The promoter of the *endo A* cytokeratin gene is activated by a 3′ downstream enhancer

Yoshihiro Takemoto+, Yasuo Fujimura, Midori Matsumoto§, Yoshitaka Tamai, Takashi Morita, Aizo Matsushiro and Masami Nozaki*
Department of Microbial Genetics, Research Institute for Microbial Diseases, Osaka University, 3-1, Yamadaoka, Suita, Osaka 565, Japan

Received December 11, 1990; Revised and Accepted April 17, 1991

ABSTRACT

Mouse cytokeratin EndoA is an intermediate filament subunit of the type II cytokeratin class which initiates expression in trophectoderm cells of blastocyst during embryogenesis. To identify the regulatory elements of the *endo A* gene, we constructed a series of CAT expression vectors and transfected them into PYS-2 cells. We found an enhancer element locating 1 kb downstream from the *endo A* gene which acts on both the *endo A* and SV40 promoters. This enhancer consists of six direct repeated sequences with homology to the PEA3 motif in polyoma virus α enhancer core. In undifferentiated F9 embryonal carcinoma cells, expression of the construct containing the enhancer was not detected. These results indicate that one of the regulatory mechanisms of *endo A* gene expression is the 3′ downstream enhancer.

INTRODUCTION

Mouse embryos retain their developmental totipotency until the 8-cell stage. During blastocyst formation they differentiate into two distinct cell types; one is the trophectoderm and the other is the inner cell mass. This is the first differentiation of mouse development. To clarify the molecular mechanism of the first differentiation processes, the expression of some differentiation markers has been investigated. A type II cytokeratin EndoA, first identified as a trophectoderm specific marker (1), makes intermediate filaments with a type I cytokeratin partner, EndoB (2−6). m-RNA of the *endo A* gene is detected at the eight cell stage and is expressed in the trophectodermal cell lineage (7−9). Therefore, analyses of the regulatory mechanisms of *endo A* and *endo B* gene expression will shed light on the understanding of the first differentiation of mouse development.

In vitro studies of these developmentally and tissue specifically regulated genes utilize embryonal carcinoma (EC) cells, which express neither *endo A* nor *endo B* until the cells are induced to differentiate to extraembryonic endoderm (1,2,10). So far, it has been suggested that *endo B* gene expression is regulated by two different mechanisms. In nonepithelial differentiated cells that do not express *endo B*, expression is restricted by methylation in the CpG rich region around the first exon. Restriction in undifferentiated F9 cells may be due to either the absence of necessary transcriptional factors or the presence of negative regulatory activities (11). Moreover, the enhancer in the first intron of the human *endo B* gene is activated in F9 cells by c-*jun* and c-*fos* expression (12). In contrast, there is some indirect evidence of the regulatory mechanisms of *endo A* gene expression. A labile inhibitor blocking *endo A* gene expression in undifferentiated PCC4 EC cells has been suggested using cycloheximide, an inhibitor of protein synthesis (13). From their study, we envision that negative factors block the expression of the *endo A* gene in undifferentiated cells and after differentiation this gene is activated in endoderm cells resulting in the removal of the negative factors. In addition, methylation may block the expression of this gene, since EndoA synthesis has been reported to be induced in teratocarcinoma derived fibroblasts by 5-Azacytidine, which is an inhibitor of DNA methylation (14,15).

So far, *cis*-acting elements specifically involved in the regulation of the *endo A* gene expression have not been found, although a promoter has been identified in the −900 to +20 region (16). They used the α-2 collagen promoter, which has strong activity in all cell types, and did not examine downstream from +2000 of the gene. Therefore, specific regulatory elements have not been identified and the molecular mechanisms of *endo A* gene expression is still obscure.

In this paper we show that the *endo A* promoter is activated by the enhancer located downstream of the *endo A* gene depending on the cell type. We believe that these findings will be helpful in understanding the molecular mechanisms by which differentiation determinations occur.

