MOLECULAR AND CELLULAR BIOLOGY, Apr. 1996, p. 1714–1721 Vol. 16, No. 4
0270-7306/96/$04.00+0


# An Activation Domain of the Helix-Loop-Helix Transcription Factor E2A Shows Cell Type Preference In Vivo in Microinjected Zebra Fish Embryos

FRANCESCO ARGENTON,[1] YOAV ARAVA,[2] AMI ARONHEIM,[2]† AND MICHAEL D. WALKER[2]*

*Department of Biology, University of Padua, Padua, Italy,[1] and Biochemistry Department, Weizmann Institute of Science, Rehovot 76100, Israel[2]*

Received 6 November 1995/Returned for modification 5 December 1995/Accepted 18 December 1995

**The E2A protein is a mammalian transcription factor of the helix-loop-helix family which is implicated in cell-specific gene expression in several cell lineages. Mouse E2A contains two independent transcription activation domains, ADI and ADII; whereas ADI functions effectively in a variety of cultured cell lines, ADII shows preferential activity in pancreatic beta cells. To analyze this preferential activity in an in vivo setting, we adapted a system involving transient gene expression in microinjected zebra fish embryos. Fertilized one- to four-cell embryos were coinjected with an expression plasmid and a reporter plasmid. The expression plasmids used encode the yeast Gal4 DNA-binding domain (DBD) alone, or Gal4 DBD fused to ADI, ADII, or VP16. The reporter plasmid includes the luciferase gene linked to a promoter containing repeats of UASg, the Gal4-binding site. Embryo extracts prepared 24 h after injection showed significant luciferase activity in response to each of the three activation domains. To determine the cell types in which the activation domains were functioning, a reporter plasmid encoding β-galactosidase and then in situ staining of whole embryos were used. Expression of ADI led to activation in all major groups of cell types of the embryo (skin, sclerotome, myotome, notochord, and nervous system). On the other hand, ADII led to negligible expression in the sclerotome, notochord, and nervous system and much more frequent expression in the myotome. Parallel experiments conducted with transfected mammalian cells have confirmed that ADII shows significant activity in myoblast cells but little or no activity in neuronal precursor cells, consistent with our observations in zebra fish. This transient-expression approach permits rapid in vivo analysis of the properties of transcription activation domains: the data show that ADII functions preferentially in cells of muscle lineage, consistent with the notion that certain activation domains contribute to selective gene activation in vivo.**

Cell-specific patterns of eukaryotic gene expression are controlled in large part via transcriptional control mechanisms mediated by characteristic sets of transcription factors (30). A major unresolved issue is to understand precisely how these sets of factors orchestrate the highly specific patterns of spatial and temporal expression observed in multicellular organisms (20). Although many biologically important transcription factors have been cloned and characterized, it remains unclear how they exert their cell- and gene-specific properties in the developing organism. For example, homeodomain proteins play a key role in developmental decisions in most if not all eukaryotic organisms (22). However, the in vivo distribution and in vitro DNA-binding specificity of these proteins is in most cases insufficient to explain the in vivo selectivity of these proteins and leads to the view that additional mechanisms are involved (15, 29). Such mechanisms may include interactions with additional proteins, leading to alteration of DNA-binding specificity (13).

The helix-loop-helix (HLH) family of transcription factors has been implicated in cell-specific control in a variety of cell lineages in both *Drosophila melanogaster* and vertebrates (19, 31, 32). It appears that cell-specific transcription in several systems is mediated by heterodimeric complexes containing a lineage-restricted HLH protein in complex with a widely distributed HLH protein (24, 31, 32). The dimeric HLH complexes apparently function by recognizing specific *cis* elements (E boxes, consensus sequence CAXXTG) located within enhancers of cell-specifically expressed genes (7). Characteristic E-box-binding complexes are observed in distinct lineages: muscle cell factors binding to the MEF1 E box contain either MyoD or myogenin in association with E2A (24). The lymphoid factor BCF1 contains E2A protein, probably as a homodimer (2, 33, 43), whereas the pancreatic beta-cell factor IEF1 (36) appears to contain the E2A protein (2, 6, 9, 44, 51) in complex with a lineage-restricted HLH protein BETA2 (IESF1) (2, 34, 37). As in the case of homeodomain proteins, the in vitro DNA-binding specificity of these complexes cannot by itself explain the in vivo expression pattern of the target gene: IEF1, for example, is capable of in vitro binding to E-box sequences located within immunoglobulin- and muscle-specific genes, yet in beta cells the insulin gene is transcribed and the immunoglobulin and muscle genes are not. Thus, additional levels of control are clearly operating to fine tune in vivo patterns of transcription.

The role of a constitutive HLH protein such as E2A in the functional heterodimer is complex. Lineage-restricted HLH proteins show limited ability to bind DNA efficiently as homodimers, yet upon heterodimerization with E2A, efficient E-box binding is observed (47). In addition, ubiquitous HLH proteins appear to directly activate transcription: transactivation capacity has been demonstrated in the N-terminal two-thirds of the human E2A protein (16) and subsequently mapped to a 108-amino-acid fragment with a putative loop–α-helical (LH) structure (40). We have shown that the mouse E2A protein possesses two distinct transcription activation do-

* Corresponding author. Phone: (972) 8 934 3597. Fax: (972) 8 934 4118.

† Present address: Department of Pharmacology, School of Medicine, University of California, San Diego, La Jolla, CA 92093-0636.

