Sequence-specific DNA binding of the proto-oncoprotein *ets-1* defines a transcriptional activator sequence within the long terminal repeat of the Moloney murine sarcoma virus

Cathy V. Gunther, Julie A. Nye, Roger S. Bryner, and Barbara J. Graves

Department of Cellular, Viral, and Molecular Biology, University of Utah School of Medicine, Salt Lake City, Utah 84132 USA

The *ets* proto-oncogene family is a group of sequence-related genes whose normal cellular function is unknown. In a study of cellular proteins involved in the transcriptional regulation of murine retroviruses in T lymphocytes, we have discovered that a member of the *ets* gene family encodes a sequence-specific DNA-binding protein. A mouse *ets-1* cDNA clone was obtained by screening a mouse thymus cDNA expression library with a double-stranded oligonucleotide probe representing 20 bp of the Moloney murine sarcoma virus (MSV) long terminal repeat (LTR). The cDNA sequence has an 813-bp open reading frame (ORF) whose predicted amino acid sequence is 97.6% identical to the 272 carboxy-terminal amino acids of the human *ets-1* protein. The ORF was expressed in bacteria, and the 30-kD protein product was shown to bind DNA in a sequence-specific manner by mobility-shift assays, Southwestern blot analysis, and methylation interference. A mutant LTR containing four base pair substitutions in the *ets-1* binding site was constructed and was shown to have reduced binding in vitro. Transcriptional efficiency of the MSV LTR promoter containing this disrupted *ets-1* binding site was compared to the activity of a wild-type promoter in mouse T lymphocytes in culture, and 15- to 20-fold reduction in expression of a reporter gene was observed. We propose that *ets-1* functions as a transcriptional activator of mammalian type-C retroviruses and speculate that *ets*-related genes constitute a new group of eukaryotic DNA-binding proteins.

[*Key Words*: MSV LTR; *ets-1*; transcription factor; DNA-binding protein; proto-oncogene]

An important first step in understanding the transforming capability of an oncogene is the discovery of the proto-oncogene's role in normal cellular function. The *ets* proto-oncogene family is a collection of sequence-related genes whose function is unknown. *ets* family members have been characterized in chicken, *ets-1* (Gegonne et al. 1987) and *ets-2* (Boulukos et al. 1988); mouse, *ets-2* (Watson et al. 1988a); human, *ets-1* and *ets-2* (Watson et al. 1985), as well as *erg-1* (Rao et al. 1987; Reddy et al. 1987) and *elk-1* and *elk-2* (Rao et al. 1989); and Drosophila, *ets-2* (Pribyl et al. 1988) and E74 (Burtis et al. 1990). These genes are related, in part, on the basis of approximately 300 nucleotides of highly conserved sequence. It has been speculated that this region of similarity encodes an important functional domain (Watson et al. 1988a). Clues pertaining to the function of *ets*-related proteins include evidence for nuclear localization (Fujiwara et al. 1988a; Boulukos et al. 1989; Pognonec et al. 1989) and phosphorylation (Fujiwara et al. 1988b; Pognonec et al. 1988) of *ets-1* and *ets-2*. Additional information about this gene family comes from

the association of *ets*-related genes with cell transformation. The founding member of this gene group, the avian *ets-1* gene, is one of the proto-oncogenes transduced by the avian leukosis virus, E26 (Leprince et al. 1983; Nunn et al. 1983). This virus causes both erythroid and myeloid leukemias (Radke et al. 1982). The *ets* sequences are required for the induction of erythroblastosis (Nunn and Hunter 1989) and appear to also affect the myeloid transformation process (Golay et al. 1988). Furthermore, the human *ets* family members *ets-1*, *ets-2*, *erg*, and *elk-1*, and *elk-2* have been mapped near chromosomal sites involved in translocations that are diagnostic of certain leukemias, lymphomas, and sarcomas (Sacchi et al. 1986; Rao et al. 1989). Here we report important new insights into the function of the *ets* gene family that should help to elucidate the oncogenic potential and normal cellular activites of *ets*-related proteins.

We demonstrate that the mouse *ets-1* gene product functions as a sequence-specific DNA-binding protein. Our discovery was made in the context of a study of the cellular transcription factors used by murine retro-

viruses. First, we mapped the binding sites of thymus proteins that interact with the transcriptional control elements in the long terminal repeat (LTR) of the Moloney murine sarcoma virus (MSV). Next, to clone cDNAs encoding such proteins, a mouse thymus cDNA expression library was probed with double-stranded DNA representing repeats of the LTR binding sites (Singh et al. 1988; Vinson et al. 1988). We report the cloning of an *ets-1* cDNA using this approach. The data demonstrate that an *ets-1*-encoded polypeptide binds DNA in a sequence-specific manner. Furthermore, a promoter that bears a mutated *ets-1* binding site and displays reduced *ets-1* binding activity in vitro also has a lower transcriptional efficiency than the wild-type LTR promoter in vivo, suggesting that *ets-1* functions as a transcriptional activator.

Results

Transcriptional elements used by MSV LTR in mouse T lymphocytes

Transcriptional control elements of the MSV LTR have been mapped previously in mouse L cells and frog oocytes, using deletion mutants of the LTR and transient expression assays (Graves et al. 1985a,b). More recently, we have begun to study transcriptional regulation of

MSV in T lymphocytes, because the replication-competent (parental) Moloney virus, the Moloney murine leukemia virus (MLV), replicates to high levels in the mouse thymus (Jaenisch 1979; Evans and Morrey 1987). To begin to identify lymphoid proteins that are used as retroviral transcription factors, nuclear extracts were prepared from calf thymus glands (an abundant source of T-lymphocyte nuclei) and assayed for LTR-binding activities by DNase I protection assays. Figure 1 shows one such activity that protects nucleotides between positions -54 and -33 in the MSV LTR. This region lies between the CAAT and TATA homologies, two elements that have been shown to function as LTR transcriptional control elements in mouse fibroblasts in culture.

Thymus cDNA library screen

To characterize the thymus protein responsible for the DNA-binding activity described above, we screened a cDNA expression library for a clone that produced a similar DNA-binding activity. A λgt11 cDNA library was constructed from polyadenylated RNA of mouse thymus glands and screened with a concatemerized 25-bp oligonucleotide that contains the LTR binding site (wild-type probe, Wt, in Fig. 1B). In a primary screen of
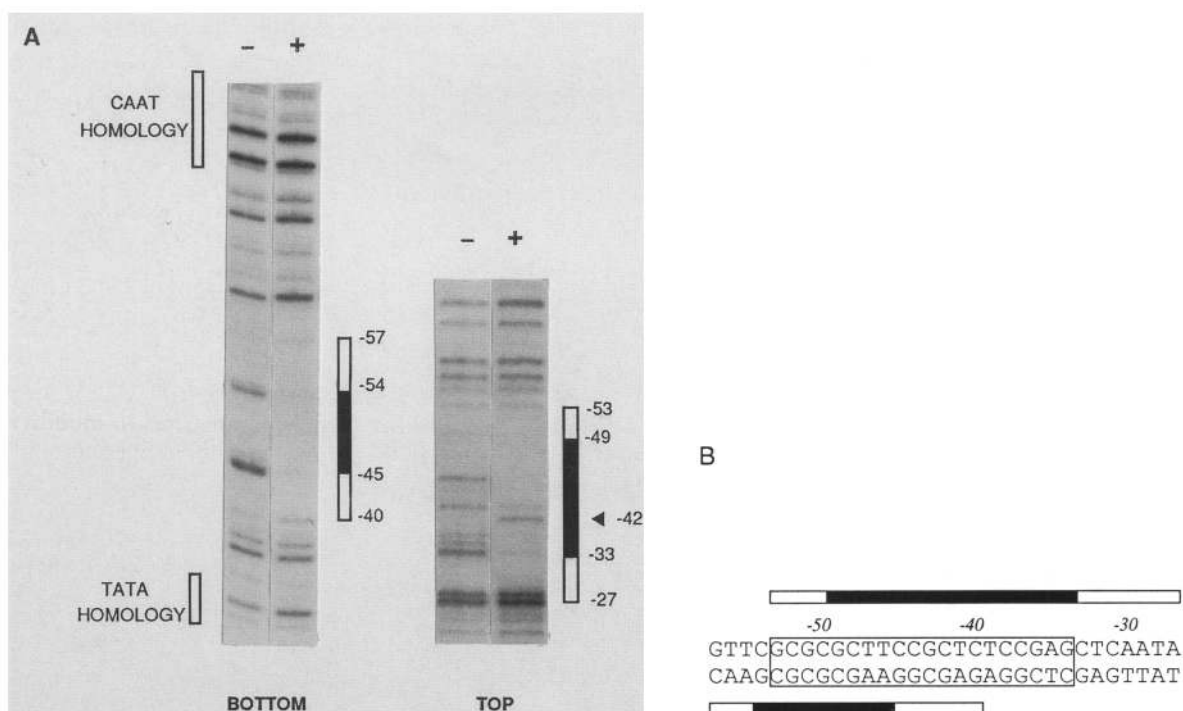
Figure 1. MSV LTR binding activity in thymus extracts. (*A*) DNase I cleavage patterns on the top and bottom strand of an MSV LTR promoter fragment in the presence (+) and absence (−) of fractionated calf thymus extracts. Positions of the LTR CAAT and TATA transcriptional control elements are noted. The arrowhead, indicates the DNase I hypersensitive site. (*B*) Sequences that underlie the thymus extract footprint. Boxed area indicates the LTR sequences included in the 25-bp oligonucleotide wild-type probe (Wt). In *A* and *B*, coordinates are promoter nucleotides numbered with respect to the start site of transcription in the LTR. The solid bar designates the protection zone. The open boxes indicate the distance between the last protected cleavage site and the first unprotected cleavage site.