MATERIALS AND METHODS

Isolation and sequencing of *endo A* genomic clone

Construction of a genomic library was as described previously (17). A total of 1.6×10^6 phage clones were screened by

* To whom correspondence should be addressed

Present addresses: +Tsukuba Life Science Center, The Institute of Physical and Chemical Research, 3-1-1, Koyadai, Tsukuba, Ibaraki 305 and §Department of Life Science, Faculty of Life Science and Technology, Tokyo Institute of Technology, Meguroku, Tokyo 152, Japan

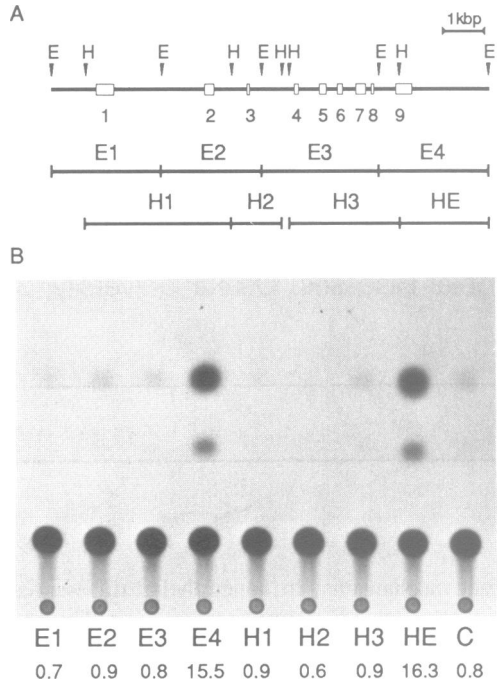

Figure 1. Identification of the *cis*-regulatory element of the *endo A* gene. (A) Restriction fragments from *endo A* gene. The *endo A* genomic DNA was sheared with *Eco*RI or *Bam*HI restriction enzymes and the resulting fragments (E1 to E4, H1 to H3, and HE) were inserted after the termination site of pendoA-CAT. Exons are indicated as open boxes. E, *Eco*RI; H, *Bam*HI. (B) Transient expression of CAT activity by pendoA-CAT containing fragments from the *endo A* gene. PYS-2 cells were cotransfected with 12 μg of pendoA-CAT containing each fragment as shown in (A) and 3 μg of β-galactosidase (β-gal) expression vector as the internal control. C, pendoA-CAT as a control. Conversion of chloramphenicol to the acetylated form is shown under the panel. The degree of conversion of CAT activity was normalized with that of β-gal activity.

hybridization with ^{32}P-labeled 1.8 kb *endo A* cDNA (6). The 10 kb fragment of one phage isolate which hybridized to the cDNA probe was digested with *Eco*RI, the restriction fragments were subcloned into the M13 phage vectors (40) and sequenced by the chain termination method (18).

Cells and transfections

All cells were maintained in Dulbecco's modified Eagle Medium (DMEM) supplemented with 10% heat-inactivated fetal calf serum. The medium was changed twice or three times every week. Transfections of cells were performed by calcium phosphate precipitation as described by Gorman (19) with a minor modification.

Plasmid constructions

The basic CAT plasmid used here, pendoA-CAT, was constructed as follows: The Klenow repaired *Xma*I-*Xma*I (−163 to +129) fragment, which contains the *endo A* promoter (16), was fused to the *Hind*III site filled in with Klenow polymerase in the coding region of the CAT gene of pSV0CAT(20). *Endo A* genomic DNA sheared with some restriction enzymes were inserted into the *Bam*HI site after the termination site of the CAT gene of pendoA-CAT.

Chloramphenicol acetyltransferase (CAT) assay

The plasmid pRSV-gal (21) was used as an internal control. Cell extraction was performed essentially as described by Sleigh (22).