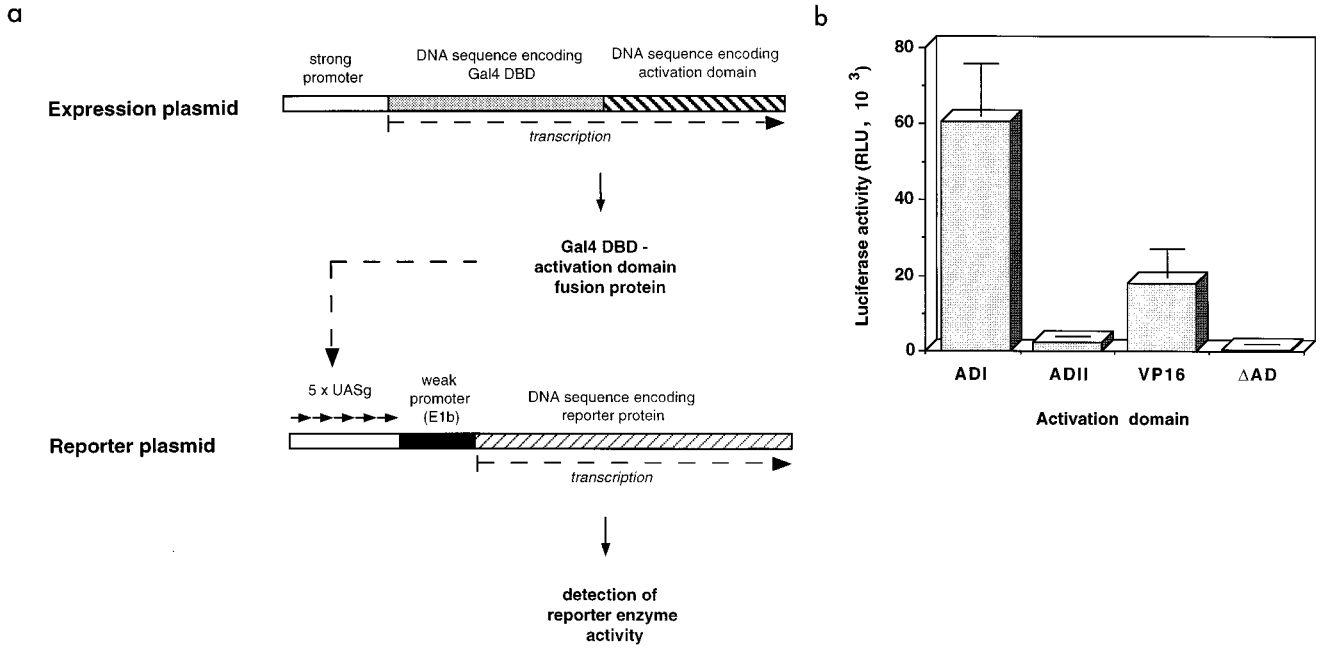



FIG. 1. (a) Schematic representation of plasmids used in this study. Expression plasmids used a strong enhancer-promoter (RSV or CMV) upstream of DNA encoding the Gal4 DBD (amino acids 1 to 147) linked to DNA encoding a transcription activation domain (ADI, ADII, or VP16) or no activation domain (ΔAD). Reporter plasmids contained the DNA encoding luciferase or β-galactosidase downstream of a promoter (E1b) linked to five copies of the Gal4-binding site (UASg). The presence of a functional *trans*-activation domain in the protein product of the expression plasmid is expected to lead to production of the reporter enzyme. (b) *trans*-activation in zebra fish embryos. Embryos (one to four cells) were injected with the indicated RSV expression plasmids together with a luciferase reporter plasmid, under the control of five copies of the UASg, and CMV-TK-CAT control plasmid at a ratio of 2:2:1. Embryos were grouped into pools of three, and whole-embryo extracts from each pool were assayed for luciferase and CAT activities as described in Materials and Methods. CAT activity measurements were used to exclude from the analysis extract pools derived from embryos which inefficiently express the injected DNA (perhaps because injections were not optimal). Data are expressed as means ± standard errors of the means of luciferase activity in relative light units (RLU). The analyses were performed with the following numbers of embryo pools: ADI, 14; ADII, 13; VP16, 16; and ΔAD, 5.

mains, ADI and ADII (3). ADI functioned efficiently in a wide range of cultured cells; ADII, on the other hand, activated transcription preferentially in cultured pancreatic beta cells (3). These observations raised the possibility that additional specificity is contributed via interactions with this activation domain. Conceivably, this may occur as a result of cell-specific posttranslational modifications of the E2A protein or via interactions with cell-specific coactivator molecules (27).

To determine whether the cell preference of ADII was observed also in an in vivo setting, we examined whether a zebra fish transient-expression system could be used for analyzing the function of transcription activation domains. Recently, it has been shown that microinjection of reporter plasmids into fertilized zebra fish embryos leads to transient expression of reporter genes in the resulting mosaic transgenics; testing the promoters of mammalian homeobox-containing genes by this procedure led to the conclusion that there is strong conservation of promoter function between fish and mammals (54). In the present study, we show that mammalian transcription activation domains function efficiently in developing zebra fish embryos. Furthermore, the preferential activity observed in cultured cells was also seen in vivo: ADII functioned efficiently in cells of the myotome and relatively inefficiently in cells of the sclerotome, notochord, and nervous system. Thus, transcription activation domains may play an important role in directing cell-specific patterns of gene expression in vivo.

## MATERIALS AND METHODS

**Plasmids.** Expression plasmids contained the strong transcription control regions of cytomegalovirus (CMV) (8) or Rous sarcoma virus (RSV) (11). CMV-TK-CAT, derived by subcloning of the CMV enhancer in pBL-CAT5 (5), and CMV-TK-Luc, derived from the former by replacement of chloramphenicol acetyltransferase (CAT) by luciferase coding regions, were gifts from M. Schartl. RSV-Gal4-ADI, RSV-Gal4-ADII, RSV-Gal4-VP16, and RSV-Gal4-ΔAD expression plasmids have been described previously (3). They include sequences encoding the DNA-binding domain (DBD) of the yeast Gal4 protein (amino acids 1 to 147 [20]). The CMV-Gal4-ΔAD expression plasmid was constructed by blunt-end cloning of the *Hin*dIII-*Bam*HI fragment of RSV-Gal4-ΔAD, comprising the Gal4 (amino acids 1 to 147) DBD coding region, a *Sma*I-*Bgl*II polylinker, and the simian virus 40 splice-polyadenylation signal, to *Bgl*II-*Sma*I sites of CMV-TK-CAT. CMV-Gal4-ADI and CMV-Gal4-VP16 were constructed by cloning the corresponding activation domains to the *Bgl*II site of CMV-Gal4-ΔAD. CMV-Gal4-ADII was derived by cloning the *Hin*dIII-*Bam*HI fragment of RSV-Gal4-ADII to the *Bgl*II-*Sma*I sites of CMV-TK-CAT. 5×UASg-E1b TATA-Luciferase (UASg-Luc) was a gift of I. Haviv, and 5×UASg-E1b TATA-LacZ (UASg-LacZ) was prepared by subcloning the 5×UASg-E1b promoter fragment to the *Bam*HI site of pGA307, a β-galactosidase expression vector (1).

**Transfections.** CHO fibroblasts, HIT T15 M2.2.2 beta cells, and $C_2C_{12}$ myoblast cells were transfected by the calcium phosphate procedure (55) as previously described (3). Briefly, the cells were exposed to 8 μg of plasmid DNA comprising 4 μg of expression plasmid (CMV.Gal4.ADI, CMV.Gal4.ADII, or CMV.Gal4.ΔAD), 2 μg of reporter plasmid (UASg-Luc), and 2 μg of internal control plasmid CMV-beta Gal (28) encoding β-galactosidase. At 8 h after transfection, the growth medium of $C_2C_{12}$ cells was replaced with medium containing 2% horse serum and 0.5 μg of insulin per ml to permit differentiation. PC12 pheochromocytoma cells were transfected with Lipofectin (Gibco-BRL) in the presence of 10 μg of plasmid comprising 4 μg of expression plasmid (RSV.Gal4.ADI, RSV.Gal4.ADII, or RSV.Gal4.ΔAD), 4 μg of reporter plasmid (UASg-Luc), and 2 μg of the internal control plasmid RSV-CAT (11). At 48 h after transfections, cell extracts were prepared and assays for luciferase (by using Promega assay reagent as specified by the manufacturer), β-galactosidase (1), and CAT (12) activity were performed.

**Microinjection.** All solutions for microinjection were prepared by diluting preparations of plasmid DNA (in 10 mM Tris-HCl [pH 8.0] plus 1 mM EDTA) to a final concentration of 40 ng/μl. Phenol red (0.25%) was also included to monitor the injections. Zebra fish were raised and bred by standard methods (53). Immediately after spawning, the bottoms of the aquaria were siphoned. The harvested eggs were washed and placed in round, depressed grooves created in

FIG. 2. β-Galactosidase activity in embryos injected with ADI expression plasmids. Embryos (one to four cells) were injected with CMV-Gal4-ADI together with a β-galactosidase reporter plasmid (UAS-LaZ) at a ratio of 1:1. After 24 h, the embryos were fixed and incubated with X-Gal to visualize β-galactosidase activity. Representative embryos are shown at ×55 (a and b) and ×400 (c and d) magnifications. Blue cells were identified according to their location and morphology. Skin cells were the most superficial, with irregular structure. Sclerotome cells were identified by their close packing and ventral medial positions in the somites. Cells of the myotome were identified by their elongated morphology and striations. Cells were counted as notochord cells if they were located in the notochord. Nervous system cells were identified on the basis of their location in the spinal cord or brain. Extensive staining of nervous system cells is seen in panel c. The rounded cells in the lower right of panel d are skin cells.

agarose microinjection plates (53). Microinjection plates were prepared by pouring 1.2% agarose, containing 1 ppm of methylene blue, into a 9-cm-diameter petri plate containing several glass capillaries (1 mm by 5 to 6 cm) aligned in parallel. Upon hardening, the agarose disc was cut out and inverted and the capillaries were removed. Microinjection needles were made from 1-mm-diameter capillaries (World Precision Instruments; thin wall, with inner filament) with a Sutter P87 horizontal needle puller and backfilled with the DNA solution for injection. The needles were oriented with a Leitz micromanipulator, and DNA solution was introduced into the eggs (one- to four-cell stage) by using a pedal-controlled Eppendorf microinjector. The eggs were incubated at 28.5°C until appropriate analyses were performed. The injection procedure was monitored in a standard binocular microscope.

