MOLECULAR AND CELLULAR BIOLOGY, Aug. 1993, p. 4556–4571
0270-7306/93/084556-16$02.00/0

Vol. 13, No. 8

A Novel Repression Module, an Extensive Activation Domain, and a Bipartite Nuclear Localization Signal Defined in the Immediate-Early Transcription Factor Egr-1

ANDREA L. GASHLER,[1]† SOWMYA SWAMINATHAN,[2] AND VIKAS P. SUKHATME[2]†*

Department of Biochemistry and Molecular Biology[1] and Departments of Medicine and Molecular Genetics and Cell Biology, Howard Hughes Medical Institute,[2] University of Chicago, Chicago, Illinois 60637-1963

Received 19 April 1993/Accepted 26 April 1993

***Egr-1* is an immediate-early response gene induced transiently and ubiquitously by mitogenic stimuli and also regulated in response to signals that initiate differentiation. The *Egr-1* gene product, a nuclear phosphoprotein with three zinc fingers of the Cys_2His_2 class, binds to the sequence CGCCCCCGC and transactivates a synthetic promoter construct 10-fold in transient-transfection assays. We have analyzed the structure and function of the Egr-1 protein in detail, delineating independent and modular activation, repression, DNA-binding, and nuclear localization activities. Deletion analysis, as well as fusions to the DNA-binding domain of GAL4, indicated that the activation potential of Egr-1 is distributed over an extensive serine/threonine-rich N-terminal domain. In addition, a novel negative regulatory function has been precisely mapped 5′ of the zinc fingers: amino acids 281 to 314 are sufficient to confer the ability to repress transcription on a heterologous DNA-binding domain. Specific DNA-binding activity was shown to reside in the three zinc fingers of Egr-1, as predicted by homology to other known DNA-binding proteins. Finally, nuclear localization of Egr-1 is specified by signals in the DNA-binding domain and basic flanking sequences, as determined by subcellular fractionation and indirect immunofluorescence. Basic residues 315 to 330 confer partial nuclear localization on the bacterial protein β-galactosidase. A bipartite signal consisting of this basic region in conjunction with either the second or third zinc finger, but not the first, suffices to target β-galactosidase exclusively to the nucleus. Our work shows that Egr-1 is a functionally complex protein and suggests that it may play different roles in the diverse settings in which it is induced.**

Extracellular signals for growth and differentiation must be translated into long-term changes in gene expression, and this process is likely to involve the sequential expression of regulatory proteins. Primary-response or immediate-early genes are those genes induced by mitogenic or other stimuli in the absence of de novo protein synthesis and thus constitute the first step in such a cascade. One such primary-response gene, *Egr-1* (69), also known as *NGFI-A* (45), *Krox24* (35), *zif268* (7), and *TIS8* (40), encodes a protein with three zinc finger motifs, suggesting that Egr-1 may mediate growth response by regulating distal gene expression. In this respect, Egr-1 is like other immediate-early transcription factors of the Fos (18, 34) and Jun (59) families. The evolutionary conservation of the *Egr-1* gene (69) as well as the broad spectrum of induction—by tetradecanoyl phorbol acetate and growth factors (7, 35, 40, 45, 69), by neuronal stimuli (9, 45, 69), by ischemic injury (16, 51), and in some contexts in response to differentiation signals (69)—implicate Egr-1 as an important nuclear intermediary in signal transduction.

In contrast to the universal and immediate response of *Egr-1* to mitogenic factors, in several cell types *Egr-1* expression is correlated with differentiative processes, in particular in cardiac, neural, osteoblast, and myeloid cell differentiation. For example, *Egr-1* induction accompanies differentiation of embryonal carcinoma P19 cells to cardiac muscle or to nerve and glial cells (69). *Egr-1* expression in adult mouse, highest in brain and heart, is consistent with these results and points to a role for sustained *Egr-1* expression in these differentiated cell types. In the developing mouse, the predominant expression of *Egr-1* in cartilage and bone, coordinate with *c-fos* induction, suggests a role for these coregulated genes in skeletal development (68). Recently, Nguyen et al. have shown that *Egr-1* is essential for macrophage differentiation of the myeloid leukemia cell line HL60 and further that *Egr-1* functions to restrict differentiation, since constitutive *Egr-1* expression is incompatible with the differentiation of HL60 cells along the granulocyte lineage (50).

Characterization of the *Egr-1* gene product has shown that this serine/threonine/proline-rich nuclear phosphoprotein (6, 12, 73) activates transcription through the sequence CGC CCCCGC in transient transfection assays (6, 8, 36, 53). Recently, Pavletich and Pabo have cocrystallized the three zinc fingers of Egr-1 with its cognate sequence, defining a DNA-binding domain (54). However, to date the regions responsible for activation and nuclear localization have not been mapped. A body of work suggests the modular nature of transcription factors, in which functional domains are structurally independent and able to confer activity on heterologous proteins (reviewed in reference 55). We have used deletion analysis and gene fusions to dissect the functional domains of Egr-1. We report here that the serine/threonine-rich N terminus of Egr-1 encodes its transactivation function. Remarkably, a compact 34-amino-acid region, residues 281 to 314, represses transcription in gene fusion experiments. The zinc finger domain encodes the DNA-binding activity, and this region in conjunction with a basic sequence 5′ of the fingers is responsible for nuclear localization of Egr-1.

* Corresponding author.
† Present address: Renal Division, Department of Medicine, Beth Israel Hospital and Harvard Medical School, Boston, MA 02215.

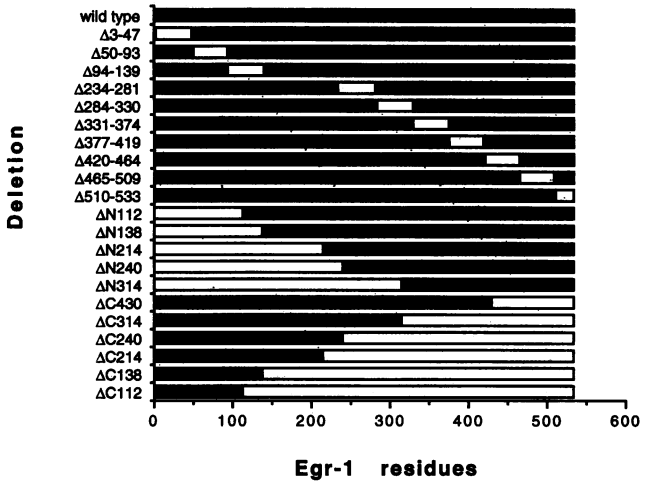


FIG. 1. Egr-1 deletion derivatives. Site-directed mutagenesis was used to generate internal deletions of the Egr-1 expression vector pCB.Egr-1. N- and C-terminal deletions were based on an Egr-1 expression vector modified to include 20 exogenous amino acids of an antigenic tag at the N terminus. The nomenclature is that the internal deletions are designated by the first and last amino acids removed. N-terminal deletions are denoted by ΔN and the last amino acid excised; C-terminal deletions are named by ΔC and the last amino acid remaining.

MATERIALS AND METHODS

Construction of Egr-1 deletion derivatives. Each of the Egr-1 deletion derivatives is expressed from the vector pCB6+, a modified version of pCB6 (M. Roth's laboratory) including *Eco*RI, *Not*I, and *Eco*RV sites. Vector pCB6+ contains the cytomegalovirus (CMV) early promoter/enhancer followed by a polylinker and human growth hormone polyadenylation signal as well as the gene for neomycin resistance expressed from the simian virus 40 (SV40) early promoter/enhancer. The parent expression vector for all deletion derivatives, pCB.Egr-1, consists of the full-length murine *Egr-1* cDNA cloned into the *Eco*RI site of pCB6+ as described previously (42). Site-directed mutagenesis was used to produce nine in-frame internal deletions of 43 to 48 amino acids and one deletion of 24 amino acids (Fig. 1). These internal deletions span the entire coding sequence except for the region between amino acids 140 and 234, which did not produce any stable derivatives. In general, each of the primers for the mutagenesis contained 20 to 25 nucleotides complementary to the regions immediately 5′ and 3′ of the desired deletion, as shown in Table 1. Clones were typically screened by *Eco*RI-*Pvu*I restriction; the deletion endpoints and reading frame were confirmed by dideoxy sequencing (62) with reagents from U.S. Biochemical. The nomenclature of these internal deletion derivatives designates the first and last amino acids excised.

To insert a 20-amino-acid antigenic tag at the N terminus of Egr-1, oligonucleotide-directed mutagenesis was used to create a unique *Nhe*I site at nucleotides 267 to 269 of pCB.Egr-1 with the primer 5′-GAGACATCAATTGCATC TCGGCCTTGCTAGCTGCCATCCCGGACCAGCGAGCT GGA-3′. Annealed synthetic oligonucleotides coding for a portion of CMV glycoprotein A followed by an *Xho*I site were inserted at the newly generated *Nhe*I site (amino acids 3 and 4 of Egr-1). The synthetic DNA encodes the 20 amino acids of the epitope, KGQKPNLLDRLRHRKNGYRH, according to optimal human codon usage. The resulting plasmid was analyzed by sequencing and tested for protein expression, nuclear localization, and the ability to transactivate compared with wild-type Egr-1 in cotransfection assays. This plasmid, which was designated pCB.Egr-1.tag and found to be equivalent to wild-type Egr-1 in all assays, was the parent for the subsequent construction of N- and C-terminal deletions.

The construction of C-terminal deletions utilized site-directed mutagenesis to insert a stop codon followed by an *Xho*I site after amino acids 112, 138, 214, 240, 314, and 430. N-terminal deletions were produced simply by excision of the internal *Xho*I fragment and religation in frame. The nomenclature is that deletions from the N terminus are designated by ΔN followed by the last amino acid deleted; C-terminal deletions are designated by ΔC and the last amino acid remaining. The reporters pΔ56foscat, containing the minimal murine c-*fos* promoter, and EBS1^{3}foscat, with three copies of a high-affinity Egr-1 binding site, have been described previously (16, 42).