300,000 phage plaques of the unamplified library, one plaque bound the probe.

Characterization of the phage DNA indicated that the single positive cDNA clone represents sequences of the mouse *ets-1* gene. A partial DNA sequence of the cDNA clone was determined by dideoxy sequencing. The deduced amino acid sequence indicated that the λgt11 cDNA insert encoded a 271-amino-acid polypeptide (excluding the adapter codons) whose sequence is in frame with the β-galactosidase coding region of λgt11. A terminator codon, TAG, was noted at position 272. Comparison of the open reading frame (ORF) sequence with the GenBank DNA sequence data base demonstrated a striking similarity to the human *ets-1* gene (Fig. 2). There is 92.6% identity of nucleotide sequences within an 813-bp region; the ORF translates into a polypeptide sequence showing 97.4% identity to the 272 carboxy-terminal amino acids of the 54-kD human *ets-1* protein. In contrast, strong similarity to the related mouse *ets-2* gene spans only 112 amino acids, which are part of the strongly conserved domain of the entire *ets* family (Fig. 2; see also Fig. 8, below). Although it is possible that the uncharacterized amino-terminal portion of this new mouse gene product will differ from the human *ets-1* amino-terminal sequences, the data base comparisons suggest that the isolated λgt11 clone with LTR-binding properties contains a cDNA of the mouse *ets-1* homolog.

The DNA-binding activity of the ets-1 polypeptide

The DNA-binding properties of the *ets-1* polypeptide clone were analyzed further by subcloning the sequences of the *ets-1* ORF into the bacterial expression vector, pET3a (Rosenberg et al. 1987). In this construction, pET-*ets-1*, the *ets-1* ORF is in-frame with the sequence encoding 11 amino acids of the T7 φ10 gene, including the initiator ATG codon. Cultures of BL21(DE3) bacteria were induced to express the *ets-1* ORF, and total bacterial cell extracts were prepared and subjected to SDS-PAGE. Figure 3A shows an abundant polypeptide in the expected 30-kD size range whose appearance was dependent on the presence of the *ets-1* ORF sequences in the expression vector. Total cell extracts resolved by SDS-PAGE were also electroblotted onto nitrocellulose and probed with the same concatemerized oligonucleotide, probe Wt, used to screen the cDNA library. The multimeric probe binds the gel replica at exactly the position of the highly expressed 30-kD protein (Fig. 3B, lane 4). No binding activity was detected in the control extracts (lane 5). Concatemerized versions of a nonspecific control probe (a 25-bp oligonucleotide which contains the binding site of an unrelated protein) did not bind the 30-kD species. A smaller bacterial protein bound the control probe, providing a useful internal control for bacterial protein concentration (lanes 6 and 7).

DNA binding of the *ets-1* polypeptide was analyzed further in gel mobility-shift assays (Fried and Crothers 1981). Cell extracts were prepared from *ets-1*-expressing cultures and control bacterial cultures, incubated with the Wt 25-bp oligonucleotide probe, and then electro-phoresed on nondenaturing polyacrylamide gels. A unique nucleoprotein complex was resolved from extracts prepared from *ets-1*-expressing cells but not from control extracts (Fig. 4, lanes 2 and 8). A 400-fold molar excess of unlabeled probe Wt almost completely eliminated this binding activity (Fig. 4, lane 6) while inclusion of an equivalent amount of the nonspecific control probe had no effect on binding (Fig. 4, lane 7). These data demonstrate that the *ets*-dependent nucleoprotein complexes are specific for the oligomer probe Wt, which represents LTR sequences.

Specific nucleotide contacts of the *ets-1* polypeptide were analyzed by methylation interference experiments performed on an 85-bp, LTR-derived restriction fragment. Figure 5A shows the relative depletion and enrichment of specific methylation sites in the protein-bound and unbound DNA fractions. Methylation of three guanine residues on the bottom strand of the binding site interfered with binding. This finding of a precise set of DNA close contacts necessary for *ets-1* DNA binding demonstrates that the *ets-1* cDNA sequence encodes a polypeptide with sequence-specific DNA-binding activity. We also performed this analysis with the LTR-binding activity in calf thymus extracts (Fig. 5B). The identical contact points were mapped, suggesting that *ets*-related proteins are responsible for the LTR-binding activity in the thymus extract.

ets-1 binding site functions as a LTR transcriptional activator sequence

The binding of the *ets-1* polypeptide fragment to the MSV LTR promoter element suggests that *ets-1* or related proteins that share its DNA-binding specificity may have a role in eukaryotic transcriptional control. We decided to investigate this possibility further by testing the transcriptional efficiency of MSV LTR-driven templates bearing a mutant *ets-1* binding site. We synthesized a mutant binding site by changing four base pairs within the DNA contact zone (Fig. 5C). Before testing the effects of these mutations on LTR-mediated transcription, we compared the affinity of the *ets-1* polypeptide for the wild-type and mutant sites in mobility-shift assays (Fig. 6). Reactions in which higher amounts of extract were included showed a higher percentage of shifted probe in both the wild-type and mutant series. The nucleoprotein complex formed on the wild-type probe (Fig. 6A, lanes 2–5) had the same mobility as the specific complex analyzed by competition and methylation interference analysis (Figs. 4 and 5). In contrast, reactions performed with the mutant probe displayed no nucleoprotein complexes with the mobility of the *ets-1*-dependent complex. Instead, nonspecific complexes of slower mobility were observed (Fig. 6A, lanes 7–10). Unlabeled oligomeric DNA of any sequence composition competed with this mutant probe binding activity (data not shown). The quantity of extract that shifted the minimally detectable amount of the wild-type probe (Fig. 6A, lane 2) is 18-fold lower than the quantity that
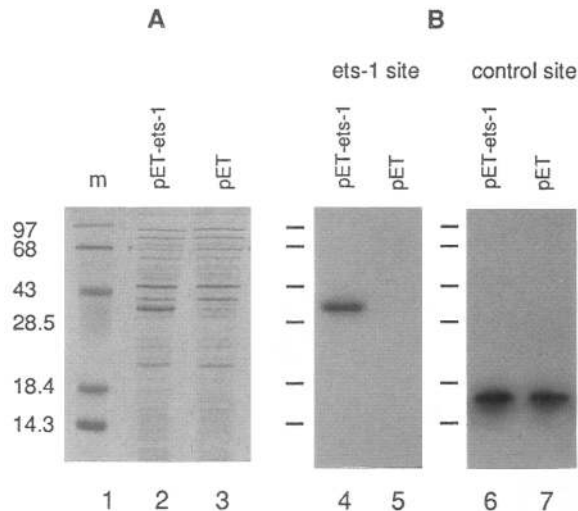
```
                    HisTrpIleAsn   AsnThrLeuGlyPhe    MetGluGlnAlaPro   GlyMetGlnAla                                      AspSerMetCysProProSerAlaThrProAlaAla
mu-ets2  232  ValProSerGluPheSerGluPro    SerPheIleThrGluSerGlnThrLeuHisPro        IleSerSerGluLeuLeuSerLeuLysTyrGluAsnAspTyrProSerValIle        AsnTyrProLysAspAsn
mu-ets1    1  GTTCCTCCCTCAGAGTTCTCAGAGCCC------AGCTTCATCACAGAGTCCTATCAGACGGTGCATCCT---ATCAGCTCGGAAGAACTCCTGTCCCTCAAGTATGAGAACGACTACCCTTCTGTCATT
mu-ets1    1                      A  A  G                                      C  T                     G      C                   T        C  G
hu-ets1  627
hu-ets1  159

                 GlySerPhe   MetLeuProLysSerArgLeuAsn        ValAsnValAsn    CysSer   Ser   Asp   PheProSerSer   ValAsn        LeuAsnAsnAsnSer                      ProLys
mu-ets2       LeuGlnAspProLeuGlnThr            AspThrLeuGlnThrAspTyrPheAlaAlaIleLysGlnGluValLeuThrProAspAspAsnMetCysLeuGlyArgAlaSerArgGlyLysLeuGly
mu-ets1  121  CTCCAGGACCCTCTCCAGACA-----------GACACCTTGCAGACAGACTACTTTGCCATCAAGCAAGAGGTGTTAACTCCAGACAACATGTGCCTGGGGAGAGCCAGTCGTGGTAAACTCGGG
mu-ets1  747                     GA                AT              A  CG  C                          A     GA
hu-ets1                          Arg              Val                                                Met  Thr
hu-ets1

              AspHis       Pro   AsnGlyCysAsp   PheGlu   Ser   Ser   LeuArg       Asn           LeuLeuAspVal               PheGlu             Glu
mu-ets2  235  GlyGlnAspSerPheGlu      SerValGluSerTyrAspSerCysAspSerArgLeuThrGlnSerTrpSerSerGlnSerSerGlnSerSerPheAsnSerLeuGlnArgValProSerPheTyrAspSerPheAspTyrGluAsp
mu-ets1       GGCCAGGACTCTTTTGAG---AGCGTAGAGAGCTACGATAGTTGTGACCCGGGTCCTCTGGAGCAGCCAGTCATCCTTCAACAGCCTGCAGCGGGTCCCTCCTATGACAGCTTCGACTACGAGGAT
mu-ets1  861                  A      A                   T                                              T      T               T                        CA   C
hu-ets1                       Ile                                                                                                               Ser
hu-ets1

              CysSerGlnSer   CysLeuSer   LeuThrMetSer                          IleGlnGlu   Ser   ProValGluGlnGly                    Val          Phe
mu-ets2  364  TyrProAlaAlaLeuProAsnHisLysProLysSerGlyGlyThrPheLysAspTyrValArgAspArgAlaAspLeu    AsnLysAspLysProValIleProAlaAlaLeuAlaLeuAlaGlyTyrThrGlySerGly
mu-ets1       TATCCGGCTGCCCTGCCCAACCACAAGCCCAAGTCTGGAGGCACCTTCAAGGACTATGTGCGTGACCGTGCTGACCTC---AACAAGGACAAGCCTGTCATTCCTGCTGCCCTGGCTCTGGCTACACAGGAAGTGGG
mu-ets1  990                  G                                                                  T                                               A          C     A
hu-ets1
hu-ets1

                                                                      Ser
mu-ets2  493  ProIleGlnLeuTrpGlnPheLeuLeuGluLeuLeuThrAspLysSerCysGlnSerPheIleSerTrpThrGlyAspGlyTrpGluPheLysLeuSerAspProAspGluValAlaAlaArgArgTrpGlyLysLys
mu-ets1       CCGATCCAGCTGTGGCAGTTTCTTCTGGAATTACTCACTGATAAGTCTTGTCAGTCCTTTATCAGCTGGACAGGAGATGGCTGGGAATTCAAGCTTTCTGACCCAGATGAGGTGGCCGCCAGGAGATGGGGAAAG
mu-ets1 1119                  A    A                                                                A  C      T                      A
hu-ets1
hu-ets1

                                                                      Ser
mu-ets2  625  ArgLysAsnLysProLysMetAsnTyrGluLysLeuSerArgGlyLeuArgTyrTyrTyrAspLysAsnIleIleHisLysThrAlaGlyLysArgTyrValTyrArgPheValCysAspLeuGlnSerLeu
mu-ets1       AGGAAAAACAAACCTAAGATGAATTATGAGAAACTGAGCCGTGGCCTTCGCTACTATTATGACAAGAAATATCATCCAAGACGGGCAAGCCTACGTTACCGCTTTGTGTGCGACCTCCAGAGCCTG
mu-ets1 1251                                            A                               C        C                  A      G  A                   T
hu-ets1
hu-ets1

              LeuGlyTyrThrProGluGluLeuHisAlaMetLeuAspValLysProAspAlaAsp       Gly   Gln   ThrGluAsp  *   514
mu-ets2  757  CTGGGGATACACCCCCTGAAGAGCTGCACGCGCCATGCTGGAGTGTAAAGCCGGATGCGACTAGTCA                       Gln   Ser  *  273
mu-ets1  757                                                       C   C   A      C      G        G         ProAspAlaAsp  *  822
hu-ets1  514                                                                                                           G  1444
hu-ets1  273                                                                                               Glu  *  431
```

