

Molecular cloning and characterization of PEA3, a new member of the *Ets* oncogene family that is differentially expressed in mouse embryonic cells

Ji-Hou Xin,¹ Alison Cowie,¹ Paul Lachance, and John A. Hassell²

Institute for Molecular Biology and Biotechnology, McMaster University, Hamilton, Ontario L8S 4K1 Canada

The PEA3 motif, first recognized in the polyomavirus enhancer, is an oncogene, serum growth factor, and phorbol ester-responsive element. An activity capable of binding to this sequence, termed PEA3 (polyomavirus enhancer activator 3), was identified in mouse 3T6 cell nuclear extracts. We have cloned cDNAs that encode PEA3 from a mouse FM3A cell cDNA library. A continuous open reading frame in the longest cDNA predicts a 555-amino-acid protein with a calculated molecular mass of 61 kD. Recombinant PEA3 binds to DNA with the same sequence specificity as that endogenous to FM3A cells and activates transcription through the PEA3 motif in HeLa cells. Deletion mapping of the protein revealed that the DNA-binding domain is located within a stretch of 102 amino acids near the carboxyl terminus. This region shares extensive sequence similarity with the ETS domain, a conserved protein sequence common to all *ets* gene family members. PEA3 is encoded by a 2.4-kb mRNA that is expressed to differing extents in fibroblastic and epithelial cell lines but not in hematopoietic cell lines. In the mouse, PEA3 expression is highly restricted; only the epididymis and the brain contain readily detectable amounts of its mRNA. Interestingly, the amount of PEA3 mRNA is down-regulated during retinoic acid-induced differentiation of mouse embryonic cell lines. These findings suggest that PEA3 plays a regulatory role during mouse embryogenesis.

[Key Words: Transcriptional activator; *ets* oncogene; PEA3 motif; developmental gene expression]

Received October 24, 1991; revised version accepted January 15, 1992.

The polyomavirus (Py) enhancer, one of the first such promoter elements to be described, acts in *cis* to enhance early and late gene transcription and viral DNA replication (Tyndall et al. 1981; Muller et al. 1983; Mueller et al. 1984; de Villiers et al. 1984; Hassell et al. 1986). It occupies ~200 bp within the noncoding region of the viral genome and is composed of multiple, functionally redundant sequence elements (Veldman et al. 1985; Mueller et al. 1988; Muller et al. 1988) that serve as binding sites for cellular transcriptional regulatory proteins (Piette et al. 1985; Bohnlein and Gruss 1986; Fujimura 1986; Ostapchuk et al. 1986; Piette and Yaniv 1986, 1987; Johnson et al. 1987; Martin et al. 1988; Asano et al. 1990; Kamachi et al. 1990). The activity of the enhancer, and by inference a number of its cognate factors, is regulated by serum, phorbol esters, and the products of a variety of nonnuclear oncogenes (C. Wasyluk et al. 1987, 1988; Imler et al. 1988; Satake et al. 1988; Yamaguchi et al. 1989; Asano et al. 1990; Murakami et al. 1990). In addition, the enhancer is differentially active

both between different cell types and within a given cell type, depending on its state of differentiation (Herbomel et al. 1984). These observations suggest that the activity and/or the expression of many of the proteins that bind to the viral enhancer is regulated by signals that govern cell proliferation and differentiation.

Although many sequence-specific DNA-binding activities that recognize sequences within the enhancer have been detected in mouse cells (the natural host of Py), with few exceptions, none of them has been unambiguously identified. To learn more about the nature, mechanism of action, regulation, and functional significance of the various activities that bind to the Py enhancer we are cloning mouse cDNAs that encode them as a first step in their analysis. Initially, we targeted PEA3 (polyomavirus enhancer activator 3), a putative transcriptional activator, for molecular cloning. PEA3 was originally described as a DNA-binding activity capable of protecting the sequence 5'-AGGAAG-3' (the PEA3 motif) from digestion by DNase I (Martin et al. 1988). The PEA3 motif, and by inference PEA3, activates transcription and Py DNA replication in murine cells in culture (Martin et al. 1988; Mueller et al. 1988; Muller et al. 1988; Murakami et al. 1990). The DNA-binding and transcrip-

¹These investigators contributed equally to the research described herein.

²Corresponding author.

tional activity of endogenous mouse PEA3 is increased by serum, tumor promoters, such as 12-*O*-tetradecanoylphorbol-13-acetate, and the products of several non-nuclear oncogenes including Py middle-T antigen, *v-src*, *Ha-ras*, *v-mos*, and *v-raf* (Wasylyk et al. 1989; Asano et al. 1990). PEA3 endogenous to human HeLa cells binds to the collagenase and *c-fos* promoters (Gutman and Wasylyk 1990), and putative PEA3-binding sites occur in the upstream regions of a number of other genes whose products are involved in cell growth, migration, and differentiation (Wasylyk et al. 1989; Gutman and Wasylyk 1990). Interestingly, the PEA3 motif forms part of the adenovirus enhancer core element (5'-A/C,GGAAGT-GAA-3') one of the first such elements originally identified by Hearing and Shenk (1983).

Here, we report the isolation of mouse cDNA clones that code for a protein capable of binding to the PEA3 motif. In keeping with the nomenclature established by Martin et al. (1988), we refer to this protein as PEA3. The amino acid sequence deduced from the PEA3 cDNA revealed that it is a member of the *ets* gene family. Recombinant PEA3 has the same DNA-binding specificity as endogenous PEA3 and, as with the latter, it activates transcription in mammalian cells in culture. Unlike other *ets* genes, PEA3 is not expressed in mouse hematopoietic cells and organs but, instead, in the epididymis and the brain. This, coupled with the observation that PEA3 expression is down-regulated following retinoic acid-induced differentiation of embryonal carcinoma cell lines, suggests that the protein plays a regulatory role in mammalian development and differentiation.

Results

Molecular cloning of cDNAs encoding PEA3

A cDNA capable of encoding PEA3 DNA-binding activity was isolated by screening a murine cDNA expression library in phage λ gt11 with a probe containing multimers of the PEA3 motif using the in situ filter detection method (Singh et al. 1988; Staudt et al. 1988; Vinson et al. 1988). The cDNA library was prepared from poly(A)-containing mRNA from mouse FM3A cells, a mammary epithelial cell line that expresses high levels of PEA3 DNA-binding activity. A screen of 500,000 plaques yielded two isolates expressing proteins that bound with specificity to the probe. Sequence analysis of these isolates suggested that one was a truncated version of the other. The phage (λ PEA3.1) with the longest open reading frame (ORF) was chosen for further analysis.

To study the DNA-binding properties of the protein encoded by the cDNA, we chose to overexpress it in *Escherichia coli* using the pGEX.3X expression vector (Smith and Johnson 1988). This vector allows for the synthesis of fusion proteins between glutathione-S-transferase (GST) and the protein encoded by the cDNA. The 1.6-kb cDNA insert was excised from λ PEA3.1 DNA by *EcoRI* cleavage and cloned into the *EcoRI* site of

pGEX.3X to yield pGEX.PEA3 (Fig. 1A). Two independent isolates of *E. coli* harboring this plasmid synthesized an abundant 65-kD fusion protein after induction with IPTG, as visualized by Coomassie brilliant blue staining (Fig. 1B, lanes 2,3). The molecular mass of the GST portion of this protein is 27 kD, yielding an estimated molecular mass of 38 kD for the PEA3-specific portion of the molecule. In contrast, a strain of *E. coli* transformed with pGEX.3X expressed only the 27-kD GST protein, which migrated slightly faster than the 31-kD molecular mass marker in this gel (Fig. 1B, lane 1).

The GST-PEA3 fusion protein was tested for its capacity to bind to the PEA3 motif after transfer to a nitrocellulose filter by the Southwestern technique (Miskimins et al. 1985). As illustrated in Figure 1B, lanes 5 and 6, a protein with the same size as the GST-PEA3 fusion protein is bound to the probe. Because this DNA-binding activity was not detected in cells that express GST (Fig. 1B, lane 4), we infer that the PEA3 portion of the fusion protein binds to the PEA3 motif.

To determine the molecular mass of endogenous mouse PEA3, we also subjected nuclear extracts prepared from FM3A cells to Southwestern analysis. Two different amounts of nuclear extract were fractionated by SDS-PAGE and either stained with Coomassie brilliant blue (Fig. 1C, lanes 7,8) or analyzed by the Southwestern technique (lanes 9,10). A major species of ~68 kD and two minor species of ~85 and 30 kD bound to the PEA3 motif. Repetition of this experiment with a probe containing a double point mutation in the PEA3 motif that abrogates PEA3 binding failed to detect these proteins (data not shown).

Recombinant and endogenous PEA3 bind to DNA with similar sequence specificity

To determine whether the bacterially produced fusion protein bound to DNA with the same sequence specificity as the DNA-binding activity in mouse cell nuclear extracts, we compared their DNA-binding properties by a number of assays. First, the sequence specificity of the mouse protein and the *E. coli* GST fusion protein were compared by electrophoretic mobility shift assays (EMSAs). Figure 2A shows the capacity of various mutant versions of the PEA3 motif to act as competitors in these assays. The wild-type oligonucleotide and mutant oligonucleotides A and E functioned as effective competitors in assays using either the mouse nuclear extract (lanes 2,3,7) or the bacterial extract (lanes 9,10,14). In contrast, mutant oligonucleotides B, C, and D functioned poorly as competitors in binding reactions with both the nuclear (lanes 4–6) and bacterial extracts (lanes 11–13). These observations indicate that the two activities bind to DNA with very similar sequence specificity.

Inspection of the data in Figure 2A allows derivation of an optimal PEA3-binding motif. Mutation of the sequences 5'-GGAA-3' in the center of the PEA3 motif (5'-AGGAAG-3') interfered with the capacity of the oligonucleotides to act as competitors in EMSAs, implying that these residues are critical for PEA3 binding. Simi-

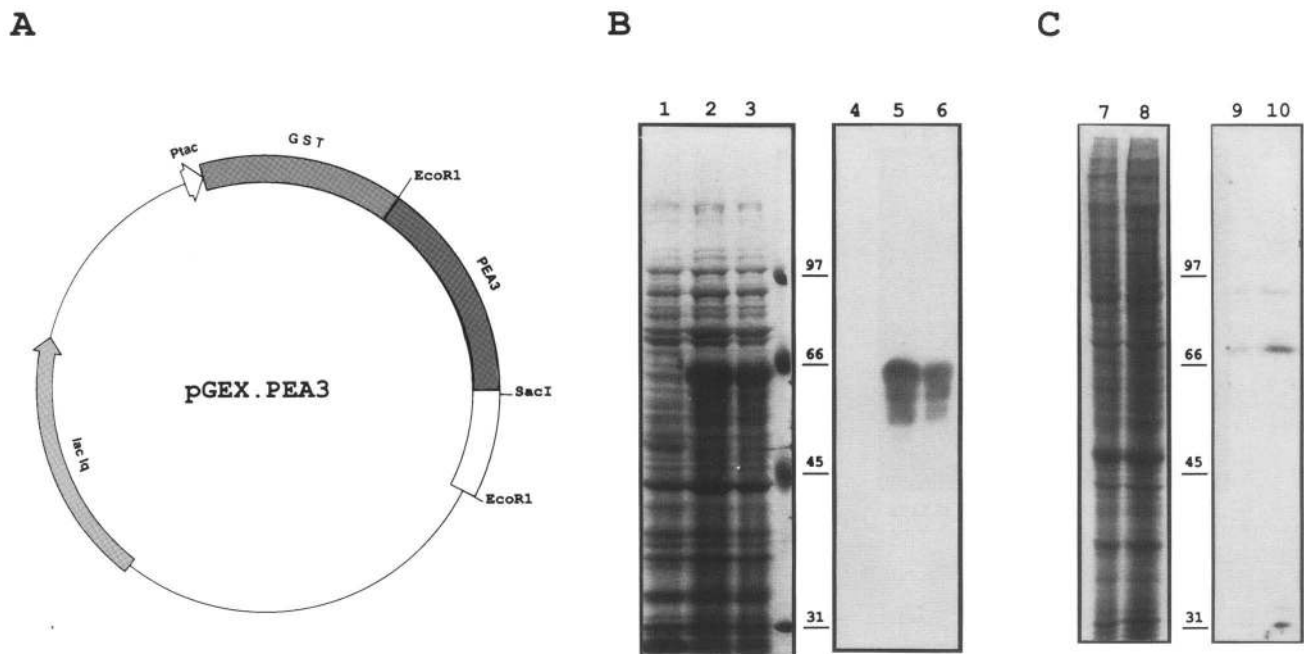


Figure 1. Overexpression of the PEA3.1 cDNA in *E. coli* and Southwestern blot analysis. (A) Map of the plasmid pGEX.PEA3. The shaded box depicts GST-coding sequences; the PEA3-coding region is shown by the darker box; its 3'-untranslated region by the open box. The gene encoding the *lac* repressor (*lac I^q*) is also shown. (B) SDS-PAGE and Southwestern analyses of bacterial extracts. Lanes 1–3 contain 10 μ g of total bacterial extract stained with Coomassie brilliant blue that had been prepared from one strain of *E. coli* harboring pGEX.3X (lane 1) and from two independent isolates carrying pGEX.PEA3 (lanes 2,3). The bacteria were treated with IPTG to induce expression of GST (lane 1) or the GST–PEA3 fusion proteins (lanes 2,3). Lanes 4–6 are the same as lanes 1–3, which had been transferred to nitrocellulose and reacted with an oligonucleotide containing the PEA3-binding site. (C) SDS-PAGE and Southwestern analyses of mouse nuclear extract. FM3A nuclear extract, 10 and 20 μ g (lanes 7 and 8, respectively), fractionated by SDS-PAGE and stained with Coomassie brilliant blue. Lanes 9 and 10 are the same as lanes 7 and 8, whose contents had been transferred to nitrocellulose and reacted as in B. Molecular weight standards are indicated between the lanes.

larly, mutation of the G and/or T at the 3' extremity of the motif (5'-AGGAAGT-3') also dramatically reduced the capacity of the oligonucleotide to act as a competitor. In contrast, mutation of the 5' A residue (5'-AGGAAG-3') to a C did not interfere with the capacity of the oligonucleotide to compete in EMSA, nor did mutation of sequences 3' of the PEA3 motif affect competition. These results suggest that the consensus PEA3 motif is closely related to the sequence 5'-A/C,GGAAGT-3'.