Construction of GAL4–Egr-1 chimeras. Plasmid pSG424 (60), encoding the DNA-binding domain of GAL4 driven by the SV40 early promoter/enhancer and followed by a polylinker and stop codons in all three reading frames, provided the starting point for all GAL4–Egr-1 chimeras. Several domains of Egr-1 (amino acids 3 to 281, 3 to 138, 138 to 281, 420 to 533, 240 to 330, 281 to 330, and 281 to 314) were amplified by the polymerase chain reaction (PCR), digested with *Bam*HI and *Xba*I, and cloned into the corresponding sites of pSG424. The specified Egr-1 coding sequence is fused in frame C terminal to GAL4 amino acids 1 to 147 with seven synthetic amino acids at the junction. The amplified region and the junction of the construct were verified by dideoxy sequencing (62). To assay activation, a reporter with minimal basal activity and five GAL4 binding sites, pGAL4$_5$E1bCAT (39), was chosen. The reporter selected for

TABLE 1. Primers used in the construction of internal deletion derivatives of Egr-1

Construct	Primer
pCB.Egr-1Δ3-47	5′-GTTCCGGCAG CACCGAGGAA TGCCATCCCG GACCAGCGAG-3′
pCB.Egr-1Δ50-93	5′-AGGACTCTGT GGTCAGGTGC TCATAGAGGA ACTGGGGAGC CCCGTTGCTC-3′
pCB.Egr-1Δ94-139	5′-AAAGTGTTGC CACTGTTGGG GGGTTGTTCG CTCGGCTCCC-3′
pCB.Egr-1Δ234-281	5′-GATAGTGGAG TGAGCGAAGG GTACTGCAAG GCTGTGCCTG-3′
pCB.Egr-1Δ284-330	5′-GGCAAGCATA TGGGCGTTCA TGGGGCGAAG GCTGCTGGGT ACGGTTCTCC-3′
pCB.Egr-1Δ331-374	5′-AGGTGGTCAC TACGACTGAA GGGTGTCTTG CTGGGCCGGI-3′
pCB.Egr-1Δ377-419	5′-TTGTCTGCTT TCTTGTCCTT ACTGAAGTTA CGCATGCAGA-3′
pCB.Egr-1Δ420-464	5′-GTGGAGGAGC CAGGAGAGGA CTGTCTTAAA TGGATTTTGG-3′
pCB.Egr-1Δ465-509	5′-GTTGAGGTGC TGAAGGAGCT GTAGGAAGTG GGCACAGGGG-3′
pCB.Egr-1Δ510-533	5′-TTTTATTCCC TTTAGCAAAT GCTGACGCCC GCAGACGGGA-3′

the repression assays, pGAL4$_5$tkCAT (64), has a higher basal activity.

Construction of Egr-1–β-galactosidase chimeras. To create a β-galactosidase expression vector suitable for constructing N-terminal fusions, the 3.0-kb *Bam*HI β-galactosidase fragment from pMC1871 was subcloned into the corresponding site of the CMV expression vector pCB6+. This fragment contains all of the coding sequence beginning with the 10th amino acid of β-galactosidase and is sufficient for its enzymatic activity. Second, a synthetic fragment containing an ATG in a Kozak context followed by the unique sites *Bgl*II, *Kpn*I, *Hin*dIII, and *Xba*I was cloned into the *Bgl*II and *Xba*I sites to create pCB.β-GAL such that (i) the 5′-most *Bgl*II site was destroyed and (ii) the β-galactosidase coding sequence was in frame with the ATG. To create the fusion proteins, Egr-1 coding sequences were amplified by PCR and cloned in frame into the unique *Bgl*II and *Xba*I sites of pCB.β-GAL, and the resulting plasmids were sequenced through the amplified regions.

To create pCB.Egr-1(315-330).β-GAL, complementary synthetic oligonucleotides were annealed with *Bgl*II and *Xba*I overhangs and cloned into the corresponding sites. Constructs containing the 5′ basic region (amino acids 315 to 330) in conjunction with either finger 2, finger 3, fingers 2 and 3, or the H-C link of finger 3 were generated by cloning in frame into the *Xba*I site of pCB.Egr-1(315-330).β-GAL. The construct containing the 5′ basic flanking sequence with the H-C link of finger 3 and body of finger 1 was made such that there were no artificial residues (at the restriction enzyme sites). Two annealed synthetic 90-mers encoding amino acids 315 to 330, 389 to 395, and 338 to 363 were amplified with outside primers by PCR and cloned into the *Bgl*II and *Xba*I sites of pCB.β-gal. The construct with the 5′ basic region, the H-C link of finger 1, and body of finger 3 was made by PCR amplification of two fragments such that no exogenous residues were introduced at the junction. The 5′ fragment amplified residues 315 to 337, and amino acids 396 to 401 were included as a noncomplementary tail in the 3′ primer. This DNA was annealed with the amplified fragment encoding residues 396 to 419, and the mixture was reamplified with outside primers and cloned into the *Bgl*II and *Xba*I sites of pCB.β-GAL.

Transfections and CAT assays. NIH 3T3 (provided by M. Wigler, Cold Spring Harbor Laboratory) and HeLa cells were maintained in Dulbecco's modified Eagle's medium with 10% calf serum. Cells were seeded at 8×10^5 cells per 100-mm-diameter plate the day before transfection. The media was replaced 2 to 6 h prior to transfection by calcium phosphate-mediated precipitation. Each precipitate included 1 μg of the internal reference pON260, a CMV-driven β-galactosidase plasmid (67), or 3 μg of pCH110, an SV40-driven β-galactosidase plasmid (Pharmacia), and remained on the cells for 16 to 20 h. Forty-eight hours after transfection, cell extracts were prepared by freeze-thaw lysis in 0.25 M Tris-HCl (pH 7.8) with protease inhibitors (1 mM phenylmethylsulfonyl fluoride, 1 μg of pepstatin A per ml, and 1 μg of leupeptin per ml). Chloramphenicol acetyltransferase (CAT) assays were performed according to Gorman et al. (17) with equal amounts of β-galactosidase activity in order to normalize for any variation in transfection efficiency. β-Galactosidase activity was assayed with the substrate *o*-nitrophenyl-β-D-galactopyranoside (ONPG) as described previously (57).

Gel retardation assays. Extracts prepared for CAT assays were used for subsequent gel shift analyses. Egr-1 binding was assayed as described in Cao et al. (6) with the synthetic oligonucleotide EBS-1 (5′-CGCCCTCGCCCCCGCGCCGG G-3′) (6) labeled with Klenow enzyme and [α-^{32}P]CTP. To assay binding of GAL4–Egr-1 chimeras, the 140-bp *Hin*dIII-*Xba*I fragment from pGAL4$_5$tkCAT, containing five copies of the 17-bp GAL4 binding site, was isolated on 3% NuSieve–1% SeaPlaque agarose and purified with Mermaid (Bio 101). The fragment was labeled with [α-^{32}P]CTP and Klenow enzyme, and unincorporated nucleotides were removed with Stratagene's Nuctrap column. Complexes were formed by incubating probe with 10 μl of normalized extract in 20 mM Tris-HCl–80 mM NaCl–1 mM dithiothreitol with 1 μg poly(dI-dC) (Pharmacia 27-7880) and 10 μg of bovine serum albumin (Calbiochem 12659) in a total of 20 μl for 30 min at room temperature. For cold competition experiments, a 50-fold molar excess of unlabeled pGAL4$_5$tkCAT *Hin*dIII–*Xba*I fragment was added.

Western immunoblot analysis. Samples of transiently transfected cells were prepared by freeze-thaw lysis and were normalized both for β-galactosidase activity (as an internal control for transfection efficiency) and for total protein, using the Bio-Rad microprotein assay. Protein samples were separated by sodium dodecyl sulfate (SDS)–10% polyacrylamide gel electrophoresis (PAGE) (71) and transferred to polyvinylidene difluoride membranes (Millipore). For the rabbit polyclonal anti-Egr-1 antiserum, a 1:4,000 dilution of R5232-2 (6) and a 1:5,000 dilution of horseradish peroxidase donkey anti-rabbit immunoglobulin (Amersham NA 934) were used. Monoclonal antibody CH28-2, a gift from L. Pereira, reacts against a short region of human CMV glycoprotein A. CH28-2 was used at a 1:20,000 dilution, and the horseradish peroxidase sheep anti-mouse immunoglobulin (Amersham NA 931) was used at a 1:3,000 dilution. Analysis using Amersham's enhanced chemiluminescence Western procedure (Amersham RPN 2106) required 5- to 50-min exposures. Autoradiograms were scanned with a Hewlett-Packard Scan Jet Plus Scanner, and protein bands were analyzed by calculating the ratio of the band of interest to full-length tagged Egr-1 protein, pCB.Egr-1.tag, using ImageQuant software.

Indirect immunofluorescence microscopy. NIH 3T3 cells were plated on Permanox chamber slides (Lab-Tek 177429) and transfected as described above. Cells were fixed in 4% formalin in 1× phosphate-buffered saline (PBS) (10 min, room temperature), permeabilized with acetone (7 min, −20°C), and blocked in diluted normal goat serum (1 h, room temperature). Cells were further incubated with a 1:500 dilution of anti-Egr-1 rabbit polyclonal antibody R5232-T (1 h, 37°C) and a 1:200 dilution of fluorescein-conjugated goat anti-rabbit antiserum (1 h, room temperature, in the dark) from Caltag Laboratories (M30301). Coverslips were mounted with antifade mounting medium as described by Adams and Pringle (2). Photographs were taken at ×40 magnification with Hypertech film.

