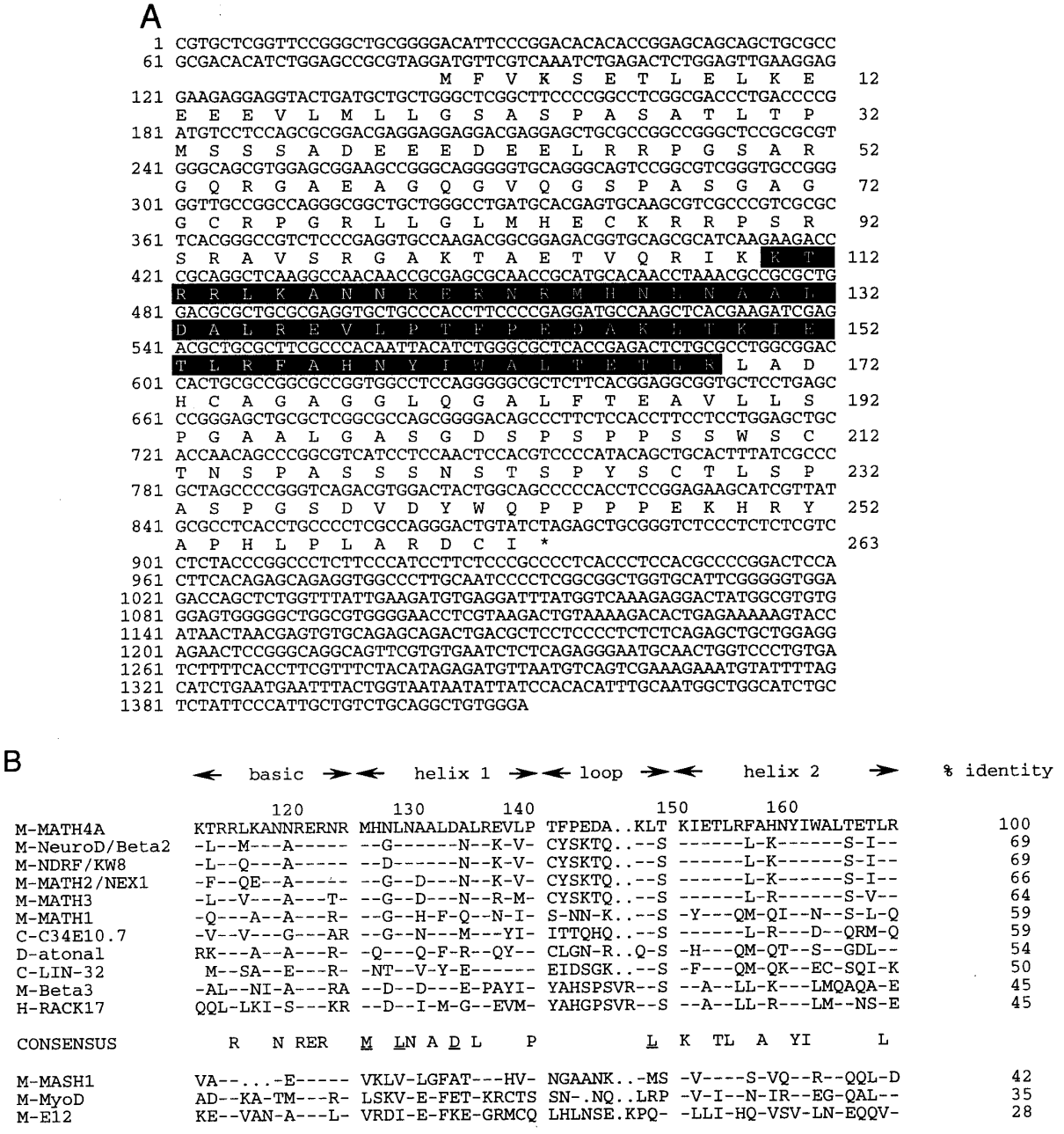



**FIG. 1.** *Math4A* cDNA sequence and sequence similarity with other bHLH proteins. (A) Nucleotide sequence and predicted open reading frame of *Math4A*. The deduced amino acid sequence is shown as a single-letter code. The bHLH region is boxed. (B) Alignment of the MATH4A bHLH region with those of other tissue-specific bHLH proteins. Residues specific for atonal-related bHLH domains within the consensus sequence are underlined. A dash indicates amino acid identity at this position. Database Accession numbers: MATH1, D43694; MATH2/Nex1, D44480, U29086; kw8, D82868; atonal, A40708; NeuroD/BETA2, U28068, U24679; C34E10.7, U10402; RACK17, U48250; Lin32, U15418.

$4 \times 10^6$ transformants, 626 colonies grew on selective medium, 200 of which have been analyzed. Of these clones, 196 were also $\beta$-galactosidase positive. To test for the specificity of the interactions with MASH1, these clones were isolated, and their plasmids rescued and retransformed into the L40 strain with the original bait or with an expression plasmid encoding LexA alone or the LexA-Lamin fusion protein as negative controls. Only 22 clones failed to inter-

act with LexA-MASH1 in the control experiment or activated reporter gene transcription with LexA alone or with LexA-Lamin. To identify the 178 remaining clones, systematic sequence analysis was performed. The sequence analysis revealed that 15 of these clones encoded bHLH proteins, the remaining ones corresponding to unrelated known genes or novel genes. Analysis of these last two groups of clones is in progress.

Among the clones encoding bHLH proteins, we isolated cDNAs for the ubiquitously expressed proteins E12, E47, ITFII, and HEB. The *E2A* gene product E12 has been shown previously to dimerize with MASH1 by coimmunoprecipitation (Johnson *et al.*, 1992). Another clone encoded the tissue-specific bHLH protein Thing1/Hxt/eHand, recently isolated in protein interaction screens with ubiquitous bHLH proteins (Cross *et al.*, 1995; Cserjesi *et al.*, 1995; Hollenberg *et al.*, 1995). In addition, we identified a novel bHLH protein which we named MATH4A (see below).

## *MATH4A Is a Novel Member of the Drosophila atonal-Related bHLH Family*

The *Math4A* clone isolated in yeast contained an insert of 1413 bp (Fig. 1A), with an uninterrupted open reading frame (ORF) beginning at the first ATG codon in frame with the VP16 sequence at position 85. This putative start codon was preceded by a sequence matching the Kozak consensus (Kozak, 1986). The ORF ended at a stop codon at position 874, suggesting that the cDNA clone contains the entire coding sequence of *Math4A*. The deduced MATH4A polypeptide is 263 amino acids in length ($M_r$ = 28,143 kDa, p*I* = 7.99) and contains a bHLH region near the center of the protein (Fig. 1A). It also contains an N-terminal Glu-rich domain and a large number of Pro and Ser residues in its C-terminal portion, which are common features of transcriptional activators (Tijan and Maniatis, 1994). Putative phosphorylation sites have been identified for casein kinase II (S35, S36, T141, T149, S237), protein kinase C (S50, T112, T153, T167), and cAMP- and cGMP-dependent protein kinases (S91). Sequence alignment of different bHLH domains (Fig. 1B) showed that MATH4A belongs to a family of proteins with bHLH domains very similar to that of the *Drosophila* proneural gene *atonal* (Jarman *et al.*, 1993). Within the bHLH region, MATH4A exhibits 69% sequence identity with mouse NeuroD/BETA2 (Lee *et al.*, 1995; Naya *et al.*, 1995), 66% with mouse MATH2/NEX-1 (Bartholomä and Nave, 1994; Shimizu *et al.*, 1995), 64% with MATH3 (Isaka *et al.*, 1996), 59% with MATH1 (Akazawa *et al.*, 1995), and 54% with atonal (Jarman *et al.*, 1993). MATH4A shares with atonal and related proteins the amino acids M125, L128, and D133 in the first helix of the bHLH domain and L148 in the loop, which appear to be characteristic of this gene family. The residue E121 in the basic region, which has been shown in the bHLH proteins MyoD (Ma *et al.*, 1994) and Max (Ferré-D'Amaré *et al.*, 1993) to make direct contacts with DNA, is also conserved in MATH4A. Outside of the bHLH domain, no significant homology was found with atonal-related proteins or any other proteins in the databases.

The designation *Math4A* (*Mammalian atonal homolog 4A*) was based on the similarity of the bHLH domain of this gene with those of other *atonal*-related genes and on our subsequent identification of two genes with very similar bHLH domains (*Math4B* and *Math4C*, data not shown), constituting a novel subfamily of *atonal*-related genes.