Additional evidence that the cloned cDNA encodes a protein with the same DNA-binding specificity as PEA3 endogenous to mouse cells was obtained by methylation interference experiments. A DNA fragment containing a single PEA3 motif was used in these experiments with the mouse nuclear extract and bacterial extract containing the GST fusion protein as a source of DNA-binding activity. The results are shown in Figure 2B for the mouse nuclear extract and in Figure 2C for the bacterially produced fusion protein. The bottom strand (see sequence below Fig. 2B,C) contains no A or G residues within the confines of the PEA3-binding motif; consequently, methylation of these residues on this strand did not affect binding of either protein (Fig. 2B,C, lanes 1–3). In contrast, methylation of the two consecutive G resi-

dues in the PEA3 motif (5'-AGGAAG-3') interfered with the binding of the activity in each extract (cf. Fig. 2B, lanes 4 and 5 with these same lanes in Fig. 2C). These results taken together with those of the competition and Southwestern analyses strongly suggest that the cDNA encodes a protein with the same DNA-binding specificity as the major PEA3-binding activity detected in the mouse nuclear extracts.

DNA sequence of PEA3

The 1.6-kb insert from λ PEA3.1 was sequenced by the dideoxy chain-termination method (Sanger et al. 1977) and found to contain an ORF capable of encoding a carboxy-terminal fragment of PEA3 comprising 329 amino acids (data not shown). This fragment has a predicted molecular mass of 36 kD, which agrees well with that (38 kD) estimated by measuring the difference in size between the GST–PEA3 fusion protein and GST after analysis by SDS-PAGE (see Fig. 1B). To derive a full-length cDNA, the partial cDNA was used as a hybridization probe to screen the FM3A cell cDNA library. Several positive phage were identified, and their cDNA inserts were sequenced. One of these, λ PEA3.9, contains an insert that encompasses the sequences borne by all

Xin et al.

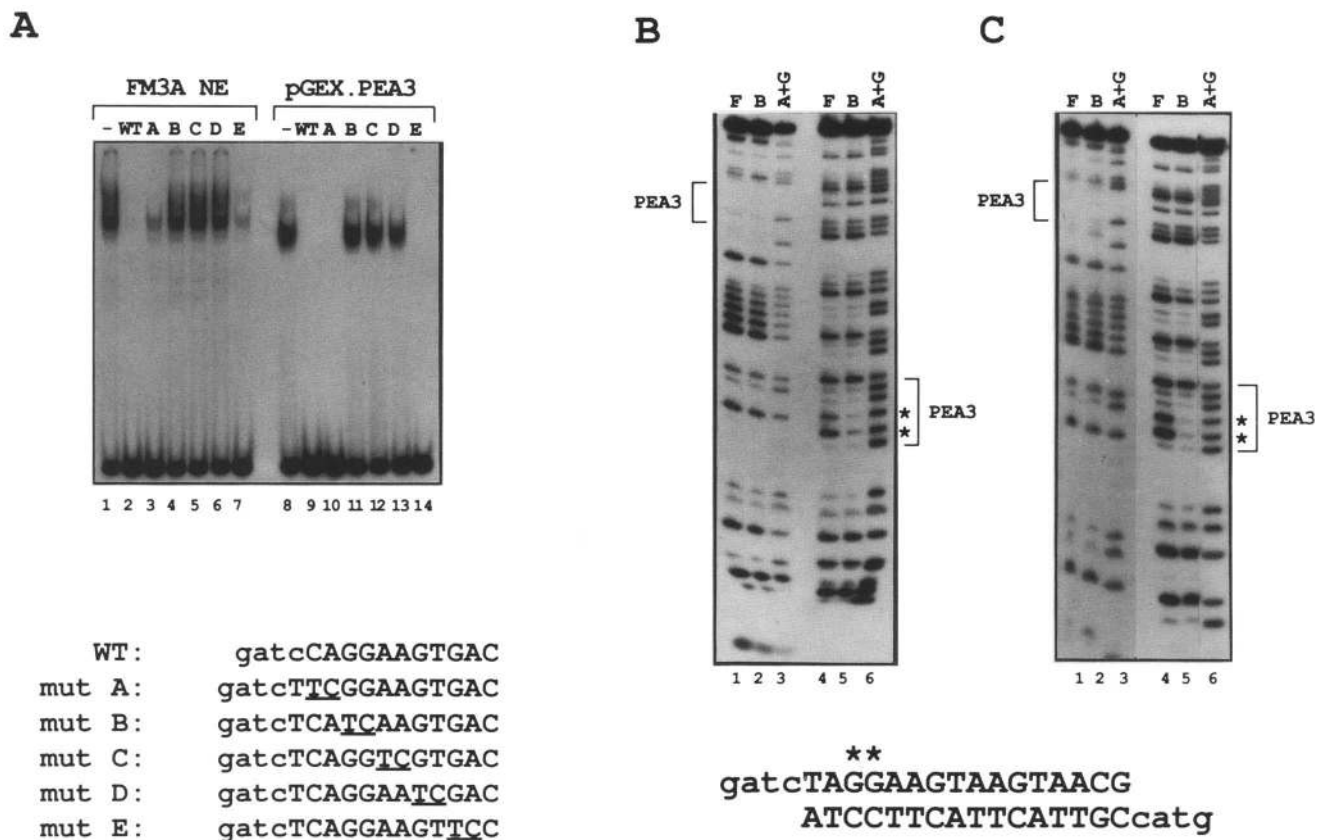


Figure 2. Comparison of the DNA-binding specificity of PEA3 endogenous to FM3A cells and the recombinant protein expressed in *E. coli*. (A) EMSA with either 5 μ g of FM3A nuclear extract (lanes 1–7) or 1 μ g of bacterial extract (lanes 8–14) in each reaction. Oligonucleotide competitors used in the reactions are indicated above the lanes by uppercase letters. The sequence of these oligonucleotides is shown below, with the mutations underlined. A 100-fold molar excess of each oligonucleotide was used in the competition experiments. (B) Methylation interference assays using FM3A nuclear extract. (Lanes 1–3) The lower DNA strand; (lanes 4–6) the upper strand. (F) The unbound DNA; (B) the bound fraction; (A + G) a chemical cleavage reaction of the probe that was used as a marker. The region of the gel corresponding to the PEA3 motif is indicated by the brackets; the G residues whose methylation affects binding are marked with an asterisk (*). (C) Methylation interference assays with bacterial extract. The symbols are the same as those used in B.

the others. The PEA3.9 cDNA comprises 2410 bp, which includes an ORF beginning with an AUG (nucleotide 126) and terminating with a stop codon beginning at nucleotide 1791 that could encode 555 contiguous amino acids (Fig. 3). The calculated molecular mass of the product of this ORF is 60,842 daltons. The putative initiation codon is preceded by three termination codons in two of the three potential reading frames. This initiation codon, however, lies in a sequence context that does not conform well to that defined by Kozak (1986) to be optimal for translation initiation. Moreover, an ORF that does not initiate with an AUG exists upstream of the putative initiation codon, which could encode an additional 28 amino acids that are continuous with and in the same frame as the major ORF beginning with the AUG. There are also two other potential translation initiation codons, an ACG at nucleotide 81 and a CUG at nucleotide 96 (Herman 1989), which are in-frame with the major ORF.

Several lines of evidence suggest that the AUG (nucle-

otide 126) is the initiation codon. First, overexpression of the cDNA in COS cells or transcription and translation of the cDNA in vitro in reticulocyte lysates yielded a 68-kD protein that was the same size as the major DNA-binding species detected by Southwestern analysis and by Western analysis of mouse FM3A cell extracts with a polyclonal antiserum (data not shown). The difference between the predicted and measured molecular mass of the protein may be the result of its high proline content (see below) or its post-translational modification. Second, deletion of all the sequences in the cDNA preceding the AUG did not alter significantly the efficiency of expression of PEA3 in COS cells or in reticulocyte lysates, nor did it affect the size of the protein product (data not shown). Therefore, it would appear that this AUG is used to initiate PEA3 translation.