Figure 2. Alignment of λ cDNA clone sequences with human *ets-1* and mouse *ets-2* sequences. Mouse *ets-1* (mu-*ets1*) nucleotide and predicted amino acid sequences are numbered with respect to the 5' end (amino terminus) of the λ clone open reading frame, excluding adapter sequences. A sequence similarity search against the GenBank data base was performed using the FASTA program (Pearson and Lipman 1988). The alignments to human *ets-1* (u-*ets1*) and mouse *ets-2* (mu-*ets2*) sequence (Watson et al. 1988b) are presented as created by the Genetics Computer Group "Bestfit" program. Only nucleotide or amino acid sequences that differ from the mouse *ets-1* cDNA are presented.

Figure 3. Expression of the *ets-1* ORF in *E. coli* and Southwestern blot analysis. SDS-lysed bacteria containing the pET-*ets-1* plasmid (lanes 2, 4, and 6) or only the pET3a vector (lanes 3, 5, and 7) were separated by SDS-PAGE. (*A*) Coomassie Blue-stained gel lanes of SDS gel-separated proteins. (Lane *m*) Molecular weight marker lane. (*B*) Southwestern blots of these same proteins probed with either the concatemerized *ets-1* binding site oligomeric probe (lanes 4 and 5) or a concatemerized, nonspecific control binding site (lanes 6 and 7). The positions of prestained molecular weight markers, transferred onto nitrocellulose, are indicated in the blot panels.

(pMSV-CAT) containing the chloramphenicol acetyltransferase (CAT) reporter gene driven by a wild-type MSV LTR. The experimental (GH) and control (CAT) plasmids were cotransfected into exponentially growing EL4 cells by electroporation. Poly(A)-selected RNA from cells transfected with the pMSV-GH construction (Fig.7A, line 1) was isolated 40 hr post-transfection and assayed by primer extension to determine the position of the cap site being used in the pMSV-GH transcription templates. Figure 7B shows the extension product of the size expected if expression from the pMSV-GH used the native MSV cap site. For quantitative comparisons of transcriptional efficiencies, the culture media were assayed for growth hormone levels and cellular proteins were assayed for CAT enzyme activity at 40 hr post-transfection. The transcriptional efficiency of the pMSV-GH template bearing the mutant *ets-1* binding site was at least 23-fold lower than the expression of the Δ5'-111 pMSV-GH template (Fig. 7A, lines 2 and 3). Within the sensitivity limits of this assay, the mutation of the *ets-1* site was equally detrimental to GH expression as the deletion of all sequences between −111 and −31 (Fig. 7A, line 4). Expression of the internal control reporter plasmid bearing the CAT gene showed a less than twofold variation between transfected cell populations. These data demonstrate that the *ets-1* binding site in the MSV LTR functions as a positive transcriptional control element. Furthermore, the results indicate that *ets-1* or a related gene that binds this positive control
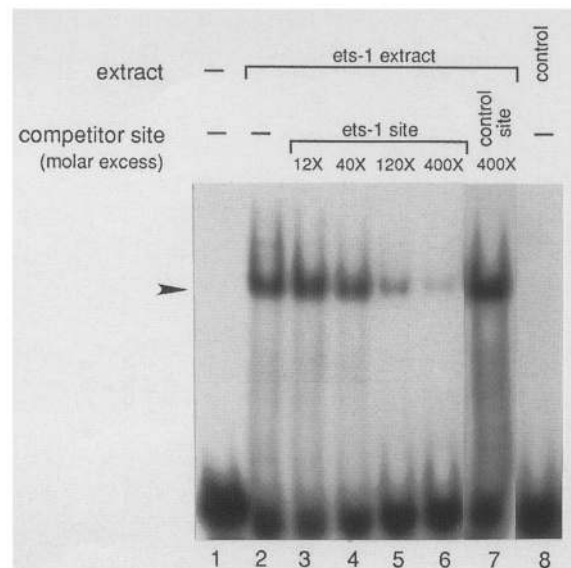
failed to cause a detectable shift of mutant probe into an *ets-1* complex (Fig. 6A, lane 10), indicating that the mutant site has a binding affinity that is at least 18-fold lower than the affinity of the wild-type site. Additional evidence for the effectiveness of the site-directed mutagenesis in destroying the *ets-1* binding site is shown in Figure 6B. The mutant probe failed to compete with the wild-type site binding activity in experiments in which unlabeled mutant probe was added as a competitor. In addition, the 30-kD bacterially expressed *ets-1*, when transferred to nitrocellulose, exhibited no binding to a concatemerized mutant probe (data not shown). These experimental results demonstate that the site-directed mutagenesis of the *ets-1* binding site effectively eliminated the *ets-1* binding capability of the MSV LTR.

To test the transcriptional efficiency of the MSV LTR containing the mutant *ets-1* binding site, transient expression assays were performed in mouse EL4 cells, a T-lymphocyte cell line. Transcription templates were constructed by cloning wild-type and mutant versions of the MSV LTR upstream of the human growth hormone (GH) reporter gene (Fig. 7A). Two mutant constructions had either the enhancer (pMSV-GH; Δ5'-111) or the enhancer plus promoter sequences (pMSV-GH; Δ5'-31) deleted. In the third mutant construction, the mutant *ets-1* binding site was placed in pMSV-GH; Δ5'-111. We chose to study the mutant binding site in this context because MSV enhancer elements have been shown to mask the effect of promoter mutations (Graves, 1985b; B. Graves, unpubl.). An internal control was provided by a plasmid