## *Interaction of MATH4A with MASH1 and E12*

To test the specificity of the interaction between MATH4A and MASH1 revealed in the yeast screen, we compared the interaction of MATH4A with MASH1, the ubiquitous bHLH protein E12, and an unrelated bHLH protein, Twist, which is expressed in mesoderm and neural crest cells (Wolf *et al.*, 1991). Interactions were analyzed in

**TABLE 1**

MATH4A Specifically Interacts with E12 and MASH1 in Yeast

| | $\beta$-Galactosidase activity[a] (U) | | | | |
|---|---|---|---|---|---|
| | VP16[b] | VP16-MATH4A | VP16-MASH1 | VP16-E12 | VP16-TWIST |
| LexA[b] | 2.7 ± 0.1 | 15.7 ± 0.1 | 6.1 ± 0.4 | 1.8 ± 0.4 | 2.5 ± 0.8 |
| LexA-MATH4A | 2.7 ± 0.1 | 18.9 ± 4 | 4.3 ± 1.5[c] | 65.0 ± 9.8[c] | <1 |
| LexA-MASH1 | <1 | 186.1 ± 21.2[c] | 20.0 ± 3 | 43.7 ± 10.9 | 2.2 ± 1.2 |
| LexA-E12 | 2.4 | 21.6 ± 0.2[c] | 12.3 ± 1.4 | 2.3 ± 0.2 | 830.0 ± 23.1 |
| LexA-TWIST | <1 | <1 | 4.2 ± 0.8 | 173.6 ± 20 | <1 |

[a] $\beta$-Galactosidase activity resulting from dimerization of fusion proteins containing bHLH motifs. LexA-bHLH fusion genes were transformed in yeast L40 containing an appropriate LacZ reporter and various VP16 fusion plasmids. $\beta$-Galactosidase activity was measured on three independent transformants grown in liquid culture.

[b] Interactions with unfused LexA and VP16 proteins were used as negative controls.

[c] MATH4A-E12 and MATH4A-MASH1 interactions were sensitive to the direction in which they were tested, a phenomenon which has been observed for numerous interactions in the two-hybrid system (Estojak *et al.*, 1995).

the yeast two-hybrid system, where they were detected by growth in selective medium or measurement of $\beta$-galactosidase activity and in an *in vitro* assay.

Interactions between bHLH proteins were analyzed in a semiquantitative manner in yeast cells by transformation with expression plasmids and measurement of $\beta$-galactosidase activity using a colorimetric assay (Table 1). In this experiment, MASH1 and MATH4A interacted with each other and with E12, but not with Twist or the negative controls. When the second helix of the bHLH domain of MASH1 was deleted in the LexA fusion protein, the interaction with MATH4A, measured by growth in synthetic medium lacking His, was abolished suggesting that the HLH domain of MASH1 mediates the dimerization with MATH4A (data not shown).

Interaction properties of MATH4A and other bHLH proteins were also examined *in vitro*, using GST fusion proteins produced in bacteria and $^{35}$S-labeled, *in vitro*-translated proteins (Fig. 2). Results obtained in this assay were consistent with those from the yeast interaction assay. MATH4A and MASH1 formed complexes with each other and with E12, but not with Twist. In addition, MASH1 and MATH4A bound to themselves in this assay. Thus, MATH4A interacts with the ubiquitous protein E12 and the neural-specific protein MASH1, but not with a bHLH protein from another family, both in yeast and *in vitro*.

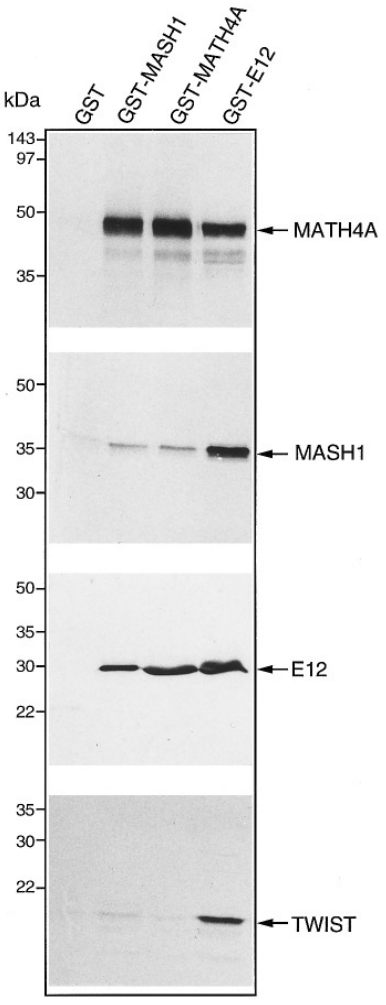



**FIG. 2.** MATH4A specifically interacts with E12 and MASH1 *in vitro*. *In vitro*-translated and radiolabeled proteins were incubated with GST fusion proteins immobilized on glutathione–Sepharose beads. Unfused GST was used as a negative control. Retained radiolabeled proteins were examined by SDS–PAGE and autoradiography. The arrowheads indicate the positions of $^{35}$S-labeled bound proteins.

### *DNA Binding Properties of MATH4A*

Tissue-specific bHLH proteins have been shown to bind E-box-containing sequences as heterodimers with ubiquitous bHLH proteins such as E12 or E47 (Murre *et al.*, 1989). In particular, NeuroD/BETA2 regulates transcription of the insulin gene by binding with high affinity to an E-box-containing sequence of the insulin gene promoter in a complex with E47 (Naya *et al.*, 1995). Since the basic regions of MATH4A and NeuroD/BETA2 are very similar, we examined whether MATH4A binds to the insulin E-box sequence by using an electrophoretic mobility shift assay (EMSA). Figure 3 shows that when NeuroD/BETA2 and E12 were cotranslated *in vitro*, a slower migrating complex formed (lane 17), suggesting that NeuroD/BETA2-E12 heterodimers bind the insulin E-box sequence, as described in Naya *et al.* (1995). Similarly, when MATH4A and E12 or MASH1 and E12 were cotranslated, new complexes were formed, suggesting that MATH4A-E12 and MASH1-E12 dimers can also bind to the insulin E-box sequence (lanes 3–5, 18, and 19). In contrast, no specific bandshift was observed when MATH4A and MASH1 were cotranslated and incubated with a labeled insulin E-box oligonucleotide (lane 20). Also, MATH4A homodimers, which can form *in vitro*, were not found to bind the E-box sequence (lane 15). These results may suggest that heterodimers between the tissue-specific bHLH proteins MASH1 and MATH4A do not bind E-box sequences efficiently. An alternative interpretation is that the composition of a bHLH dimer could influence its DNA-binding specificity. For example, replacement of a ubiquitous partner by a tissue-specific bHLH protein may produce a dimer recognizing a different range of E-box sequences (Brown and Baer, 1994; see Discussion).

### *Math4A Is Expressed in Undifferentiated Neural Precursors*

The *Math4A* clone was isolated from a cDNA library generated from 9.5- to 12.5-d.p.c. mouse embryos, suggesting that the gene is expressed during embryonic development. Northern blot analysis on total RNA isolated from 10.5- to 14.5-d.p.c. mouse embryos showed that *Math4A*