The 5'-untranslated region of the cDNA is composed of 125 nucleotides, whereas its 3'-untranslated region comprises 620 nucleotides that do not include a poly(A) tail. There is a putative 3'-end formation-poly(A) signal

```

GAATTCGGTGGATGTGCTTTAAACCACACCTAACGTTTGAGCACAAGTCTCACGAAGTGGCGCTACAACCTCA 74
                                     a q v s r t g a t t s
TCGAAACGAAGTCTCCAAATCTGTCCAACGCAAAAACCAAAAGGAGTCTAATGACTAAGTCTTCCAACCACAAC 149
s e t k s p n l s n a k t q r s l M T K S S N H N      8
TGTCTGTGCCCGGAAAAACAAGCCGGGGTCTGGGGACCCGGGGCTCAGGCCGCTCGCTCCGGCTAGCCCC 224
C L L R P E N K P G L W G P G A Q A A S L R P S P    33
GCCACCTAGTTGTGTATCCCCGGGCATGCTGAGCATCCCCCGGGCTCCGGCAGAGACCCCGGACCTCAG 299
A T L V V S S P G H A E H P P A A P A Q T P G P Q   58
GTCTTGCCTCCGCGGGGGCCCGCCCTGTGGCCGGAGCGGCCGATGGAGCGGAGGATGAAAGGGGGA 374
V S A S A R G S P T P V A G G S G R M E R R M K G S  83
TACTTGGACCAGGAGTGCCTACACCTTCTGCAGAAATCTCCCGAAATGGGAGCTTGGGCGAAGCGCTGATG 449
Y L D Q R V P Y T F C S K S P G N G S L G E A L M   108
GTCCCAGGGAAAGTCCATGGACCCGGGTCCTGCCGCCCTCCGACTCAGAAGATCTCTTCCAGGACCTCAGT 524
V P Q G K L M D P G S L P P S D S E D L F Q D L S   133
CACTTCCAAGAGACGTGGCTCGCAGAAGCTCAGGTACCGGACAGTGTAGCAGTTTGTTCCTGATTTCATTCA 599
H F Q E T W L A E A Q V P D S D E Q F V P D F H S   158
GAAACTTAGCTTTCATAGCCCCACCACAGGATCAAGAAGGAACCCAGAGTCCCCGCACAGACCCCGCCCTG 674
E N L A F H S P T T R I K K E P Q S P R T D P A L   183
TCCTGCAGCAGGAAGCCACTCCCCTACCACCATGGAGAGCAGTGCCTTACTCCAGACAAATCGCCATCAAG 749
S C S R K P P L P Y H H G E Q C L Y S R Q I A I K   208
TCCCCCGCTCCCGGTGCCCTGGACAGTCGCCCTGCAGCCCTTTCCAGGGCAGAACAGCAGCAGACCTCCTG 824
S P A P G A P G Q S P L Q P F S R A E Q Q Q S L Q   233
AGAGCCTCCAGCTCTTCCCAGTCCCACCCTGGCCACGGGTACCTTGGTGAGCACAGTCCGCTTCTCCAGCAGCCC 899
R A S S S S Q S H P G H G Y L G E H S S V F Q Q P   258
GTGGACATGTGCCACTCTCACATCTCCTCAGGAGGGGGCCGGGAACCTTCCCAGCCCCCTATCAACACCAA 974
V D M C H S F T S P Q G G R E P L P A P Y Q H Q   283
CTGTCCGGCCCTGCCACCCTACCCCAAGCAACTTCAAGCAGGAGTACCATGACCCCTGTACGAACAGCCT 1049
L S E P C P P Y P Q Q N F K Q E Y H D P L Y E Q A   308
GGCCAGCCCCTTCAAGCCAGGGTGGGGTTCAGTGGCCACAGTACCAGGGCGGGGGTGGTGATCAAACAGGAG 1124
G Q P A S S Q G G V S G H R Y P G A G V V I K Q E   333
CGCACAGACTTCGCTACGACTCAGATGTCCTTGGATGTGCATCAATGTACTCCACCCAGAGGGTCTCTGGGA 1199
R T D F A Y D S D V P G C A S M Y L H P E G F S G   358
CCCTCTCCAGGTGATGGAGTATGGGTATGAAAATCCCTTCGACCATTCCCAGATGATGTCTGCATT 1274
P S P G D G V M G Y G Y E K S L R P F P D D V C I   383
GTCCCTAAAAAATTGAAGGACATCAAGCAGGAAGGGATGGAGCTTCCGGGAGGGGGCCACCTACAGCGC 1349
V P K K F E G D I K Q E G I G A F R E G P P Y Q R   408
CGCGTGTGCTTACAAGTGGCAGTTTCTGGTGGCCCTGCTGGATGACCCAAATAATGCTCATTTTATTGCTTGG 1424
R G A L Q L W Q F L V A L L D D P T N A H F I A W   433
ACAGGCGGGGAATGGAGTTAAACTAATTGAACCTGAAGAGGTTGCCAGGCTCTGGGGTATCCAGAAGAACCGG 1499
T G R G M E F K L I E P E E V A R L W G I Q K N R   458
CCAGCCATGAATTATGACAAGCTGAGCCGCTCGCTGCATATAATTATGAGAAAGGCATCATGCAGAAGTGGCT 1574
P A M N Y D K L S R S L R Y Y Y E K G I M Q K V A   483
GGCGAAGCTACGTGTACAAGTTGTGTGCGAGCCGGAGGCCCTGTTCTCTTGGCCTTCCAGATAATCAACGT 1649
G E R Y V Y K F V C E P E A L F S L A F P D N Q R   508
CCAGCTCTGAAGGCTGAGTTTGACCGCCAGTCAAGTGGAGGACACAGTCCCTTTGTCCTTGGATGAGAGT 1724
P A L K A E F D R P V S E E D T V P L S H L D E S   533
CCTGCCACTCCCAGAATCACTGGCCCGCTCCGCCCCTCGGCCACAGAGGTGGATATTTCTTACTAGGCACCA 1799
P A Y L P E L T G P A P P F G H R G G Y S Y       555
GTGGCTTCCCCTTGACATGGTGGGGTGGCTCAGTGTATATCAACTGATTGGTATTGGTGAAGCCCTCTTTT 1874
TGATGCCTGTAGAAGTCTCTGGGGTCCAGCTCCACTATCCCATCTGATACTCTGGCCAGACTCAGCTGCTAAC 1949
CAGAGTCTCGGGAAAACAGTGGAGGCAGGCCAAATCTAAAGGCAGTAGCTGAAGTTCCGCTGTGGCTCAGCTGT 2024
ACCTTCACTTCACTTGGCCTCTGCTAGGTCTTGTCTCAGAGGCAAGTTCCCTCACCCCAACAGAGATCCAG 2099
TGTTTATTTCTGGGACATACAGGACTTCCCTTGTATTATGGCAACAGGCGCAAGGGATTCTCAGAACACC 2174
CTGTGTCTCCCCTCTCCAACCCCATGGGAGACAAAGTTCGCTGGCTTCTGCCCTGAACAGGGGGGCTCTG 2249
TGTTCTGGTGTCTGTCTTGGGAGGCAGGAGATGTGGCGGCAGCTGGGGGGGGTGTGGAAGTAGAGATGGC 2324
TCTTGCCTTAGCCCTACCCAGGCGTAATTCACCTTTCCTTATGCCAGACCTTAATAAAGCCTCTGCTTCT 2399
CCCCGAATTC
                                                                                                                                                               2410

```

Figure 3. Nucleotide and deduced amino acid sequence of PEA3. Nucleotide numbers are shown in roman type, and amino acid numbers (with the Met at nucleotide 126 as amino acid number 1) are in italics. The lowercase letters depict the amino acid sequence corresponding to a predicted ORF, which does not initiate with an AUG. The three boxed regions from the amino terminus to the carboxyl terminus include an acidic region, a glutamine-rich region, and the ETS domain. A putative polyadenylation signal (5'-AATAAA-3') is underlined.

(5'-AATAAA-3'), 23 nucleotides from the end of the cDNA, suggesting that the 3' region of the cDNA includes most of the 3'-untranslated sequences in the mRNA. However, the 5' end of the cDNA is incomplete.

Primer extension analyses suggest that the cDNA lacks ≥ 90 nucleotides from its 5' end (data not shown).

Translation of the ORF predicts a protein rich in proline (12%), glycine (10%), and serine (9%) residues.

These amino acids do not appear to be clustered in any particular region of the molecule. In contrast, the region between amino acids 124 and 150 is rich in acidic amino acids (33%) and that between 217 and 315 is rich in glutamine residues (17%). Acidic and glutamine-rich regions have been found previously in the activation domains of transcription factors (Johnson and McKnight 1989; Mitchell and Tjian 1989). Comparison of the PEA3 DNA and amino acid sequence with those in the EMBL and Swiss Prot data bases revealed a region of sequence similarity at its carboxyl terminus to the Ets family of proteins. This region, termed the ETS domain, comprises ~85 amino acids that define the DNA-binding domain of several Ets proteins (Karim et al. 1990). Alignment of the ETS domain of PEA3 with that of other members of the family revealed ~60% sequence identity with most of the Ets family members (Fig. 4). PEA3 is least related to *Drosophila melanogaster* E74 protein (49%) and murine PU.1, the product of the *Spi-1* gene (36%). Comparison of the ETS domains of the 18 members of this family reveals that there are 19 invariant residues and 6 residues of similar structure within this region. A consensus ETS domain can be derived, which appears bipartite in structure: The amino-terminal half of the domain contains a conserved leucine cluster and a tryptophan repeat, whereas the carboxy-terminal half is rich in basic residues and tyrosine. The tryptophan repeat, three tryptophans spaced by 18–19 amino acids, has been recognized previously in the v-Myb oncoprotein as well as in Ets-related proteins (Anton and Frampton 1988; Karim et al. 1990).

DNA-binding domain of PEA3

To define the DNA-binding domain of PEA3, we derived separately a set of amino-terminal and carboxy-terminal deletion mutants of the GST-PEA3.1 fusion protein. These mutants were expressed in *E. coli* and the bacterial lysates used in EMSA with a radiolabeled probe bearing the PEA3 motif. We used this approach because PEA3 synthesized in vitro was incapable of binding to DNA. The results of this analysis are illustrated in Figure 5. Each of the mutant proteins was expressed to approximately the same extent in *E. coli*, and this was about the same as native GST (Fig. 5, cf. lane 1 with lanes 2–7 and 8–12). The results of the EMSA with each of the mutated proteins are illustrated in Figure 5B. Deletion of amino acids between the end of GST and PEA3 amino acid 331 had no effect on the capacity of the resulting fusion protein to bind to the PEA3 motif (cf. lanes 2 and 3). A deletion extending to amino acid 410 reduced, but did not abolish, the DNA-binding capacity of the fusion protein (cf. lanes 3 and 4). In contrast, deletion of 6 or more amino acids beyond this point completely abolished DNA binding at all concentrations of extract tested (lanes 5,6). This establishes the amino-terminal border of the DNA-binding domain between amino acids 410 and 416. Carboxy-terminal deletion mutants with end points in noncoding sequences (Δ Sac) or in coding sequences up to amino acid 511 bound the DNA probe

efficiently (lanes 11,12). In contrast, all mutants with deletions larger than this failed to bind to the probe (lanes 8–10). These results place the carboxyl terminus of the PEA3 DNA-binding domain between amino acids 475 and 511. This analysis suggests that the DNA-binding domain of PEA3 is located within the region bound by amino acids 410 and 511, a segment that encompasses the ETS domain, which has been shown previously to function as the DNA-binding domain of several other Ets-related proteins (Karim et al. 1990).

Transcriptional activation by recombinant PEA3

The PEA3 motif in the Py enhancer functions to activate transcription and viral DNA replication, implying that the factor that binds to this element augments both processes. To determine whether the cDNA borne by λ PEA3.9 encodes a protein capable of enhancing transcription, we expressed it in HeLa cells and measured its capacity to effect expression of a reporter plasmid containing a minimal promoter composed of multimers of the PEA3 motif and a TATA element juxtaposed to the chloramphenicol acetyltransferase (CAT) gene and 3' processing/polyadenylation signals from SV40 (Fig. 6C). As a control, we also tested the capacity of recombinant PEA3 to enhance CAT synthesis in the same cell line from a reporter bearing a mutated version of the PEA3 motif known to have much reduced affinity for PEA3 (mutant B, see Fig. 2A). The reporter carrying the wild-type PEA3 motifs was expressed in HeLa cells, whereas that bearing a mutant version of this motif was not (Fig. 6A,B, cf. lanes 1 and 3). This illustrates that efficient CAT gene expression is dependent on the PEA3 motif in these cells. Presumably, PEA3 activity endogenous to HeLa cells acts through the PEA3 motif to activate CAT transcription. Cotransfection of the reporters with an expression vector capable of encoding recombinant PEA3 enhanced expression of CAT by the reporter bearing the wild-type PEA3 elements by a factor of ≥ 10 -fold (Fig. 6A,B, lane 2). Expectedly, recombinant PEA3 was unable to act through the mutant elements to activate CAT gene expression (Fig. 6A,B, lane 4). These observations indicate that the cDNA encodes a transcriptional activator that functions through a wild-type, but not a mutant, PEA3 motif to enhance gene expression.

Expression of PEA3 mRNA is cell specific

To establish the size and examine the cell specificity of expression of PEA3 RNA, equal amounts of total cellular RNA from various mouse, rat, monkey, and human cell lines of fibroblastic and epithelial origin were analyzed by Northern blot analysis with the PEA3.1 probe (Fig. 7). Two RNA species (2.4 and 4.1 kb) of differing abundance were detected in all of the cell lines with the exception of the NS1 mouse B-cell line. The failure to detect PEA3 RNA in this cell line was not the result of RNA degradation, because subsequent analysis of the same blot with a β -actin gene probe (Ponte et al. 1983) revealed the presence of essentially equal amounts of its transcript in