Figure 4. DNA-binding activity of *ets-1* polypeptide analyzed by mobility-shift assays. Lysates of BL21(DE3) cells containing either the pET-*ets-1* plasmid (*ets-1* extract) or only the pET3a vector (control extract) were incubated with ³²P-labeled oligomeric probe, Wt, in the presence or absence of unlabeled competitor DNA, then analyzed on a native 5% polyacrylamide gel. (Lane *1*) free probe, (lanes 3–6) competition by unlabeled Wt probe. (Lane *7*) lack of competition by a nonspecific control probe. The arrowhead indicates position of *ets-1*-dependent nucleoprotein complex.
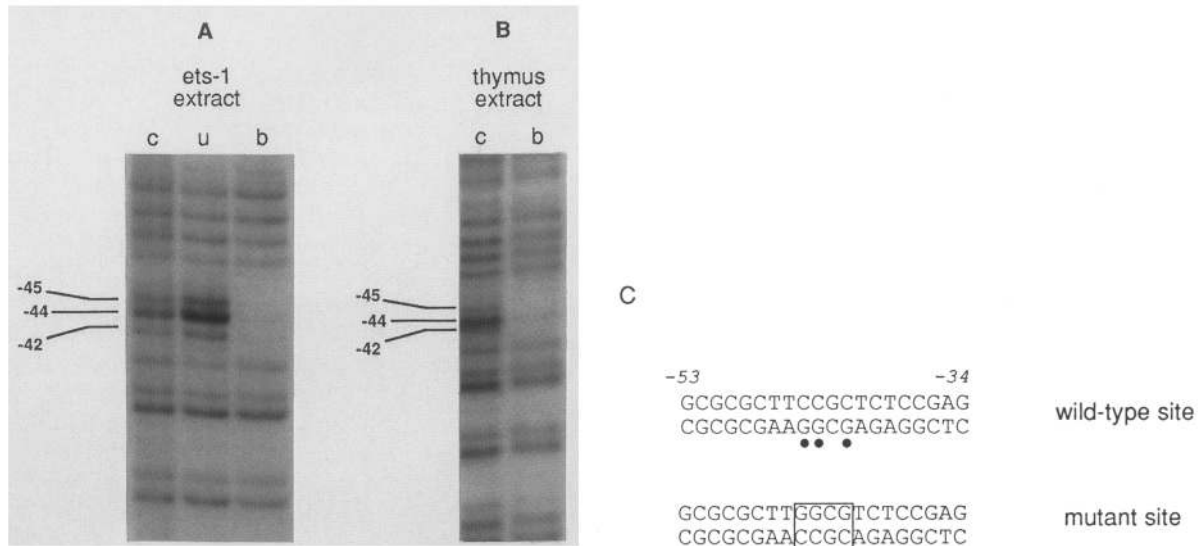
Figure 5. Methylation interference analysis of *ets-1* DNA-binding activity. (*A*) Methylation interference was performed with an 85-bp restriction fragment of the peMSV plasmid that was end-labeled on the bottom strand 50 nucleotides downstream of the *ets-1* binding site. The control lane (*c*) contains DNA that was methylated and cleaved without being exposed to protein. The second lane (*u*) contains unbound DNA that was incubated with *ets-1*-containing extracts but had the mobility of free DNA on nondenaturing, 5% acrylamide gels. The third lane (*b*) contains cleavage products of bound DNA recovered from slower migrating nucleoprotein complexes. (*B*) The same analysis performed with fractionated calf thymus extracts. (*C*) Summary of methylation interference data on the LTR *ets-1* binding site sequence bottom strand. Boxed base pairs are positions of transversions introduced in the LTR *ets-1* binding site. (●) Guanine residue at which methylation interfered with binding. Coordinates are promoter nucleotides numbered with respect to the start site of transcription in the LTR.
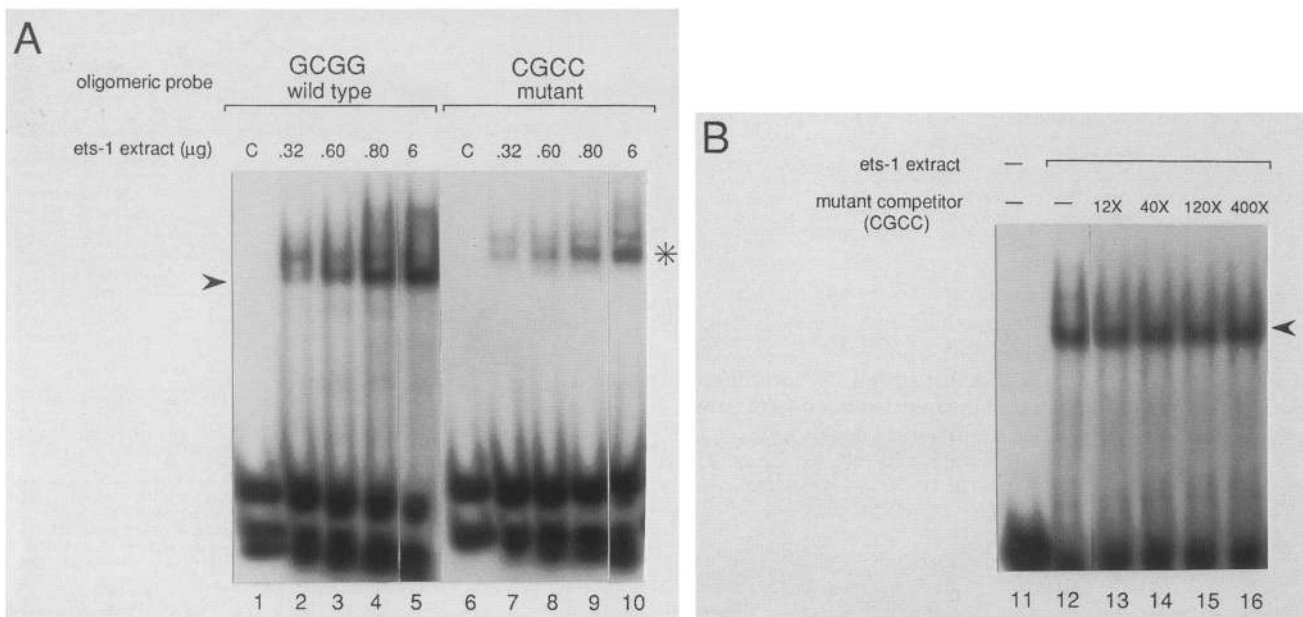

Figure 6. Comparison of *ets-1* binding affinity to the wild-type and mutant *ets-1* binding sites. The arrowhead indicates the *ets-1*-dependent nucleoprotein complex. The asterisk indicates nonspecific complexes that are sensitive to competition by oligonucleotides of any sequence composition. (*A*) Wild-type and mutant oligomeric probes were ³²P-labeled, incubated with increasing amounts of *ets-1*-containing bacterial extracts (as indicated), and analyzed in mobility-shift assays. Lanes *1* and *6* display probe only. The lower band observed in the probe-only lanes as well as in the experimental lanes represents single-stranded oligonucleotides in the probe preparation. The slight (and unexplained) difference in mobility of the wild-type and mutant oligomeric probes does not significantly affect the mobility of the observed nucleoprotein complexes. (*B*) Competition experiments in which ³²P-labeled wild-type probe is incubated with *ets-1*-containing extract in the presence or absence of unlabeled mutant site probe in the indicated molar excess.
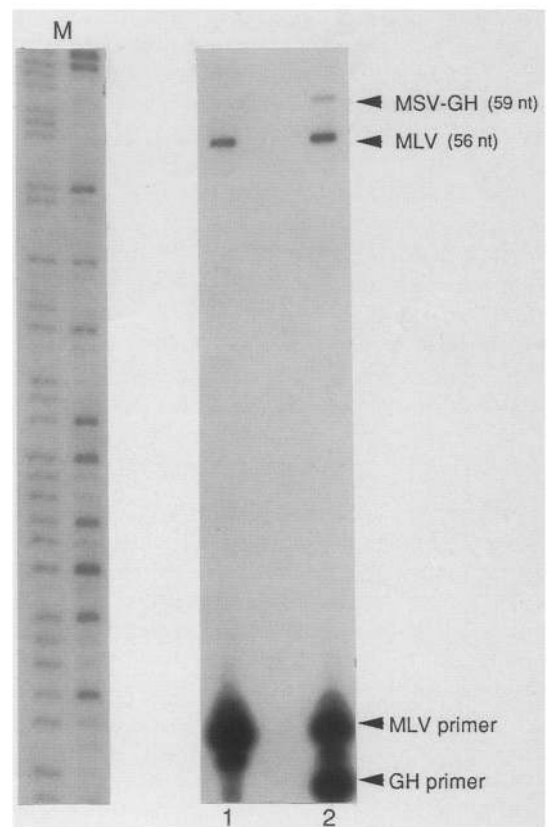
A
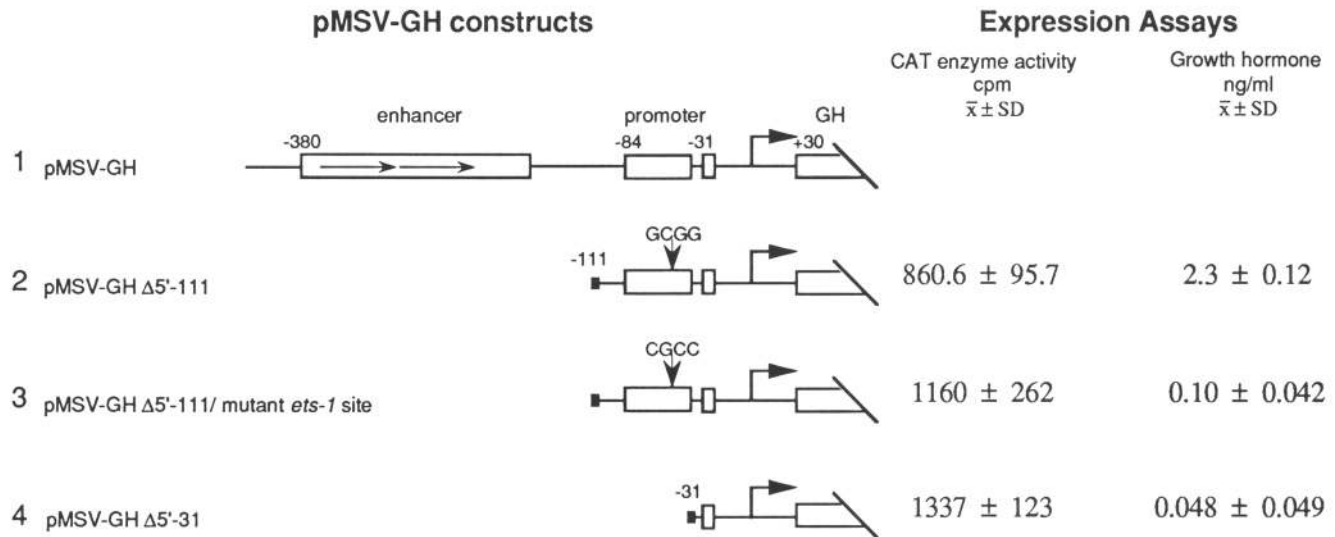
Figure 7. Effect of *ets-1* binding site mutations on the transient expression of MSV LTR promoter–growth hormone gene fusions. (*A*) MSV LTR regions, as defined in mouse L cells and frog oocytes (Graves et al. 1985b), fused to human growth hormone (GH) reporter gene, pMSV-GH constructions. The pMSV-GH construct, line 3, bears the indicated mutations in the *ets-1* binding site. Transient expression assays were performed (in triplicate) in EL4 cells cotransfected with pMSV-GH and control pMSV-CAT plasmids. The means and standard deviations are presented. The mean background value for CAT activity (998 cpm) was subtracted from the presented cpm values. The mean background value for the growth hormone assay (503 cpm) was subtracted from experimental cpm values before conversion to nanograms per milliliter growth hormone. The GH concentrations in lines 3 and 4 were computed from cpm values that are near the level of sensitivity of the assay. (*B*) Right panel shows primer extension mapping of transcripts synthesized in EL4 cells. (Lane *1*) RNA from untransfected EL4 cells, assayed with a control primer that detects endogenous retroviral RNA transcripts (MLV); (lane *2*) RNA from cells transfected with the pMSV-GH construct bearing the intact MSV LTR (Fig. 6A, line *1*), assayed with MLV and GH primers. The GH primer is specific for the transcripts synthesized from the transfected templates. In the left panel (*M*) chemical sequencing tracts (G + A and G) were co-electrophoresed as molecular weight markers.