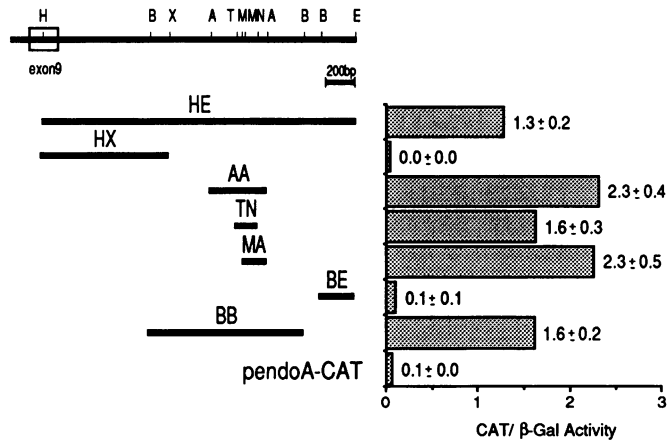

Figure 2. Identification of the minimal *cis*-acting element located at the 3′ end of the *endo A* gene. Various restriction fragments in the HE were inserted into pendoA-CAT and transfected into PYS-2 cells. The 9th exon is shown as an open box. The values shown are the means and standard deviation (± SD) of the CAT activity per β-gal activity (12) of three independent experiments. A, *Alu*I; B, *Bal*I; E, *Eco*RI; H, *Bam*HI; M, *Mae*I; N, *Nla*III; T, *Taq*I; X, *Xba*I.

Cell extracts were normalized with regard to β-galactosidase activity and the CAT assay was according to the method of Gorman et al. (20). Acetylated products were separated on a TLC-plate and the radioactivity was measured by β-scanning (AMBIS systems, San Diego) the TCL-plate.

RESULTS

The 3′ downstream element of the *endo A* gene has a positive effect on the *endo A* promoter

To identify the *cis*-acting element, we constructed the pendoA-CAT plasmid containing the *endo A* promoter and a part of the first exon fused to the chloramphenicol acetyltransferase (CAT) gene. Fragments of *endo A* genomic DNA sheared with *Eco*RI and *Bam*HI restriction enzymes (Fig. 1A) were inserted into the *Bam*HI site located downstream of the CAT gene. These constructs were introduced into PYS-2 cells expressing the *endo A* gene constitutively. Transient CAT assays showed that E1, E2, E3, H1, H2, and H3 fragments had no effect on the *endo A* promoter. On the other hand, the 2.5 kb E4 and 2.1 kb HE fragments stimulated CAT activity 19 to 20 fold (Fig. 1B). Since there are no significant differences between the CAT activities of the E4 and HE constructs, the *cis*-element should be located in the HE fragment (Fig. 1A). To define the *cis*-acting element more precisely, several HE restriction fragments were tested for *cis*-acting activity. As shown in Figure 2, 1.0 kb BB, 0.4 kb AA, 144 bp TN and 151 bp MA as well as HE stimulated CAT activity 27 to 38 fold, in which there are no significant differences, but the 0.9 kb HX and 0.2 kb BE fragments did not. These data indicated that the 204 bp sequence from *Taq*I to *Alu*I, localizing 1 kb downstream of the *endo A* gene, is sufficient to activate the *endo A* promoter. To analyze the *cis*-acting element in detail, a minimal fragment was sequenced. Figure 3 shows a part of the sequence, which was made of six direct repeated units of 22 bp with several mismatches (shadowed in Fig. 3). The TN fragment contains all six repeats and the MA fragment has four almost completely consensus repeat units, and these two fragments had similar CAT activity as described in Figure 2. These results indicated that the essential region of the *cis*-acting

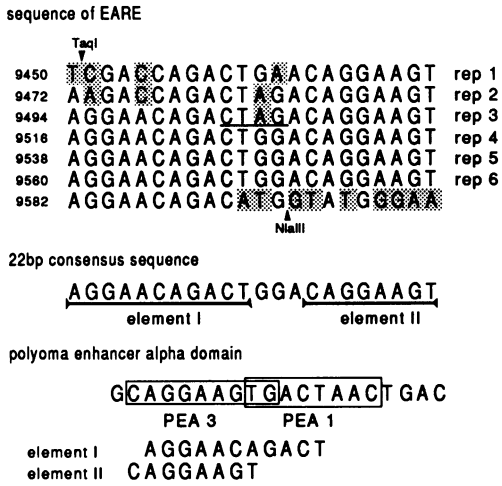

Figure 3. Nucleotide sequence of the *TaqI/NlaIII* EARE enhancer. Numbering of the sequence is relative to the *Eco*RI site 5′ upstream of the 1st exon shown in Figure 1 (A) as 1. Thin underlined sequence indicates *MaeI* site used to construct the MA fragment in Fig. 2. Shadowed nucleotides indicate the site mismatched to the 22 bp consensus sequence. Elements I and II are compared to the homologous sequence of the PEA1 and PEA3 motifs of the polyoma virus α enhancer. Boxed sequences indicate PEA1 and PEA3 motifs.