% Identity

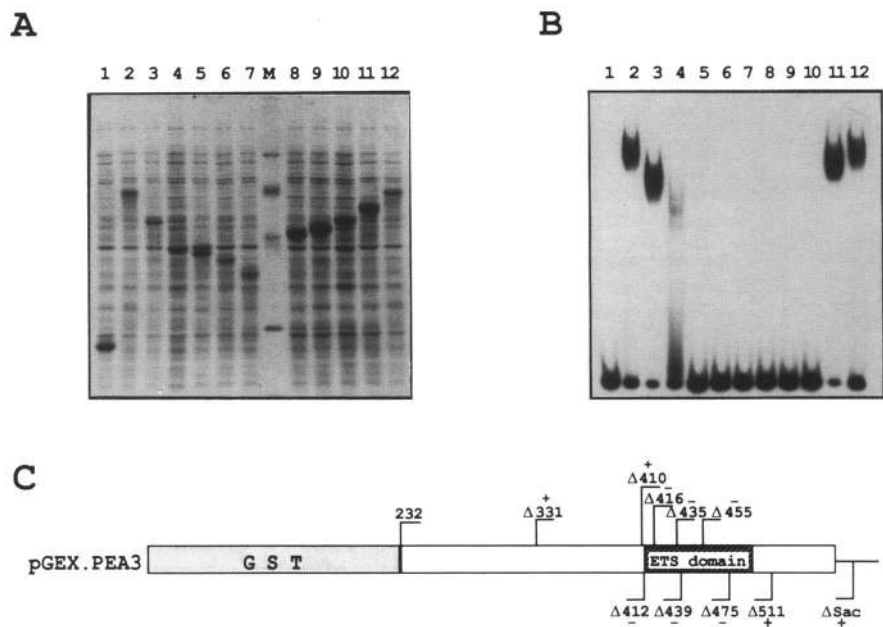
PEA3	100%	CAIQMWFILVALLDPTNAHFIAWVG-RGMEFKLI--EPEKVARLWGIQK-NRPAMNVDKLSHSGIARYYEKGINQKVGAGERVYTK	491
D-Ets-2	66%	GPIQLWQFILLEHLDK-CQSFISWVG-DGWEFKLT--DDEKVARWNGIRK-NKPKMAYEKLSSGIRFYIDKNI IHKTAGKRYVYRF	150
SU-Ets-2	65%	GPIQLWQFILLEHLTKK-CQHII SWVG-DGWEFKLS--DEDEVARRWGRK-NKPKMAYEKLSSGIRFYIDKNI IHKTAGKRYVYRF	88
M/H/C/X-Ets-1	63%	GPIQLWQFILLEHLTDKS-CQSFISWVG-DGWEFKLS--DEDEVARRWGRK-NKPKMAYEKLSSGIRFYIDKNI IHKTAGKRYVYRF	245/414/458/411
M/H-Ets-2	62%	GPIQLWQFILLEHLSDKS-CQSFISWVG-DGWEFKLIA--DEDEVARRWGRK-NKPKMAYEKLSSGIRFYIDKNI IHKTAGKRYVYRF	441/442
C-Ets-2	62%	GPIQLWQFILLEHLTDKS-CQSFISWVG-DGWEFKLA--DDEVARRWGRK-NKPKMAYEKLSSGIRFYIDKNI IHKTAGKRYVYRF	452
X-Ets-2	62%	GPIQLWQFILLEHLTDKS-CQSFISWVG-DGWEFKLT--DDEVARRWGRK-NKPKMAYEKLSSGIRFYIDKNI IHKTAGKRYVYRF	325
M-Fli-1	62%	GQIQMWFILLEHLSDSA-NASCIWEG-TNGEFKMT--DDEVARRWGRK-NKPKMAYEKLSSGIRFYIDKNI IHKTAGKRYVYRF	360
V-E26	61%	GPIQLWQFILLEHLTDKS-CQSFISWVG-DGWEFKLS--DDEVARRWGRK-NKPKMAYEKLSSGIRFYIDKNI IHKTAGKRYVYRF	639
H-Erg-1/2	61%	GQIQMWFILLEHLSDSS-NSSCIWEG-TNGEFKMT--DDEVARRWGRK-NKPKMAYEKLSSGIRFYIDKNI IHKTAGKRYVYRF	373
H-Elk-1	60%	--VTMWFILQREQGSCHII SWVSRDGEFKLV--DADEVARRWGLR-NKTNMAYDKLSRALRIFYIDKNI IRKYSQKRFVYRF	85
D-E1g	56%	GQVQWQFILLEHLTACE-HTDVIEAVG-TEGEFKLT--DDEVARRWGEKK-NKPKMAYEKLSSGIRFYIDKNI IHKTAGKRYVYRF	105
R-GABPα	56%	GQIQMWFILLEHLTDKD-ARDCSIWVG-DEGEFKLN--QPELVQKNGQRM-NKPKMAYEKLSSGIRFYIDKNI IHKTAGKRYVYRF	399
D-E74	49%	----LMEFILLKLEQREYCPRFKFNREKGYEKLIV--DSKAVSRWGMHR-NKPKMAYEKLSSGIRFYIDKNI IHKTAGKRYVYRF	868
M/H-PU.1 (Spi-1)	36%	KKIRIYQFILLDLRSQDMKDS-IMWVDKDKGTQFSSKHEKALHRRGIGKGRKKTQFMABALRNNGKTYEVKVK-KKLTQSF	248/246

CONSENSUS
 ----LW-FLV-LL-----I-W-----F-----V--WG--K-R--M-Y----R-LR-Y-----K-----R--Y-F
 Y L I L K K

Figure 4. Comparison of the ETS domain of PEA3 with all known ets gene family members. Percent identity with PEA3 is indicated, as is the amino acid number of the amino-terminal residue of each sequence (right). Conserved amino acids are shaded. A derived consensus sequence is shown at the bottom. Sequence information for the ets family members was obtained from either the EMBL or Swiss-Prot data bases using the FSTNSCAN or FSTPSCAN or FSTPSCAN programs by IntelliGenetics. The letters preceding each protein refer to the species of origin: (D) *Drosophila*; (SU) sea urchin; (R) rat; (M) mouse; (H) human; (C) chicken; (X) *Xenopus*; (V) virus E26. D-Ets-2 (Pribyl et al. 1988); Su-Ets-2 (Chen et al. 1988); C-Ets-1 (Leprince et al. 1988); M-Ets-1 (Chen 1990); X-Ets-1 (Steigler et al. 1990); H-Ets-1, M/H-Ets-2 (Watson et al. 1988); C-Ets-2 (Boulukos et al. 1988); X-Ets-2 (Wolff et al. 1990); M-Fli-1 (Ben-David et al. 1991); V-E26 (Nunn et al. 1983); H-Erg-1/2 (Reddy et al. 1987); H-Elk-1 (Rao et al. 1989); D-Elg (Pribyl et al. 1991); R-GABPα (LaMarco et al. 1991); D-E74 (Burtis et al. 1990); M-PU.1 (M-Spi-1) (Klemz et al. 1990); H-PU.1 (H-Spi-1) (Ray et al. 1990).

Xin et al.

Figure 5. Deletion mapping of the PEA3 DNA-binding domain. (A) Coomassie-stained gel of bacterial extracts. Ten micrograms of protein was loaded per lane. (B) EMSA using bacterial extract with the radiolabeled PEA3 motif. One microgram of extract was used per assay. Lanes 1–12 in A correspond to lanes 1–12 in B. Extracts were prepared from *E. coli* harboring pGEX.3X (lane 1), pGEX.PEA3 (lane 2), or various mutants of the latter (lanes 3–12) and analyzed by EMSA. The results of analysis of mutants with amino-terminal deletions between amino acid 232 and 331 (lane 3), 410 (lane 4), 416 (lane 5), 435 (lane 6), and 455 (lane 7) are shown. The results of analysis of mutants with carboxy-terminal deletions from the *EcoRI* site (Fig. 3A) to the *SacI* site (lane 12) and amino acid 511 (lane 11), 475 (lane 10), 439 (lane 9), and 412 (lane 8) are also shown. (C) Linear representation of the GST-PEA3 fusion protein with the position of the ETS domain is shown. Amino-terminal deletion end points are indicated above the line, and carboxy-terminal deletion end points are shown below it. The + and - signs refer to DNA-binding activity of each mutant.



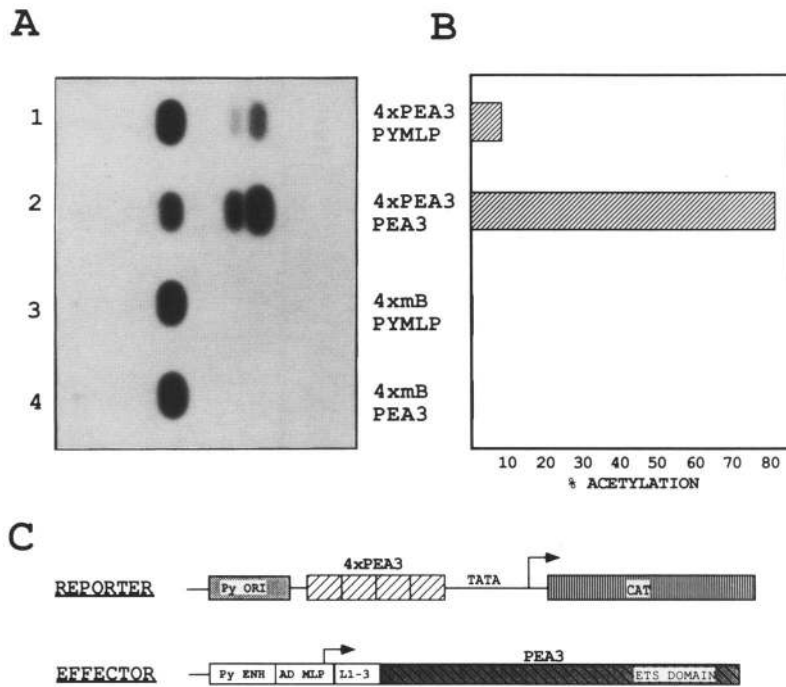
each lane (data not shown). FM3A cells, from which the PEA3 cDNAs were isolated, expressed abundant quantities of the 2.4-kb RNA, as did mouse 3T3 and C127 cells. In general, the murine cell lines preferentially expressed the 2.4-kb RNA, whereas the primate cell lines expressed the larger RNA in greater quantities than they did the smaller RNA. (The 4.1-kb RNA from human and monkey cells is somewhat larger than that from the mouse and rat cell lines.) It is not clear whether the 4.1-kb RNA is an alternate PEA3 RNA (derived by differential initiation, splicing, or termination/polyadenylation of transcription), or a PEA3-related species. Among the *ets*-related genes, only the 4.0-kb mouse *Fli-1* mRNA, which is ubiquitously expressed, is about the same size as the larger PEA3 transcript (Ben-David et al. 1991). To learn whether the failure to detect PEA3 RNA in the mouse B-cell line was unique to these cells, we also assayed other mouse T-cell lines, as well as numerous mouse erythroleukemic cell lines, for PEA3 RNA. We did not detect PEA3 transcripts in any of these cell lines (data not shown). These observations indicate that PEA3 RNA generally is not expressed in murine cells of lymphoid or erythroid origin.

To examine the tissue distribution of PEA3 mRNA, we isolated RNA from various mouse organs and used gene-specific, antisense RNA probes in RNase protection assays (Fig. 8). Hybridization of the PEA3 probe with the tissue RNA followed by RNase digestion is expected to yield a radiolabeled product of 305 nucleotides, whereas hybridization of the RNA with the ribosomal protein gene probe (Dudov and Perry 1984), used as a

control, is expected to yield a protected fragment of 195 nucleotides (Fig. 8). Among the tissues examined, only epididymis (lane 10) and brain (lane 2) expressed appreciable amounts of PEA3 RNA in comparison to the FM3A cell line (lane 14). Mammary glands (lane 13) also contained PEA3 RNA, but in amounts much reduced compared with these other organs. In keeping with data obtained from analyses of hematopoietic cell lines, neither the spleen (lane 6), the thymus (lane 8), nor the bone marrow (lane 11) expressed PEA3 RNA. These results indicate that the PEA3 gene is expressed in a tissue-restricted manner quite unlike that of any other *ets* gene.

PEA3 expression is developmentally regulated

Previous analyses of expression of *ets*-related genes in *D. melanogaster* (Burtis et al. 1990) and *Xenopus laevis* (Chen et al. 1990) have suggested that their protein products play a role in the early development of these organisms. To learn whether PEA3 might act similarly during early mammalian development, we used embryonic cell lines to investigate the expression of PEA3 RNA before and after the induction of their differentiation. The F9 and P19 lines of mouse embryonic carcinoma cells were induced to differentiate with retinoic acid. Under the conditions used, the F9 cells differentiated into cells akin to parietal extraembryonic endoderm, whereas the P19 cells became neurons and astrocytes (Rudnicki and McBurney 1987). Total cellular RNA was isolated from the cells before and for various periods during retinoic acid treatment and examined for PEA3 transcripts by



carrying the PEA3 cDNA (lane 2), the mutant reporter with the control effector (lane 3), and the mutant reporter with the effector encoding cloned PEA3 (lane 4) are shown. (B) Graphic representation of the data displayed in A. (C) Illustration of the general structure of the reporter and effector plasmids.

Northern blot analysis. The results are illustrated in Figure 9, and they clearly show that retinoic acid treatment of both F9 (Fig. 9A) and P19 embryonal carcinoma cells (Fig. 9B) resulted in a marked reduction of the steady-

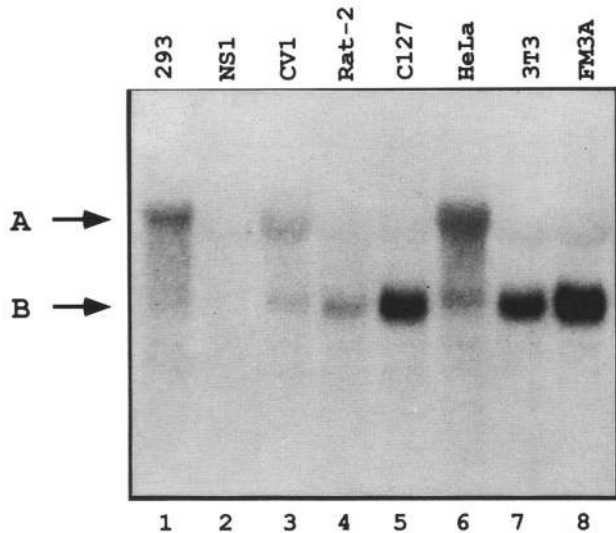


Figure 7. Northern blot analysis of PEA3 RNA in different cell lines. Ten micrograms of total RNA was used for each sample, and the blot was probed with the radiolabeled PEA3.1 cDNA. The cellular source of RNA is indicated above the lanes. (A) The 4.1-kb RNA; (B) the 2.4-kb RNA species.