element can function as a transcriptional activator of LTR-mediated gene expression.

Discussion

We isolated a cDNA clone of the mouse *ets-1* gene from a mouse thymus cDNA expression library by screening bacteriophage plaques with a protein binding site sequence derived from the MSV LTR. The cDNA clone encodes a 271-amino-acid polypeptide that shows strong similarity (97.4% identity) to the carboxy-terminal portion of the human *ets-1* gene product. The 30-kD polypeptide encoded by the cDNA was expressed in bacteria, visualized by SDS-PAGE, and shown to bind the conca-

```
                                                      RR  KRK K K      K  R  R    K  HK  KR   R
mu-ets1 149 KPVIPAAALAGYTGSGPIQLWQFLLELLTDKS-CQSFISWTG-DGWE-FKLSDPDEVARRWGKRKNKPKMNYEKLSRGLRYYYDKNIIHKTAGKRYVYRFV 246
   v-ets     KPVIPAAALAGYTGSGPIQLWQFLLELLTDKS-CQSFISWTG-DGWE-FKLSDPDEVARRWGKRKNKPKMDYEKLSRGLRYYYDKNVIHKTAGKRYVYRFV
 ch-ets1     KPVIPAAALAGYTGSGPIQLWQFLLELLTDKS-CQSFISWTG-DGWE-FKLSDPDEVARRWGKRKNKPKMNYEKLSRGLRYYYDKNIIHKTAGKRYVYRFV
 Hu-ets1     KPVIPAAALAGYTGSGPIQLWQFLLELLTDKS-CQSFISWTG-DGWE-FKLSDPDEVARRWGKRKNKPKMNYEKLSRGLRYYYDKNIIHKTAGKRYVYRFV

 mu-ets2     KPVIPAAVLAGFTGSGPIQLWQFLLELLSDKS-CQSFISWTG-DGWE-FKLADPDEVARRWGKRKNKPKMNYEKLSRGLRYYYDKNIIHKTSGKRYVYRFV
 ch-ets2     KPVIPAAILAGFTGSGPIQLWQFLLELLTDKS-CQSFISWTG-DGWE-FKLADPDEVARRWGRRKNKPKMNYEKLSRGLRYYYDKNIIHKTSGKRYVYRFV
 Hu-ets2     KPVIPAAVLAGFTGSGPIQLWQFLLELLSDKS-CQSFISWTG-DGWE-FKLADPDEVARRWGKRKNKPKMNYEKLSRGLRYYYDKNIIHKTSGKRYVYRFV

 Hu-erg      ILGPTSSRLANP-GSGQIQLWQFLLELLSDSS-NSSCITWEG-TNGE-FKMTDPDEVARRWGERKSKPNMNYDKLSRALRYYYDKNIMTKVHGKRYAYKFD
 Hu-elk      KPVIPAAALAGYTGSGPVTLWQFLLQLLREQG-NGHIISWTSRDGGE-FKLVDAEEVARLWGLRKNKTNMNYDKLSRALRYYYDKNIIRKVSGQKFVYKFV
 Dr-E74      KPKLEMGVKRRSREGSTTYLWEFLLKLLQDREYCPRFIKWTN-REKGVFKLVDSKAVSRLWGMHKNKPDMNYETMGRALRYYYQRGILAKVDGQRLVYQFV
                                                       R   bK K       R  R   b   K    R
```

Figure 8. *ets*-family conserved domain. The predicted amino acid sequence of viral (Nunn et al. 1983), chicken (Watson et al. 1988b), and human (Watson et al. 1988a) *ets-1*; mouse (Watson et al. 1988a), chicken (Boulukos et al. 1988), and human (Watson et al. 1988a) *ets-2*; human *erg* (Rao et al. 1987) and *elk-1* (Rao et al. 1989) and *Drosophila* E74 (Burtis et al. 1990) over a 97-amino-acid domain that shows striking conservation. Basic amino acids in cellular genes of the *ets-1* group are indicated at *top* of figure. The extent of conservation of these basic amino acid positions throughout the whole family is indicated at the bottom of the Figure. *b* indicates either H (histidine), K (lysine), or R (arginine). Conserved tryptophans are boxed. (v) Viral; (mu) mouse; (Hu) human; (ch) chicken; (Dr) *Drosophila*.

temerized oligonucleotide probe used to screen the cDNA library. These data demonstrate that the *ets-1* ORF of the λ clone encodes a DNA-binding activity.

However, two pieces of data suggest that this cDNA does not encode the entire mouse *ets-1* polypeptide. First, there is no initiator methionine in the amino-terminal portion of the ORF. If initation occurred at the first AUG in the cDNA sequence, 198 bp that are highly conserved in both human and chicken *ets-1* gene products would be untranslated. Second, immunoprecipitation studies indicate that the mouse *ets-1*-related polypeptides have molecular weights in the range of 63,000–70,000 (Pognonec et al. 1988), whereas the polypeptide encoded by the isolated λgt11 clone is only 30 kD. Minimally, the *ets-1* cDNA clone used in these experiments is missing the amino-terminal half of the mouse *ets-1* gene product. It will be important to determine whether the full-length mouse *ets-1* protein has the same DNA-binding properties as the polypeptide produced from the cDNA clone.

Nevertheless, the study of a fragment of the *ets-1* protein indicates that the 271 carboxy-terminal amino acids of the mouse *ets-1* protein are sufficient for sequence-specific DNA binding. The sequence of this region bears no resemblance to well-characterized DNA-binding domains (zinc fingers or helix-turn-helix motifs). However, the 271 amino acids do include the region of the *ets-1* gene that is highly homologous among all members of the *ets* gene family. Figure 8 shows the amino acid sequence of the highly conserved region of five *ets* family members, *ets-1* (mouse, chicken, human, and virus), *ets-2* (mouse, chicken, human), *erg* and *elk-1* (human), and E74 (*Drosophila*). Several features of the primary predicted amino acid sequence in this conserved domain are striking and suggest that this domain might be involved in DNA binding. A set of three tryptophans, as noted by Anton and Frampton (1988), demonstrate spacing similar to the tryptophan repeats of the recently described *myb* motif (Tice-Baldwin et al. 1989). Proteins bearing this repeat exhibit sequence-specific DNA-

binding activities; however, a role for the tryptophans in DNA-binding activity of the proteins has not been determined. A high density of basic amino acids is observed between amino acids 203 and 246 (16 of 43 amino acids are lysine, histidine, or arginine). Clusters of positively charged amino acids have been shown to be critical for the DNA binding by the leucine zipper/basic domain (bZIP) group of DNA-binding proteins (Landschulz et al. 1989). (No obvious conservation of the bZIP motif is seen in the *ets* family.) Additional experiments are in progress to delineate the DNA-binding domain of *ets-1* and to determine whether the conserved domain of the *ets* family is sufficient for sequence-specific DNA binding.

The DNA sequence specificity of *ets-1* binding was confirmed by methylation interference experiments. A set of close DNA contacts important to the stability of the *ets-1* nucleoprotein complexes was identified. The importance of these specific contacts was tested further by the construction of a mutant oligonucleotide probe in which the nucleotides involved in putative close contacts were altered. In mobility-shift assays, the mutant probe showed no specific binding; in competition experiments, the mutant site oligonucleotide failed to compete with the wild-type binding site probe. We have not analyzed any additional *ets-1* binding sites from which to derive a consensus binding sequence. However, the *ets-1* binding site is in a region of the MSV LTR that shows strong conservation in the LTRs of all type-C murine, feline, and simian retroviruses (Golemis et al. 1990). The conserved sequences nearest the binding site include the CGCGCTTNNGC between positions −52 and −42. The NNGC nucleotides were altered in our site-directed mutagenesis. Additional experiments are in progress to describe more completely the sequence requirements for *ets-1* binding.