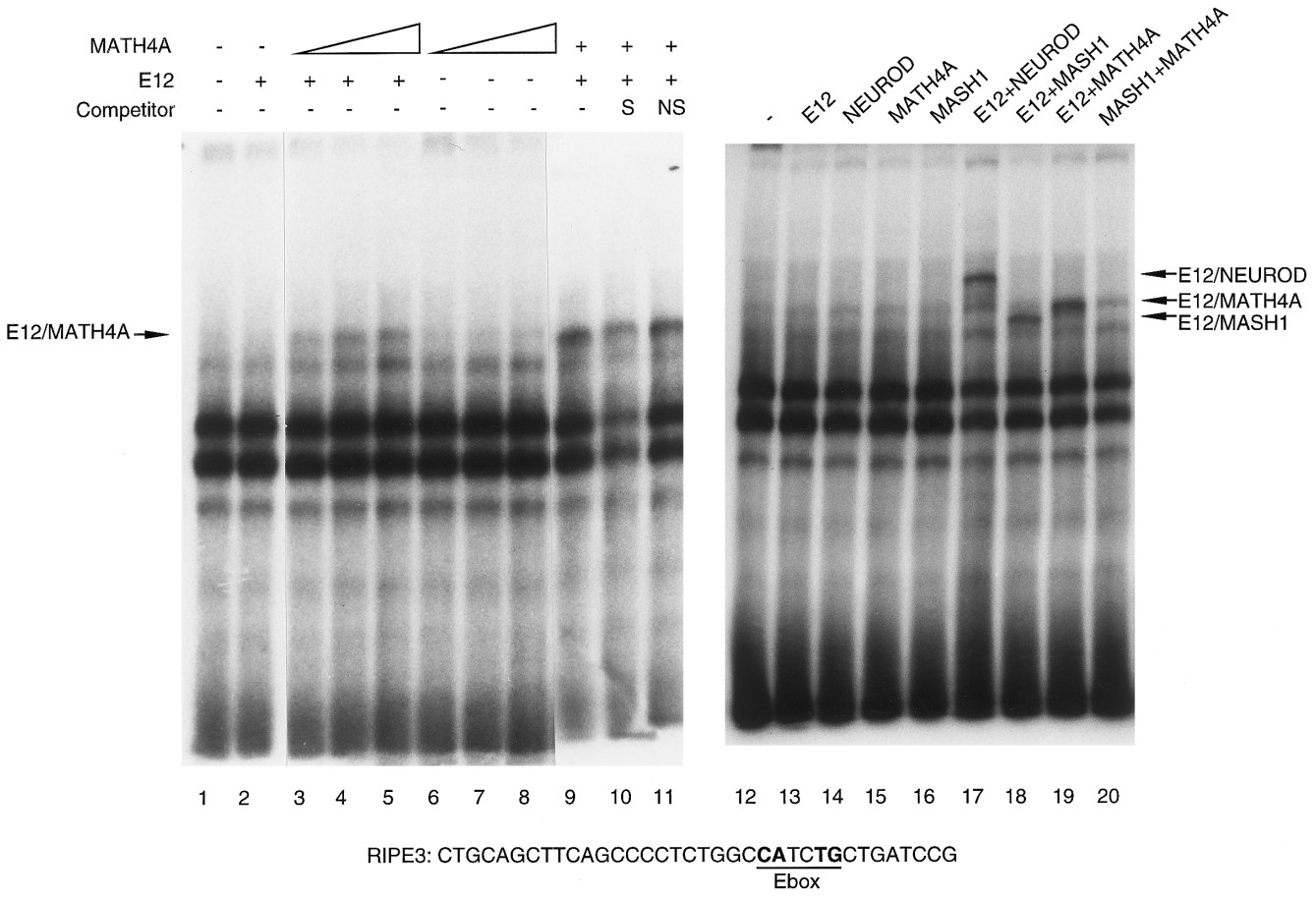
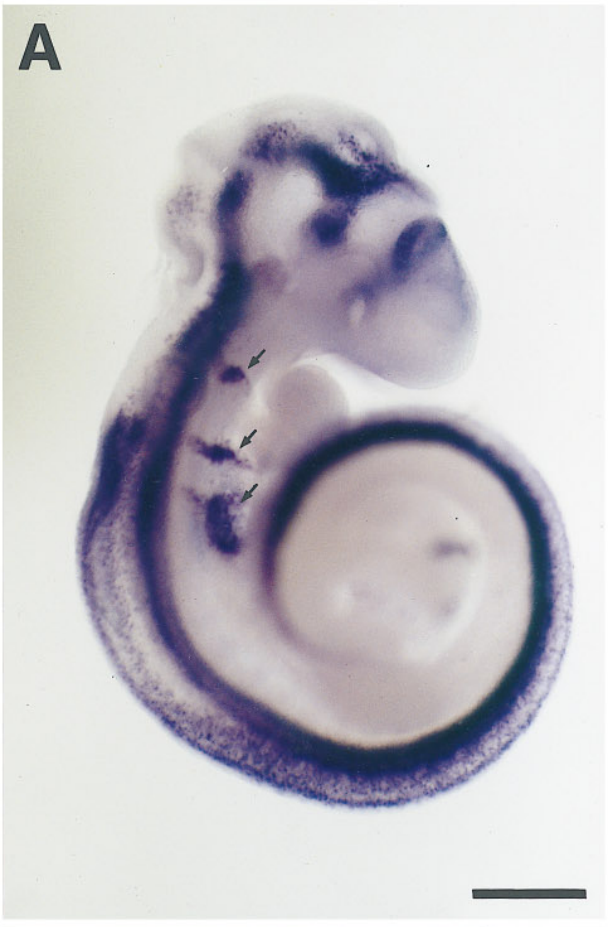
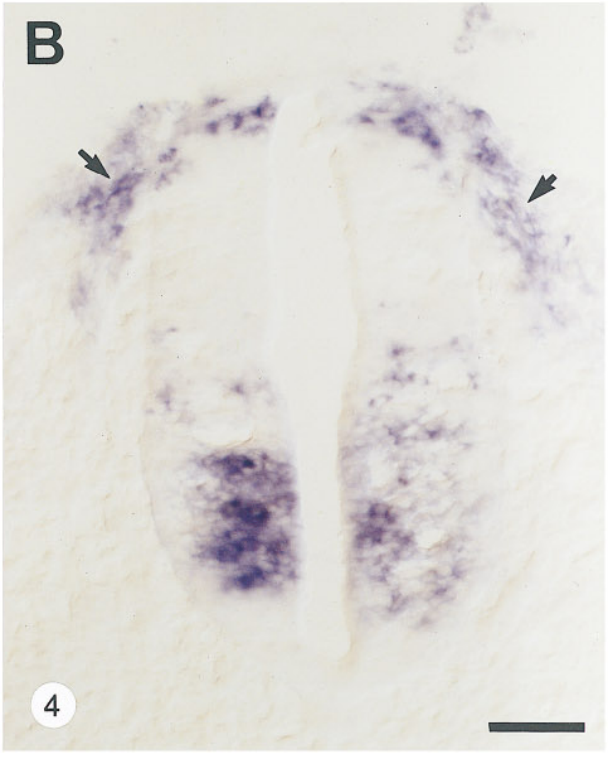


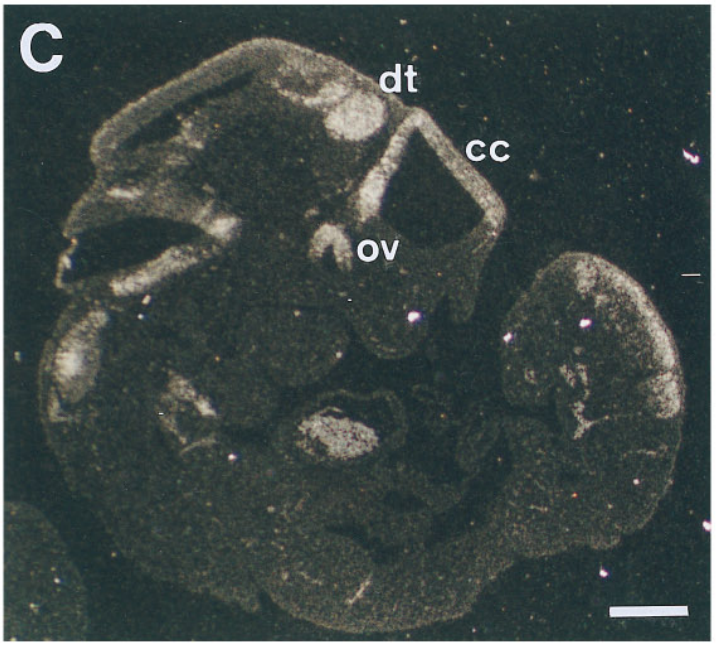

**FIG. 3.** DNA-binding activity of MATH4A. *In vitro*-translated products were assayed by gel mobility shift assay for binding to the insulin E-box containing RIPE3 probe (Naya *et al.*, 1995). Two microliters of each translation product was used for the binding reactions, except for lanes 4 and 7 (4 $\mu$l of MATH4A translation product) and lanes 5 and 8 (6 $\mu$l of MATH4A). A 50 M excess of unlabeled double-stranded RIPE3 probe (S) competed for the MATH4A/E12 heterodimer, whereas an unrelated sequence (NS) had no effect, suggesting that the binding was specific (lanes 9–11). The background bands present in all lanes represent complexes from the reticulocyte lysate binding to the probe. Competition for MATH4A-E12 binding with the unlabeled probe was only partial, probably because the probe also competed for binding to these other complexes, as shown by the reduction of background bands. The RIPE3 sequence is shown with the E-box underlined. Positions of the MATH4A-E12, NeuroD-E12, and MASH1-E12 complexes are indicated.