element is located in the repeat unit. We have termed this *cis*-regulatory element EARE for *endo A* repeat element. The consensus sequence of EARE is notable. Element I and II were homologous with a polyoma virus enhancer α domain protein 3 (PEA3) motif, AGGAAG (23) (Fig. 3).

EARE has enhancer activity

To test for enhancer activity, we inserted EARE upstream of the simian virus (SV) 40 promoter of pA10CAT2 plasmid (24) and downstream of the termination site of the CAT gene with both orientations (Fig. 4A). Plasmid pA10CAT2-EARE was transfected into PYS-2 cells. The EARE fragment remarkably activated the SV40 promoter independent of position and orientation (Fig. 4B, UN, UI, DN and DI). These results showed that EARE has enhancer activity. However, it is clear that EARE was more effective upstream of the promoter than downstream (about three to five fold).

Cell type specificity

During teratocarcinoma differentiation, *endo A* is expressed in differentiated endodermal cells (2,25). To compare the relative activities of the EARE enhancer in different cell types, pendoA-CAT-EARE plasmids were transfected into F9, PYS-2 and 3T3 (A31) cells. As shown in Figure 5, pendoA-CAT-EARE was predominantly active in PYS-2, weakly so in 3T3 cells but was not active in F9 cells and retinoic acid treated F9 cells (data not shown). The *endo A* enhancer, EARE, is thus inactive in undifferentiated and differentiated EC cells.

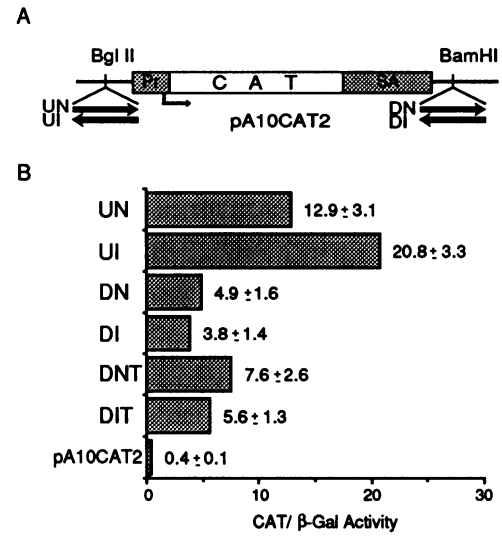

Figure 4. Enhancer activity of EARE. (A) Construction of pA10CAT2-EARE. EARE shown by bold arrows was introduced into *BglII* site (upstream of the SV40 promoter) and the *Bam*HI site (downstream of poly A additional signal) of pA10CAT2 with either orientation. Pr, SV40 early promoter; SA, SV40 splice and poly(A) signals. (B) CAT activity of pA10CAT2 containing EARE. U, upstream of SV40 promoter; D, downstream of poly A additional signal of CAT gene; N, Native direction; I, inverted direction; T, two copies of EARE. The values presented are the means and standard deviation (±SD) of the CAT activity per β-gal activity (12) of three different experiments.