**Reporter function.** For quantitative reporter analyses, embryos were injected with a solution containing an RSV expression vector encoding the Gal4 DBD fused to an activation domain, a reporter plasmid (UASg-Luc), and an internal control plasmid (CMV-TK-CAT) at a ratio of 2:2:1. At 24 h later, embryos were collected, grouped in pools of three, frozen in dry ice, and disrupted in 100 μl of 0.1 M potassium phosphate (pH 7.8)–1 mM dithiothreitol by using the plunger of a 1-ml disposable syringe. The homogenate was subjected to three cycles of freeze-thawing and centrifuged for 1 min in an Eppendorf centrifuge. The supernatant solution was then assayed for reporter function. Luciferase activity was measured (with 40 μl of embryo extract) in a Lumac luminometer with a buffer containing 20 mM Tricine (pH 7.8), 20 mm KCl, 5 mM $MgSO_4$, 1 mM EDTA, 0.05 mM luciferin, 2.5 mM ATP, 0.25 mM coenzyme A, and 25 mM dithiothreitol. The CAT assay (with 30 μl of embryo extract) was performed as previously described (3).

For in situ analysis, embryos were injected with CMV expression vectors together with the β-galactosidase reporter plasmid UASg-LacZ (1:1 ratio). In situ staining for β-galactosidase activity was performed as described previously (54), except that the concentrations of both $K_4[Fe_3(CN)_6]$ and $K_3[Fe_2(CN)_6]$ were 5 mM, the concentration of X-Gal (5-bromo-4-chloro-3-indoly-β-D-galactoside) was 0.2%, and embryos were stained for 16 to 24 h. Cells were identified under the microscope following manual dechorionation. For each embryo, blue cells were classified into five major groups (skin, sclerotome, myotome, notochord, and nervous system) according to the criteria of cell location and morphology established by Westerfield et al. (54) and the percentage of cells in each category was calculated. Data were expressed as means ± standard errors of the means for all embryos in each group. Approximately 5 to 10% of blue cells could not be unambiguously identified and were not included in the subsequent analysis.

**Nuclear extracts and EMSA.** Embryos were collected 24 h after injection, manually dechorionated, pooled (30 to 100 embryos per pool) in 3-cm-diameter plates with a minimal amount (1 drop) of phosphate-buffered saline solution (Ca and Mg free), and squeezed with a cover slip. Cells were collected in Eppendorf tubes by rinsing the plates and recovered by centrifugation. Nuclear extracts were prepared by a modification (2) of published procedures (41). Double-stranded oligonucleotides were labeled by the Klenow fill-in reaction in the presence of [α-$^{32}$P]dATP. Electrophoretic mobility shift assay (EMSA) was performed as described previously (2). The sequences of the oligonucleotides used are as follows:

```
UASg:  5′ GATCCCGGAGTACTGTCCTCCGA
          GGCCTCATGACAGGAGGCTTTCG 5′
AP-1:  5′ GATCCGTTGCTGACTAATTGAGA
          GCAACGACTGATTAACTCTCTAG 5′
```

## RESULTS

**Functional test of mammalian activation domains in developing zebra fish.** To test whether mammalian activation domains can function in the developing zebra fish, expression plasmids encoding the Gal4 DBD (21) fused to the activation

FIG. 3. β-Galactosidase activity in embryos injected with ADII expression plasmids. Embryos (one to four cells) were injected with CMV-Gal4-ADII together with a β-galactosidase reporter plasmid as described in the legend to Fig. 2. After 24 h, the embryos were fixed and incubated with X-Gal to visualize β-galactosidase activity. Representative embryos are shown at ×55 (a and b) and ×400 (c and d) magnifications. Blue cells were classified as described in the legend to Fig. 2. The elongated cells in panels c and d are cells of the myotome.

domains of ADI, ADII, or VP16, controlled by the RSV transcription control region (11), were injected into one- to four-cell embryos. Each expression plasmid was coinjected with a reporter plasmid containing the luciferase gene controlled by a promoter containing five copies of the UASg sequence (the recognition sequence for the Gal4 DBD). A schematic representation of the expression and reporter plasmids is shown in Fig. 1a. The injected solution included an internal control plasmid containing the CAT gene under the control of the herpes simplex virus thymidine kinase promoter and the CMV enhancer.

After 24 h of development, whole-embryo extracts were prepared and the activities of luciferase and CAT were measured. CAT activity was used as an internal control to monitor efficient expression of injected genes; embryo extracts showing insignificant CAT activity (ineffective injection, nonviable embryos, etc.) were excluded from the analysis. Injection of the expression plasmid encoding the Gal4 DBD alone (ΔAD) produced very low luciferase activity (Fig. 1b). On the other hand, plasmids encoding Gal4 DBD fused to ADI, ADII, or VP16 produced significantly higher activities; of the three activation domains, ADII showed the least activity (Fig. 1b). Interestingly, ADI showed significantly higher activity than VP16, itself a highly potent activation domain in mammalian cells (50). Similar results were obtained when the CAT gene was used as a reporter and luciferase was used as a control for injection efficiency (data not shown).

**Cell specificity of activation domains in vivo.** To define the cell types in which the activation domains function, we chose to use the CMV transcriptional regulatory region in expression vectors, since it appears to function efficiently in all major cell types of the developing zebra fish (54). A reporter plasmid containing five copies of UASg linked to the gene encoding β-galactosidase was injected together with each expression plasmid. At 24 h after the injections, embryos were fixed and stained for β-galactosidase activity with the chromogenic substrate X-Gal. Embryos were examined under the microscope, and blue cells were assigned to one of five categories according to their morphology and location as described previously (54). Blue cells were not observed following staining of uninjected embryos (data not shown). Photographs of representative embryos are shown for embryos injected with expression plasmids encoding ADI (Fig. 2) and ADII (Fig. 3), although it should be borne in mind that classification is performed under the microscope and only a fraction of cells can be identified from the photographs. To illustrate the appearance of specific cell types, parallel injections were performed with the plasmid CMV-beta Gal; examples of specific cell types are shown to illustrate their appearance on photography (Fig. 4). In Fig. 2a and c, prominent staining of nervous system cells in the brain is apparent. The elongated cells in Fig. 3c and d are cells of the myotomes according to their characteristic morphology. Embryos injected with CMV expression vector-encoding VP16 frequently

FIG. 4. β-Galactosidase activity in embryos injected with the CMV-beta Gal plasmid (28). After 24 h, the embryos were fixed and incubated with X-Gal to visualize β-galactosidase activity. Embryos were examined under the microscope, and blue cells were identified as described in the legend to Fig. 2. Examples of the various cell categories are shown (magnifications: ×300 [a to h], ×750 [i]). (a) Skin. (b) Sclerotome. (c) Nervous system (top), sclerotome (bottom). (d) Nervous system. (e) Notochord. (f) Notochord. (g) Myotome. (h) Skin. (i) Myotome.

showed abnormal development and were therefore excluded from the analysis.