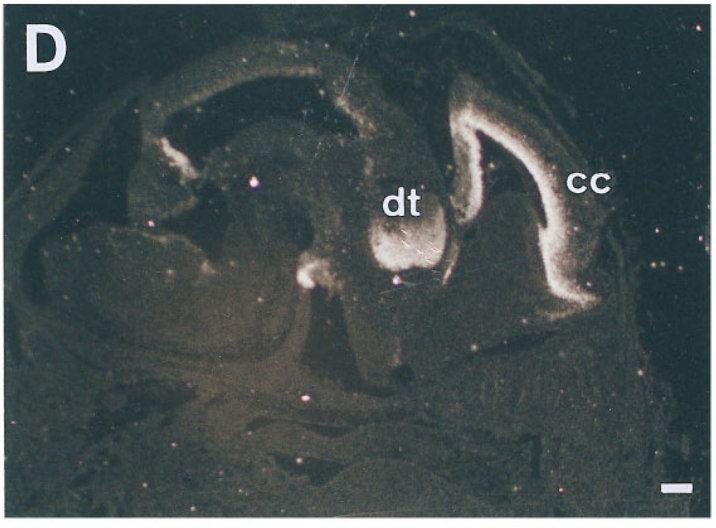
cDNA hybridizes with a 2.8-kb mRNA transcript (data not shown). *Math4A* expression was examined by RNA *in situ* hybridization on whole embryos and tissue sections (Fig. 4). *Math4A* expression appeared restricted to the developing CNS and PNS. *Math4A* transcripts were first detected at 8.5 d.p.c. and were localized initially to the neuroepithelium of the ventral mesencephalon and to the first epibranchial placode (data not shown). A few hours later, transcripts were also present in cells located in the basal plate of the spinal cord and in the ventral hindbrain, with the highest expression reached at midbrain level (data not shown).

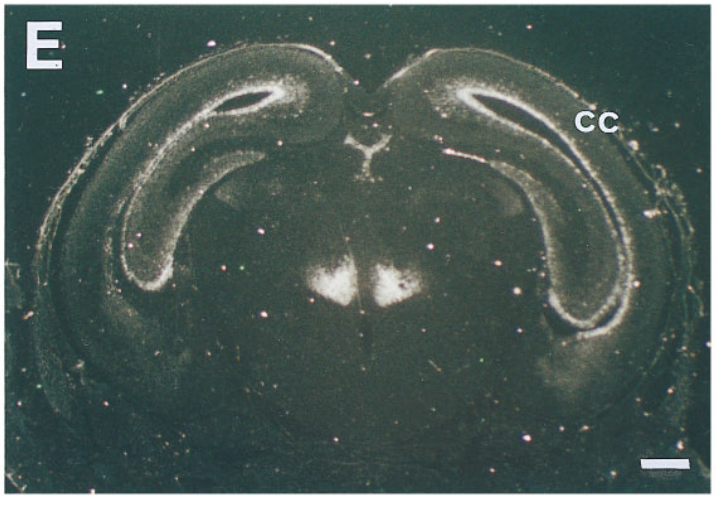
At 9.5 d.p.c., *Math4A* was expressed in the ventral neural tube, from the caudal spinal cord to the basal diencephalon and in the three epibranchial placodes (Fig. 4A). *Math4A* transcripts were also found in the dorsal-most region of the spinal cord and in cells emanating from it, which most likely correspond to early migrating neural crest cells (Fig. 4B). By 10.5 d.p.c., *Math4A* was transcribed more rostrally, in the dorsal telencephalon, the dorsal thalamus, and the optic vesicles (Fig. 4C). At this and later embryonic stages, expression domains in the telencephalon and diencephalon presented sharp borders which corresponded to boundaries between major divisions of the forebrain. In particular, *Math4A* was expressed at high levels in the dorsal telencephalon and excluded from the ventral telencephalon and was expressed in the dorsal thalamus and absent from the ventral thalamus (Figs. 6A and 6D at 12.5 d.p.c.; Fig. 4D at 15.5 d.p.c.). *Math4A* transcripts were also present in the pretectum and the mammilary and preoptic areas of the hypothalamus (Figs. 6A and 6D at 12.5 d.p.c. and data not shown). Expression in the cerebral cortex, dorsal thalamus, and retina remained at later embryonic stages. (Figs. 4E and


A


B
4


C
dt
cc
ov


D
dt
cc


E
cc

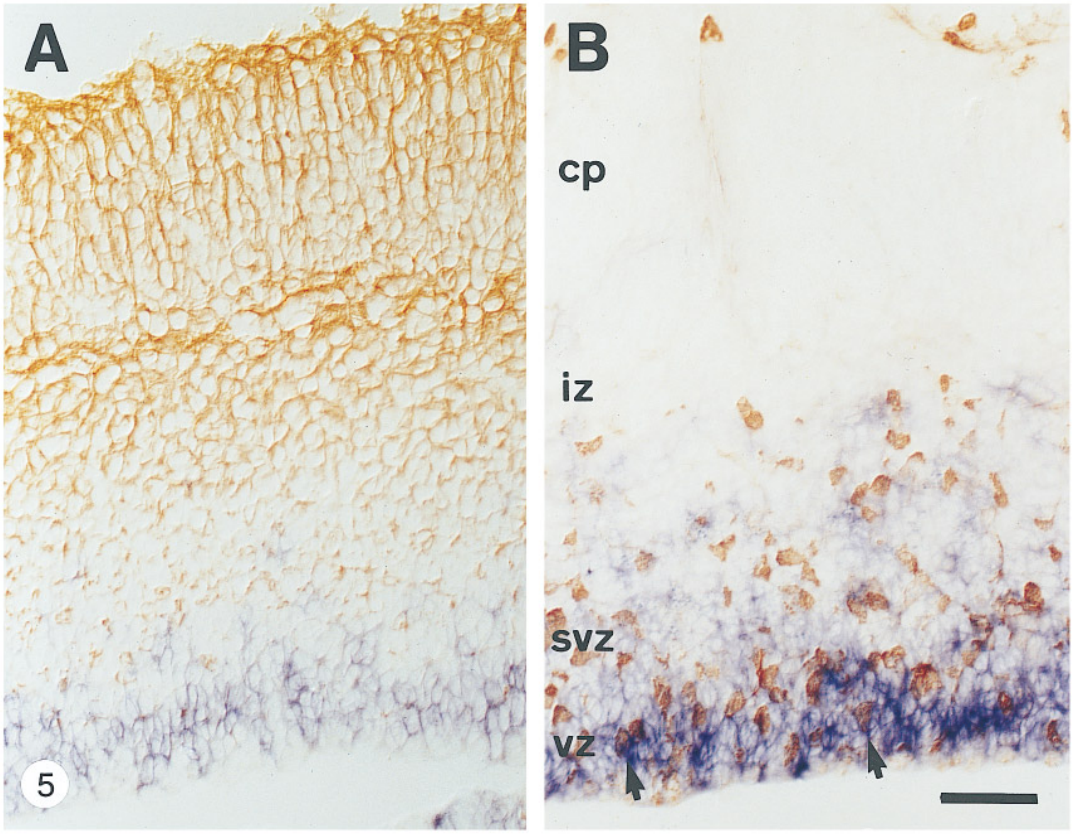



**FIG. 5.** Expression of *Math4A* in the cerebral cortex at 15.5 d.p.c. *In situ* hybridization of adjacent parasagittal sections with digoxygenin-labeled RNA probes for *Math4A*. The section in (A) was also processed for immunoperoxidase staining of class III $\beta$ tubulin, and the section in (B) was processed for immunoperoxydase staining to detect BrdU incorporation. The BrdU was injected intraperitoneally to the mother 2 hr before fixation. *Math4A* expression is restricted to the ventricular and subventricular zones where cells are undifferentiated ($\beta$ tubulin−) and proliferating (BrdU+). Some *Math4A*-expressing cells have incorporated BrdU (arrows) and are thus dividing. cp, cortical plate; iz, intermediate zone; svz, subventricular zone; vz, ventricular zone. Scale bar, 50 $\mu$m.

7A and data not shown). At birth, *Math4A* transcripts were still present along the lateral ventricles of the telencephalic vesicles and in the cerebellum. No expression was found in the adult brain (data not shown).