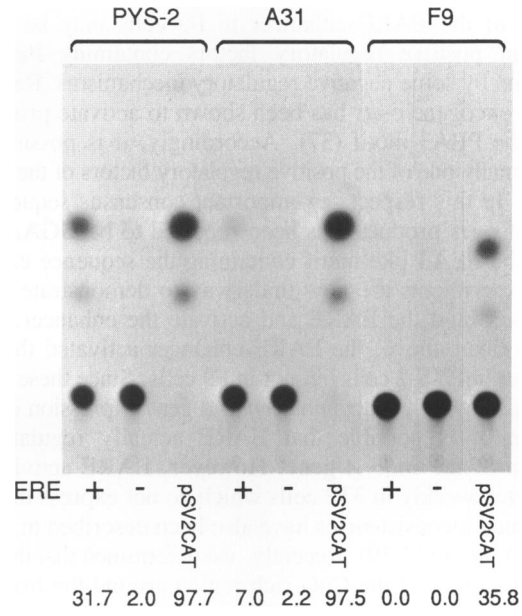

Figure 5. Cell type specific activity of EARE. pendoA-CAT-EARE, pendoA-CAT or pSV2CAT constructs were transfected into PYS-2, 3T3 (A31) and F9 cells as shown in the panel. Conversion of chloramphenicol to the acetylated form is shown under the panel.

DISCUSSION

We identified an enhancer element (*endo A* repeat element; EARE) located 3′ downstream of the *endo A* gene, which activated *endo A* promoter dependent on cell types. The EARE enhancer was made of six direct repeat units containing PEA3 like motif of polyoma virus enhancer α.

So far, a *cis*-acting element of the *endo A* gene has not been found, although the −1850 to +2000 region of the gene has been analysed using an α-2 collagen gene promoter (16). We found the *endo A* enhancer, that is the position of EARE, located 3′ downstream of the *endo A* gene, a region which has not been analyzed.

Most enhancers of eukaryotic genes described previously are found at the 5′ upstream region or the intron near the promoter. Recently, however, downstream enhancers have been found in β-globin (26,27), γ-globin (28), Histone H5 (29), T-cell receptor β-chain (30), T cell receptor α (31), TCR/CD3-ε (32), and immunoglobulin κ (33) genes as well as EARE. Of these enhancers, that of the chicken adult β-globin gene also had a positive effect on the regulation of the embryonic ε-globin gene positioned downstream of the β-globin gene (27). Moreover, it is suggested that the enhancer located at the 3′ end of the T cell receptor α locus can act over a 69 kb span and that any cellular oncogenes translocated to this locus in T-CLL and T-ALL cells will be under α enhancer regulation (31). Judging from these data, the EARE may act on a gene located down stream from the *endo A* gene. In this regard, the *endo A* gene usually coexpresses with a keratin partner, the *endo B* gene, in trophectodermal cell lineage, endodermal cells and many epithelial tissues (2,3,9,34). The possibility that the EARE plays some roles in *endo B* gene expression cannot be excluded, although the positional relationships of both genes have not yet been clarified.

The EARE has a homologous sequence to the PEA3 motif found in polyoma virus enhancer α, in which differentiation dependent activity has been detected in differentiated endodermal cells but not in undifferentiated F9 (EC) cells (35,36). In this report, we also showed that the EARE had enhancer activity in PYS-2 cells but not in F9 cells. These data indicate that the PEA3 like motif in EARE is controlled by the same regulatory mechanisms as the PEA3 motif in the polyoma enhancer. The inactivity of the EARE enhancer in F9 cells may be due to insufficient positive regulatory factors containing PEA3 or suppression by some negative regulatory mechanisms. Recently, the proto-oncogene c-*ets* has been shown to activate promoters through the PEA3 motif (37). Accordingly, it is possible that c-*ets* is actually one of the positive regulatory factors of the EARE enhancer. In this respect, an important consensus sequence of binding of c-*ets* products has been reported to be GGAA (38). Because the PEA3 like motif containing the sequence exists in EARE, experiments are now underway to demonstrate that c-*ets* products bind the EARE and activate the enhancer.

As described above, the EARE enhancer activated the *endo A* promoter in PYS-2 cells but not in F9 cells. Since these results were consistent with endogenous *endo A* gene expression in these cell types, it is possible that EARE actually regulates the expression of the *endo A* gene. However, EARE activity was also detected weakly in 3T3 cells which do not express the *endo A* gene. Such inconsistencies have also been described in human *endo B* in L cells (12,39). Recently, we determined that the *endo A* gene structure and the CpG rich region around the first exon of the gene was highly methylated in 3T3 cells but undermethylated in PYS-2 cells (40) as well as in *endo B* gene (11). These data suggested that *endo A* gene expression is controlled by two mechanisms. In differentiated nonepithelial cells such as fibroblasts, it may be suppressed by methylation of the DNA sequence. In contrast, in endodermal cells such as PYS-2, the *endo A* gene is undermethylated and expression is controlled by positive regulatory factors.