Figure 6. Transcriptional activation by recombinant PEA3 in HeLa cells. (A) Autoradiograph of a thin-layer chromatography plate illustrating the activation of CAT gene expression effected by the cloned PEA3 cDNA. The reporter constructs (C) contain four copies of the wild-type PEA3 motif (4x PEA3) or a mutant version thereof (4x mB) carrying a double-point mutation that blocks PEA3-binding juxtaposed to a synthetic TATA box, which directs transcription initiation of the downstream CAT gene. The presence of the Py origin (Py ori) in this reporter is not relevant in these experiments because FM3A cells do not support replication of the reporter unless Py large tumor antigen is provided in *trans*. These reporter DNAs were cotransfected separately with either an effector plasmid comprising the Py enhancer, adenovirus major late promoter and tripartite leader, and the PEA3 cDNA (PEA3) or a control effector carrying the same sequences but lacking the PEA3 cDNA (PYMLP). CAT enzymatic activity was measured and normalized by reference to the total amount of protein in each sample. The results of cotransfections of the reporter carrying the wild-type PEA3 motif (4x PEA3) with the control effector (PTYMLP) lacking the PEA3 cDNA (lane 1); the wild-type reporter with the effector

carrying the PEA3 cDNA (lane 2); the mutant reporter with the control effector (lane 3), and the mutant reporter with the effector encoding cloned PEA3 (lane 4) are shown. (B) Graphic representation of the data displayed in A. (C) Illustration of the general structure of the reporter and effector plasmids.

state amount of PEA3 RNA. The decline in PEA3 RNA was more rapid in P19 cells than in F9 cells. Under identical conditions, the amount of β -actin RNA per cell remained unchanged. These observations suggest that retinoic acid treatment of embryonal carcinoma cells reduces the synthesis or enhances the instability of PEA3 RNA, and raise the prospect that PEA3 expression is developmentally regulated during mouse embryogenesis.

Discussion

We have cloned mouse cDNAs that code for a transcriptional activator capable of binding to the PEA3 motif. Several lines of evidence indicate that the protein encoded by these cDNAs corresponds to PEA3, the DNA-binding activity initially identified in mouse cells that binds to the Py enhancer (Martin et al. 1988). First, the cDNA-encoded protein bound to DNA with the same sequence specificity as that endogenous to mouse FM3A cells. Second, the apparent molecular size of the protein product encoded by the ORF in the cDNA, synthesized either in vitro in reticulocyte lysates or in vivo in monkey COS cells, was the same as that of the endogenous mouse PEA3 identified by Western or Southwestern analyses (data not shown). The discrepancy between the calculated molecular mass of PEA3 (61 kD) and its molecular mass measured by SDS-PAGE (68 kD) may be a consequence of its high proline content or could be the result of post-translational modifications that cause its

Xin et al.

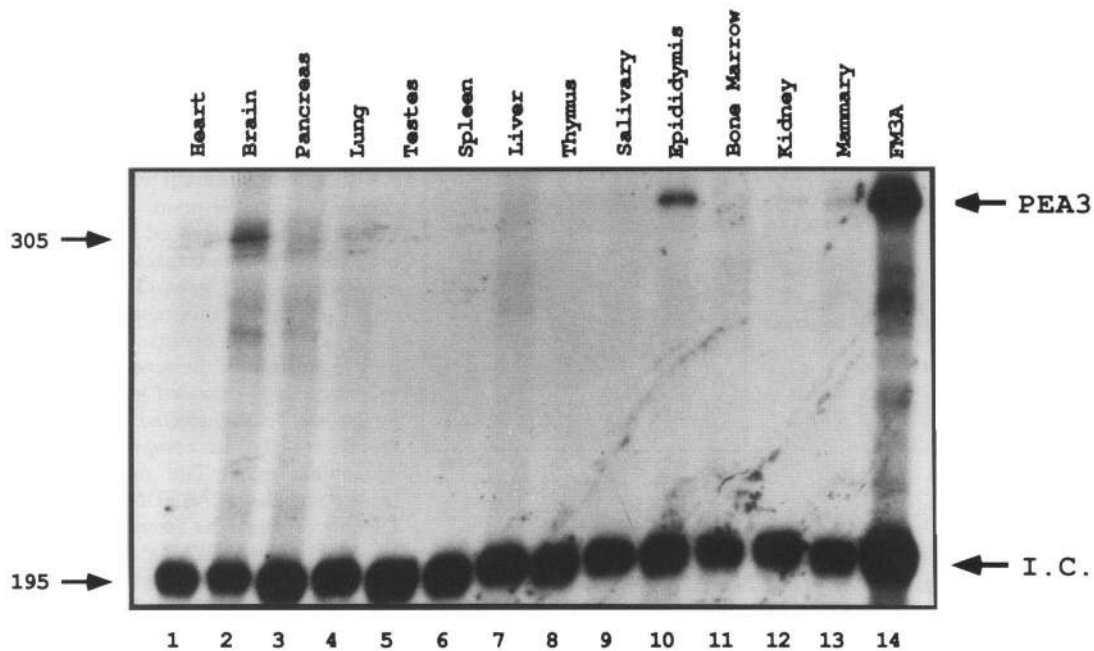
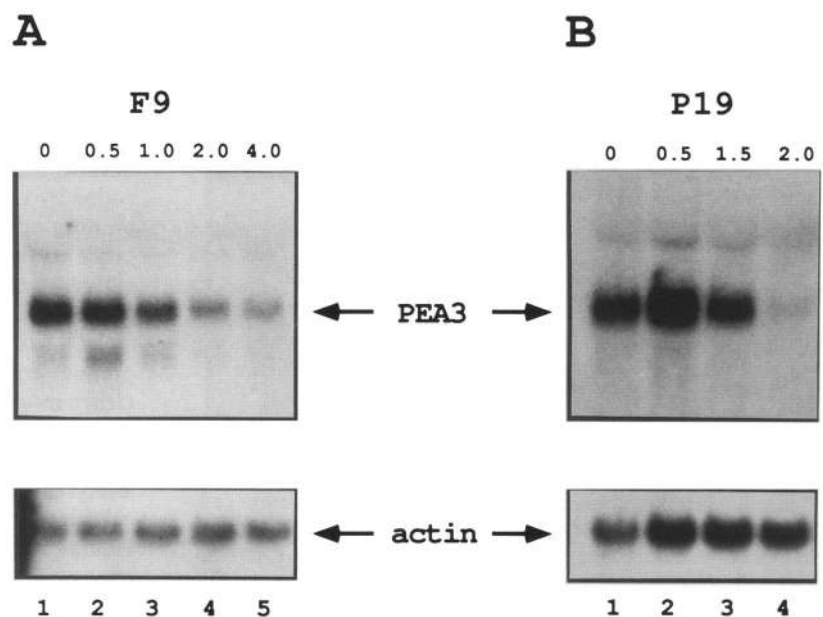


Figure 8. PEA3 RNA in mouse tissues analyzed by RNase protection. Total RNA (15 μ g) was isolated from various mouse organs and analyzed by RNase protection as described in Materials and methods. The integrity and amount of RNA in the samples was monitored by using a ribosomal protein gene probe (I.C.). The expected size of the protected PEA3 RNA is 305 nucleotides, whereas that for the ribosomal protein gene is 195 nucleotides.

anomalous migration in polyacrylamide gels. In this regard we have recently found that PEA3 is phosphorylated (data not shown). Whereas Southwestern analysis identified PEA3 as the principal PEA3-specific DNA-binding activity in the mouse nuclear extracts, minor activities migrating with apparent molecular masses of 85 and 30 kD were also detected. These other PEA3 motif-binding proteins do not correspond in size to any of the known

mouse Ets proteins, which are all significantly smaller than PEA3. Bhat et al. (1987) have described a 30-kD protein that cross-reacts with antibodies to Ets-2 and is expressed preferentially in proliferating hepatocytes. This protein may be a stable Ets-2 degradation product or an Ets-related protein and could be the 30-kD species that we detected by Southwestern analysis. The nature of the 85-kD species is unknown; it might be a novel

Figure 9. Northern blot analysis of PEA3 RNA in embryonal carcinoma cell lines. Ten micrograms of total RNA was used for each lane. (A) RNA prepared from F9 cells at different times after addition of retinoic acid, probed with the PEA3.1 cDNA. Time, in days, is indicated above the lanes. The blot was re-probed with β -actin to ensure equal loading of the samples (*bottom*). (B) RNA prepared from P19 cells after retinoic acid treatment, with time, in days, indicated above the lanes. This blot was also hybridized with a radiolabeled β -actin DNA probe to ensure that equal amounts of RNA were in each lane.



member of the *ets* gene family or an unrelated protein with similar DNA-binding specificity.

The sequence of PEA3 has a number of interesting features. The protein is rich in proline, glycine, and serine residues, which are for the most part distributed throughout the protein. It may be noteworthy that the ETS domain contains few of these amino acids. PEA3, which functions as a transcriptional activator, possesses a 27-amino-acid region rich in acidic residues and a 99-amino-acid region with a high concentration of glutamine residues. The activation domains of several transcriptional activators are known to contain such motifs, but we currently do not know whether either of these regions in PEA3 functions in this fashion. Although it is thought that PEA3 acts in the cell nucleus, we have failed to find a nuclear localization signal by inspection of its sequence. Chicken, human, and mouse Ets-1 and Ets-2 are found predominantly in the nucleus (Boulukos et al. 1988, 1989; Pognonec et al. 1989; Koizumi et al. 1990), and this location is specified, in the case of the chicken Ets-1 protein, by a region in its ETS domain that includes the sequence GKRKNKPK (Boulukos et al. 1989). The homologous segment in PEA3 (GIQKNRPA) contains basic residues, but otherwise is related to that of chicken Ets-1 in only four of the eight positions. Whether PEA3 is found exclusively in the nucleus and whether this sequence specifies nuclear localization remains to be established.

PEA3 possesses a stretch of ~80 amino acids that shares extensive sequence similarity with the ETS domain, a conserved region common to the Ets family of proteins (Karim et al. 1990). In a number of other Ets proteins this region constitutes their DNA-binding domain (Karim et al. 1990). Mapping experiments showed that the DNA-binding domain of PEA3 comprises no more than 102 amino acids and includes the entirety of the ETS domain. More recently, we have shown that this region by itself is sufficient for sequence-specific binding of PEA3 to DNA (data not shown). Comparison of the ETS domains of various proteins reveals several conserved features. The amino-terminal half of the ETS domain includes a leucine-rich region potentially capable of forming an α -helix, a partially overlapping 40-amino-acid stretch containing three tryptophans spaced 18 or 19 residues apart, and a region of ~40 amino acids rich in basic residues that comprises the carboxy-terminal half of this domain. The first tryptophan in the repeat is common to all of the Ets proteins with the exception of PU.1, which contains tyrosine, an amino acid of similar structure, at this location. The tryptophan repeat, but not any of the intervening amino acids, has also been identified as a component of the Myb DNA-binding domain (Anton and Frampton 1988; Saikumar et al. 1990). This arrangement of amino acids suggests that the ETS domain is bipartite; the amino-terminal region may provide a structural framework to orient the positively charged carboxy-terminal amino acids to make base-specific contacts with DNA.

Recently, six members of the *ets* gene family have been shown to encode sequence-specific DNA-binding

proteins. These include mouse PU.1 (Klemz et al. 1990), chicken (Wasylyk et al. 1990), mouse (Gunther et al. 1990), and human (Ho et al. 1991) Ets-1, *Drosophila* E74 (Burtis et al. 1990), chicken (Wasylyk et al. 1990) and human Ets-2 (Bosselut et al. 1990), and rat GABP α (LeMarco et al. 1991; Thompson et al. 1991). Each of these proteins recognizes similar motifs in DNA that share a centrally located 5'-GGAA-3' element. Although each of the Ets proteins known to bind to DNA appears to recognize a common core sequence, preliminary experiments suggest that these proteins may not bind to identical sequences in vitro. For example, PEA3 does not bind to the sequence recognized by PU.1 (5'-AGGAAC-3'), although PU.1 binds to this sequence as well as that bound by PEA3 (5'-AGGAAG-3') (S. Bowman and J.A. Hassell, unpubl.). It is possible that all of the Ets proteins recognize the same central core sequence but that each protein interacts with unique bases that flank this core. Interestingly, the binding of GABP α to its cognate element is increased and stabilized by another protein, termed GABP β . Structurally unrelated domains in each protein mediate their interaction. The ETS domain and 37 carboxy-terminal flanking amino acids in GABP α and repeats of a 33-amino-acid motif in GABP β are required for heterodimerization of these proteins (Thompson et al. 1991). The 33-amino-acid repeats are also found in ankyrin, a component of red blood cell membranes, in the plasma membrane protein products of the *Notch* gene of *D. melanogaster* and the *glp-1* and *lin-12* genes of *Caenorhabditis elegans*, and in the transcription regulatory proteins (the precursor to NF- κ B and I- κ B) (Thompson et al. 1991 and references therein). Whether other or all Ets proteins dimerize with GABP β -like subunits and whether this serves as a means to regulate their activity remains to be determined. In this regard, we have been unable to detect an activity that increases or stabilizes the binding of recombinant PEA3 to DNA. We have demonstrated, however, that several cellular proteins are capable of binding with specificity to PEA3 affinity matrices (data not shown). More rigorous experiments need to be carried out to learn whether PEA3 and other Ets proteins act with partners to bind avidly to DNA.