Nuclear/cytoplasmic fractionation. Two transiently transfected 100-mm-diameter plates of NIH 3T3 cells were pooled for this analysis. Cells were lysed on ice in 100 μl of hypotonic buffer (25 mM Tris-HCl [pH 7.4], 1 mM $MgCl_2$, 5 mM KCl, 1 mM phenylmethylsulfonyl fluoride) with 0.5% Nonidet P-40 (NP-40) by pipetting briefly up and down; 2 M sucrose was immediately added to a final concentration of 0.25 M, and nuclei were pelleted at $1,000 \times g$ in a microcentrifuge for 1 to 2 min at 4°C. Nuclei were washed in hypotonic buffer–0.5% NP-40 before lysis in 100 μl of hypotonic buffer–1% SDS. After addition of 2× protein sample buffer, nuclear fractions were sonicated.

β-Galactosidase histochemical analysis. NIH 3T3 cells tran-

siently transfected with 2 μg of the Egr-1–β-galactosidase expression vector and 23 μg of pCB6+ vector were stained by a variation of the procedure described by Sanes et al. (61). After being washed with 1× PBS at room temperature, the cells were fixed with 4% paraformaldehyde in PBS for 20 min at 4°C, rinsed briefly with 1× PBS, and incubated with PBS for 30 min at 4°C. Cells were stained with a 0.4-mg/ml concentration of 5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside (X-Gal) in 4 mM $K_3Fe(CN)_6$–4 mM $K_4Fe(CN)_6$–2 mM $MgCl_2$. Cells were typically stained for 1 to 2 h; however, expression of some constructs required staining for up to 24 h. To stop the reactions, cells were rinsed with 1× PBS and stored in 30% sucrose in PBS at 4°C.

RESULTS

General strategy. Wild-type Egr-1 protein is known to be localized to the nucleus (6, 12, 73), bind specifically to the site CGCCCCCGC (6, 8, 36), and activate transcription through this sequence in transient cotransfection assays (36, 53). We have mutagenized the 533-amino-acid coding region of the murine *Egr-1* cDNA in a CMV expression vector in order to determine the functional domains of Egr-1. Site-directed mutagenesis generated a series of in-frame internal deletions of 43 to 48 amino acids. To accurately determine protein levels of various constructs, the parent expression plasmid was modified to include a 20-amino-acid antigenic tag as well as a unique *Xho*I site at the N terminus. This construct was tested for protein expression, nuclear localization, and the ability to transactivate and was equivalent to wild-type Egr-1 in all assays. Construction of C-terminal deletions utilized oligonucleotide-directed mutagenesis to insert a stop codon followed by an *Xho*I site at six positions spanning the length of the coding sequence. N-terminal deletions were created by excision of the internal *Xho*I fragment from the corresponding C-terminal deletion and religation in frame.

In general, the approach that we have taken is initially to define domains important for transcriptional activity, DNA binding, and nuclear localization by looking for loss of function with deletion derivatives of Egr-1. A constructive approach, in which segments of Egr-1 are fused to a heterologous DNA-binding domain and assayed for transcriptional activity or to the bacterial protein β-galactosidase and assayed for nuclear localization, has substantiated the results of deletion analysis and refined the map of Egr-1's functional domains.

Egr-1 functions as a transcriptional activator. Egr-1 has been shown to bind to several related sequences in its own 5′ flanking region (6, 8). To assay transcriptional activity in transient cotransfections, a CAT reporter was constructed to include three copies of a high-affinity Egr-1 binding site, CGCCCCCGC, in front of the minimal murine c-*fos* promoter (Fig. 2A). The Egr-1 expression vector pCB.Egr-1 contains the full-length murine *Egr-1* cDNA under control of the CMV early promoter/enhancer. Increasing amounts of Egr-1 expression vector were cotransfected with CAT reporter and an internal control for transfection efficiency, CMV.β-galactosidase. Figure 2B shows that Egr-1 activates transcription of the synthetic reporter containing three Egr-1 binding sites 10-fold in NIH 3T3 cells. Transactivation is absolutely dependent on the presence of Egr-1 binding sites; Egr-1 has no effect on the minimal *fos* promoter alone (Fig. 2C). As a negative control, an Egr-1 derivative deleted for part of zinc fingers 1 and 2 has been included; this mutant, pCB.Egr-1Δ331-374, does not activate transcription. Similar data obtained for the HeLa cell line show that the ability of Egr-1 to transactivate is not cell type specific (see below).

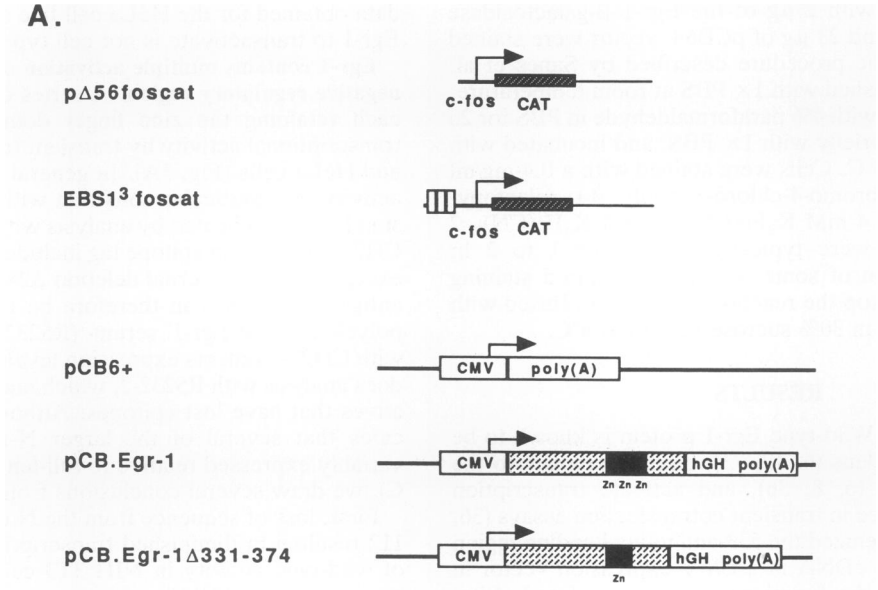
Egr-1 contains multiple activation domains and a potential negative regulatory region. A series of deletion derivatives, each retaining the zinc finger domain, was assayed for transcriptional activity by transient transfection in NIH 3T3 and HeLa cells (Fig. 3A). In general, we interpret the CAT activity of a particular construct with respect to its expression level as indicated by analysis with monoclonal antibody CH28-2 against an epitope tag included in the construct. The exception is the internal deletion Δ284–330, which lacks the antigenic tag and can therefore be recognized only with a polyclonal anti-Egr-1 serum (R5232-2). Western analysis with CH28-2 reflects expression levels more accurately than does analysis with R5232-2, which may underestimate derivatives that have lost epitopes. Although densitometry indicates that several of the larger N-terminal deletions are variably expressed relative to full-length Egr-1 (Fig. 3B and C), we draw several conclusions from these data.

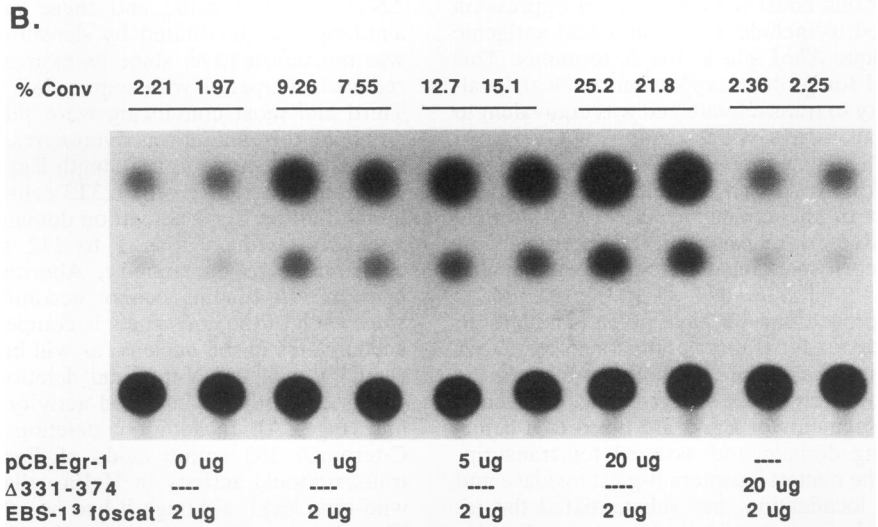
First, loss of sequence from the N terminus to amino acid 112 resulted in diminished transcriptional activity to 51.1% of wild-type activity in NIH 3T3 cells (Fig. 3A), although mutant protein ΔN112 was expressed as well as full-length Egr-1 was expressed (Fig. 3B). Second, deletion ΔN214 had only 14.7% activity, compared with 38.1% activity for ΔN138, in 3T3 cells, and these mutants are expressed similarly, as quantitated by densitometry. Mutant ΔN240 was not informative, since its expression was dramatically reduced compared with expression of full-length protein. Third and most convincing were the results with ΔN314. Although this mutant was overexpressed substantially (four- to fivefold) relative to full-length Egr-1, it activated to only 17.2% of wild-type levels in 3T3 cells. In sum, these results imply that the Egr-1 activation domain maps to amino acids 1 to 314, with regions 1 to 112 and 138 to 214 both contributing to the activity. Aberrant localization or impaired DNA binding cannot account for loss of activity, since each of the constructs is competent to bind DNA and accumulates in the nucleus, as will be discussed below.