The transcriptional activity of an LTR promoter bearing a mutant binding site for *ets-1* sheds new light on the biological function of the *ets* gene family. In transient expression assays in EL4 cells, multiple mutations

in the *ets-1* binding site strongly debilitated the promoter activity. These data suggest that a protein with the DNA-binding activity of *ets-1* functions as a transcriptional activator. It is reasonable to propose that *ets-1* is the EL4 cell protein that functions as an activator of the LTR-mediated expression, since RNA and protein products of the mammalian and chicken *ets-1* genes have been reported to be expressed preferentially in lymphoid cells (Chen 1985; Ghysdael et al. 1986; Bhat et al. 1987, 1989), including mouse EL4 cells. However, if the DNA-binding domain of *ets-1* is encoded by the highly conserved region of the *ets* gene family (as hypothesized above), the data cannot eliminate other *ets* family members as possible transcriptional activators in EL4 cells. For example, *ets-2*, which has a more general tissue distribution than *ets-1* (Bhat et al. 1987), also could be active in the EL4 transient expression assays.

The differential expression of *ets-1* in the thymus and spleen of adult animals has led to attempts to determine a biological role for the proto-oncogene *ets-1* in T-lymphocyte development and function. Chicken and mouse *ets-1* proteins become phosphorylated during Ca^{2+}-induced T-cell activation (Pognonec et al. 1988), suggesting a role in T-lymphocyte function. More recently, a role of mouse *ets-1* in thymocyte development has been proposed based on the coordinate expression of *ets-1* and CD4 in early thymocytes (Bhat et al. 1989). The finding that an *ets-1* polypeptide binds a promoter element in a sequence-specific manner suggests that the *ets-1* gene product functions as a transcription factor during thymocyte development and/or mature T-lymphocyte function.

It has long been speculated that the ability of Moloney MLV to cause lymphomas is due to the successful replication of the virus in lymphoid tissue (Jaenisch 1979; Chatis et al. 1983; Evans and Morrey 1987; Speck et al. 1990). The ability of the retroviral LTR to exploit a lymphoid transcription factor would explain in part this success. As described by Golemis and colleagues (1990), the regions of the MSV LTR bearing the *ets-1* binding site is highly conserved in the LTRs of all murine, simian, and feline type-C retroviruses. We speculate that *ets-1* or other members of the *ets* proto-oncogene family function as transcription factors for all of these viruses during their replication in mammalian cells.

Chicken *ets-1* is transduced by the avian leukosis virus E26 as part of the p135 transforming protein. The p135 protein also contains a fragment of the chicken *myb* gene. Since *myb* also binds DNA in a sequence-specific manner (Biedenkapp et al. 1988; Ness et al. 1989), the sequence-specific DNA-binding properties of the viral p135 must now be tested. If the *ets-1* and/or *myb* DNA-binding domain(s) are intact in the p135 oncogene, altered utilization of the native DNA-binding activities of these two cellular gene products could be responsible for p135-mediated transformation. For example, a transcriptional activation domain on *myb* could regulate promoters bearing a *ets-1* binding site (or vice versa).

Finally, the discovery that *ets-1* is a sequence-specific DNA-binding protein and that a cellular protein with

this DNA-binding activity is a transcriptional activator has implications for the function of the entire *ets* family. Two provocative features of the *ets* gene family have been reported previously. First, there is differential tissue distribution of various *ets*-related proteins in human, chicken, and mouse. For example, as noted above, *ets-1* expression is highest in lymphoid lineages, whereas *ets-2* has a more general tissue distribution (Bhat et al. 1987). A recently described family member, *elk-1*, is more abundant in lung and testis than in hematopoietic lineages (Rao et al. 1989). Second, as described above, there is strong conservation of a 97-amino-acid domain that has some resemblance to DNA-binding proteins. We speculate that the *ets* family of proteins is a new class of transcriptional activators. Furthermore, we propose that the members of the family share a conserved structural domain designed for DNA binding. This phenomenon has been well described in the case of the *Oct-1* and *Oct-2* proteins, which show different tissue distribution and regulation, yet share a common DNA-binding domain and bind an identical octameric DNA sequence (Herr et al. 1988). The broader class of homeo domain proteins (Scott et al. 1989) also illustrate this emerging feature of eukaryotic transcriptional regulators. As in other groups of related DNA-binding proteins, an important question to be resolved for the *ets* family will be how differential utilization of a common DNA-binding structure is regulated.

Methods

Binding site probes

DNase I protection assays were performed on the 150-bp *Hind*III–*Bgl*II restriction fragment of pMSV-tk (Δ5'-111) (Graves et al. 1985b), end-labeled on the top strand at the *Hind*III site or on the bottom strand at the *Bgl*II site with T4 polynucleotide kinase and [γ-^{32}P]ATP. An LTR subclone, peMSV(−74/−31), which contains the MSV LTR promoter sequences between positions −74 and −31 (with respect to the transcription start site) was used in the *ets-1* 30-kD binding studies. It was constructed from a 5'-deletion mutant of the MSV LTR, whose −74 end point bears a *Bam*HI linker. This plasmid was restricted with *Sst*I and treated with S1 nuclease to expose position −31 to the addition of a *Bgl*II linker. The 50-bp *Bam*HI–*Bgl*II restriction fragment was then ligated into the *Bam*HI site of pEMBL 18 polylinker. The 85-bp *Hind*III–*Eco*RI restriction fragment of peMSV(−74/−31) was end-labeled at either the *Eco*RI polylinker site (bottom strand) or the *Hind*III site (top strand) with T4 polynucleotide kinase and [γ-^{32}P]ATP. Radiolabeled fragments were resolved by native polyacrylamide gel electrophoresis, electrophoresed onto NA-45 paper, then eluted with 1 M NaCl. Alternatively, 25-nucleotide complementary oligonucleotides were synthesized and annealed for binding studies. The sequence of the synthetic oligomeric probe, termed Wt, which represents wild-type LTR sequences between position −53 and −34 is as follows:

```
GATCTGCGCGCTTCCGCTCTCCGAG
    ACGCGCGAAGGCGAGAGGCTCCTAG
```

The mutant *ets-1* binding site probe, termed Mt, was also generated from two synthetic complementary oligonucleotides.

It has the identical sequence as the Wt probe except for four base pair substitutions (indicated in lower case):

GATCTGCGCGCTTggcgTCTCCGAG
ACGCGCGAAccgcAGAGGCTCCTAG

The nonspecific control probe, which contains a binding site for the mammalian transcription factor, C/EBP (Vinson et al. 1988), was also synthesized.

AATTGCCTGATTGCCCAATTGT
CGGACTAACGGGTTAACATTAA

End-labeling of all oligonucleotide probes with T4 polynucleotide kinase and $[\gamma\text{-}^{32}\text{P}]\text{ATP}$ preceded annealing. Concatemerized versions of oligomer probes were constructed by ligating radiolabeled, annealed oligonucleotides with T4 DNA ligase for 12 hr. Gel electrophoresis of ligation products indicated that 10–20 copies of the repeat unit were present in each multimeric DNA molecule.

Thymus cDNA library construction

Polyadenylated RNA was prepared from thymus glands of 4-week-old mice (BALB/c). A bacteriophage λgt11 cDNA expression library was constructed following standard protocols (Gubler and Hoffman 1983; Huynh et al. 1985), as detailed by Klickstein and Neve (1989). The following modifications were made: An equal molar mixture of poly(dT) and random hexamer primers were used for first-strand cDNA synthesis. Blunt-ended double-stranded cDNA was engineered for cloning into EcoRI-restricted and phosphatased λgt11 arms (Promega) by the addition of EcoRI adaptors:

AATTCCCGGGATCC
GGGCCCTAGG

Recombinant phage were packaged in Gigapack Gold II extracts (Stratagene). The library contained 2.5×10^6 independent recombinants.

Screening of λgt11 plaques with DNA-binding site probes

The thymus cDNA expression library was screened with the Wt oligomeric probe by the protocol of Vinson et al. (1988) except for the following alterations. No guanidinium hydrochloride denaturation was performed. Recombinant phage (from the unamplified library) were plated at 3×10^4 pfu/150-mm plate for the primary screening. Concatemerized oligonucleotide probes, labeled to a specific activity of 10^9 cpm/μg, were used at a concentration of 2–5 ng/ml of binding buffer. All filter manipulations were performed at 4°C.

λgt11 clone subcloning and sequencing

Two EcoRI restriction fragments (547 bp and 2.1 kbp) of the ets-1 cDNA insert were cloned into the Bluescript pKS vector (Stratagene) and subjected to dideoxy sequencing using Sequenase (US Biochemical). Both strands of the 822 bp (Fig. 2) were sequenced using either the M13 universal primer, the T3 primer, or synthetic primers designed to hybridize to internal sites in the EcoRI subclones.

Bacterial expression of ets-1 ORF

The pET-ets-1 plasmid contains 886 bp of mouse ets-1 sequence. It was constructed by converting the AflII site of the ets-1 cDNA at position +886 (with respect to the cloning adapter) to a BamHI site by addition of a synthetic linker sequence. The BamHI site of the cDNA cloning adapter and BamHI linker site were then restricted to generate an 886-bp restriction fragment for insertion in the pET3a vector (Rosenberg et al. 1987). BL21(DE3) pLysS cells (Rosenberg et al. 1987) containing either the control pET3 plasmid or the pET-ets-1 plasmid were grown at 37°C to an OD_{600} of 0.5 units. IPTG was then added to 1 mM and cells were incubated for an additional hr at 37°C. Cells were harvested at 3000g for 5 min at 23°C, resuspended in lysis buffer [50 mM Tris-HCl (pH 7.5), 1 mM EDTA, 1 mM DTT, 1 mM PMSF], quick-frozen, and thawed. The lysate was incubated on ice for 15 min to allow lysis from endogenous lysozyme, then adjusted to 1 M NaCl. Samples were incubated for 20 min on a rotator at 4°C, then centrifuged at 10,000g for 30 min at 4°C. The supernatant was dialyzed against binding buffer TGMEK [25 mM Tris-HCl (pH 7.9), 10% glycerol, 5 mM $MgCl_2$, 0.1 mM EDTA] containing 100 mM KCl and supplemented with 1 mM DTT, 1 mM PMSF. The ets-dependent 30-kD polypeptide comprised 3–5% of the protein mass in the extracted cell lysates.