To determine at which stage in the differentiation of neural lineages *Math4A* was expressed, double-labeling experiments were performed with markers of postmitotic neurons and neural precursors. Double labeling of the telencephalic vesicles at 15.5 d.p.c. for *Math4A* RNA and class III $\beta$ tubulin showed that *Math4A* expression was excluded from the cell layers where differentiating neurons migrate and accumulate (Fig. 5A). Furthermore, double labeling for *Math4A* RNA and for BrdU incorporation showed that *Math4A* expression in the cerebral wall was restricted to the ventricular zone (VZ) at high levels and the subventricular zone (SVZ) at lower levels. Some of the *Math4A*-expressing cells

**FIG. 4.** Expression pattern of *Math4A* in the embryonic mouse nervous system. RNA *in situ* hybridization on a whole mount 9.5-d.p.c. embryo (A), a transverse section of the spinal cord of a 9.5-d.p.c. embryo at trunk level (B), a parasagittal section of a 10.5-d.p.c. embryo (C), a sagittal section of the head of a 15.5-d.p.c. embryo (D), and a coronal section of the head of a 17.5-d.p.c. embryo (E). (A) At 9.5 d.p.c., *Math4A* is expressed in the basal diencephalon, mesencephalon, hindbrain and spinal cord, in epibranchial placodes (arrows), and cells in the dorsal spinal cord. (B) *Math4A* is expressed in the basal plate of the spinal cord, in cells on both sides of the roof plate, and in cells detached from the dorsal neural tube which probably correspond to migrating neural crest cells (arrows). (C–E) From 10.5 d.p.c. until birth, *Math4A* is expressed in the presumptive cerebral cortex (cc), dorsal thalamus (dt), and retina (optic vesicle, ov). Transcripts are also found in the midbrain (not shown). Scale bars, 500 $\mu$m (A, C–E), 50 $\mu$m (B).

in the VZ had incorporated BrdU after a 2-hr pulse and were thus dividing (Fig. 5B). A similar restriction of expression to undifferentiated zones was observed in the dorsal thalamus, retina, tectum, hindbrain, and spinal cord (Figs. 6A, 6D, and 7A and data not shown).

Examination of the distribution of *Math4A* transcripts using a nonradioactive RNA probe revealed that they were not present uniformly in VZ and SVZ cells. In the cerebral cortex (Fig. 5) and the retina (Fig. 7A), expressing cells were grouped in clusters, among which a few cells presented transcripts levels above the expression in the rest of the cluster, while other cells in the VZ showed no detectable expression. The labeled clusters spanned the thickness of the VZ, as shown in the retina at E15.5 (Fig. 7A). A similar confinement of expression to cell clusters in the retina has been reported for CASH1, the chicken homolog of MASH1 (Jasoni *et al.*, 1994).

### *Limited Overlap of Math4A and Mash1 Expression Domains*

Both *Math4A* and *Mash1* are expressed in neural precursors of the embryonic CNS and PNS (Lo *et al.*, 1991; Guillemot and Joyner, 1993; this study). To determine whether the ability of MATH4A to interact with MASH1 *in vitro* and in yeast could reflect an *in vivo* situation, the expression patterns of the two genes were compared. Adjacent sections of the head region of embryos between 10.5 d.p.c. and birth were hybridized with $^{35}$S- and digoxygenin-labeled RNA probes (Figs. 6 and 7 and data not shown). *Math4A* and *Mash1* expression patterns were for the most part distinct, with overlaps limited mainly to posterior regions of the brain. For instance, the two genes were coexpressed in the preoptic area of the hypothalamus, the retina, tectum, cerebellar primordium, and hindbrain (Figs. 6 and 7). In contrast, complementarity in expression of the two genes was highest in the forebrain. *Mash1* was highly expressed in the ventral telencephalon, in particular in the medial and lateral ganglionic eminences, and in the ventral thalamus (Figs. 6B and 6E), where *Math4A* was not expressed (Figs. 6A and 6D). Conversely, *Math4A* was highly expressed in the VZ of the cerebral cortex and dorsal thalamus, areas where *Mash1* transcripts were only detectable at low levels. In the developing PNS, expression of *Math4A* and *Mash1* appeared to be mutually exclusive. *Mash1* is transiently and exclusively expressed in autonomic ganglia (Lo *et al.*, 1991; Guillemot and Joyner, 1993). In contrast, *Math4A* was expressed in epibranchial placodes, which produce sensory neurons contributing to some of the cranial ganglia (Fig. 4A).

## DISCUSSION

In vertebrate and invertebrate organisms, families of bHLH genes have been shown to regulate multiple steps in the development of neural tissues (Campuzano and Modolell, 1992; Guillemot *et al.*, 1993; Lo *et al.*, 1994; Turner and Weintraub, 1994; Lee *et al.*, 1995). The expression of the novel bHLH gene *Math4A* in undifferentiated neural precursors suggests that this gene controls early steps in the development of neural lineages in the CNS and PNS of mouse embryos.

### *MATH4A Belongs to a Family of Tissue-Specific bHLH Proteins Homologous to Drosophila atonal*

Sequence analysis has revealed that *Math4A* belongs to a family of vertebrate genes expressed mainly in the nervous system and sharing high sequence similarity in a bHLH-type domain with the *Drosophila* proneural gene *atonal*. This family currently includes the genes *MATH1* (Akazawa *et al.*, 1995), *MATH2/NEX-1* (Bartholomä and Nave, 1994; Shimizu *et al.*, 1995), *NeuroD/BETA2* (Lee *et al.*, 1995; Naya *et al.*, 1995), *kw8/NDRF* (Kume *et al.*, 1996; Yasunami *et al.*, 1996), and *BETA3* (Peyton *et al.*, 1996). Sequence conservation within the family is highest in the basic region, helix1, the C-terminal part of the loop, and helix2 of the bHLH domain. Indeed, a number of residues in helix1 (M125, H126, L128, L132, D133, V138) and the C-end of the loop (K147, L148) are shared by most or all members of the family and distinguish them from other bHLH gene families. Some of these residues (M125, L128, L148) are also shared with *Drosophila atonal* and with two bHLH genes of *Caenorhabditis elegans*, C34E10.7 and *lin-32* (Jarman *et al.*, 1993; Zhao and Emmons, 1995; Fig. 1). Sequence divergence in the family is highest in the six N-terminal residues of the loop, except between MATH2/NEX1, NeuroD/BETA2, and kw8/NDRF, which also share additional similarity in helix1 and thus constitute a subfamily of more closely related genes (93 to 96% similarity in the bHLH domain). Of the murine genes, *MATH1* is most similar to *Drosophila atonal*, owing to the conservation of nonconsensus residues in helix1 and helix2 (66% similarity with *atonal* in the bHLH domain versus 59 to 61% for other members of the family).

### *Math4A Expression in the Nervous System*

Genes of the *atonal*-related family are mainly expressed in the nervous system. However, they are expressed in different regions and more strikingly at different stages of neural differentiation in their respective expression domains. *MATH1* is transiently expressed in some of the cranial ganglia and in the dorsal spinal cord and hindbrain (Akazawa *et al.*, 1995). *NeuroD* expression is more widespread, with transcripts detected in cranial ganglia, the olfactory epithelium, and in the brain, where it is transiently present in differentiating neural cells (Lee *et al.*, 1995). In contrast, *MATH2/Nex1* expression in the developing spinal cord and telencephalon marks postmitotic neurons and transcripts are maintained in adult neurons of the cerebellum, hippocampus, and cerebral cortex (Bartholomä and Nave, 1994; Shimizu *et al.*, 1995). *Math4A* expression clearly differs from these patterns in that it is restricted in the CNS to

undifferentiated cells located in the ventricular and subventricular zones of the developing brain and spinal cord, at least some of which are dividing. The stage-specific expression of *Math4A*, *NeuroD*, and *MATH2* in regions such as the cerebral cortex (Bartholomä and Nave, 1994; Lee *et al.*, 1995; Shimizu *et al.*, 1995; this article) suggest that the three genes are sequentially expressed in neuronal lineages, where they are likely to perform distinct functions. The expression of *Math4A* is restricted to cell clusters in the VZ and the SVZ, suggesting that this gene could establish in precursor cells a particular state required for their subsequent differentiation. In contrast, the expression of *NeuroD* and *Math2* suggest that these genes promote and maintain terminal differentiation, respectively. The activity of related bHLH genes at successive stages of development of neural lineages would be strikingly parallel to the sequential roles of the myogenic bHLH genes in muscle development (Jan and Jan, 1993; Olson and Klein, 1994).

*Math4A* is first expressed in cells of the neural tube which are presumably already committed to a neural fate. It is thus unlikely to have a function in this tissue analogous to that of *atonal*, which is a proneural gene promoting internal sensory organ or photoreceptor cell development in ectoderm cells in *Drosophila* (Jarman *et al.*, 1993, 1994). However, *Math4A* is also expressed in the epibranchial placodes (Fig. 4A and data not shown). The ectoderm cells of these placodes generate sensory neurons which contribute to cranial ganglia, and *Math4A* is expressed in these cells before the onset of neurogenesis (data not shown). Further expression and functional studies will be necessary to determine whether *Math4A* contributes to the specification of a neural fate in these cells.