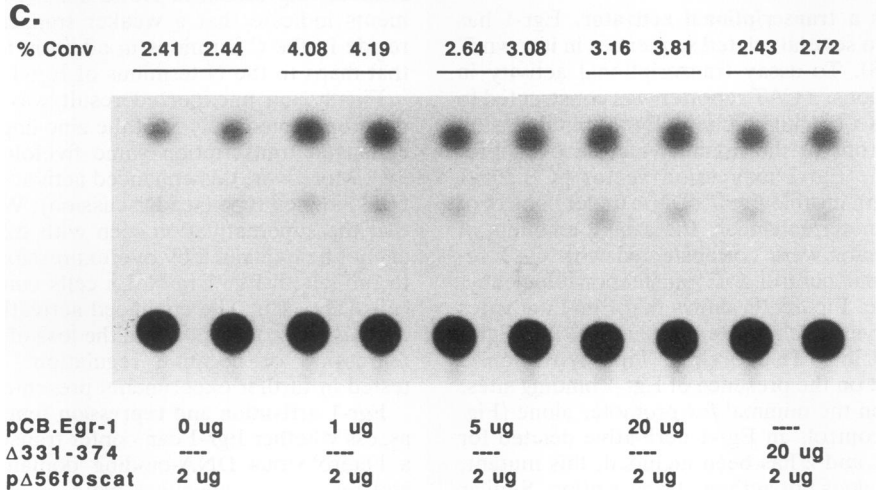
Analysis of the N-terminal deletions in HeLa cells was consistent with the impaired activity seen in the NIH 3T3 line (Fig. 3A). In addition, deletion ΔC430, removing the C-terminal 103 amino acids of Egr-1, revealed reduced transcriptional activity in HeLa cells to about 20% of the wild-type level, although it had no effect in NIH 3T3 cells. This decrease in transcriptional activity was not due to reduced expression in HeLa cells (Fig. 3C). These experiments indicate that a weaker transcriptional activity may reside in the C terminus in addition to the strong activator that maps to the N terminus of Egr-1.

Finally, an unexpected result was that a small internal deletion immediately 5′ of the zinc finger domain (Δ284–330) enhanced transcription some fivefold in HeLa cells (Fig. 3A). Moreover, this enhanced activation was not seen in the NIH 3T3 cell type (see Discussion). Western analysis shows that the superactivation seen with Δ284–330 in HeLa cells cannot be explained by overexpression of Δ284–330 relative to full-length Egr-1 in HeLa cells compared with NIH 3T3 cells (Fig. 3D). The enhanced activation seen with deletion Δ284–330 is consistent with the loss of a region important for repression or negative regulation. This hypothesis was tested in further experiments presented below.

Egr-1 activation and repression functions are portable. To assess whether Egr-1 can confer transcriptional activities on a heterologous DNA-binding domain, putative activation and repression sequences were fused to the DNA-binding


A.
pΔ56foscat
c-fos CAT
EBS1[3] foscat
c-fos CAT
pCB6+
CMV poly(A)
pCB.Egr-1
CMV hGH poly(A)
Zn Zn Zn
pCB.Egr-1Δ331-374
CMV hGH poly(A)
Zn


B.
% Conv 2.21 1.97 9.26 7.55 12.7 15.1 25.2 21.8 2.36 2.25
pCB.Egr-1 0 ug 1 ug 5 ug 20 ug ----
Δ331-374 ---- ---- ---- ---- 20 ug
EBS-1[3] foscat 2 ug 2 ug 2 ug 2 ug 2 ug


C.
% Conv 2.41 2.44 4.08 4.19 2.64 3.08 3.16 3.81 2.43 2.72
pCB.Egr-1 0 ug 1 ug 5 ug 20 ug ----
Δ331-374 ---- ---- ---- ---- 20 ug
pΔ56foscat 2 ug 2 ug 2 ug 2 ug 2 ug

domain of the yeast transcription factor GAL4. Regions of Egr-1 were amplified by PCR and cloned into expression vector pSG424 (60), encoding GAL4 amino acids 1 to 147 (Fig. 4A). Resulting chimeras were tested for the ability to transactivate a reporter containing five GAL4 binding sites in front of the E1b minimal promoter (Fig. 4B, left). Egr-1 amino acids 3 to 281 function to activate transcription about 100-fold in this assay; when subdivided into amino acids 3 to 138 or 138 to 281, these segments function as well as does the intact domain. In addition, transcription is activated some fivefold by the C-terminal region 420 to 533. In all cases, activation is dependent on the presence of the GAL4 binding sites (Fig. 4B, right). Results of gel shift assays with the transiently transfected 3T3 cell extracts ensure that all of the fusion proteins are expressed at comparable levels and are competent for DNA binding (Fig. 4C). For example, the weak activity of GAL4–Egr-1(420-533) is not a result of low protein levels, since this chimera is expressed as well as is GAL4–Egr-1(3-281). Also, GAL4–Egr-1(138-281) is overexpressed compared with GAL4–Egr-1(3-281), accounting for its superlative transactivation function.

To determine whether Egr-1 encodes a repressor function, we constructed a fusion of GAL4(1-147) and the region of Egr-1 which, when deleted, enhanced activation. A GAL4 reporter with a high basal activity suitable for examining transcriptional repression was chosen (pGAL4$_5$tkCAT). In this assay, the GAL4–Egr-1(240-330) chimera repressed transcription 10-fold in a GAL4 binding site-dependent manner (Fig. 5A). A fusion containing only 34 amino acids of Egr-1, GAL4–Egr-1(281–314), repressed transcription similarly. Gel shift analyses (Fig. 5B) indicate that this repression is not due to overexpression of the GAL4–Egr-1 fusion proteins. Titration of the chimeric repressor (Fig. 5C) shows that as little as 1 μg of GAL4–Egr-1(281-314) represses transcription more than fivefold; even 10 μg of the GAL4 DNA-binding domain does not recapitulate this effect. Therefore, in contrast to an extensive redundant activation domain, Egr-1 contains a repression function that can be precisely localized.

The DNA-binding activity of Egr-1 maps to the zinc finger domain. By homology to other zinc finger proteins such as TFIIIA (63) and Sp1 (30), the DNA-binding activity of Egr-1 was expected to reside in the zinc finger domain. Gel mobility shift assays with extracts from HeLa cells transiently transfected with the series of internal deletion derivatives show that only amino acids 331 to 419 of Egr-1 (encoding the three zinc fingers) are required for specific DNA binding (Fig. 6A). The deletion immediately N terminal of the zinc finger domain (eight amino acids 5′ of the first cysteine) has no effect on DNA binding. The deletion C terminal of the zinc fingers (four amino acids 3′ of the last histidine) may slightly impair DNA binding. Western analysis with polyclonal anti-Egr-1 antisera indicates that the loss of DNA-binding activity with deletions within the zinc finger domain is not due to a reduction in protein expression (Fig. 6B).

The bipartite nuclear localization signal (NLS) of Egr-1 consists of a basic sequence and a region of the DNA-binding domain. Earlier work has shown that the Egr-1 gene product is localized to the nucleus (6, 12, 73). Several independent methods contributed to our delineation of the nuclear localization signal of Egr-1: subcellular fractionation and immunofluorescence of cells transfected with Egr-1 deletion derivatives, as well as in situ staining of Egr-1–β-galactosidase fusion proteins. Nuclear/cytoplasmic fractionation and Western analysis of C-terminal deletions of Egr-1 revealed that while ΔC430 remained nuclear, further deletion in from the C terminus produced derivatives that were predominantly cytoplasmic, for example, ΔC314 (Fig. 7). These results were corroborated by indirect immunofluorescence microscopy. Deletion from the N terminus to 314 or from the C terminus to 430 produced derivatives that were nuclear (Fig. 8A, panels a to d). In contrast, C-terminal deletions past 430 were expressed throughout the cell (Fig. 8A, panels e to f). From these analyses, amino acids 315 to 430 appeared essential for proper nuclear localization. This region includes the three zinc fingers and adjacent sequences. Both the 5′ flanking sequence Lys-315 Pro Ser Arg Met Arg Lys Tyr Pro Asn Arg Pro Ser Lys Thr Pro-330 and the 3′ flanking sequence Lys-420 Asp Lys Lys Ala Asp Lys Ser Val Val-429 are characterized by a preponderance of basic residues. Immunofluorescence of internal deletion derivatives showed that each was properly targeted to the nucleus, suggesting that no single signal in Egr-1 directs nuclear accumulation (Fig. 8B and data not shown).

To further delimit the residues within region 315 to 429 required for nuclear localization, segments of Egr-1 were fused to the large bacterial protein β-galactosidase (Fig. 9). The CMV promoter and a synthetic ATG drive expression of β-galactosidase coding sequence beginning at codon 10. This protein retains its enzymatic activity, and staining of transiently transfected 3T3 cells with X-Gal show it to be distributed throughout the nucleus and cytoplasm as seen by others with similar constructs (31) (Fig. 10A). However, when the extensive finger domain with both 5′ and 3′ basic flanking sequences (aa 315 to 429) was fused N terminal to β-galactosidase, the resulting chimera was found exclusively in the nucleus of transfected cells (Fig. 10B). This result indicates that the NLS of Egr-1 can function to confer nuclear localization on a heterologous bacterial protein. Because the C terminus of Egr-1 had never been assayed for nuclear localization in the absence of amino acids 315 to 430 and might contain an additional NLS, pCB.420-533.β-GAL was constructed. Subsequent analysis showed that the C terminus of Egr-1 does not contain a redundant nuclear localization signal (data not shown). Further analysis of residues 315 to 429 demonstrated that while the zinc finger domain alone or in conjunction with the 3′ basic sequence was not sufficient for nuclear targeting (Fig. 10C and data not shown), the finger domain with the adjacent 5′ basic region localized β-galactosidase precisely to the nucleus (Fig. 10D). Yet the 5′ basic sequence alone directs only partial nuclear

FIG. 2. Evidence that Egr-1 functions as a transcriptional activator. (A) CAT reporters and expression vectors for cotransfection assays. The reporter pΔ56foscat contains 56 nucleotides of 5′ sequence of the murine c-*fos* promoter. EBS1^{3}foscat includes three high-affinity Egr-1 binding sites (CGCCCCCGC) in front of the c-*fos* TATA box. The expression vector pCB.Egr-1 consists of the full-length murine *Egr-1* cDNA driven by the CMV early promoter/enhancer. pCB.Egr-1Δ331-374 contains an in-frame deletion removing zinc finger 1 and part of finger 2. hGH, human growth hormone. (B and C) Titration of wild type Egr-1 transactivation. NIH 3T3 cells were transiently transfected with 0, 1, 5, or 20 μg of pCB.Egr-1 or 20 μg of Δ331–374 and 2 μg of EBS1^{3}foscat reporter (B) or 2 μg pΔ56foscat (C). CAT activity expressed as percent conversion is shown at the top.