There is considerable qualitative and quantitative variation in the pattern of blue cells among embryos injected with a given plasmid (54) or combination of plasmids (Fig. 2 and 3). This presumably results from differences in the distribution of injected DNA from embryo to embryo (54). The injected DNA undergoes only limited replication and chromosomal integration (46). As a result, within a given embryo, a relatively small number of cells express the reporter even with a promoter (such as CMV) which is active in multiple cell types (8, 54). The simplicity and rapidity of the procedures, however, permit a representative pattern of distribution to be obtained for each activation domain by analysis of multiple embryos.

Such an analysis reveals a substantial difference in the distribution of positive cells resulting from injection of ADI vector compared with ADII; thus, ADI showed significant activity in all major cellular compartments whereas ADII showed little or no activity in the sclerotome, notochord, and nervous system (Fig. 5). ADII was strongly active in the myotome and skin (Fig. 5). The ratio of ADII to ADI ranged from 1.7 for the myotome to 0.1, 0, and 0 for the notochord, sclerotome, and nervous system, respectively (Fig. 5). A similar preference for the myotome was observed with expression vectors containing the RSV transcription control region replacing that of CMV (data not shown).

**Time course.** We also examined whether the preference of ADII is conserved during further development of the embryo. Embryos were injected with CMV-Gal4-ADII and stained after 24, 48, or 72 h of development (Fig. 6). After 48 h, all blue cells were either cells of the myotome or skin, and after 72 h, the preference for the myotomes became even higher, with over 90% of blue cells being cells of the myotome. Clearly, therefore, the selectivity of ADII is maintained up to 72 h. However, beyond 24 h, there was a significant decrease in the average number of positive cells per embryo, probably because of degradation of the plasmid DNA (54). As a result, analysis of expression patterns becomes more difficult at these later time points, and investigation of expression patterns at times beyond 72 h will require analysis of substantially larger numbers of embryos.

**EMSA.** To verify that the different activating plasmids were expressed at comparable levels in the embryo, EMSA was performed (Fig. 7). Nuclear extracts from embryos injected 24 h previously with expression plasmids encoding ADI, ADII, or VP16 were incubated with UASg probe and fractionated on a nondenaturing gel. With uninjected embryos, no DNA-binding complexes were seen (Fig. 7, lanes 10 to 12), whereas with extracts from injected embryos, several bands appeared. The faster-migrating band (band c) (Fig. 7, lanes 1, 3, 4, and 6) probably results from binding of monomer Gal4 fusion protein to the probe, whereas the two slower bands (band a and b) probably result from binding of dimers and tetramers. In lanes 7 and 9, the two major bands probably represent dimer and tetramer binding. The DNA-binding specificity of the observed complexes was tested: all binding was inhibited efficiently with nonradioactive UASg oligonucleotide (lanes 2, 5, and 8) whereas nonrelated oligonucleotide (AP1-binding sequence) failed to inhibit binding (lanes 3, 6, and 9). No dramatic differences in the intensity of the bands were seen on comparing

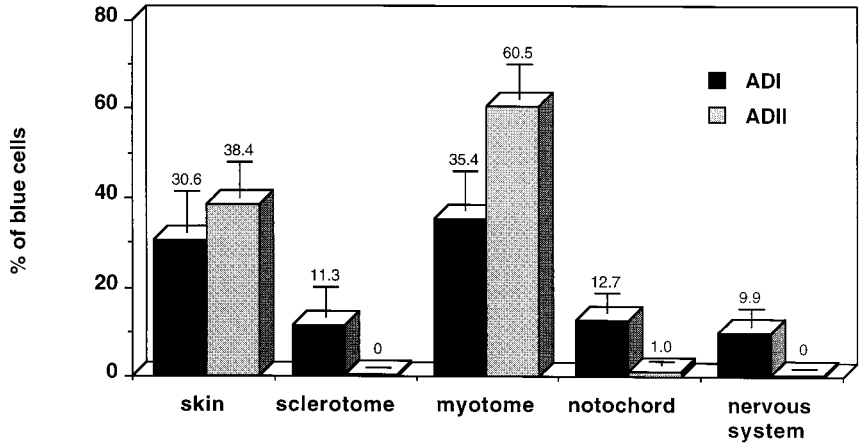



FIG. 5. Distribution of blue cells in embryos injected with CMV expression plasmids encoding Gal4-ADI or Gal4-ADII. Embryos were injected with the indicated CMV expression plasmid and the β-galactosidase reporter plasmid as described in the legend to Fig. 2. After 24 h, embryos were fixed and incubated with X-Gal to visualize β-galactosidase activity. For each embryo, blue cells were classified into the five major groups (see the legend to Fig. 2) and the percentage of cells in each category was calculated. Data are expressed as means ± standard errors of the means for each category. The analysis of ADI included 15 embryos (total of 179 blue cells), and the analysis of ADII included 22 embryos (total of 145 blue cells).

extracts from embryos injected with ADI or ADII expression plasmids. This indicates that the total DNA-binding activity of these Gal4 fusion proteins in injected embryos is similar; therefore, the observed differences in reporter activity probably reflect the intrinsic activity of the activation domain rather than differences in protein expression, stability, nuclear translocation, or DNA-binding efficiency.

Our previous results showed that ADII functions effectively in pancreatic beta cells but not in a variety of other cells including fibroblasts, epithelial cells, and liver cells (3). To extend the range of cell types studied to those relevant for the present study, we transfected myoblast ($C_2C_{12}$) and neuronal precursor (PC12) cells, as well as the previously characterized CHO fibroblasts and HIT pancreatic beta cells, with ADI and ADII expression vectors. As previously shown (3), ADI functions efficiently in CHO and HIT cells whereas ADII functions efficiently in HIT but not CHO cells (Table 1). ADI functions effectively in both $C_2C_{12}$ and PC12 lines, whereas ADII shows low activity in PC12 cells but substantial activity in $C_2C_{12}$ cells (Table 1). In this experiment, ADI showed extremely high levels of activity in $C_2C_{12}$ cells, higher than ADII in these cells and higher than ADI in other cells. However, ADI does not show strong preference for muscle cells in the in vivo system. Perhaps the in vivo system is more sensitive to the difference between an inactive and a modestly active activation domain rather than between a strong activation domain and a very strong one; when a cell contains a modestly active activation domain, sufficient β-galactosidase may be generated to produce a blue cell. Taken together, these transfection results parallel the observations made with developing embryos and further emphasize the biological relevance of our findings.