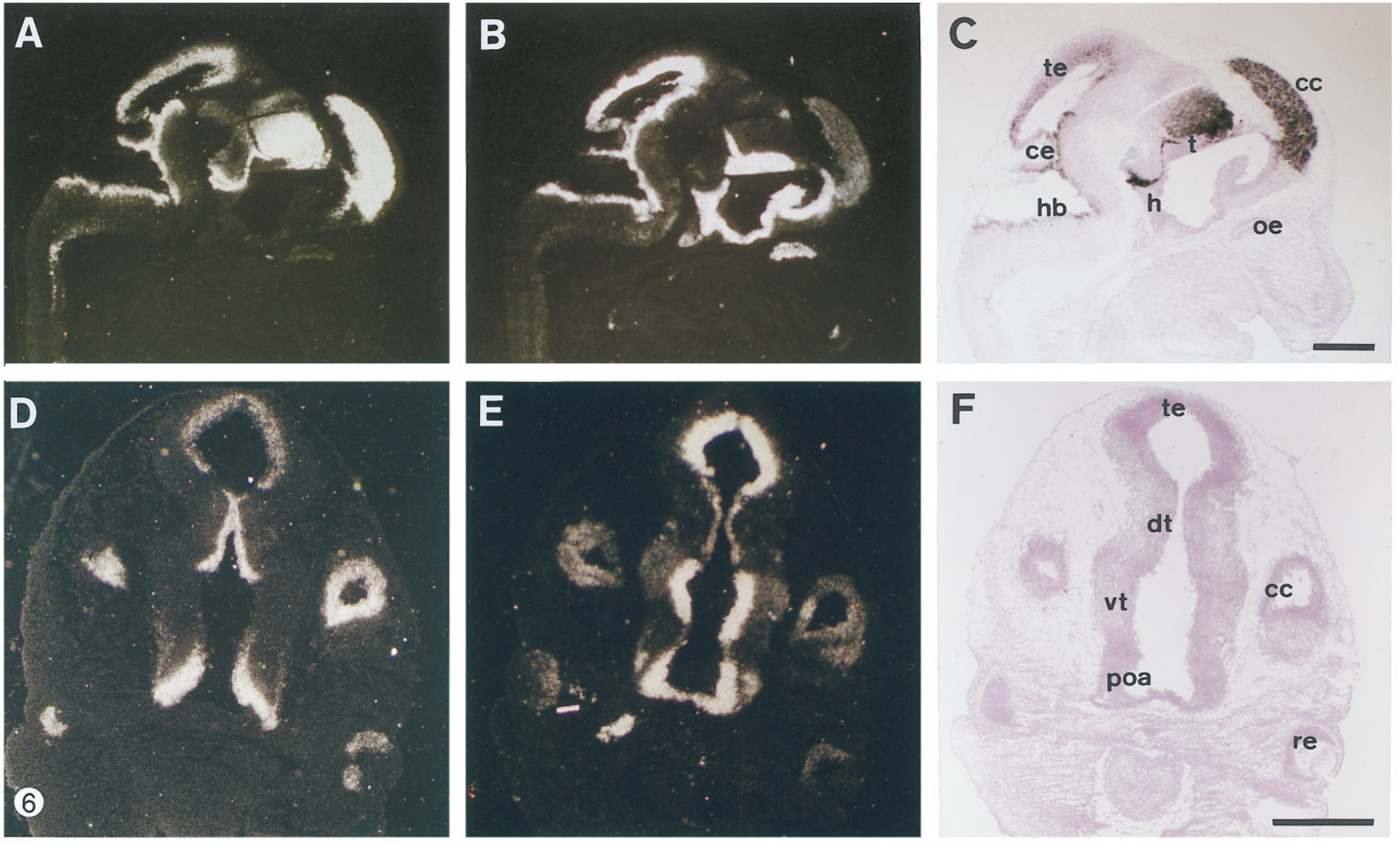
Loss-of-function and gain-of-function mutant phenotypes in the fly have led to the hypothesis that *atonal* is an important determinant of *Drosophila* internal sensory organ identity, distinguishing it from the other proneural *AS-C* genes which promote external sensory organ development (Jarman *et al.*, 1993). Interestingly, *Math4A* and *Mash1*, which are related to *atonal* and *AS-C* genes, respectively, have distinct and, in many regions of the nervous system, complementary expression patterns. In the telencephalon, *Mash1* is mostly expressed in ventral domains and *Math4A* in cortical areas. In the diencephalon, *Mash1* is highly expressed in the ventral thalamus, and *Math4A* in the dorsal thalamus. In the PNS, *Mash1* expression is restricted to autonomic precursors, whereas *Math4A* is expressed in placodes giving rise to sensory neurons. Given these different expression patterns and that both genes are expressed in undifferentiated cells, it is possible that they specify in neural precursors the production of distinct types of neurons. To address this issue, genetic experiments, and particularly the analysis of gain-of-function phenotypes, will be necessary.

## *MATH4A Can Interact with MASH1*

Interaction assays *in vitro* and in yeast have shown that MATH4A can homodimerize and heterodimerize with the ubiquitous bHLH protein E12, and also bind to the neural-specific bHLH protein MASH1. These interactions appear to be specific since there is no binding, for example, between MATH4A and Twist, an unrelated bHLH protein expressed in mesoderm and neural crest. Whether the interaction between MATH4A and MASH1 occurs *in vivo* and contributes to their function in neural precursors remains to be addressed. Analogous interactions could take place between one of these factors and other tissue-specific bHLH proteins with more similar expression patterns. We have recently isolated the genes *Math4B* and *Math4C* which are closely related to *Math4A*, and analysis of the expression and biochemical properties of their products may determine whether they constitute likely partners for interaction with MASH1 or MATH4A.

A MATH4A-E12 complex binds to the insulin E-box sequence which is bound by NeuroD/BETA2 in pancreatic $\beta$ cells (Naya *et al.*, 1995). This is not surprising given the similarity of the DNA-binding basic regions of MATH4A and NeuroD/BETA2 and the promiscuity of *in vitro* binding of bHLH proteins to E-boxes. In contrast, MATH4A homodimers and MATH4A-MASH1 heterodimers do not bind this E-box sequence. These complexes could lack DNA-binding activity altogether, as reported for some tissue-specific bHLH protein homodimers (Cabrera and Alonso, 1991). Formation of such complexes could be a mechanism to negatively regulate the transcriptional activity of these proteins. Alternatively, complexes that do not contain ubiquitous bHLH proteins could display distinct DNA-binding specificities. Along this line, specificities for different E-box sequences have been reported for HEN1-HEN1 homodimers and HEN1-E12 heterodimers (Brown and Baer, 1994). If this was the case for bHLH proteins like MATH4A and MASH1, formation of heterodimers between these proteins could substantially increase the repertoire of genes that they activate. This combinatorial activity of bHLH proteins could be a mechanism to specify distinct neuronal phenotypes in different lineages.

**FIG. 6.** Expression of *Math4A* and *Mash1* in the brain at 12.5 d.p.c. *In situ* hybridization of adjacent parasagittal sections (A–C) and coronal sections (D–F) with $^{35}$S-labeled RNA probes for *Math4A* (A, D) and *Mash1* (B, E). (C) and (F) are bright-field pictures of the sections in (A) and adjacent to (D), respectively. Expression of *Math4A* and *Mash1* overlap in some brain regions such as the tectum and preoptic area. *Math4A* is also highly expressed in regions where *Mash1* is expressed faintly or not at all, such as the dorsal thalamus and the cerebral cortex. cc, cerebral cortex; ce, cerebellar primordium; dt, dorsal thalamus; h, hypothalamus; hb, hindbrain; oe, olfactory epithelium; poa, preoptic area; re, retina; t, thalamus; te, tectum; vt, ventral thalamus. Scale bar, 500 $\mu$m.