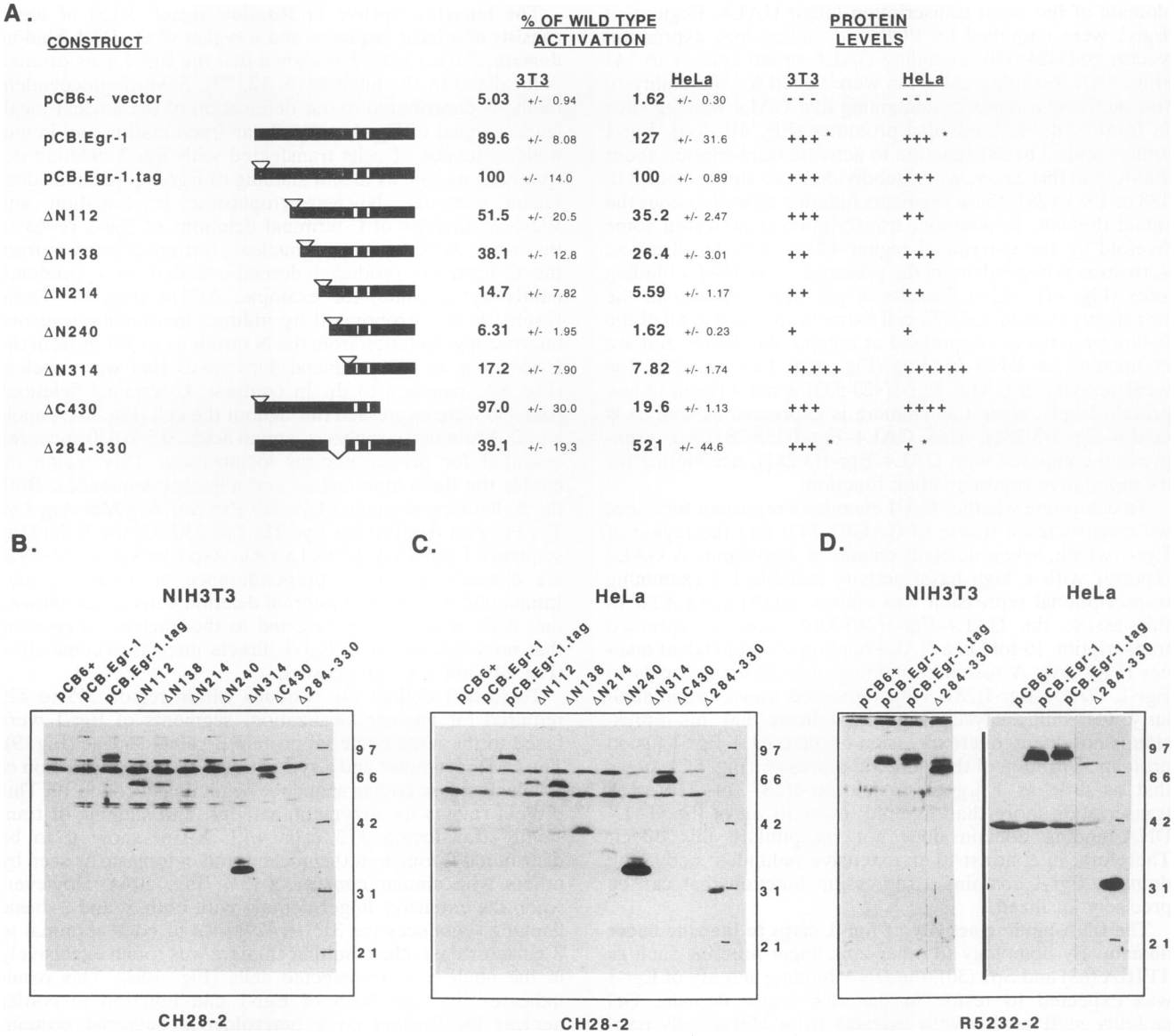


FIG. 3. Mutational analysis of Egr-1. (A) Transcriptional activity and expression levels of Egr-1 deletion derivatives. Each construct retains the zinc finger domain (depicted as three black bars), and all but the internal deletion Δ284–330 have 20 exogenous amino acids inserted in frame at the N terminus. Transient cotransfections were done in NIH 3T3 or HeLa cells with 2 μg of EBS1[3]foscat reporter and 20 μg of the indicated expression vector. Forty-eight hours after transfection, extracts were prepared and assayed for both CAT activity and protein levels. Percent of wild-type activation is the CAT activity of a given construct relative to that of pCB.Egr-1.tag. In each case, protein levels were determined by Western analysis with monoclonal antibody CH28-2 except for the wild-type protein pCB.Egr-1 and the internal deletion pCB.Egr-1Δ284-330, which lack the epitope tag recognized by CH28-2. These constructs, analyzed with R5232-2, are designated with asterisks. Densitometry of transfected NIH 3T3 cell extracts gave the following ratios of mutant protein relative to pCB.Egr-1.tag: pCB.Egr-1.tag, 1.00; ΔN112, 1.46; ΔN138, 0.59; ΔN214, 0.70; ΔN240, 0.10; ΔN314, 4.46; and ΔC430, 0.87. Densitometric quantification of mutant proteins relative to pCB.Egr-1.tag in HeLa cells gave the following values: pCB.Egr-1.tag, 1.00; ΔN112, 0.57; ΔN138, 1.55; ΔN214, 1.31; ΔN240, 0.15; ΔN314, 4.97; and ΔC430, 1.85. The expression levels of wild-type Egr-1 and pCB.Egr-1Δ284-330, which lack the epitope tag recognized by antibody CH28-2, were assessed with the anti-Egr-1 polyclonal antiserum R5232-2 as shown in panel D. (B) Western analysis of transiently transfected NIH 3T3 cells with CH28-2. (C) Western blotting of transiently transfected HeLa cells with monoclonal antibody CH28-2. (D) Western analysis of the internal deletion Δ284–330 with anti-Egr-1 polyclonal antiserum R5232-2 in both NIH 3T3 and HeLa cells. Sizes are indicated in kilodaltons.

accumulation with some residual staining in the cytoplasm, implying that the zinc fingers themselves participate in nuclear localization (data not shown).

To define the region of the DNA-binding domain that cooperates with the 5′ basic flanking sequence, chimeras with the 5′ basic region (amino acids 315 to 330) and finger 1, fingers 2 and 3, finger 2, or finger 3 were constructed. While the 5′ basic flanking sequence in combination with finger 3 (Fig. 10F) or to a lesser extent finger 2 suffices for nuclear accumulation, finger 1 is unable to provide this function (Fig. 10E). The primary amino acid sequences of fingers 1 and 3, which bind to the same three-nucleotide subsite, are quite