## DISCUSSION

Despite intensive efforts, the mechanisms whereby transcription activation domains increase the efficiency of the transcription process are incompletely understood (18). Several classes of activation domain have been distinguished, the most care-

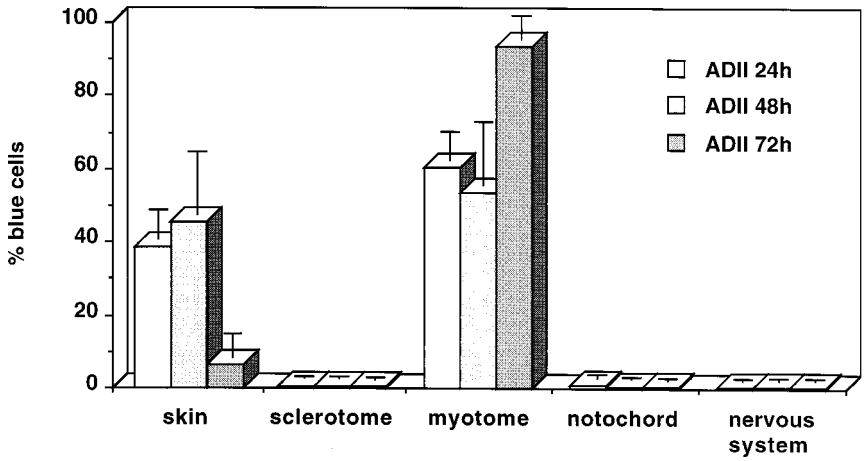



FIG. 6. Time course of *trans* activation by CMV-Gal4-ADII. Embryos (one to four cells) were injected with CMV-Gal4-ADII together with the β-galactosidase reporter plasmid as described in the legend to Fig. 3. Embryos were fixed after 24, 48, or 72 h and incubated with X-Gal to visualize β-galactosidase activity. Classification and quantitation of blue cells were as described in the legend to Fig. 2. Analysis of 48-h time points included seven embryos (total of 33 blue cells), and analysis of 72-h time points included seven embryos (total of 21 blue cells).

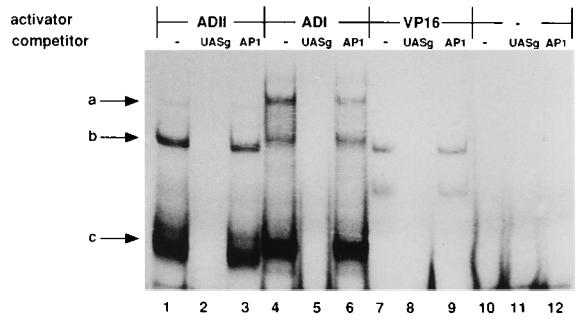



FIG. 7. EMSA with extracts of embryos injected with expression plasmids encoding Gal4 activation domain fusion proteins. Embryos were injected with CMV-Gal4-ADII (lanes 1 to 3), CMV-Gal4-ADI (lanes 4 to 6), or CMV-Gal4-VP16 (lanes 7 to 9). Lanes 10 to 12 represent extract from uninjected embryos. At 24 h after injection, nuclear extracts were prepared from whole embryos and incubated (10 μg of protein per reaction) with a [$^{32}$P]UASg probe without competitor (lanes 1, 4, 7, and 10), with a 100-fold molar excess of nonradioactive UASg competitor (lanes 2, 5, 8, and 11), or with a 100-fold molar excess of nonrelated competitor (AP1 sequence; lanes 3, 6, 9, and 12). Samples were resolved on a nondenaturing gel. Bands a, b, and c in lanes 1 to 6 probably result from binding of tetramers, dimers, and monomers, respectively, of the encoded Gal4 fusion proteins. The two bands in lanes 7 and 9 probably represent binding of tetramers and dimers to the probe.

fully studied being the acidic activation domain (14). Activation domains appear to interact with components of the transcription machinery in such a way that a rate-limiting component, e.g., TATA-binding protein (45) or TFIIB (26), is recruited to the promoter or a preassembled transcription complex is more efficiently brought to the promoter (23). There is evidence that different classes of activation domain may interact with distinct coactivator molecules or TATA-binding protein-associated factors (10, 17).

Our previous studies showed that the mouse E2A protein possesses two transcription activation domains; one of them (ADII) functioned preferentially in cultured pancreatic beta cells, whereas the other (ADI) was highly active in all cell lines tested (3). In the present study, we extend these findings to an in vivo setting. We chose to use the zebra fish because of the ease of collecting and injecting fertilized eggs, their rapid development, and the optical clarity of early embryos (53). Characterization of several zebra fish transcription factors (35, 42), including the E2A protein (56), has shown close sequence similarity to their mammalian counterparts. Furthermore, mammalian promoters function efficiently in fish cells (8). Recently, a transient-expression procedure based on microinjection to fertilized zebra fish embryos has been developed: it was demonstrated that mammalian promoters function in the fish in a region-specific fashion that is similar to that seen in mammals (54), indicating that endogenous fish transcription factors are able to recognize appropriate heterologous *cis* elements. We adapted this procedure, using appropriate expression and reporter plasmids, to permit measurement of transactivation domain function both in embryo extracts and in situ.

In agreement with the ability of some mammalian activation domains to function in a wide range of eukaryotes including yeast species (38, 39) and developing *Xenopus* embryos (57), we have shown that mammalian transcription activation domains are capable of functioning in the developing zebra fish. ADI functions effectively in all major compartments of the 24-h zebra fish embryo, in accordance with its activity in multiple mammalian cell types (3). On the other hand, ADII is more selective, being most active in cells of the myotome, the precursors of skeletal muscle cells. Given that HLH proteins in general, and E2A proteins in particular, are believed to play a key role in muscle cell determination (52), this result is intriguing. Indeed, transfection experiments performed with cultured mammalian myoblasts and neuronal precursor cells are completely consistent with the results of our in vivo experiments. It will be interesting to establish whether ADII also functions effectively in the pancreatic beta cells of developing zebra fish. Apparently, these cells are not yet present in the 24-h embryo, but, rather, the pancreatic islets begin to appear at 3 days postfertilization (36a). At this stage of development, we observed relatively few blue cells in microinjected embryos (presumably because the injected DNA becomes diluted out or degraded in the cells of the developing embryo). To examine ADII function in islets, it will therefore be necessary to analyze a larger number of mosaic transgenics or to generate transgenics bearing integrated copies of appropriate expression and reporter plasmids.

TABLE 1. Activity of ADI and ADII in multiple cell types

| Cell type | Fold transactivation for[a]: | |
|---|---|---|
| | ADI | ADII |
| HIT | 20.7 ± 6.4 | 51.1 ± 15.0 |
| $C_2C_{12}$ | 541.3 ± 127.9 | 20.1 ± 1.5 |
| PC12 | 40.9 ± 8.9 | 2.2 ± 0.4 |
| CHO | 47.7 ± 3.8 | 1.7 ± 0.3 |

[a] Cells were transfected as described in Materials and Methods. Normalized luciferase activity produced by ADI and ADII expression vectors is expressed relative to that produced by the control expression vector as described in Materials and Methods. Results shown are means ± standard errors of the means of three or four independent experiments.

It is currently believed that transcription activation involves multiple protein-protein interactions among transcription factors. Some degree of specificity may result from the steric constraints at the promoter region, dictating a unique combination of factors to bind at a given promoter (as in a jigsaw puzzle) (20, 25, 48), thus ensuring strict control of defined genes. Our previous work (3) and that of others (4, 49) with cultured cells indicate that an additional source of specificity derives from transcription activation domains that possess an intrinsic preference in particular cell types (independently of other promoter-specific transcription factors which normally function in vivo). This specificity may result from cell-specific covalent modifications of the activation domain, or, alternatively, the transcription activation domain may interact with coactivating factors distributed preferentially in particular cell types (27). Our present results show that cell-specific control of activation domain activity is apparent not only in cultured cell lines but also in vivo in a developing vertebrate. Such a mechanism may prove to be an important and general strategy for achieving appropriate control of eukaryotic transcription.