The mouse PEA3 gene encodes an mRNA of ~2.4 kb, which roughly corresponds in length (2410 bp) to the longest cDNA that we have cloned. Primer extension experiments, however, show that the cDNA lacks ~90 nucleotides at its 5' end (data not shown), and the absence of a poly(A) tract indicates that an unknown number of nucleotides are missing from this end too. There is a putative poly(A)-addition signal (5'-AATAAA-3'), 23 nucleotides from the 3' end of the cDNA, which, if functional, suggests that nearly all of the 3'-noncoding sequences are represented in the cDNA. Poly(A) addition and 3' end formation signals generally occur between 10 and 20 nucleotides upstream of the start of the poly(A) tail in the mRNA (Wickens 1990).

A survey of the presence of PEA3 RNA in various cell lines revealed that PEA3 not only was expressed in mouse cells but also in cells of rat, monkey, and human origin, demonstrating that the PEA3 gene is conserved

and expressed in other species. Southern analyses confirm that the mouse PEA3 gene is conserved among these species, and the simplicity of the hybridization pattern is consistent with it being a single-copy gene (data not shown). Whereas PEA3 RNA was readily detected in mouse cells of epithelial and fibroblastic origin, it was not found in cells of lymphoid or erythroid origin. Examination of mouse organs using a sensitive gene-specific RNase protection assay confirmed the absence of PEA3 RNA in any of the hematopoietic tissues tested. This contrasts with the other mouse *ets* genes (*ets-1*, *ets-2*, *Spi-1*, and *Fli-1*) that have been identified to date in that each of these is expressed to high levels in cells of hematopoietic origin. For example, *Ets-1* and *Ets-2* are expressed predominantly in the thymus and T lymphocytes (Bhat et al. 1989); *Spi-1* is expressed preferentially in B cells and macrophages (Klemz et al. 1990); and *Fli-1* is expressed in erythroleukemic cells as well as in the thymus and the spleen (Ben-David et al. 1991). Whereas *ets-1* and *Spi-1* transcripts have been detected only in cells of hematopoietic origin, *ets-2* and *Fli-1* are expressed in other tissues as well. PEA3 is the only member identified to date that is apparently not expressed in these hematopoietic cells. In the mouse, expression of PEA3 RNA is highly restricted; brain and epididymis, and, to a lesser degree, mammary glands, contained PEA3 RNA. Although it is not clear what, if anything, these three organs have in common, this tissue-specific pattern of expression is unlike that of any other *ets* gene. It is interesting that FM3A cells, which were established after benzopyrene treatment of mouse mammary tissue, express PEA3 RNA to higher levels than any other source of RNA examined, suggesting that PEA3 overexpression may be linked to oncogenesis. This possibility is supported by our recent finding that PEA3 RNA is overexpressed in the mammary glands of transgenic mice that carry and express the rat *c-neu* proto-oncogene (M. Trimble, W. Muller, and J.A. Hassell, unpubl.). We are currently using in situ RNA hybridization techniques to refine our analysis of the sites of expression of PEA3, and we are pursuing the possibility that overexpression of PEA3 may be linked to malignancies, particularly of the mammary epithelium.

PEA3 expression is not only restricted to certain tissues and cells, but it is also regulated during the differentiation of embryonal carcinoma cells in vitro. Induction of the differentiation of F9 and P19 cells with retinoic acid results in a sharp drop in the steady-state amounts of PEA3 RNA. Retinoic acid treatment of F9 cells directs their differentiation into cells that resemble parietal extraembryonal endoderm, whereas this agent induces P19 cells to differentiate into neurons and astrocytes. Therefore, extinction of PEA3 RNA by retinoic acid occurs independent of the differentiation pathway that is induced. How retinoic acid manifests its effect is not known; it could negatively regulate PEA3 RNA synthesis or enhance its degradation. Whatever the mechanism, the observation that PEA3 expression is negatively regulated by retinoic acid provides a correlation between expression of this protein and early developmental

events that commit stem cells to particular lineages, and raises the prospect that PEA3 plays a transcriptional regulatory role during early mammalian development. To learn whether PEA3 acts as a regulator of development, we plan to express dominant-negative mutants in transgenic mice and to disrupt the PEA3 gene in the mouse germ line.

Materials and methods

Animal cell culture and preparation of nuclear extracts

FM3A cells were grown in Dulbecco's modified Eagle medium (DMEM) supplemented with 10% bovine serum and antibiotics and maintained in a humidified CO₂ atmosphere. Nuclear extracts were prepared by the method of Dignam et al. (1983). F9 and P19 cells were grown on gelatin-coated Petri dishes in DMEM supplemented with heat-inactivated fetal bovine serum (10%). To induce their differentiation, F9 cells growing directly on gelatin-coated petri dishes were treated with 0.5 μM retinoic acid, whereas P19 cells were allowed to aggregate in bacterial petri dishes before addition of retinoic acid to 0.5 μM (Rudnicki and McBurney 1987).

RNA isolation and analysis

Total cellular RNA was extracted from mammalian cells and mouse tissues by the modified guanidinium thiocyanate method (Sambrook et al. 1989). For Northern analyses, 10 μg of total cellular RNA was denatured with glyoxal and dimethyl-sulfoxide, fractionated by electrophoresis through a 1% agarose gel containing 10 mM NaPO₄ (pH 7.0), and transferred to a GeneScreen Plus membrane (Dupont) as described previously (Sambrook et al. 1989). The membranes were prehybridized in 50% formamide, 1% SDS, 1.0 M NaCl, and 10% dextran sulfate for 2 hr at 42°C, and hybridization was carried out with a ³²P-labeled probe comprising the 1.6-kb *EcoRI* fragment of pGEX.PEA3 cDNA (1 × 10⁶ cpm/ml) in a solution containing 100 μg/ml of denatured salmon sperm DNA, 50% formamide, 10% dextran sulfate, and 1.0 M NaCl for 16 hr at 42°C. The membranes were washed sequentially in 2× SSC for 5 min at room temperature, in 2× SSC containing 1.0% SDS for 30 min at 65°C, and in 0.1× SSC for 30 min at room temperature. Thereafter, the membranes were exposed to X-ray film with an intensifying screen at -70°C. When required, the membranes were stripped of the probe by boiling in distilled water for 10 min and rehybridized with an actin probe as described above.

RNase protection experiments were carried out using 15 μg of total cellular RNA per assay as described (Melton et al. 1984). The 365-nucleotide T7 antisense PEA3 probe, labeled with ³²P and purified by polyacrylamide gel electrophoresis, spanned PEA3 nucleotides 559–864 and included 50 nucleotides from the vector, pGEM-7Zf(+). The 235-nucleotide mouse ribosomal protein L32 probe (Dudov and Perry 1984) was prepared with T3 RNA polymerase and purified as described above. The radiolabeled probes were hybridized overnight with various RNA samples, treated with RNases, and fractionated on denaturing polyacrylamide gels. RNA samples containing PEA3 RNA yielded a protected RNA species of 305 nucleotides, whereas those that carried the ribosomal protein gene transcript yielded a protected product 195 nucleotides in length.

Construction and analysis of cDNA libraries

To prepare the cDNA expression library, cytoplasmic RNA was isolated from mouse FM3A cells by the method of Favaloro et

al. (1980), as described by Sambrook et al. (1989). Poly(A)⁺ RNA was purified from total cytoplasmic RNA using a Poly(A) Quick Kit (Stratagene). A *λgt11* cDNA library was constructed from this RNA using oligo(dT) and random hexamer primers (Sambrook et al. 1989). The primary library contained 2.0×10^6 independent recombinants with insert sizes 1.0 kb or larger.

The primary *λgt11* cDNA library was screened by the in situ filter detection method (Singh et al. 1988; Staudt et al. 1988; Vinson et al. 1988). The filter-bound proteins were denatured and renatured to enhance the signal (Vinson et al. 1988). Recombinant phage (5×10^5) were screened at a density of 2×10^4 per filter after lytic infection of *E. coli* strain Y1090 (Young and Davis 1983). The probe that was used for screening the cDNA expression library was prepared by nick translation of a concatenated oligonucleotide with the following sequence: 5'-GATC-CAGGAAGTGAC-3'. A probe bearing a mutated version of the PEA3 motif, 5'-GATCTAAAAAGTGACTAACG-3', was also prepared and used as a control for nonspecific binding. Each of the probes had a specific activity of $>10^8$ cpm/ μ g. Nitrocellulose filters bearing transferred phage were subjected to a denaturation-renaturation cycle in 6.0 M guanidine hydrochloride in binding buffer [25 mM NaCl, 5 mM MgCl₂, 1 mM DTT, and 25 mM HEPES (pH 7.9)]. The binding reaction was carried out in binding buffer supplemented with 0.25% nonfat dry milk, 10 μ g/ml of denatured salmon sperm DNA, and 10^6 cpm/ml of radiolabeled probe for 12 hr. After incubation with the probe, the filters were washed with binding buffer containing 0.25% nonfat dry milk, dried, and autoradiographed. All incubations and filter manipulations were performed at 4°C.

Expression of GST-PEA3 fusion protein in *E. coli*

The 1.6-kb, PEA3.1 *EcoRI* fragment was cloned into the *EcoRI* site of pGEX.3X (Pharmacia) yielding the recombinant plasmid pGEX.PEA3, which was transferred into the *E. coli* strain XL-1 Blue (Stratagene). Overnight cultures were diluted 1 : 10 and grown for 90 min before the addition of IPTG to 1 mM. After 2 hr, bacteria were collected and extracts were prepared as described (Singh et al. 1988). Samples of the bacterial extracts were electrophoresed through SDS-PAGE and the gels were stained with Coomassie brilliant blue to ensure that the GST-PEA3 fusion protein had been expressed. Extracts were stored at -80°C until use.

Electrophoretic mobility-shift assays

Binding reactions were performed in a final volume of 20 μ l containing 0.2 ng of radiolabeled DNA, 2 μ g of poly[d(I-C)] (Pharmacia), 1–10 μ g of protein in 50 mM Tris-HCl (pH 7.5), 50 mM NaCl, 2 mM DTT, 2 mM EDTA, and 2 mM spermidine. The reactions were carried out for 20 min at room temperature. To stop the reaction, 5 μ l of 10% glycerol + 0.01% Orange G dye was added, and the samples were run on 5% polyacrylamide gels (40 : 1 ratio acrylamide/bisacrylamide) in 0.25 \times TBE [1 \times TBE = 89 mM Tris-borate (pH 8.0) and 2 mM EDTA]. When employed, competitor DNAs were added at 100-fold molar excess before the addition of the radiolabeled DNA.

The probes for DNA-binding reactions were obtained from recombinant plasmid DNAs that carried synthetic complementary oligonucleotides containing the PEA3 motif. The oligonucleotides were cloned by ligation into the *Bam*HI site of pBlue-script(KS) (Stratagene) and subsequently excised with flanking polylinker sequence to generate fragments of 40–60 bp in length. The DNA fragments were end-labeled with [γ -³²P]ATP and polynucleotide kinase. The sequence of the oligonucleotide representing the PEA3 motif used in the methylation interfer-

ence experiment is 5'-GATCTAGGAAGTAAGTAACG-3', whereas that used in all of the other experiments is 5'-GATC-CAGGAAGTGAC-3'. These oligonucleotides share a central sequence, 5'-AGGAAGT-3', that includes the PEA3 motif. When annealed to their complement the oligonucleotides carried 5'-GATC-3' projections, which facilitated their molecular cloning in different vectors. There was no difference in the activity of these oligonucleotides in the various assays in which they were used.