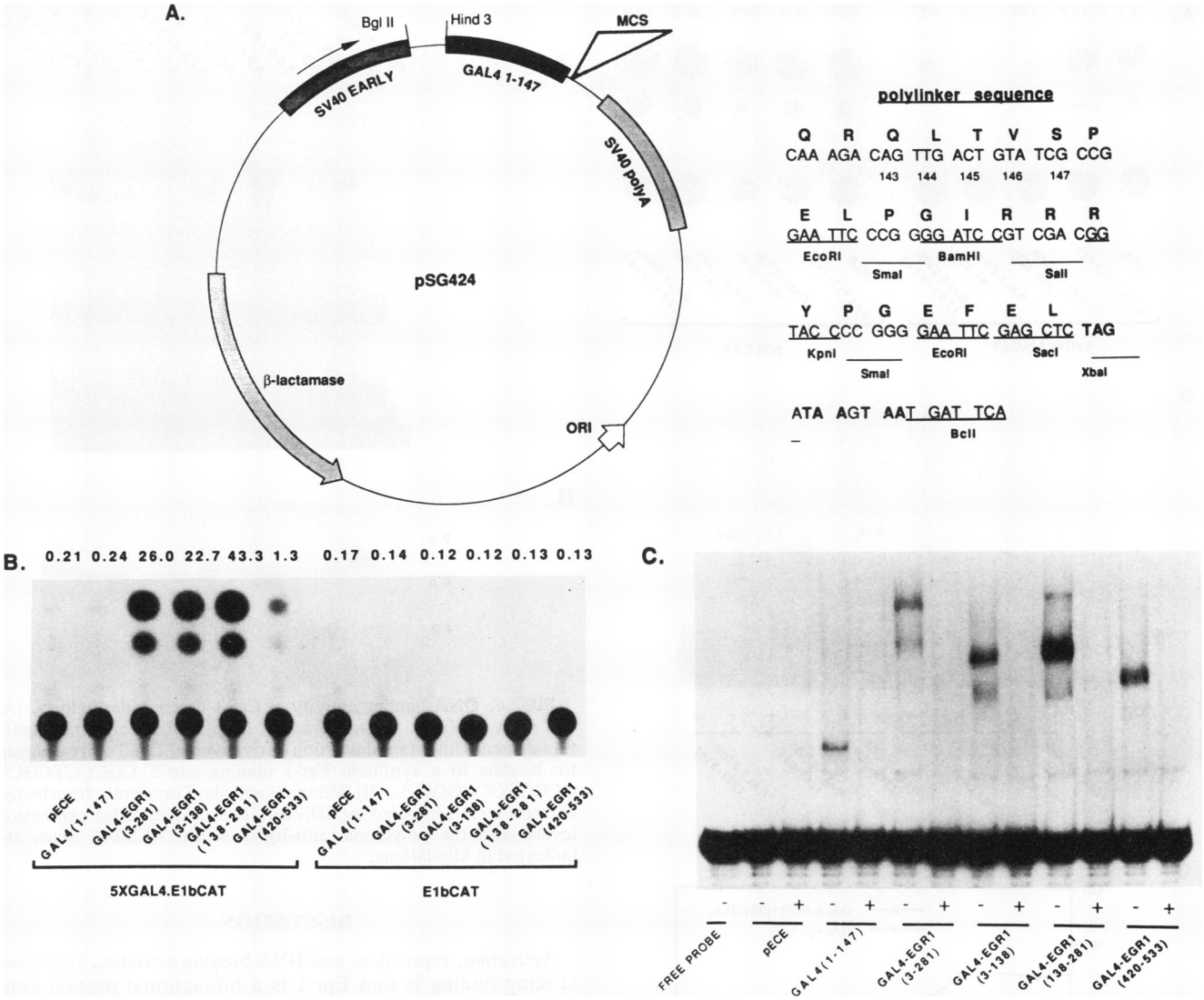


FIG. 4. Activation by GAL4–Egr-1 chimeras. (A) pSG424 expression plasmid. GAL4 amino acids 1 to 147 encoding the DNA-binding domain and NLS are under control of the SV40 early promoter/enhancer in vector pECE. MCS, multiple cloning site; ORI, origin of replication. (B) Transactivation by GAL4–Egr-1 chimeras. NIH 3T3 cells were transiently transfected with 20 μg of the indicated expression vector and 2 μg of reporter plasmid. The left portion of the panel shows transfections with the reporter 5XGAL4.E1bCAT, which contains five copies of the 17-bp GAL4 binding site in front of the minimal E1b promoter and CAT; the right portion depicts transfections with the reporter E1bCAT, which is a minimal E1b promoter construct lacking GAL4 binding sites. CAT activity is expressed as percent conversion for each construct. (C) DNA-binding activity of GAL4–Egr-1 chimeras. Extracts used for CAT assays were tested by gel shift analysis for the ability to bind a 140-bp probe containing five GAL4 sites. Protein-DNA complexes were formed in the presence (+) or absence (−) of a 50-fold molar excess of cold competitor DNA.

similar (Fig. 11A). Principal differences lie in the H-C link region (a set of seven amino acids between the histidine of one finger and the cysteine of the following finger extremely well conserved among Cys_2His_2 zinc finger proteins) and in several basic residues preceding the first histidine in finger 3. These basic residues, which are also absent in finger 2, may account for the enhanced nuclear accumulation of constructs retaining finger 3 versus finger 2 but seem unlikely to explain the striking difference between the localization of fusions with finger 2 versus finger 1. Although the H-C links preceding fingers 2 and 3 conform well to the consensus, the H-C link of finger 1 does not. We examined whether the most important nuclear determinants lie in the H-C link or within the body of finger 3 by constructing chimeras between fingers 1 and 3 (in conjunction with the 5′ basic sequence). Staining with X-Gal shows a construct with the 5′ basic region, the H-C link of finger 3, and the body of finger 1 to be cytoplasmic, while a chimera containing the 5′ basic region, the H-C link of finger 1, and the body of finger 3 accumulates somewhat in the nucleus (data not shown). From this experiment, we conclude that although the H-C link of finger 3 contributes marginally to nuclear localization, the most important nuclear determinants in finger 3 lie in the body of the finger.

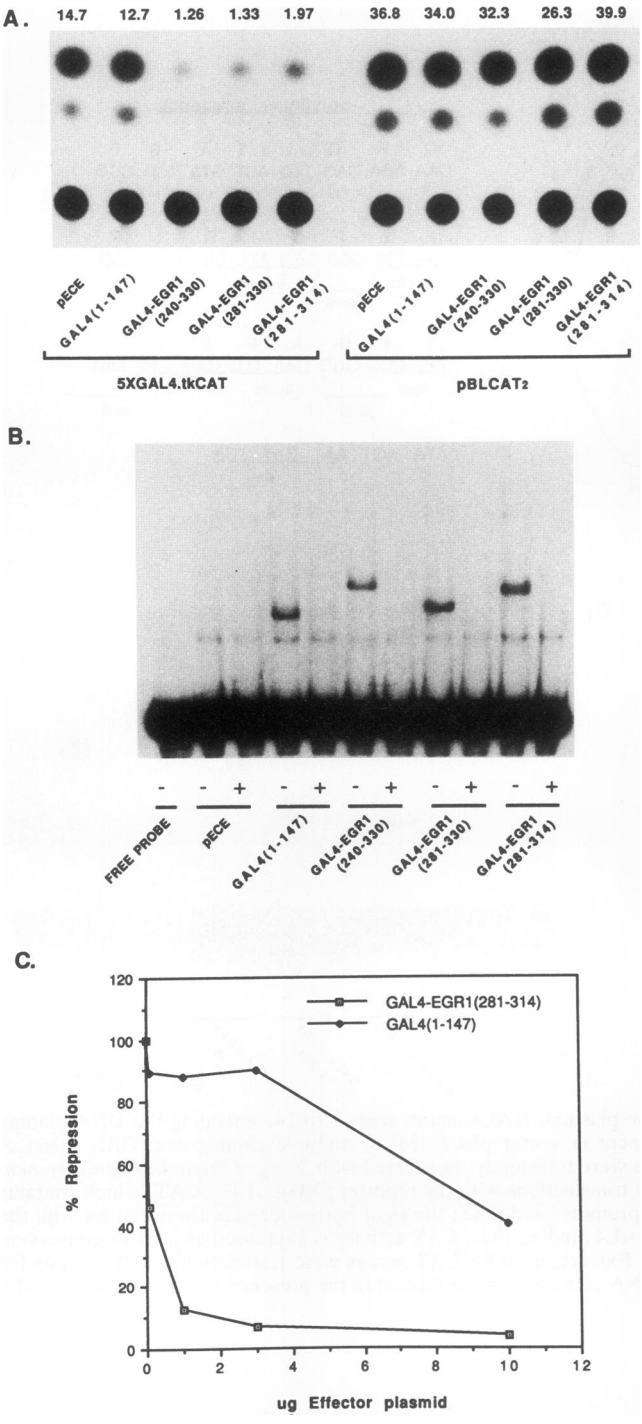


FIG. 5. Repression by GAL4–Egr-1 chimeras. (A) CAT activity of GAL4–Egr-1 fusions. NIH 3T3 cells were transiently transfected with 2 μg of the specified expression vector and 10 μg of reporter plasmid. Each of the chimeras contains GAL4 amino acids 1 to 147 under control of the SV40 early promoter/enhancer in vector pECE fused in frame to the designated residues of Egr-1. CAT activity is given as percent conversion above each construct. $pBLCAT_2$ contains sequence from −105 to +51 of the herpes simplex virus thymidine kinase gene. 5XGAL4.tkCAT is a $pBLCAT_2$-based plasmid with five copies of the 17-bp GAL4 binding site. (B) DNA-binding activity of GAL4–Egr-1 chimeras. Gel shift analysis was performed as for Fig. 4. (C) Titration of repression by GAL4–Egr-1(281-314) and GAL4(1-147). Each transient cotransfection includes 2 μg of 5XGAL4.tkCAT reporter and 100 ng, 1 μg, 3 μg, or 10 μg of effector plasmid. Percent repression is the CAT activity of reporter with the specified amount of effector plasmid relative to the activity of the reporter with control vector pECE.

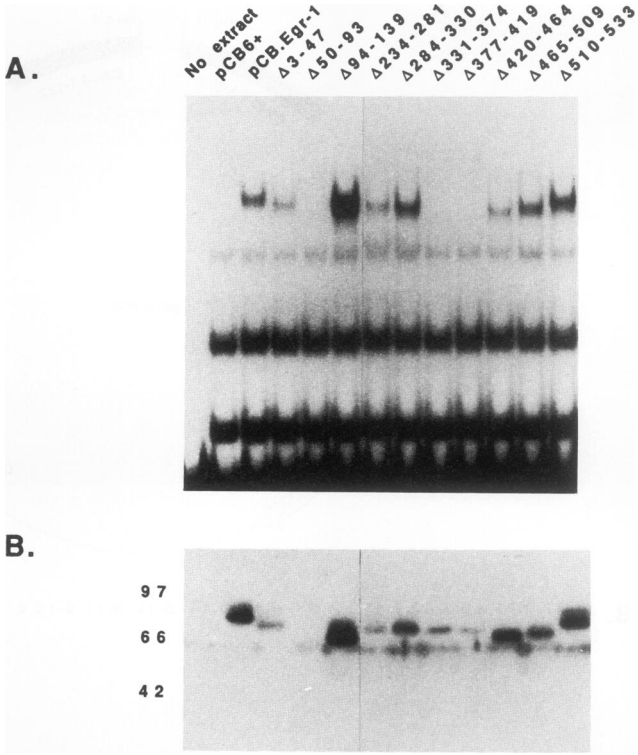


FIG. 6. DNA-binding activity of Egr-1 deletion derivatives. (A) Gel shift analysis. Extracts made from HeLa cells transiently transfected with internal deletion derivatives of Egr-1 were assayed for binding to a synthetic Egr-1 binding site 5′-CGCCCTCGCC CCCGCGCCGGG-3′. (B) Western analysis. Transiently transfected extracts were separated by SDS-PAGE and assayed for expression levels with the polyclonal anti-Egr-1 serum R5232-2. Sizes are indicated in kilodaltons.