It remains to be determined whether the EARE enhancer is the only controlling element for the expression of the *endo A* gene. Recently, it has been suggested that the human *endo B* gene is positively regulated by c-*jun* and c-*fos* through the AP-1 motif in the first intron (12). In their report, endogenous *endo A* gene is also induced by c-*fos* expression in F9 cells. In addition, we found three AP-1 motifs in the *endo A* gene (40), which are located in the first exon, first intron, and 3′ downstream. In undifferentiated F9 cells, c-*jun* and c-*fos* are either absent or extremely low in abundance (41–43), but are induced after differentiation (42,43). The changes in the levels of both factors during differentiation are consistent with *endo A* and *endo B* gene expression. It has also been suggested that the *endo A* gene is negatively regulated by *trans*-acting factors (44). Therefore, it is possible that *endo A* gene expression is regulated by multiple regulatory mechanisms.

REFERENCES

1. Brûlet,P., Babinet,C., Kemler,R. and Jacob,F. (1980) *Proc. Natl. Acad. Sci. USA*, **77**, 4113–4117.
2. Oshima,R.G. (1981) *J. Biol. Chem.*, **256**, 8124–8133.
3. Howe,W.E., Klier,F.G. and Oshima,R.G. (1986) *J. Histochem. Cytochem.*, **34**, 785–793.
4. Singer,P.A., Trevor,K. and Oshima,R.G. (1986) *J. Biol. Chem.*, **261**, 538–547.
5. Nozaki,M., Iwakura,Y. and Matsushiro,A. (1986) *Dev. Biol.*, **113**, 17–28.
6. Morita,T., Tondella,M.L.C., Takemoto,Y., Hashido,K., Ichinose,Y., Nozaki,M. and Matsushiro,A. (1988) *Gene*, **68**, 109–117.
7. Duprey,P., Morello,D., Vasseur,M., Babinet,C., Condamine,H., Brûlet,P. and Jacob,F. (1985) *Proc. Natl. Acad. Sci. USA*, 82, 8535–8539.
8. Nozaki,M., Murata,K., Morita,T. and Matsushiro,A. (1988) *Biochem. Biophys. Res. Commun.*, **154**, 890–894.
9. Hashido,K., Morita,T., Matsushiro,A. and Nozaki,M. (1991) *Exp. Cell Res.*, **192**, 203–212.
10. Oshima,R.G. (1982) *J. Biol. Chem.*, **257**, 3414–3421.
11. Oshima,R.G., Trevor,K., Shevinsky,L.H., Ryder,O.A. and Ceceña,G. (1988) *Genes Dev.*, **2**, 505–516.
12. Oshima,R.G., Abrams,L. and Kulesh,D. (1990) *Genes Dev.*, **4**, 835–848.
13. Crémisi,C. and Duprey,P. (1987) *Nucleic Acids Res.*, **15**, 6105–6116.
14. Darmon,M. (1985) *Dev. Biol.*, **110**, 47–52.
15. Semat,A., Duprey,P., Vasseur, M. and Darmon,M. (1986) *Differentiation*, **31**, 61–66.
16. Brûlet,P., Duprey,P., Vasseur,M., Kaghad,M., Morello,D., Blanchet,P., Babinet,C., Condamine,H. and Jacob,F. (1985) *Cold Spring Harbor Symp. Quant. Biol.* vol. L, pp. 51–57.
17. Ichinose,Y., Morita,T., Zhang,F., Srimahasongcram,S., Tondella,M.L.C., Matsumoto,M., Nozaki,M. and Matsushiro,A. (1988) *Gene*, **70**, 85–95.
18. Sanger,F., Nicklen,S. and Coulson,A.R. (1977) *Proc. Natl. Acad. Sci. USA*, 74, 5463–5467.
19. Gorman,C. (1985) In Glover,D.M. (ed.), DNA cloning-A Practical Approach. IRL Press, Oxford. vol.II, pp. 143–159.
20. Gorman,C.M., Moffat,L.F. and Howard,B.H. (1982) *Mol. Cell. Biol.*, **2**, 1044–1051.
21. Edlund,T., Walker,M.D., Barr,P.J. and Rutter,W.J. (1985) *Science*, **230**, 912–916.
22. Sleigh,M.J. (1986) *Anal. Biochem.*, **156**, 251–256.
23. Martin,M.E., Piette,J., Yaniv,M., Tang,W.-J. and Folk,W.R. (1988) *Proc. Natl. Acad. Sci. USA*, **85**, 5839–5843.
24. Laimins,L.A., Gruss,P., Pozzatti,R. and Khoury,G. (1984) *J. Virol.*, **49**, 183–189.
25. Vasseur,M., Duprey,P., Brûlet,P. and Jacob,F. (1985) *Proc. Natl. Acad. Sci. USA*, 82, 1155–1159.
26. Boddine,D.M. and Ley,T.J. (1987) *EMBO J.*, **6**, 2997–3004.
27. Choi,O.-R.B. and Engel,J.D. (1988) *Cell*, **55**, 17–26.
28. Kollias,G., Hurst,J., deBoer,E. and Grosveld,F. (1987) *Nucleic Acids Res.*, **15**, 5739–5747.
29. Trainor,C.D., Stamler,S.J. and Engel,J.D. (1987) *Nature*, **328**, 827–830.
30. Krimpenfort,P., de Jong,R., Uematsu,Y., Dembic,Z., Ryser,S., von Boehmer,H., Steinmetz,M. and Berns,A. (1988) *EMBO J.*, **7**, 745–750.
31. Winoto,A. and Baltimore,D. (1989) *EMBO J.*, 8, 729–733.
32. Clevers,H., Lonberg,N., Dunlap,S., Lacy,E. and Terhorst,C. (1989) *EMBO J.*, **8**, 2527–2535.
33. Meyer,K.B. and Neuberger,M.S. (1989) *EMBO J.*, **8**, 1959–1964.
34. Oshima,R.G., Howe,W.E., Klier,F.G., Adamson,E.D. and Shevinsky,L.H. (1983) *Dev. Biol.*, **99**, 447–455.
35. Kryszke,M.-H., Piette,J. and Yaniv,M. (1987) *Nature*, **328**, 254–256.