A
B
C
te
cc
ce
t
hb
h
oe
D
E
F
te
dt
vt
cc
poa
re
6

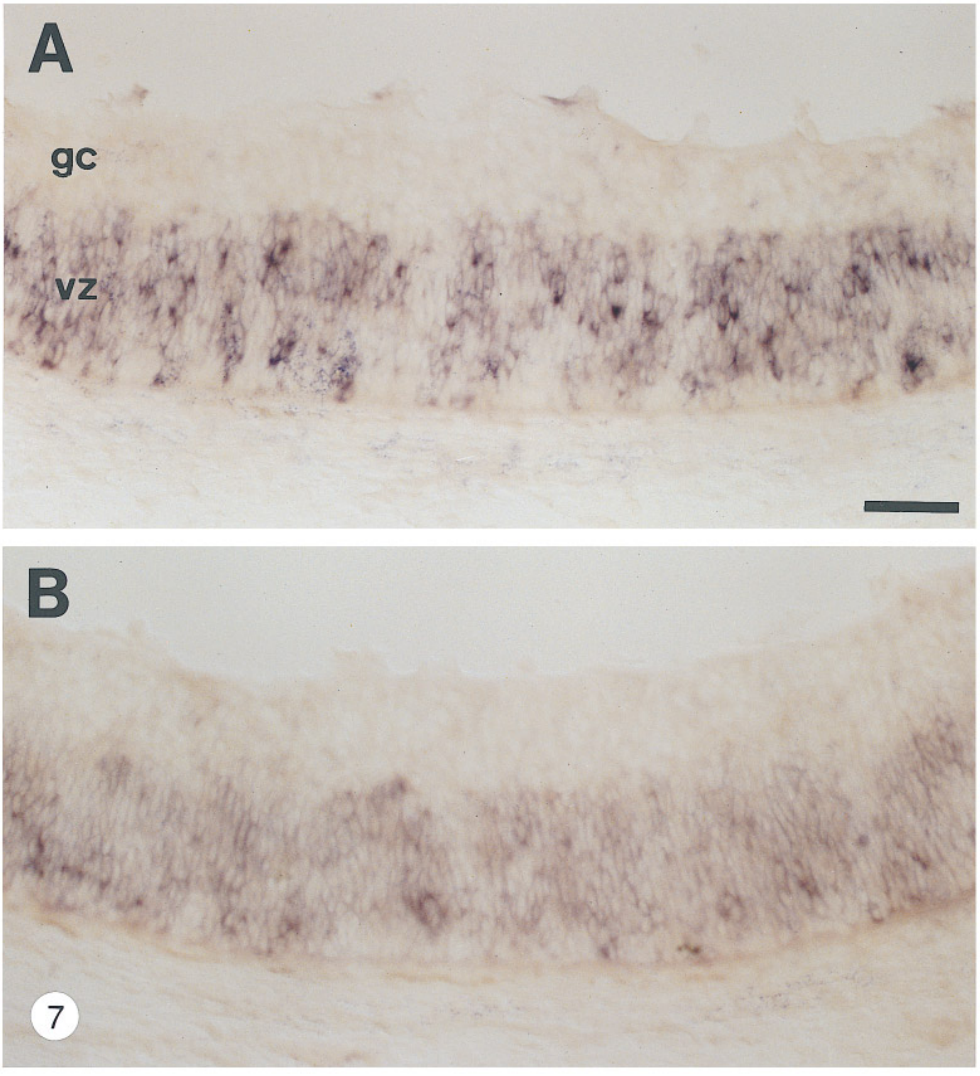



**FIG. 7.** Expression of *Math4A* and *Mash1* in the retina at 15.5 d.p.c. *In situ* hybridization of adjacent sections with digoxygenin-labeled RNA probes for *Math4A* (A) and *Mash1* (B). Both genes are expressed in subsets of cells in the ventricular zone (vz) of the retina. gc, ganglion cell layer. Scale bar, 50 $\mu$m.

## ACKNOWLEDGMENTS

We thank Bertrand LeDouarin and Jean Marie Garnier for the VP16AD mouse embryonic cDNA library, Stanley Hollenberg for the pBTM116 plasmid, Fabienne Perrin-Schmitt for the mTwist clone, Valérie Meyer for the histology, Serge Vicaire for sequencing, Adrien Staub for oligonucleotide synthesis, Bertrand LeDouarin, Claudine Gaudon, Régine Losson, and Michel Maira for advice on the yeast screen and DNA binding assay, Domingos Henrique for the section *in situ* protocol, Olivier Poch for helpful discussions, and Siew-Lan Ang for critical reading of the manuscript. This work was supported by funds from the Institut National de la Santé et de la Recherche Médicale, the Centre National de la Recherche Scientifique, and the Centre Hospitalier Universitaire Régional and by grants from the Fondation pour la Recherche Médicale and the Ministère de l'Enseignement et de la Recherche. C.F. was supported by a long-term fellowship from the Human Frontier Science Program.

## REFERENCES

Akazawa, C., Ishibashi, M., Shimizu, C., Nakanishi, S., and Kageyama, R. (1995). A mammalian helix-loop-helix factor structurally related to the product of *Drosophila* proneural gene atonal is a positive transcriptional regulator expressed in the developing nervous system. *J. Biol. Chem.* **270,** 8730–8738.

Ausubel, F. M., Brent, R., Kingston, R. E., Moore, D. D., Seidman,

J. G., Smith, J. A., and Struhl, K. (1987). ''Current Protocols in Molecular Biology.'' Wiley, New York.

Bartholomä, A., and Nave, K.-A. (1994). NEX-1: A novel brain-specific helix-loop-helix protein with autoregulation and sustained expression in mature cortical neurons. *Mech. Dev.* **48,** 217–228.

Benezra, R., Davis, R. L., Lockshon, D., Turner, D. L., and Weintraub, H. (1990). The protein Id: A negative regulator of helix-loop-helix DNA binding proteins. *Cell* **61,** 49–59.

Bengal, E., Ransone, L., Scharfmann, R., Dwarki, V. J., Tapscott, S. J., Weintraub, H., and Verma, I. M. (1992). Functional antagonism between c-Jun and MyoD proteins: A direct physical association. *Cell* **68,** 507–519.

Blackwell, T. K., and Weintraub, H. (1990). Differences and similarities in DNA-binding preferences of MyoD and E2A protein complexes revealed by binding site selection. *Science* **250,** 1104–1110.

Brown, L., and Baer, R. (1994). HEN1 encodes a 20-kilodalton phosphoprotein that binds an extended E-box motif as a homodimer. *Mol. Cell. Biol.* **14,** 1245–1255.

Cabrera, C. V., and Alonso, M. C. (1991). Transcriptional activation by heterodimers of the *achaete-scute* and *daughterless* gene products of *Drosophila*. *EMBO J.* **10,** 2965–2973.

Campuzano, S., and Modolell, J. (1992). Patterning of the *Drosophila* nervous system: The *achaete-scute* gene complex. *Trends Genet.* **8,** 202–208.

Cross, J. C., Flannery, M. L., Blanar, M. A., Steingrimsson, E., Jenkins, N. A., Copeland, N. G., Rutter, W. J., and Werb, Z. (1995). *Hxt* encodes a basic helix-loop-helix transcription factor that regulates trophoblast cell development. *Development* **121,** 2513–2523.

Cserjesi, P., Brown, D., Lyons, G. E., and Olson, E. N. (1995). Expression of the novel basic helix-loop-helix gene *eHAND* in neural crest derivatives and extraembryonic membranes during mouse development. *Dev. Biol.* **170,** 664–678.

Décimo, D., Georges-Labouesse, E., and Dollé, P. (1995). *In situ* hybridization of nucleic acid probes to cellular RNA. *In* ''The Practical Approach Series: Gene Probes 2'' (B. D. Hames and S. J. Higgins, Eds.), pp. 183–210. IRL Press, Oxford.

Ellenberg, T., Fass, D., Arnaud, M., and Harrison, S. C. (1994). Crystal structure of transcription factor E47. E-box recognition by a basic region helix-loop-helix dimer. *Genes Dev.* **8,** 970–980.

Estojak, J., Brent, R., and Golemis, E. A. (1995). Correlation of two-hybrid affinity data with in vitro measurements. *Mol. Cell. Biol.* **15,** 5820–5820.

Ferré-D'Amaré, A. R., Prendergast, G. C., Ziff, E. B., and Burley, S. K. (1993). Recognition by Max to its cognate DNA through a dimeric b/HLH/Z domain. *Nature* **363,** 38–45.

Fields, S., and Sternglanz, R. (1994). The two-hybrid system: An assay for protein-protein interactions. *Trends Genet.* **10,** 286–292.

Guillemot, F., and Joyner, A. L. (1993). Dynamic expression of the murine Achaete-Scute homologue Mash-1 in the developing nervous system. *Mech. Dev.* **42,** 171–185.

Guillemot, F., Lo, L. C., Johnson, J. E., Auerbach, A., Anderson, D. J., and Joyner, A. L. (1993). Mammalian achaete-scute homolog 1 is required for the early development of olfactory and autonomic neurons. *Cell* **75,** 463–476.

Hollenberg, S. M., Sternglanz, R., Cheng, P. F., and Weintraub, H. (1995). Identification of a new family of tissue-specific basic helix-loop-helix proteins with a two-hybrid system. *Mol. Cell. Biol.* **15,** 3813–3822.