Methylation interference analysis

A DNA fragment containing a single PEA3 motif (see above for sequence) was end-labeled using [γ -³²P]ATP and polynucleotide kinase. The radiolabeled DNA was partially methylated with dimethylsulfate (Maxam and Gilbert 1980) before adding it to the binding reactions. DNA (2×10^5 cpm) was incubated in 50 μ l with 10 μ g of poly[d(I-C)] and either 15 μ g of FM3A cell nuclear extract or 3 μ g of bacterial extract from *E. coli* harboring pGEX.PEA3 as described above for EMSA. Bands corresponding to bound and free DNA were excised from the gel, and the DNA electroeluted, cleaved with ammonium acetate and piperidine (Maxam and Gilbert 1980), and electrophoresed through a 15% acrylamide-urea denaturing gel in TBE.

Southwestern analysis

Bacterial and mouse cell extracts were electrophoresed through SDS-polyacrylamide gels and transferred to nitrocellulose (Miskimins et al. 1985). The filters were treated with 5% powdered nonfat milk in 50 mM Tris-HCl (pH 7.5), 50 mM NaCl, 1 mM EDTA, and of 1 mM DTT. The same concatenated oligonucleotides used to screen the cDNA library were employed at 10^6 cpm/ml in 50 mM Tris-HCl (pH 7.5), 75 mM NaCl, 1 mM EDTA, 1 mM DTT, and 5 μ g/ml of salmon sperm DNA. Binding reactions were carried out for 1 hr at room temperature. Thereafter, the filters were washed twice for 15 min in the same buffer lacking DNA, blotted on 3MM paper (Whatman), and exposed to X-ray film.

Sequencing of PEA3 cDNAs

PEA3 cDNAs were excised as *EcoRI* fragments from *λgt11* and cloned into pGEM-7Zf(+) (Promega). A nested series of deletions spanning the cDNA insert were generated using exonuclease III and S1 nuclease (Henikoff 1987). Plasmid DNA was prepared by a modified alkali lysis procedure (Birnboim and Doly 1979) excluding treatment with lysozyme. Double-stranded DNA templates were sequenced by the dideoxy chain-termination method (Sanger et al. 1977), using Sequenase reagents and protocols from U.S. Biochemical with modifications (Zhang et al. 1988). The complete DNA sequence was obtained from both strands of the cDNAs analyzed.

Computer analysis

Sequence entry, compilation, and ORF analysis was performed using MacVector Sequence Analysis Software (IBI). The EMBL and Swiss-Prot data bases were searched with the PEA3 DNA sequence and its predicted ORF using the PC/Gene FSTNSCAN and FSTPSCAN programs (Intelligenetics). Multiple sequence alignments of the *ets* family members was accomplished by the PC/Gene CLUSTAL program (Intelligenetics).

Construction of PEA3 deletion mutants

Amino-terminal and carboxy-terminal deletion mutants of the GST-PEA3 fusion gene were generated by use of polynucleotide chain reaction (PCR) with pairs of primers corresponding in sequence to the desired end points. The amino-terminal deletion mutants were generated with primer pairs; the variable primer "5'-primer" included sequences corresponding to an *EcoRI* site and an additional 21 nucleotides corresponding in sequence to the desired end point in PEA3 (5'-GGAATCCN21-3'), whereas the "3'-common primer" was complementary to the last 12 nucleotides in the cDNA and terminated in an *EcoRI* site (5'-GAATTCGGATCAATTTTGC-3'). Carboxy-terminal deletion mutants were generated similarly. The common 5' primer was identical in sequence to those that span the GST-PEA3 junction of pGEX-PEA3.1 and included an internal *EcoRI* site (5'-CGGGAATTCCTCCTGAGAGCCA-3'); the variable 3' primers corresponded in sequence to the desired end point and included an in-frame stop codon followed by an *EcoRI* site (5'-CGGAATTCCTAN20-3'). After PCR, the amplified DNA fragments were cleaved with *EcoRI*, separated by agarose gel electrophoresis, and the appropriate sized fragment was eluted, purified and, finally, cloned into the *EcoRI* site of pGEX-3X.

DNA transfections and CAT assays

HeLa cells growing as monolayers were transfected with 2.5 µg of reporter plasmid DNA in the presence or absence of 7.5 µg of effector plasmid DNA using the calcium phosphate technique (Mueller et al. 1988). Cell lysates were prepared 48 hr after transfection, and CAT assays were carried out as described previously (Mueller et al. 1988). The various DNAs were tested in duplicate on three separate occasions by using independent preparations of reporter and effector DNAs.

Acknowledgments

We are grateful to Winship Herr for providing recombinant bacteriophages that served as positive controls during the screening of the cDNA library. We thank Michele McWilliams for culturing and preparing RNA from the F9 and P19 cells, Malcolm Trimble for preparing mouse organ RNA, Laurie Kendry for helping to type the manuscript, Laura Hastings for technical assistance, and Sharen Bowman, Bill Muller, Malcolm Trimble, and Corrinne Lobe for critical reading of the manuscript. This research was supported by the National Cancer Institute of Canada and by the Medical Research Council of Canada.

The publication costs of this article were defrayed in part by payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 USC section 1734 solely to indicate this fact.

Note added in proof

The sequence data described in this paper have been submitted to the EMBL/GenBank data libraries under accession no. X63190.

References

- Anton, I.A. and J. Frampton, 1988. Tryptophans in myb proteins. *Nature* **336**: 719.
- Asano, M., Y. Murakami, K. Furukawa, Y. Yamaguchi-Iwai, M. Satake, and Y. Ito. 1990. A polyomavirus enhancer-binding protein, PEBP5, responsive to 12-O-tetradecanoylphorbol-

- 13-acetate but distinct from AP-1. *J. Virol.* **64**: 5927–5938.
- Ben-David, Y., E.B. Giddens, K. Letwin, and A. Bernstein. 1991. Erythroleukemia induction by Friend murine leukemia virus: Insertional activation of a new member of the *ets* gene family, *Fli-1*, closely linked to *c-ets-1*. *Genes & Dev.* **5**: 908–918.
- Bhat, N.K., R.J. Fisher, S. Fujiwara, R. Ascione, and T.S. Papas. 1987. Temporal and tissue-specific expression of mouse *ets* genes. *Proc. Natl. Acad. Sci.* **84**: 3161–3165.
- Bhat, N.K., K.L. Komschlies, S. Fujiwara, R.J. Fisher, B.J. Mathieson, T.A. Gregorio, H.A. Young, J.W. Kasik, K. Ozato, and T.S. Papas. 1989. Expression of *ets* genes in mouse thymocyte subsets and T cells. *J. Immunol.* **142**: 672–678.
- Birnboim, H.C. and J. Doly. 1979. A rapid alkaline extraction procedure for screening recombinant plasmid DNA. *Nucleic Acids Res.* **7**: 1513–1523.
- Bohnlein, E. and P. Gruss. 1986. Interaction of distinct nuclear proteins with sequences controlling the expression of polyomavirus early genes. *Mol. Cell Biol.* **6**: 1401–1411.
- Bosselut, R., J.F. Duvall, A. Gégonne, M. Bailly, A. Mémar, J. Brady, and J. Ghysdael. 1990. The product of the *c-ets-1* proto-oncogene and the related *Ets 2* protein act as transcriptional activators of the long terminal repeat of human T cell leukemia virus HTLV-1. *EMBO J.* **9**: 3137–3144.
- Boulukos, K.E., P. Pognonec, A. Gegue, F. Galibert, J.C. Gesquière, D. Stéhelin, and J. Ghysdael. 1988. Identification in chickens of an evolutionarily conserved cellular *ets-2* gene (*c-ets-2*) encoding nuclear proteins related to the products of the *c-ets* proto-oncogene. *EMBO J.* **7**: 697–705.
- Boulukos, K.E., P. Pognonec, B. Rabault, A. Begue, and J. Ghysdael. 1989. Definition of an *ets-1* protein domain required for nuclear localization in cells and DNA-binding in vitro. *Mol. Cell Biol.* **9**: 5718–5721.
- Burtis, K.C., C.S. Thummel, C. Weldon Jones, F.D. Karim, and D.S. Hogness. 1990. The *Drosophila* 74EF early puff contains E74, a complex ecdysone-inducible gene that encodes two *ets*-related proteins. *Cell* **61**: 85–99.
- Chen, J.H. 1990. Cloning, sequencing, and expression of mouse *c-ets-1* cDNA in baculovirus expression system. *Oncogene Res.* **5**: 277–285.
- Chen, Z.Q., N.C. Kan, L. Pribyl, J.A. Lautenberger, E. Moudrianakis, and T.S. Papas. 1988. Molecular cloning of the *ets* proto-oncogene of the sea urchin and analysis of its developmental expression. *Dev. Biol.* **125**: 432–440.
- Chen, Z-Q., L.A. Burdett, A.K. Seth, J.A. Lautenberger, and T.S. Papas. 1990. Requirement of *ets-2* expression for *Xenopus* oocyte maturation. *Science* **250**: 1416–1418.
- de Villiers, J., W. Schaffner, C. Tyndall, S. Lupton, and R. Kamen. 1984. Polyomavirus DNA replication requires an enhancer. *Nature* **312**: 242–246.
- Dignam, J.D., R.M. Lebovitz, and R.G. Roeder. 1983. Accurate transcription initiation by RNA polymerase II in a soluble extract from isolated nuclei. *Nucleic Acids Res.* **11**: 1475–1489.
- Dudov, K.P. and R.P. Perry. 1984. The gene family encoding the mouse ribosomal protein L32 contains a uniquely expressed intron-containing gene and an unmutated processed gene. *Cell* **37**: 457–468.
- Favaloro, J., R. Treisman, and R. Kamen. 1980. Transcription maps of polyomavirus-specific RNA: Analysis by two-dimensional nuclease S1 gel mapping. *Methods Enzymol.* **65**: 718–749.
- Fujimura, F.K. 1986. Nuclear activity from F9 embryonal carcinoma cells binding specifically to the enhancers of wild-type polyoma virus and PyEC mutant DNAs. *Nucleic Acids Res.* **14**: 2845–2861.