DISCUSSION

Activation, repression, and DNA-binding activities. Our most striking finding is that Egr-1 is a bifunctional protein containing domains that can both activate and repress transcription in transient-transfection assays. This immediate-early transcription factor possesses a robust serine/threonine-rich N-terminal activation domain and a novel repression domain distinct from the alanine-rich sequence shown to be responsible for repression in the *Drosophila* protein Krüppel (38).

Activation domains, and more recently repression domains, have been demonstrated to function as independent, modular components of transcription factors. Activation domains are not typified by a single consensus sequence but instead fall into several discrete classes; for example, acidic domains in GAL4 (41), GCN4 (26), VP16 (60), and GATA-1 (43), glutamine-rich stretches in Sp1 (11) and Oct-2/OTF2 (15, 48), proline-rich sequences in CTF/NF-1 (44), and serine/threonine-rich regions in Pit-1/GHF-1 (70) all function to activate transcription. The activation domains of Fos and Jun are rich in both acidic and proline residues (1, 4), and for other activators like the CCAAT/enhancer-binding protein

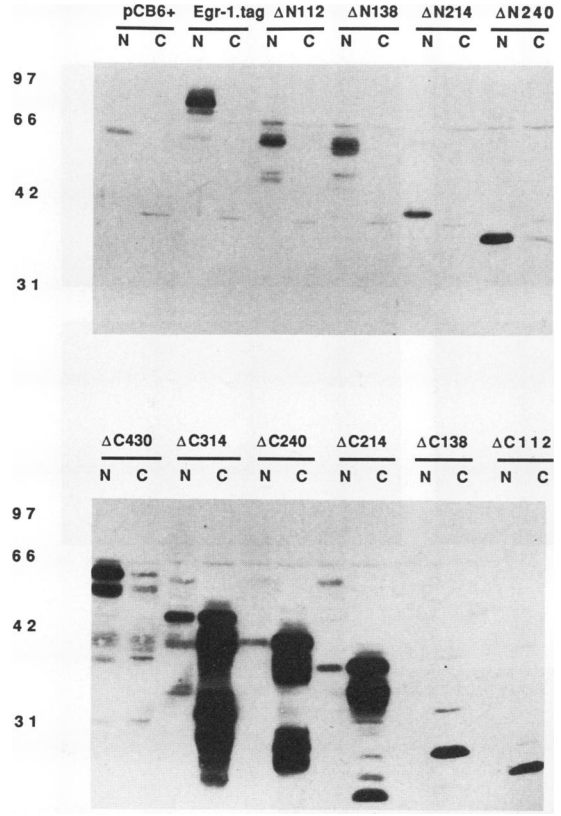


FIG. 7. Nuclear/cytoplasmic fractionation of fibroblasts transfected with Egr-1 deletion constructs. Transiently transfected NIH 3T3 cells were lysed in hypotonic buffer with 0.5% NP-40 by pipetting briefly up and down. Sucrose was added to 0.25 M to stabilize the nuclei, and the nuclear and cytoplasmic fractions were separated by pelleting briefly in a microcentrifuge at 4°C. Nuclei were washed in hypotonic buffer–0.5% NP-40 before lysis in hypotonic buffer–1% SDS. Samples were subjected to SDS-PAGE and analyzed by Western blotting with antiserum R5232-2. Sizes are indicated in kilodaltons.

C/EBP (13), no evident sequence motif has emerged. We have shown that fusions of the GAL4 DNA-binding domain and residues 3 to 281 of Egr-1 activate transcription some 100-fold. This N-terminal domain is 30% serine/threonine/tyrosine rich over a span of ~180 residues (amino acids 60 to 240) and includes several tracts of 5 to 7 consecutive serine or threonine residues. The large size of this activation domain may contribute to its potency relative to the smaller, previously described serine/threonine-rich activator Pit-1 (70). Moreover, this transactivation domain is impervious to mutation in that substantial deletions in the extensive N-terminal domain do not destroy transcriptional activity. It has been suggested that serine/threonine-rich domains may be phosphorylated and in this way function as acidic activators (27, 55); in this regard, it is interesting to note that Egr-1 is phosphorylated (7, 12, 73). A second, weaker activation domain mapped to the C terminus of Egr-1, which has octapeptide repeats reminiscent of the phosphorylated YSPTSPS reiterations in the carboxy-terminal domain of the RNA polymerase II large subunit (10).

To date, the only well-characterized repression domain is the alanine-rich sequence in the *Drosophila* gap protein Krüppel (38, 75). Other *Drosophila* proteins such as Even-skipped (3, 24) and Engrailed (24, 29) and mammalian DNA-binding proteins such as Tst-1/SCIP (46), WT1 (42), and YY1/NF-E1/δ (25, 52, 64) repress transcription. Of these, Krüppel, Engrailed, WT1, and YY1/NF-E1/δ have been shown to confer their repression function on a heterologous DNA-binding domain. However, except in the case of Krüppel, the protein sequences responsible have not been precisely delineated. The novel repression domain of Egr-1 represents an extremely compact function of 34 amino acids that has been evolutionarily conserved through zebrafish, the lowest vertebrate in which a homolog of Egr-1 has been identified (12a). The primary sequence of the Egr-1 repressor (Fig. 11A) is not alanine or glycine rich as is the case for Krüppel and suggested for SCIP and YY1 and as such represents a transcriptional motif distinct from those observed previously. We speculate that a site of potential proline-directed phosphorylation contained within this sequence (Thr-285 Pro) might regulate Egr-1's bifunctional transcriptional activity (21, 32).

Several models have been proposed as mechanisms of repression (37). Repressor proteins might act through competitive binding, either at the transcription start site or at the binding site of an upstream activating protein. Alternatively, a repressor might inhibit the activity of an activator or a component of the basal transcription apparatus without affecting binding, by direct protein-protein interactions. By a third DNA binding-independent mechanism, termed squelching, repression results from the titration of limiting factors essential for activation (37). Repression by GAL4–Egr-1 fusions is unlikely to be a result of squelching, since the effect is DNA binding site dependent (Fig. 5A). We believe that the mechanism may involve an interaction with a component of the basal transcription machinery since GAL4–Egr-1(281-314) represses minimal promoter constructs (74). It is unlikely that repression occurs through displacement of basal factors by competition for binding sites because neither the GAL4 DNA-binding domain alone nor GAL4–Egr-1(420-533) (data not shown) can mimic the repression seen with GAL4–Egr-1(281-314). In sum, these results support the hypothesis that the Egr-1 repression domain works by affecting a function other than DNA binding of a basal transcription factor.

Some cell-type-specific differences in the ability of Egr-1 to activate have been observed with constructs ΔC430 and Δ284–330. Importantly, Western analysis of transfected 3T3 and HeLa cells shows that differences in CAT activity between the two cell lines are not due to differential protein stability; construct ΔC430, which displays reduced activity in HeLa cells although not in 3T3 cells, is not expressed at a lower ratio to full-length Egr-1 in HeLa cells compared with 3T3 cells (Fig. 3B and C). Second, the dramatic enhancement in transcriptional activity seen with Δ284–330 in HeLa but not 3T3 cells cannot be explained by overexpression of this construct compared with full-length Egr-1 in the HeLa cell line (Fig. 3D). These results suggest that Egr-1 may be modified by, or act in association with, cell-type-specific factors. The fact, however, that the C terminus and the internal region 281 to 314 are competent to activate and repress, respectively, as part of GAL4 fusions in 3T3 cells (Fig. 4B and 5A) implies that the cell type specificity involves components of Egr-1 not incorporated in the GAL4 fusions. For example, the lack of cell type specificity in the GAL4 chimeras could be due to the absence of a site for a cell-specific kinase or the lack of a domain involved in cell-specific protein-protein interactions.