36. Furukawa,K., Yamaguchi,Y., Ogawa,E., Shigesada,K., Satake,M., and Ito,Y. (1990) *Cell Growth & Differentiation*, **1**, 135–147.
37. Wasylyk,B., Wasylyk,C., Flores,P., Begue,A., Leprince,D. and Stehelin,D. (1990) *Nature*, **346**, 191–193.
38. Karim,F.D., Urness,L.D., Thummel,C.S., Klemsz,M.J., McKercher,S.R., Celada,A., Van Beveren,C., Maki,R.A., Gunther,C.V., Nye,J.A. and Graves,B.J. (1990) *Genes Dev.*, **4**, 1451–1453.
39. Kulesh,D.A. and Oshima,R.G. (1988) *Mol. Cell. Biol.*, **8**, 1540–1550.
40. Tamai,Y., Takemoto,Y., Matsumoto,M., Morita,T., Matsushiro,A. and Nozaki,M. (1991) *Gene*, in press.
41. Müller,R. and Wagner,E.F. (1984) *Nature*, **311**, 438–442.
42. Rüther,U., Wagner,E.F. and Müller,R. (1985) *EMBO J.*, **4**, 1775–1781.
43. Lockett,T.J. and Sleigh,M.J. (1987) *Exp. Cell Res.*, **173**, 370–378.
44. Mason,I., Murphy,D. and Hogan,B. (1985) *Differentiation*, **30**, 76–81.
45. Onclercq,R., Lavenu,A. and Crémisi,C. (1989) *Nucleic Acids Res.*, **17**, 735–753.