- Gunther, C.V., J.A. Nye, R.S. Bryner, and B.J. Graves. 1990. Sequence-specific DNA binding of the proto-oncogene *ets-1* defines a transcriptional activator sequence within the long terminal repeat of the Moloney murine sarcoma virus. *Genes & Dev.* **4**: 667–679.
- Gutman, A. and B. Wasylyk. 1990. Nuclear targets for transcription regulation by oncogenes. *Trends Genet.* **7**: 49–54.
- Hassell, J.A., W.J. Muller, and C.R. Mueller. 1986. The dual role of the polyomavirus enhancer in transcription and DNA replication. *Cancer Cells* **4**: 561–569.
- Hearing, P. and T. Shenk. 1983. The adenovirus type 5 E1A transcriptional control region contains a duplicated enhancer element. *Cell* **33**: 695–703.
- Henikoff, S. 1987. Unidirectional digestion with exonuclease III in DNA sequence analysis. *Methods Enzymol.* **155**: 156–165.
- Herbomel, P., B. Bourachot, and M. Yaniv. 1984. Two distinct enhancers with different cell specificities co-exist in the regulatory region of polyoma. *Cell* **39**: 653–662.
- Herman, R.C. 1989. Alternatives for the initiation of translation. *Trends Biochem. Sci.* **14**: 219–222.
- Ho, I.-C., K.N. Bhat, L.R. Gottschalk, R. Lindsten, C.B. Thompson, T.S. Papas, and J.M. Leiden. 1991. Sequence-specific binding of human Ets-1 to the T cell receptor α gene enhancer. *Science* **250**: 814–818.
- Imler, J.L., C. Schatz, C. Wasylyk, B. Chatton, and B. Wasylyk. 1988. A Harvey-ras responsive transcription element is also responsive to a tumor-promoter and to serum. *Nature* **332**: 275–278.
- Johnson, P.F. and S.L. McKnight. 1989. Eukaryotic transcriptional regulatory proteins. *Annu. Rev. Biochem.* **58**: 799–839.
- Johnson, P.F., W.H. Landschulz, B.J. Graves, and S.L. McKnight. 1987. Identification of a rat liver nuclear protein that binds to the enhancer core element of three animal viruses. *Genes & Dev.* **1**: 133–146.
- Kamachi, Y., E. Ogawa, M. Asano, S. Ishida, Y. Murakami, M. Satake, Y. Ito, and K. Shigesada. 1990. Purification of a mouse nuclear factor that binds to both the A and B cores of the polyomavirus enhancer. *J. Virol.* **64**: 4808–4819.
- Karim, F.D., L.D. Urness, C.S. Thummel, M.J. Klemsz, S.R. McKercher, A. Celada, C. Van Bevern, R.A. Maki, C.V. Gunther, J.A. Nye, and B.J. Graves. 1990. The ETS domain: A new DNA-binding motif that recognizes a purine-rich core DNA sequence. *Genes & Dev.* **4**: 1451–1453.
- Klemz, M.J., S.R. McKercher, A. Celada, C. Van Bevern, and R.A. Maki. 1990. The macrophage and B cell-specific transcription factor PU.1 is related to the *ets* oncogene. *Cell* **61**: 113–124.
- Koizumi, S., R.J. Fisher, S. Fujiwara, C. Jorcyk, N.K. Bhat, A. Seth, and T.S. Papas. 1990. Isoforms of the human *ets-1* protein: Generation by alternative splicing and differential phosphorylation. *Oncogene* **5**: 675–681.
- Kozak, M. 1986. Point mutations define a sequence flanking the AUG initiator codon that modulates translation by eukaryotic ribosomes. *Cell* **44**: 283–292.
- LaMarco, K., C.C. Thompson, B.P. Byers, E.M. Walton, and S.L. McKnight. 1991. Identification of Ets- and Notch-related subunits in GA binding protein. *Science* **253**: 789–792.
- Leprince, D., M. Duterque-Coquillaud, R.P. Li, C. Henry, A. Flourens, B. Debuire, and D. Stehelin. 1988. Alternative splicing within the chicken *c-ets-1* locus: Implications for transduction within the E26 retrovirus of the *c-ets* proto-oncogene. *J. Virol.* **62**: 3233–3241.
- Martin, M.E., J. Piette, M. Yaniv, W.-J. Tang, and W.R. Folk. 1988. Activation of the polyomavirus enhancer by a murine activator protein 1 (AP-1) homolog and two contiguous proteins. *Proc. Natl. Acad. Sci.* **85**: 5839–5843.
- Maxam, A.M. and W. Gilbert. 1980. Sequencing end-labelled DNA with base specific chemical cleavages. *Methods Enzymol.* **65**: 499–560.
- Melton, D.A., P.A. Kreig, M.R. Rebagliati, T. Manitis, K. Zinn, and M.R. Green. 1984. Efficient *in vitro* synthesis of biologically active RNA and RNA hybridization probes from plasmids containing a bacteriophage SP6 promoter. *Nucleic Acids Res.* **12**: 7035–7056.
- Miskimins, W.K., M.P. Roberts, A. McClelland, and F.H. Ruddle. 1985. Use of a protein-blotting procedure and a specific DNA probe to identify nuclear proteins that recognize the promoter region of the transferrin receptor. *Proc. Natl. Acad. Sci.* **82**: 6741–6744.
- Mitchell, P.J. and R. Tjian. 1989. Transcriptional regulation in mammalian cells by sequence-specific DNA binding proteins. *Science* **245**: 371–378.
- Mueller, C.R., A.-M. Mes-Masson, M. Bouvier, and J.A. Hassell. 1984. Location of sequences in polyomavirus DNA that are required for early gene expression *in vivo* and *in vitro*. *Mol. Cell. Biol.* **4**: 2594–2609.
- Mueller, C.R., W.J. Muller, and J.A. Hassell. 1988. The polyomavirus enhancer comprises multiple functional elements. *J. Virol.* **62**: 1667–1678.
- Muller, W.J., C.R. Mueller, A.-M. Mes, and J.A. Hassell. 1983. Polyomavirus origin for DNA replication comprises multiple genetic elements. *J. Virol.* **47**: 586–599.
- Muller, W.J., D. Dufort, and J.A. Hassell. 1988. Multiple subelements within the polyomavirus enhancer function synergistically to activate DNA replication. *Mol. Cell. Biol.* **8**: 5000–5015.
- Murakami, Y., M. Asano, M. Satake, and Y. Ito. 1990. A tumor promoting phorbol ester, TPA, enhances polyomavirus DNA replication by activating the function of the viral enhancer. *Oncogene* **5**: 5–13.
- Nunn, M.F., P.H. Seeburg, C. Moscovici, and P.H. Duesberg. 1983. Tripartite structure of the avian erythroblastosis virus E26 transforming gene. *Nature* **306**: 391–395.
- Ostapchuk, P., J.F.X. Diffley, J.T. Bruder, B. Stillman, A.J. Levine, and P. Hearing. 1986. Interaction of a nuclear factor with the polyomavirus enhancer region. *Proc. Natl. Acad. Sci.* **83**: 8550–8554.
- Piette J. and M. Yaniv. 1986. Molecular analysis of the interaction between an enhancer binding factor and its DNA target. *Nucleic Acids Res.* **14**: 9595–9611.
- . 1987. Two different factors bind to the alpha domain of the polyomavirus enhancer, one of which also interacts with the SV40 and *c-fos* enhancers. *EMBO J.* **6**: 1331–1337.
- Piette, J., M.-H. Kryszke, and M. Yaniv. 1985. Specific interactions of cellular factors with the B enhancer of polyomavirus. *EMBO J.* **4**: 2675–2685.
- Pognonec, P., K.E. Boulukos, and J. Ghysdael. 1989. The *c-ets-1* protein is chromatin associated and binds to DNA *in vitro*. *Oncogene* **4**: 691–697.
- Ponte, P., P. Gunning, H. Blau, and L. Kades. 1983. Human actin genes are single copy for α -skeletal and α -cardiac actin but multicopy for β - and δ -cytoskeletal genes: 3' Untranslated regions are isotype specific but are conserved in evolution. *Mol. Cell. Biol.* **3**: 1783–1791.
- Pribyl, L.J., D.K. Watson, M.J. McWilliams, R. Ascione, and T.S. Papas. 1988. The *Drosophila ets-2* gene: Molecular structure, chromosomal localization, and developmental expression. *Dev. Biol.* **127**: 45–53.
- Pribyl, L.J., D.K. Watson, R.A. Schulz, and T.S. Papas. 1991. D-elg, a member of the *Drosophila ets* gene family: Se-

Xin et al.

- quence, expression and evolutionary comparison. *Oncogene* **6**: 1175–1183.
- Ray, D., S. Culine, A. Tavitian, and F. Moreau-Gachelin. 1990. The human homologue of the putative proto-oncogene SPI-1 characterization and expression in tumors. *Oncogene* **5**: 663–668.
- Reddy, E.S.P., V.N. Rao, and T.S. Papas. 1987. The *erg* gene: A human gene related to *ets* oncogene. *Proc. Natl. Acad. Sci.* **84**: 6131–6135.
- Rao, V.N., K. Huebner, M. Isobe, A. Ar-Rushdi, E. Croce, and S.Y. Reddy. 1989. *elk*, tissue-specific *ets*-related genes on chromosomes X and 14 near translocation breakpoints. *Science* **244**: 66–70.
- Rudnicki, M.A. and M.W. McBurney. 1987. Cell culture method and induction of differentiation of embryonal carcinoma cell lines. In *Teratocarcinomas and embryonic stem cells: A practical approach* (ed. E.J. Robertson), pp. 19–49. IRL Press, Oxford, UK.
- Saikumar, P., R. Murali, and E.P. Reddy. 1990. Role of tryptophan repeats and flanking amino acids in Myb-DNA interactions. *Proc. Natl. Acad. Sci.* **87**: 8452–8456.
- Sambrook, J., F. Fritsch, and T. Maniatis. 1989. *Molecular cloning: A laboratory manual*, 2nd. ed. Cold Spring Harbor Press, Cold Spring Harbor, New York.
- Sanger, F., S. Nicklen, and A.R. Coulson. 1977. DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci.* **74**: 5463–5467.
- Satake, M., T. Ibaraki, and Y. Ito. 1988. Modulation of polyomavirus enhancer binding proteins by Ha-ras oncogene. *Oncogene* **3**: 69–78.
- Singh, H., J.H. LeBowitz, A.S. Baldwin, Jr., and P. Sharp. 1988. Molecular cloning of an enhancer binding protein: Isolation by screening of an expression library with a recognition site DNA. *Cell* **52**: 415–423.
- Smith, D.B. and K.S. Johnson. 1988. Single-step purification of polypeptides expressed in *Escherichia coli* as fusions with glutathione S-transferase. *Gene* **67**: 31–40.
- Staudt, L.M., R.G. Clerc, H. Singh, J.H. LeBowitz, P.A. Sharp, and D. Baltimore. 1988. Cloning of cDNA encoding a B-cell restricted octamer binding factor. *Science* **241**: 577–580.
- Steigler, P., C. Wolff, M. Baltzinger, J. Hirtzlin, F. Senan, D. Meyer, J. Ghysdael, D. Stehelin, N. Befort, and P. Remy. 1990. Characterization of *Xenopus laevis* cDNA clones of the *c-ets-1* proto-oncogene. *Nucleic Acids Res.* **18**: 5298.
- Thompson, C., T. Brown, and S. McKnight. 1991. Convergence of Ets- and Notch-related structural motifs in a heterodimeric DNA binding motif. *Science* **253**: 762–768.
- Tyndall, C., G. LaMantia, C.M. Thacker, J. Favaloro, and R. Kamen. 1981. A region of the polyomavirus genome between the replication origin and the late protein coding sequences is required in *cis* for both early gene expression and viral DNA replication. *Nucleic Acids Res.* **9**: 6231–6250.
- Veldman, G.M., S. Lupton, and R. Kamen. 1985. Polyomavirus enhancer contains multiple redundant sequence elements that activate both DNA replication and gene expression. *Mol. Cell. Biol.* **5**: 649–658.
- Vinson, C.R., K.L. LaMarco, P.F. Johnson, W.H. Landshulz, and J.L. McKnight. 1988. *In situ* detection of sequence-specific DNA-binding activity specified by a recombinant bacteriophage. *Genes & Dev.* **2**: 801–806.
- Wasylyk, B., C. Wasylyk, P. Flores, A. Begue, D. LePrince, and D. Stehelin. 1990. The *c-ets* proto-oncogenes encode transcription factors that co-operate with *c-fos* and *c-jun* for transcriptional activation. *Nature* **346**: 191–193.
- Wasylyk, B.C., J.L. Imler, J. Perez-Mutul, and B. Wasylyk. 1987. The c-Ha-ras oncogene and a tumor promoter activate the polyomavirus enhancer. *Cell* **48**: 525–534.
- Wasylyk, C., J.L. Imler, and B. Wasylyk. 1988. Transforming but not immortalizing oncogenes activate the transcription factor PEA1. *EMBO J.* **7**: 2475–2483.
- Wasylyk, C., P. Flores, A. Gutman, and B. Wasylyk. 1989. PEA3 is a target for transcription activation by non-nuclear oncogenes. *EMBO J.* **8**: 3371–3378.
- Watson, D.K., M.J. McWilliams, P. Lapis, J.A. Lautenberger, C.W. Schweinfest, and T.S. Papas. 1988. Mammalian *ets-1* and *ets-2* genes encode highly conserved proteins. *Proc. Natl. Acad. Sci.* **85**: 7862–7866.
- Wickens, M. 1990. How the messenger got its tail: Addition of poly(A) in the nucleus. *Trends Biochem. Sci.* **15**: 277–281.
- Wolff, C.M., P. Stiegler, M. Baltzinger, D. Meyer, J. Ghysdael, D. Stehelin, N. Befort, and P. Remy. 1990. Isolation of two different *c-ets-2* proto-oncogenes in *Xenopus laevis*. *Nucleic Acids Res.* **18**: 4603–4604.
- Yamaguchi, Y., M. Satake, and Y. Ito. 1989. Two overlapping sequence motifs within the polyomavirus enhancer are independently the targets of stimulation by both the tumor promoter 12-O-tetradecanoylphorbol-13-acetate and the Ha-ras oncogene. *J. Virol.* **63**: 1040–1048.
- Young, R.A. and R.W. Davis. 1983. Yeast RNA polymerase II genes: Isolation with antibody probes. *Science* **222**: 778–782.
- Zhang, H., R. Scholl, J. Browse, and C. Somerville. 1988. Double stranded DNA sequencing as a choice for DNA sequencing. *Nucleic Acids Res.* **16**: 1220.



Molecular cloning and characterization of PEA3, a new member of the Ets oncogene family that is differentially expressed in mouse embryonic cells.

J H Xin, A Cowie, P Lachance, et al.

Genes Dev. 1992, **6**:

Access the most recent version at doi:[10.1101/gad.6.3.481](https://doi.org/10.1101/gad.6.3.481)

References

This article cites 80 articles, 33 of which can be accessed free at:
<http://genesdev.cshlp.org/content/6/3/481.full.html#ref-list-1>

License

Email Alerting Service

Receive free email alerts when new articles cite this article - sign up in the box at the top right corner of the article or [click here](#).

Dharmacon[™] Reagents
Custom synthesis, RNAi, and CRISPR solutions

Infinite Reliability [More](#)

horizon
a PerkinElmer company

The advertisement features a dark background with a colorful, abstract image of what appears to be a DNA double helix or a similar molecular structure. The text is arranged in a clean, modern layout.