DNA-binding assays

Southwestern blot analysis was performed on total cell lysates of BL21(DE3) cells. The pelleted cells were resuspended directly in PAGE sample buffer and subjected to SDS-PAGE (Laemmli 1970). Proteins were electroblotted onto nitrocellulose filters (Towbin et al. 1979). The filters were processed and probed using conditions identical to the bacteriophage plaque replica filters.

Mobility-shift assays for titrations, competitions and methylation interference were performed essentially as described by Fried and Crothers (1981) and modified by Nye and Graves (1990). In a total volume of 20 μl, 0.05 ng of ^{32}P-labeled oligonucleotide was incubated with 0.3–6 μg of bacterial extract, 1 μl poly[d(I-C)], and 0.25 ng of a nonspecific 25-bp oligonucleotide in binding buffer at 50 mM KCl. After 20 min at 4°C, the samples were electrophoresed in 5% (30 : 0.8) polyacrylamide gels using a 45 mM Tris-borate (pH 8.3) buffer. Methylation interference reactions were performed as described by Siebenlist and Gilbert (1980) and modified by Nye and Graves (1990).

DNase I footprint assays were performed as previously described (Graves et al. 1986). In a total volume of 50 μl, ^{32}P-end-labeled DNA (~1.5 ng) was incubated with 1 μg of calf thymus nucleus extract (fractionated as described below) and 100 ng of poly[d(I-C)] in binding buffer.

pMSV-GH constructs and ets-1 binding site mutagenesis

The pMSV-GH plasmids were constructed by cloning the LTR-bearing, HindIII–BglII restriction fragments of wild-type, Δ5'-111, and Δ5'-31 MSV-tk plasmids (Graves et al. 1985b) upstream of the human growth hormone gene in pOGH (Nichols Institute). At the 3' terminus of LTR sequences in these constructs, position +30, a synthetic BglII linker is fused to the BamHI site of pOGH. The LTR bearing a disrupted ets-1 binding site was constructed by synthesizing complementary oligonucleotides containing LTR sequences between a native BssHI site at position −52 and a native SstI site at position −32, except for the four base substitutions between positions −42 and −45. This synthetic DNA was ligated into an MSV LTR-bearing plasmid that had been restricted at the unique SstI site and partially restricted with BssHI to generate a BssHI cloning site at −52. Bacteria were transformed and the resulting colonies were screened by hybridization to the mutant oligonucleotide. The presence of the mutated sequence and the

absence of other changes in the LTR promoter region were confirmed by sequencing.

Transient expression assays

EL4 cells (kindly provided by N. Speck) were maintained in suspension culture in Dulbecco's modified minimal essential medium, supplemented with 10% fetal calf serum. A total of 6.5×10^6 EL4 cells from logarithmically growing cultures were collected by centrifugation, washed once with calcium- and magnesium-free Tris-buffered saline, and resuspended in 0.65 ml of HEPES-buffered saline (Wigler et al. 1977) adjusted to 0.19 mM dithiothreitol. Plasmid DNA (40 μg of pMSV-GH and 5 μg of pMSV-CAT) was then added. The electroporator (Hoeffer) was charged at 190 volts and 490 μF and subsequently discharged across the cell suspension. Immediately after electroporation, a quarter of the cell suspension was diluted into 4 ml of growth medium and placed directly in a growth chamber. Human growth hormone concentration of the media was determined at 40 hr post-transfection using a radioimmune assay and standard growth hormone solutions (Selden et al. 1986) according to manufacturer's specifications (Nichols Institute). CAT assays were performed as described by Clevers et al. (1988). Primer extension assays were performed as previously described (Graves et al. 1985b). The GH primer, a 22-nucleotide synthetic oligonucleotide, is complementary to the coding strand of the most promoter-proximal growth hormone sequences in the pMSV-GH constructs. The extension products include the sequences of a *Bgl*II linker (7 nucleotides) and the LTR between position + 30 and the cap site. The 25-nucleotide MLV primer (Graves et al. 1985a), which detects expression from endogeneous MLV proviruses, served as an internal control for RNA recovery.

Calf thymus extracts

Thymus glands from newly slaughtered calves (12–18 months) were defatted, cut into chunks, frozen in liquid nitrogen, and stored at −80°C. Nuclei were isolated in buffers similar to those described for the isolation of at liver nuclei (Gorski et al. 1986; Graves et al. 1986). Frozen thymus glands were thawed for 5 min in buffer A [0.25 M sucrose, 15 mM Tris-Cl (pH 7.9), 120 mM KCl, 15 mM NaCl, 2 mM EDTA] and pieces were then mixed with cold buffer B [2 M sucrose, 10 mM Tris-Cl (pH 7.9), 25 mM KCl, 1 mM EDTA, 0.15 mM spermine, 0.5 mM spermidine, 10% glycerol] at a ratio of 2 ml/gram of tissue. The mixture was homogenized for 20 sec at high speed in an Oster commercial blender. Homogenates were sieved three times through two layers of cheesecloth and the filtrate was then centrifuged at 2500g for 30 min at 4°C. The supernatant was removed by aspiration and the nuclear pellet was resuspended in buffer C [10 mM Tris-Cl (pH 7.9), 100 mM KCl, 2 mM MgCl$_2$, 0.1 mM EDTA, 0.15 mM spermine, 0.5 mM spermidine, 10% glycerol] to a concentration of approximately 3.5×10^8 nuclei/ml of buffer (estimate 2.5×10^8 nuclei per gram thymus tissue). Buffers B and C contained the following protease inhibitors: 1 mM DTT, 2.5 mM benzamidine, 0.5 mM PMSF, 1 μg/ml leupeptin, 1 μg/ml pepstatin. Nuclear extracts were prepared from isolated thymic nuclei as previously described for preparation of rat liver nuclear extracts (Graves et al. 1986). The final ammonium sulfate precipitate was redissolved in TGMEK$_{100}$ (supplemented with 1 mM DTT, 0.5 mM PMSF, 2.5 mM benzamidine) to a protein concentration of approximately 8 mg/ml, dialysed in TGMEK$_{100}$, and centrifuged at 10000g for 10 min at 4°C.

The extract was fractionated by two chromatography steps conducted in TGMEK. Nuclear extracts were loaded onto a 40-ml heparin agarose (Bethesda Research Laboratories) column equilibrated with TGMEK$_{100}$ and bound protein was eluted by buffer salt steps of 300 mM and 500 mM KCl. The 500 mM KCl fraction contained LTR binding activity as assayed by DNaseI protection experiments. After dialysis to 100 mM KCl, the active fraction was loaded onto a 5-ml double-stranded DNA Sepharose column (Graves et al. 1986). The bound fraction was eluted with a 500 mM KCl and used, undialysed, in the DNA-binding assays.

Acknowledgments

Note added in proof

Sequence data described in this paper have been submitted to the EMBL/GenBank Data Libraries.

References

Anton, I.A. and J. Frampton. 1988. Tryptophans in *myb* proteins. *Nature* **336**: 719.

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.

Biedenkapp, H., U. Borgmeyer, A.E. Sippel, and K.-H. Klempnauer. 1988. Viral *myb* ocogene encodes a sequence-specific DNA-binding activity. *Nature.* **335**: 835–837.

Boulukos, K.E., P. Pognonec, A. Begue, F. Galibert, J.C. Gesquiere, D. Stehelin, 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 *ets1* protein domain required for nuclear localization in cells and DNA-binding activity in vitro. *Mol. Cell. Biol.* **9**: 5718–5721.

Burtis, K.C., C.S. Thummel, C.W. 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* (in press)

Chatis, P.A., C.A. Holland, J.W. Hartley, W.P. Rowe, and N. Hopkins. 1983. Role for the 3' end of the genome in determining disease specificity of Friend and Moloney murine leukemia viruses. *Proc. Natl. Acad. Sci.* **80**: 4408–4411.

Chen, J.H. 1985. The proto-oncogene c-*ets* is preferentially expressed in lymphoid cells. *Mol. Cell. Biol.* **5**: 2993–3000.

Clevers, H.C., S. Dunlap, T.E. Wileman, and C. Terhorst. 1988. Human Cd3-Epsilon gene contains three miniexons and is transcribed from a non-TATA promoter. *Proc. Natl. Acad. Sci.* **85**: 8156–8160.

Evans, L.H. and J.D. Morrey. 1987. Tissue specific replication of

Friend and Moloney murine leukemia viruses in infected mice. *J. Virol.* **61**: 1350–1357.

Fried, M. and D.M. Crothers. 1981. Equilibria and kinetics of lac repressor-operator interactions by polyacrylamide gel electrophoresis. *Nucleic Acids Res.* **9**: 6505–6525.

Fujiwara, S., R.J. Fisher, A. Seth, N.K. Bhat, S.D. Showalter, M. Zewig, and T.S. Papas. 1988a. Characterization and localization of the products of the human homologs of the v-*ets* oncogene. *Oncogene* **2**: 99–103.