## ACKNOWLEDGMENTS

We thank B. Geiger, A. Yarden, and S. Bitzur for many helpful discussions and for generously sharing with us the use of their zebra fish facility. We also thank L. Colombo for his support and encouragement; I. Haviv, Z. Kam, C. W. Park, A. Rosen, and S. Thor for useful discussions; M. Pack and L. Moss for communication of unpublished results; and Y. Citri, Y. Shaul, and A. Yarden for constructive suggestions on the manuscript.

This work was supported by the Juvenile Diabetes Foundation, the Minerva Foundation, the Israel Cancer Research Fund, and the Forschheimer Center.

## REFERENCES

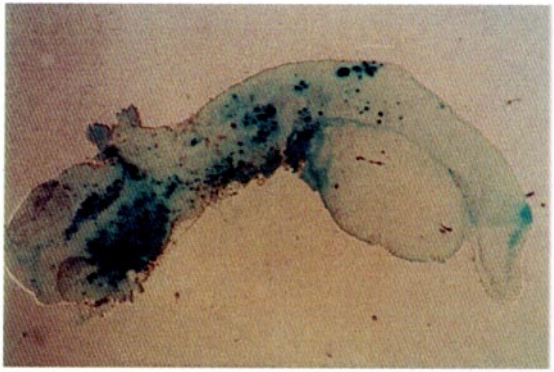
1. **An, G., K. Hidaka, and L. Siminovitch.** 1982. Expression of bacterial β-galactosidase in animal cells. Mol. Cell. Biol. **2:**1628–1632.
2. **Aronheim, A., H. Ohlsson, C. W. Park, T. Edlund, and M. D. Walker.** 1991. Distribution and characterization of helix-loop-helix enhancer-binding proteins from pancreatic β cells and lymphocytes. Nucleic Acids Res. **19:**3893–3899.

3. **Aronheim, A., R. Shiran, A. Rosen, and M. D. Walker.** 1993. The E2A gene product contains two separable and functionally distinct transcription activation domains. Proc. Natl. Acad. Sci. USA **90:**8063–8067.
4. **Baichwal, V. R., A. Park, and R. T. Tjian.** 1992. The cell-type-specific activator region of c-Jun juxtaposes constitutive and negatively regulated domains. Genes Dev. **6:**1493–1502.
5. **Boshart, M., M. Kluppel, A. Schmidt, G. Schutz, and B. Luckow.** 1992. Reporter constructs with low background activity utilizing the cat gene. Gene **110:**129–130.
6. **Cordle, S. R., E. Henderson, H. Masuoka, P. A. Weil, and R. Stein.** 1991. Pancreatic β-cell-type-specific transcription of the insulin gene is mediated by basic helix-loop-helix DNA-binding proteins. Mol. Cell. Biol. **11:**1734–1738.
7. **Ephrussi, A., G. M. Church, S. Tonegawa, and W. Gilbert.** 1985. B lineage-specific interactions of an immunoglobulin enhancer with cellular factors in vivo. Science **227:**134–138.
8. **Friedenreich, H., and M. Schartl.** 1990. Transient expression directed by homologous and heterologous promoter and enhancer sequences in fish cells. Nucleic Acids Res. **18:**3299–3305.
9. **German, M. S., M. A. Blanar, C. Nelson, L. G. Moss, and W. J. Rutter.** 1991. Two related helix-loop-helix proteins participate in separate cell-specific complexes that bind the insulin enhancer. Mol. Endocrinol. **5:**292–299.
10. **Goodrich, J. A., T. Hoey, C. J. Thut, A. Admon, and R. Tjian.** 1993. Drosophila TAFII40 interacts with both a VP16 activation domain and the basal transcription factor TFIIB. Cell **75:**519–530.
11. **Gorman, C. M., G. T. Merlino, M. C. Willingham, I. R. Pastan, and B. H. Howard.** 1982. The Rous sarcoma virus long terminal repeats is a strong promoter when introduced into a variety of eukaryotic cells by DNA-mediated transfection. Proc. Natl. Acad. Sci. USA **79:**6777–6781.
12. **Gorman, C. M., L. F. Moffat, and B. H. Howard.** 1982. Recombinant genomes which express chloramphenicol acetyltransferase in mammalian cells. Mol. Cell. Biol. **2:**1044–1051.
13. **Goutte, C., and A. D. Johnson.** 1993. Yeast a1 and alpha 2 homeodomain proteins form a DNA-binding activity with properties distinct from those of either protein. J. Mol. Biol. **233:**359–371.
14. **Hahn, S.** 1993. Structure and function of acidic transcription activators. Cell **72:**481–483.
15. **Hayashi, S., and M. P. Scott.** 1990. What determines the specificity of action of Drosophila homeodomain proteins. Cell **63:**883–894.
16. **Henthorn, P., M. Kiledjian, and T. Kadesch.** 1990. Two transcription factors that bind the immunoglobulin enhancer mE5/kE2 motif. Science **247:**467–470.
17. **Hoey, T., R. O. Weinzierl, G. Gill, J. L. Chen, B. D. Dynlacht, and R. Tjian.** 1993. Molecular cloning and functional analysis of Drosophila TAF110 reveal properties expected of coactivators. Cell **72:**247–260.
18. **Hori, R., and M. Carey.** 1994. The role of activators in assembly of RNA polymerase II transcription complexes. Curr. Opin. Genet Dev. **4:**236–244.
19. **Jan, Y. N., and L. Y. Jan.** 1993. HLH proteins, fly neurogenesis, and vertebrate myogenesis. Cell **75:**827–830.
20. **Johnson, P. F., and S. L. McKnight.** 1989. Eukaryotic transcriptional regulatory proteins. Annu. Rev. Biochem. **58:**799–839.
21. **Keegan, L., G. Gill, and M. Ptashne.** 1986. Separation of DNA binding from the transcription-activating function of a eukaryotic regulatory protein. Science **231:**699–704.
22. **Kenyon, C.** 1994. If birds can fly, why can't we—homeotic genes and evolution. Cell **78:**175–180.
23. **Koleske, A., and R. A. Young.** 1994. Holoenzyme composed of RNA polymerase II and general transcription factors is responsive to an activator. Nature (London) **368:**466–469.
24. **Lassar, A. B., R. L. Davis, W. E. Wright, T. Kadesch, C. Murre, A. Voronova, D. Baltimore, and H. Weintraub.** 1991. Functional activity of myogenic HLH proteins requires hetero-oligomerization with E12/E47-like proteins in vivo. Cell **66:**305–315.
25. **Lehming, N., D. Thanos, J. M. Brickman, J. Ma, T. Maniatis, and M. Ptashne.** 1994. An HMG-like protein that can switch a transcriptional activator to a repressor. Nature (London) **371:**175–179.
26. **Lin, Y.-S., and M. R. Green.** 1991. Mechanism of action of an acidic transcriptional activator in vitro. Cell **64:**971–981.
27. **Luo, Y., H. Fujii, T. Gerster, and R. G. Roeder.** 1992. A novel B cell-derived coactivator potentiates the activation of immunoglobulin promoters by octamer-binding transcription factors. Cell **71:**231–241.
28. **MacGregor, G. R., A. E. Mogg, J. F. Burke, and C. T. Caskey.** 1987. Histochemical staining of clonal mammalian cell lines expressing E coli β-galactosidase indicates heterogenous expression of the bacterial gene. Somatic Cell Mol. Genet. **13:**253–265.
29. **Manak, J. R., and M. P. Scott.** 1993. Able assistants for homeodomain proteins. Curr. Biol. **3:**318–320.
30. **Mitchell, P. J., and R. T. Tjian.** 1989. Transcriptional regulation in mammalian cells by sequence-specific DNA binding proteins. Science **245:**371–378.
31. **Murre, C., G. Bain, M. A. Vandijk, I. Engel, B. A. Furnari, M. E. Massari, J. R. Matthews, M. W. Quong, R. R. Rivera, and M. H. Stuiver.** 1994. Structure and function of helix-loop-helix proteins. Biochim. Biophys. Acta **1218:**129–135.
32. **Murre, C., P. S. McCaw, H. Vaessin, M. Caudy, L. Y. Jan, Y. N. Jan. C. V. Cabrera, J. N. Buskin, S. D. Hauschka, A. B. Lassar, H. Weintraub, and D. Baltimore.** 1989. Interactions between heterologous helix-loop-helix proteins generate complexes that bind specifically to a common DNA sequence. Cell **58:**537–544.
33. **Murre, C., A. Voronova, and D. Baltimore.** 1991. B-cell and myocyte-specific E2-box-binding factors contain E12/E47-like subunits. Mol. Cell. Biol. **11:**1156–1160.
34. **Naya, F. J., C. M. M. Stellrecht, and M. J. Tsai.** 1995. Tissue-specific regulation of the insulin gene by a novel basic helix-loop-helix transcription factor. Genes Dev. **9:**1009–1019.
35. **Njolstad, P. R., A. Molven, I. Hordvik, J. Apold, and A. Fjose.** 1988. Primary structure, developmentally regulated expression and potential duplication of the zebrafish homeobox gene ZF-21. Nucleic Acids Res. **16:**9097–9111.
36. **Ohlsson, H., O. Karlsson, and T. Edlund.** 1988. A beta-cell-specific protein binds to the two major regulatory sequences of the insulin gene enhancer. Proc. Natl. Acad. Sci. USA **85:**4228–4231.