Hsu, H. L., Huang, L., Tsan, J. T., Funk, W., Wright, W. E., Hu, J. S., Kingston, R. E., and Baer, R. (1994). Preferred sequences for DNA recognition by the TAL1 helix-loop-helix proteins. *Mol. Cell. Biol.* **14,** 1256–1265.

Isaka, F., Shimizu, C., Nakanishi, S., and Kageyama, R. (1996). Genetic mapping of four mouse bHLH genes related to *Drosophila* proneural gene *atonal*. *Genomics*, in press.

Jan, Y. N., and Jan, L. Y. (1993). HLH proteins, fly neurogenesis, and vertebrate myogenesis. *Cell* **75,** 827–830.

Jarman, A. P., Grau, Y., Jan, L. Y., and Jan, Y. N. (1993). atonal is a proneural gene that directs chordotonal organ formation in the Drosophila peripheral nervous system. *Cell* **73,** 1307–1321.

Jarman, A. P., Grell, E. H., Ackerman, L., Jan, L. Y., and Jan, Y. N. (1994). Atonal is the proneural gene for Drosophila photoreceptors. *Nature* **369,** 398–400.

Jasoni, C. L., Walker, M. B., Morris, M. D., and Reh, T. A. (1994). A chicken achaete-scute homolog (CASH-1) is expressed in a temporally and spatially discrete manner in the developing nervous system. *Development* **120,** 769–783.

Johnson, J. E., Birren, S. J., Saito, T., and Anderson, D. J. (1992). DNA binding and transcriptional regulatory activity of mammalian achaete-scute homologous (MASH) proteins revealed by interaction with a muscle-specific enhancer. *Proc. Natl. Acad. Sci. USA* **89,** 3596–3600.

Kozak, M. (1986). Point mutations define a sequence flanking the AUG initiator codon that modulates translation by eukaryotic ribosomes. *Cell* **44,** 283–292.

Kume, H., Maruyama, K., Tomita, T., Iwatsubo, T., Saido, T. C., and Obata, K. (1996). Molecular cloning of a novel basic helix-loop-helix protein from the rat brain. *Biochem. Biophys. Res. Commun.* **219,** 526–530.

Larson, R. C., Lavenir, I., Larson, T. A., Baer, R., Warren, A. J., Wadman, I., Nottage, K., and Rabbitts, T. H. (1996). Protein dimerization between Lmo2 (Rbtn2) and Tal1 alters thymocyte development and potentiates T cell tumorigenesis in transgenic mice. *EMBO J.* **15,** 1021–1027.

Le Douarin, B., Pierrat, B., vom Baur, E., Chambon, P., and Losson, R. (1995). A new version of the two-hybrid assay for detection of protein-protein interactions. *Nucleic Acids Res.* **23,** 876–878.

Lee, J. E., Hollenberg, S. M., Snider, L., Turner, D. L., Lipnick, N., and Weintraub, H. (1995). Conversion of *Xenopus* ectoderm into neurons by NeuroD, a basic helix-loop-helix protein. *Science* **268,** 836–844.

Lo, L., Guillemot, F., Joyner, A. L., and Anderson, D. J. (1994). MASH-1: A marker and a mutation for mammalian neural crest development. *Perspect. Dev. Neurobiol.* **2,** 191–201.

Lo, L. C., Johnson, J. E., Wuenschell, C. W., Saito, T., and Anderson, D. J. (1991). Mammalian achaete-scute homolog 1 is transiently expressed by spatially restricted subsets of early neuroepithelial and neural crest cells. *Genes Dev.* **5,** 1524–1537.

Ma, P. C., Rould, M. A., Weintraub, H., and Pabo, C. O. (1994). Crystal structure of MyoD bHLH domain-DNA complex: Perspectives on DNA recognition and implications for transcriptional activation. *Cell* **77,** 451–459.

Molkentin, J. D., Black, B. L., Martin, J. F., and Olson, E. N. (1995). Cooperative activation of muscle gene expression by MEF2 and myogenic bHLH proteins. *Cell* **83,** 1125–1136.

Murre, C., McCaw, P. S., Vaessin, H., Caudy, M., Jan, L. Y., Jan, Y. N., Cabrera, C. V., Buskin, J. N., Hauschka, S. D., Lassar, A. B. *et al.* (1989). Interactions between heterologous helix-loop-helix proteins generate complexes that bind specifically to a common DNA sequence. *Cell* **58,** 537–544.

Myat, A., Henrique, D., Ishhorowicz, D., and Lewis, J. (1996). A chick homologue of Serrate and its relationship with Notch and Delta homologues during central neurogenesis. *Dev. Biol.* **174,** 233–247.

Naya, F. J., Stellrecht, C. M. M., and Tsai, M. J. (1995). Tissue-specific regulation of the insulin gene by a novel basic-helix-loop-helix transcription factor. *Genes Dev.* **9,** 1009–1019.

Olson, E. N., and Klein, W. H. (1994). bHLH factors in muscle development: Dead lines and commitments, what to leave in and what to leave out. *Genes Dev.* **8,** 1–8.

Patrick, D. R., Oliff, A., and Heimbrook, D. C. (1993). Identification of a novel retinoblastoma gene product binding site on human papillomavirus type 16 E7 protein. *J. Biol. Chem.* **269,** 6842–6850.

Peyton, M., Stellrecht, C. M. M., Naya, F. J., Huang, H. P., Samora, P. J., and Tsai, M. J. (1996). BETA3, a novel helix-loop-helix protein, can act as a negative regulator of BETA2 and MyoD-responsive genes. *Mol. Cell. Biol.* **16,** 626–633.

Shieh, S.-Y., and Tsai, M.-J. (1991). Cell-specific and ubiquitous factors are responsible for the enhancer activity of the rat insulin II gene. *J. Biol. Chem.* **266,** 16707–16714.

Shimizu, C., Akazawa, C., Nakanishi, S., and Kageyama, R. (1995). MATH-2, a mammalian helix-loop-helix factor structurally related to the product of *Drosophila* proneural gene atonal, is specifically expressed in the nervous system. *Eur. J. Biochem.* **229,** 239–248.

Sommer, L., Shah, N., Rao, M., and Anderson, D. J. (1995). The cellular function of MASH1 in autonomic neurogenesis. *Neuron* **15,** 1245–1258.

Sun, X. H., and Baltimore, D. (1991). An inhibitory domain of E12 transcription factor prevents DNA binding in E12 homodimers but not in E12 heterodimers. *Cell* **64,** 459–470.

Takebayashi, K., Sasai, Y., Sakai, Y., Watanabe, T., Nakanishi, S., and Kageyama, R. (1994). Structure, chromosomal locus, and promoter analysis of the gene encoding the mouse helix-loop-helix factor HES-1. Negative autoregulation through the multiple N box elements. *J. Biol. Chem.* **269,** 5150–5156.

Tjian, R., and Maniatis, T. (1994). Transcriptional activation: A complex puzzle with few easy pieces. *Cell* **76,** 5–8.

Turner, D. L., and Weintraub, H. (1994). Expression of *achaete-scute homolog 3* in *Xenopus* embryos converts ectodermal cells to a neural fate. *Genes Dev.* **8,** 1434–1447.

Vojtek, A. B., Hollenberg, S. M., and Cooper, J. A. (1993). Mammalian Ras interacts directly with the serine/threonine kinase Raf. *Cell* **74,** 205–214.

vom Baur, E., Zechel, C., Heery, D., Heine, M., Garnier, J.-M., Vivat, V., Le Douarin, B., Gronemeyer, H., Chambon, P., and Losson, R. (1996). Differential ligand-dependent interactions between the AF-2 activating domain of nuclear receptors and the putative transcriptional intermediary factors mSUG1 and TIF1. *EMBO J.* **15,** 110–124.

Wolf, C., Thisse, C., Stoetzel, C., Thisse, B., Gerlinger, P., and Perrin-Schmitt, F. (1991). The *M-twist* gene of *Mus* is expressed in subsets of mesodermal cells and is closely related to the *Xenopus X-twi* and the *Drosophila twist* genes. *Dev. Biol.* **143,** 363–373.

Yasunami, M., Suzuki, K., Maruyama, H., Kawakami, H., Nagai, Y., Hagiwara, M., and Ohkubo, H. (1996). Molecular cloning and characterization of a cDNA encoding a novel basic helix-loop-helix protein structurally related to NeuroD/BHF1. *Biochem. Biophys. Res. Commun.* **220,** 754–758.

Zhao, C., and Emmons, S. W. (1995). A transcription factor controlling development of peripheral sense organs in *C. elegans*. *Nature* **373,** 74–78.

Received for publication July 5, 1996
Accepted August 20, 1996