Fujiwara, S., R.J. Fisher, N.K. Bhat, S. Espina, and T.S. Papas. 1988b. A short-lived nuclear phosphoprotein encoded by the human *ets-2* proto-oncogene is stabilized by activation of protein kinase C. *Mol. Cell. Biol.* **8**: 4700–4706.

Gegonne, A., D. Leprince, M. Duterque-Coquillaud, B. Vandenbunder, A. Flourens, J. Ghysdael, B. Debuire, and D. Stehelin. 1987. Multiple domains for the chicken cellular sequences homologous to the v-*ets* oncogene of the E26 retrovirus. *Mol. Cell. Biol.* **7**: 806–812.

Ghysdael, J., A. Gegonne, P. Pognonec, D. Dernis, D. Leprince, and D. Stehelin. 1986. Identification and preferential expression in thymic and bursal lymphocytes of a c-*ets* oncogene-encoded Mr 54,000 cytoplasmic protein. *Proc. Natl. Acad. Sci.* **83**: 1714–1718.

Golay, J., M. Introna, and T. Graf. 1988. A single point mutation in the v-*ets* oncogene affects both erythroid and myelomonocytic cell differentiation. *Cell* **55**: 1147–1158.

Golemis, E.A., N.A. Speck, and N. Hopkins. 1990. Alignment of U3 region sequences of mammalian type C viruses: Identification of highly conserved motifs and implications for enhancer design. *J. Virol.* **64**: 534–542.

Gorski, K., M. Carneiro, and U. Schibler. 1986. Tissue-specific in vitro transcription from the mouse albumin promoter. *Cell* **47**: 767–776.

Graves, B.J., S.P. Eisenberg, D.M. Coen, and S.L. McKnight. 1985a. Alternate utilization of two regulatory domains within the Moloney murine sarcoma virus long terminal repeat. *Mol. Cell. Biol.* **5**: 1959–1968.

Graves, B.J., R.N. Eisenman, and S.L. McKnight. 1985b. Delineation of transcriptional control signals within the Moloney murine sarcoma virus long terminal repeat. *Mol. Cell. Biol.* **5**: 1948–1958.

Graves, B.J., P.F. Johnson, and S.L. McKnight. 1986. Homologous recognition of a promoter domain common to the MSV LTR and the HSV *tk* gene. Cell. **44**: 565–576.

Gubler, U. and B.J. Hoffman. 1983. A simple and very effective method for generating cDNA libraries. *Gene* **25**: 263–269.

Herr, W., R.A. Sturm, R.G. Clerc, L.M. Corcoran, D. Baltimore, P.A. Sharp, H.A. Ingraham, M.G. Rosenfeld, M. Finney, G. Ruvkun, and H.R. Horvitz. 1988. The POU domain: a large conserved region in the mammalian *pit-1*, *oct-1*, *oct-2*, and *Caenorhabditis elegans unc-86* gene products. *Genes Dev.* **2**: 1513–1516.

Huynh, T.V., R.A. Young, and R.W. Davis. 1985. Constructing and screening cDNA libraries in λgt10 and λgt11. In *DNA cloning: A practical approach* (ed. D. Glover), pp.49–78. IRL Press. Oxford.

Jaenisch, R. 1979. Moloney leukemia virus gene expression and gene amplification in preleukemic and leukemic BALB/Mo mice. *Virology.* **93**: 80–90.

Klickstein, L.B. and R.L. Neve. 1989. Construction of recombinant DNA libraries. In *Current protocols in molecular biology* (ed. F. M. Ausbel, R. Brent, R. Kingston, D. Moore, J. G. Serdman, J. A. Smith, and K. Struhl). John Wiley and Sons. New York.

Laemmli, U.K. 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage. *Nature* **227**: 680–685.

Landschulz, W.H., P.F. Johnson, and S.L. McKnight. 1989. The DNA-binding domain of the rat liver nuclear protein C/EBP is bipartite. *Science* **243**: 1681–1699.

Leprince, D., A. Gegonne, J. Coll, C. de Taisne, A. Schneeberger, C. Lagrou, and D. Stehelin. 1983. A putative second cell-derived oncogene of the avian leukaemia retrovirus E26. *Nature* **306**: 395–397.

Ness, S.A., Å. Marknell, and T. Graf. 1989. The v-*myb* oncogene product binds to and activates the promyelocyte-specific *mim-1* gene. *Cell* **59**: 1115–1125.

Nunn, M.F. and T. Hunter. 1989. The *ets* sequence is required for induction of erythroblastosis in chickens by avian retrovirus E26. *J. Virol.* **63**: 398–402.

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.

Nye, J.A. and B.J. Graves. 1990. Alkylation interference identifies essential DNA contacts for sequence-specific binding of C/EBP. *Proc. Natl. Acad. Sci.* (in press)

Pearson, W.R. and D.J. Lipman. 1988. Improved tools for biological sequence comparison. *Proc. Natl. Acad. Sci.* **85**: 2444–2448.

Pognonec, P., K.E. Boulukos, J.C. Gesquiere, D. Stehelin, and J. Ghysdael. 1988. Mitogenic stimulation of thymocytes results in the calcium-dependent phosphorylation of c-*ets-1* proteins. *EMBO J.* **7**: 977–983.

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.

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.

Radke, K., H. Beug, S. Kornfeld, and T. Graf. 1982. Transformation of both erythroid and myeloid cells by E26, an avian leukemia virus that contains the *myb* gene. *Cell* **31**: 643–653.

Rao, V.N., T.S. Papas, and E.S.P. Reddy. 1987. *erg*, a human *ets*-related gene on chromosome 21: Alternative splicing, polyadenylation, and translation. *Science* **237**: 635–639.

Rao, V.N., K. Huebner, M. Isobe, A. Ar-Rushdi, C.M. Croce, and E.S.P. Reddy. 1989. *elk*, tissue-specific *ets*-related genes on chromosomes X and 14 near translocation breakpoints. *Science* **244**: 66–70.

Reddy, E.S.P., V.N. Rao, and T.S. Papas. 1987. The *erg* gene: A human gene related to the *ets* oncogene. *Proc. Natl. Acad. Sci.* **84**: 6131–6135.

Rosenberg, A.H., B.N. Lade, D. Chui, S. Lin, J.J. Dunn, and F.W. Studier. 1987. Vectors for selective expression of cloned DNAs by T7 RNA polymerase. *Gene* **56**: 125–135.

Sacchi, N., D.K. Watson, A.H.M. Guerts Van Kesel, A. Hagemeijer, J. Kersey, H.D. Drabkin, D. Patterson, and T.S. Papas. 1986. Hu-*ets-1* and Hu-*ets-2* genes are transposed in acute leukemias with (4;11) and (8;21) translocation. *Science* **231**: 379–382.

Scott, M.P., J.W. Tamkun, and G.W.I. Hartzell. 1989. The structure and function of the homeodomain. *Biochim. Biophys. Acta* **989**: 25–48.

Selden, R.F., K.B. Howie, M.E. Rowe, H.M. Goodman, and D.D. Moore. 1986. Human growth hormone as a reporter gene in regulation studies employing transient gene expression. *Mol. Cell. Biol.* **6**: 3173–3179.

Siebenlist, U. and W. Gilbert. 1980. Contacts between Esche-

richia coli RNA polymerase and an early promoter of phage T7. *Proc. Natl. Acad. Sci.* **77**: 122–126.

Singh, H., J.H. LeBowitz, A.S. Baldwin, and P.A. 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.

Speck, N.A., B. Renjifo, E. Golemis, T. Fredrickson, J.W. Hartley, and N. Hopkins. 1990. Mutation of the core or adjacent LVb elements of the Moloney murine leukemia virus alters disease specificity. *Genes Dev.* **4**: 233–242.

Tice-Baldwin, K., G.R. Fink, and K.T. Arndt. 1989. BAS1 has a *myb* motif and activates HIS4 transcription only in combination with BAS2. *Science* **246**: 931–935.

Towbin, H., T. Staehelin, and J. Gordon. 1979. Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: Procedure and some applications. *Proc. Natl. Acad. Sci.* **76**: 4350–4354.

Vinson, C.R., K.L. LaMarco, P.F. Johnson, W.H. Landschulz, and S.L. McKnight. 1988. In situ detection of sequence-specific DNA-binding activity specified by a recombinant bacteriophage. *Genes Dev.* **2**: 801–806.

Watson, D.K., M.J. McWilliams-Smith, M.F. Nunn, P.H. Duesberg, S.J. O'Brien, and T.S. Papas. 1985. The *ets* sequence from the transforming gene of avian erythroblastosis virus, E26, has unique domains on human chromosomes 11 and 21: Both loci are transcriptionally active. *Proc. Natl. Acad. Sci.* **82**: 7294–7298.

Watson, D.K., M.J. McWilliams, P. Lapis, J.A. Lautenberger, C.W. Schweinfest, and T.S. Papas. 1988a. Mammalian *ets-1* and *ets-2* genes encode highly conserved proteins. *Proc. Natl. Acad. Sci.* **85**: 7862–7866.

Watson, D.K., M.J. McWilliams, and T.S. Papas. 1988b. Molecular organization of the chicken *ets* locus. *Virology* **164**: 99–105.

Wigler, M., S. Silverstein, L. Lee, A. Pellicer, Y. Cheng, and R. Axel. 1977. Transfer of purified herpes virus thymidine kinase gene to cultured mouse cells. *Cell* **11**: 223–232.

References	This article cites 52 articles, 29 of which can be accessed free at: **http://genesdev.cshlp.org/content/4/4/667.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.**