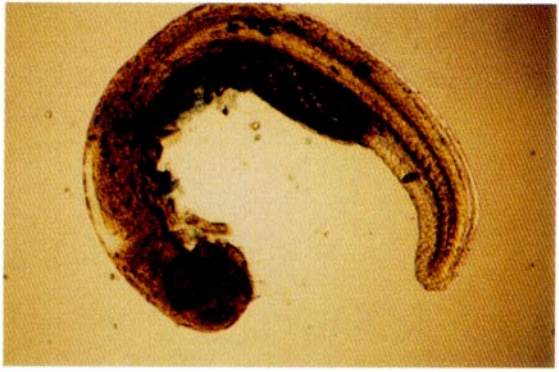
36a. **Pack, M., and L. Moss.** Personal communication.

37. **Park, C. W., and M. D. Walker.** 1992. Subunit structure of cell-specific E box binding proteins analyzed by quantitation of electrophoretic mobility shift. J. Biol. Chem. **267:**15642–15649.
38. **Ptashne, M.** 1988. How eukaryotic transcriptional activators work. Nature (London) **335:**683–689.
39. **Ptashne, M., and A. A. F. Gann.** 1990. Activators and targets. Nature (London) **346:**329–331.
40. **Quong, M. W., M. E. Massari, R. Zwart, and C. Murre.** 1993. A new transcriptional-activation motif restricted to a class of helix-loop-helix proteins is functionally conserved in both yeast and mammalian cells. Mol. Cell. Biol. **13:**792–800.
41. **Schreiber, E., P. Matthias, M. M. Muller, and W. Schaffner.** 1989. Rapid detection of octamer binding proteins with 'mini-extracts', prepared from a small number of cells. Nucleic Acids Res. **17:**6419.
42. **Schreiber-Agus, N., J. Horner, R. Torres, F. C. Chiu, and R. A. DePinho.** 1993. Zebra fish *myc* family and *max* genes: differential expression and oncogenic activity throughout vertebrate evolution. Mol. Cell. Biol. **13:**2765–2775.
43. **Shen, C. P., and T. Kadesch.** 1995. B-cell-specific DNA binding by an E47 homodimer. Mol. Cell. Biol. **15:**4518–4524.
44. **Shieh, S.-Y., and M.-J. Tsai.** 1991. Cell-specific and ubiquitous factors are responsible for the enhancer activity of the rat II insulin gene. J. Biol. Chem. **266:**16708–16714.
45. **Stringer, K. F., C. J. Ingles, and J. Greenblatt.** 1990. Direct and selective binding of an acidic transcriptional activator domain to the TATA-box factor TFIID. Nature (London) **345:**783–786.
46. **Stuart, G. W., J. V. McMurray, and M. Westerfield.** 1988. Replication, integration, and stable germ line transmission of foreign sequences injected into early zebrafish embryos. Development **103:**403–412.
47. **Sun, X.-H., and D. Baltimore.** 1991. An inhibitory domain of E12 transcription factor prevents DNA binding in E12 homodimers but not in E12 heterodimers. Cell **64:**459–470.
48. **Thompson, C. C., and S. L. McKnight.** 1992. Anatomy of an enhancer. Trends Genet. **8:**232–236.
49. **Tora, L., J. White, C. Brou, D. Tasset, N. Webster, E. Scheer, and P. Chambon.** 1989. The human estrogen receptor has two independent nonacidic transcriptional activation functions. Cell **59:**477–487.
50. **Triezenberg, S. J., R. C. Kingsbury, and S. L. McKnight.** 1988. Functional dissection of VP16, the trans-activator of herpes simplex virus immediate early gene expression. Genes Dev. **2:**718–729.
51. **Vierra, C. A., Y. Jacobs, L. Ly, and C. Nelson.** 1994. Patterns of Pan expression and role of Pan proteins in endocrine cell type-specific complex formation. Mol. Endocrinol. **8:**197–209.
52. **Weintraub, H., R. Davis, S. Tapscott, M. Thayer, M. Krause, R. Benezra, T. K. Blackwell, D. Turner, R. Rupp, S. Hollenberg, Y. Zhuang, and A. Lassar.** 1991. The *myoD* gene family: nodal point during specification of the muscle cell lineage. Science **251:**761–766.
53. **Westerfield, M.** 1993. The zebrafish book (edition 2.1). A guide for the laboratory use of zebrafish. University of Oregon Press, Eugene.
54. **Westerfield, M., J. Wegner, B. G. Jegalian, E. M. DeRobertis, and A. W. Puschel.** 1992. Specific activation of mammalian Hox promoters in mosaic transgenic zebrafish. Genes Dev. **6:**591–598.
55. **Wigler, M., A. Pellicer, S. Silverstein, R. Axel, G. Urlaub, and L. Chasin.** 1979. DNA-mediated transfer of the adenine phosphoribosyltransferase locus into mammalian cells. Proc. Natl. Acad. Sci. USA **76:**1373–1376.
56. **Wulbeck, C., C. Fromental-Ramain, and J. A. Campos-Ortega.** 1994. The HLH domain of a zebrafish hE12 homologue can partially substitute for functions of the HLH domain of Drosophila daughterless. Mech. Dev. **46:**73–85.
57. **Xu, L. C., D. Rungger, O. Georgiev, K. Seipel, and W. Schaffner.** 1994. Different potential of cellular and viral activators of transcription revealed in oocytes and early embryos of Xenopus-laevis. Biol. Chem. Hoppe-Seyler **375:**105–112.

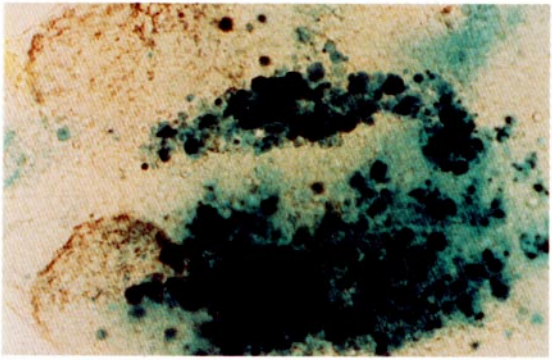
a



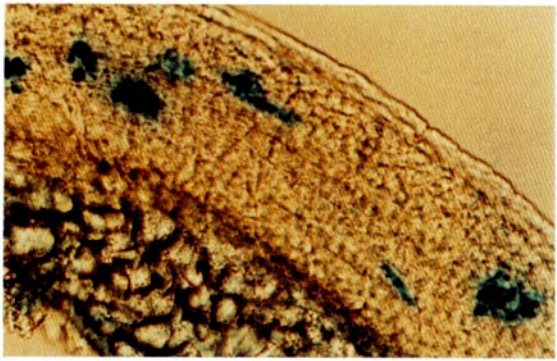
b

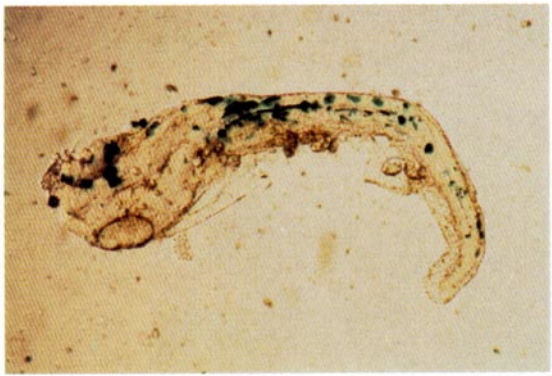


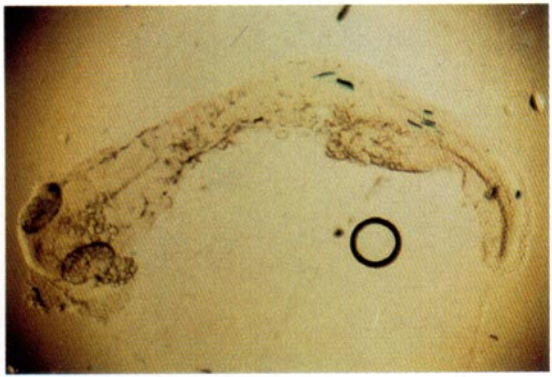
c



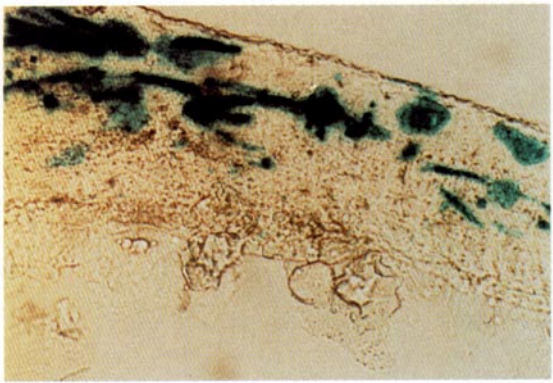
d



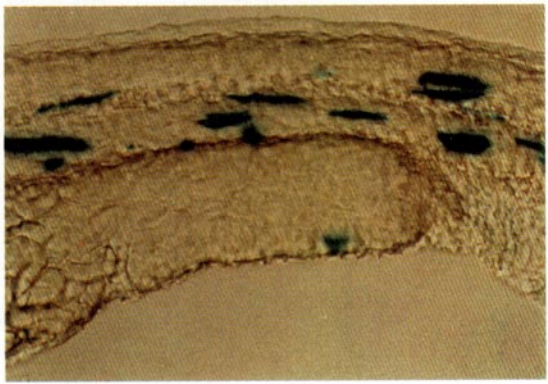


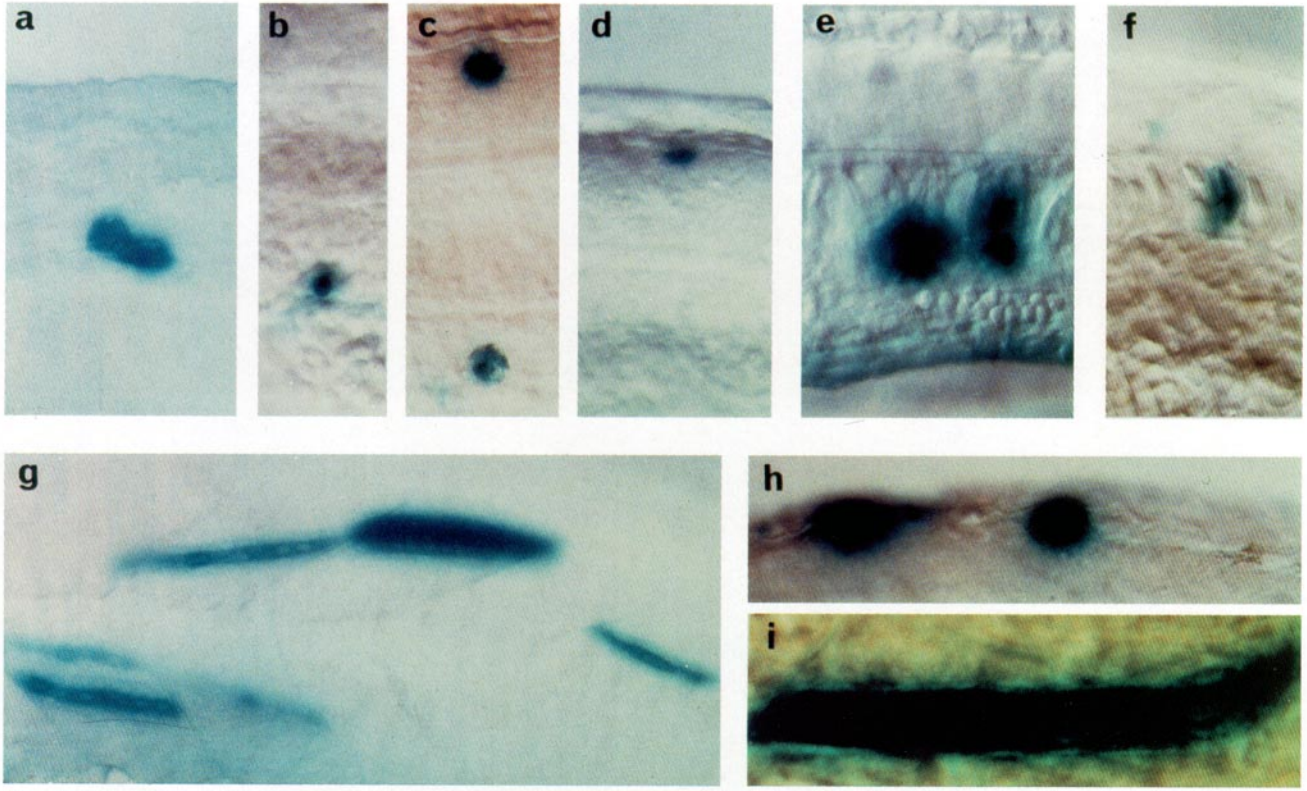


c



d




a
b
c
d
e
f
g
h
i