Egr-1 is one of only a small number of factors that contain modular domains capable of regulating transcription both

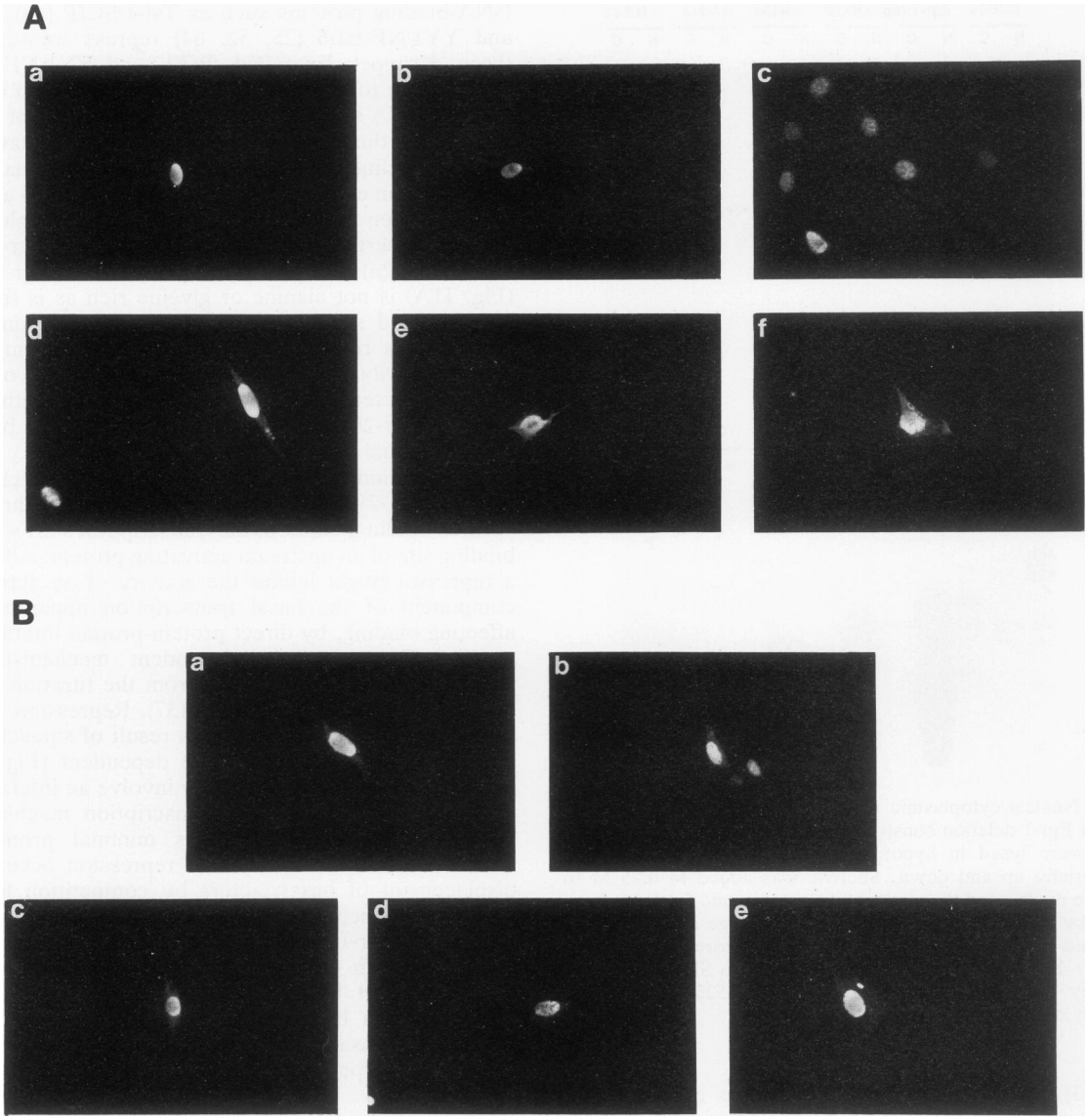


FIG. 8. Immunofluorescence staining for Egr-1 in NIH 3T3 cells transfected with Egr-1 deletion constructs. At 40 to 48 h after transfection, NIH 3T3 cells were fixed, permeabilized, and incubated with anti-Egr-1 polyclonal antibody R5232-2 and a fluorescein-conjugated goat anti-rabbit secondary antibody. (A) N- and C-terminal deletion derivatives. a, pCB.Egr-1.tag; b, pCB.Egr-1ΔN112; c, pCB.Egr-1ΔN314; d, pCB.Egr-1ΔC430; e, pCB.Egr-1ΔC314; f, pCB.Egr-1ΔC112. (B) Internal deletion constructs. a, pCB.Egr-1; b, pCB.Egr-1Δ284-330; c, pCB.Egr-1Δ331-374; d, pCB.Egr-1Δ377-419; e, pCB.Egr-1Δ420-464.

positively and negatively. Other examples include Krüppel (75), YY1/NF-E1/δ (reviewed in reference 20), and the immediate-early factors Fos and Jun (1). The ability to work as either an activator or a repressor may be common to immediate-early transcription factors to allow for versatility of effector functions. Posttranslational modifications or interactions with regulating factors may enable these complex factors to regulate transcription either positively or negatively. Specifically, we can postulate that (i) Egr-1 may alternatively activate or repress transcription, depending on whether it is induced in response to positive growth signals or differentiation cues, or (ii) Egr-1 may activate and repress different target genes in response to a single inducing agent in order to fulfill several functions in a specific cell type. For example, in the case of HL60 cell differentiation with TPA, Egr-1 promotes macrophage differentiation while preventing granulocytic differentiation. In particular, we can speculate that Egr-1 positively regulates genes essential for macrophage differentiation while negatively regulating genes required for specialized granulocytic functions.

As predicted by homology to other members of the class of Cys_2His_2 zinc finger proteins, the DNA-binding activity of Egr-1 resides in the zinc finger domain. Our results are consistent with the work of Pavletich and Pabo (54) showing that an Egr-1 zinc finger domain fragment purified from bacteria (amino acids 333 to 421) binds DNA specifically. We have extended these results by examining regions outside the zinc finger domain for their effects on DNA binding. We have no evidence for a second domain that may modulate the affinity of DNA binding, as is the case for Sp1 (30).

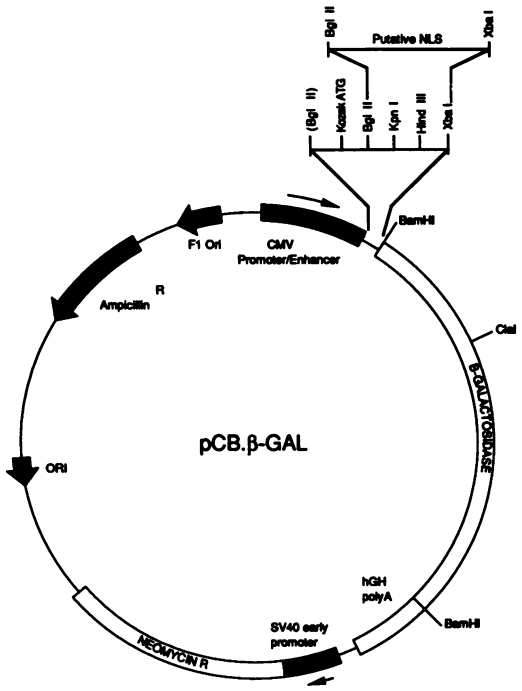
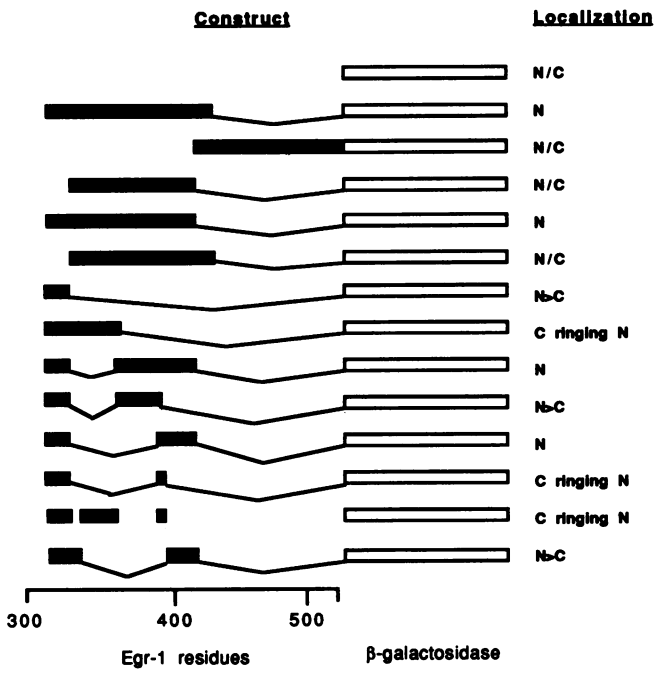


FIG. 9. Egr-1–β-galactosidase fusion proteins. (A) Construction of β-galactosidase fusions. pCB.β-GAL contains β-galactosidase coding sequence (lacking the first nine amino acids) expressed from the CMV early promoter/enhancer and a synthetic ATG. Unique cloning sites permit N-terminal fusions to codon 10 of β-galactosidase. (B) Localization of Egr-1–β-galactosidase fusions. Distribution of chimeras was determined by staining transiently transfected NIH 3T3 cells with X-Gal. N, nuclear; C, cytoplasmic.

B.

Name of construct	Region of Egr-1
pCB.β-GAL	none
pCB.315-429.β-GAL	5'basic/finger domain/3'basic
pCB.420-533.β-GAL	C-terminus
pCB.331-419.β-GAL	finger domain
pCB.315-419.β-GAL	5'basic/finger domain
pCB.331-429.β-GAL	finger domain/3'basic
pCB.315-330.β-GAL	5'basic
pCB.315-367.β-GAL	5'basic/finger1
pCB.315-330/361-419.β-GAL	5'basic/fingers 2 and 3
pCB.315-330/361-392.β-GAL	5'basic/finger 2
pCB.315-330/389-419.β-GAL	5'basic/finger 3
pCB.315-330/389-395.β-GAL	5'basic/HClink3
pCB.315-330/389-395/338-363.β-GAL	5'basic/HClink3/finger1
pCB.315-337/396-419.β-GAL	5'basic/HClink1/finger3



Nuclear localization. NLSs are generally short stretches of 8 to 10 amino acids characterized by basic residues as well as proline. NLS sequences are retained in the mature protein, may be found at any position as long as it is exposed on the protein surface, and can be present in multiple copies. Proteins targeted to the nucleus enter through nuclear pores by a two-step process; the first step is a rapid, signal-dependent binding to the nuclear pore periphery, while the second step is a slower, ATP- and temperature-dependent translocation across the pore (reviewed in references 14 and 65).

In Egr-1, basic residues cluster only in the finger domain and adjacent sequences (Fig. 11B), suggesting that the karyophilic signal of Egr-1 resides here. We have demonstrated that the basic region immediately 5′ of the finger domain in combination with finger 2 or 3 is sufficient to target

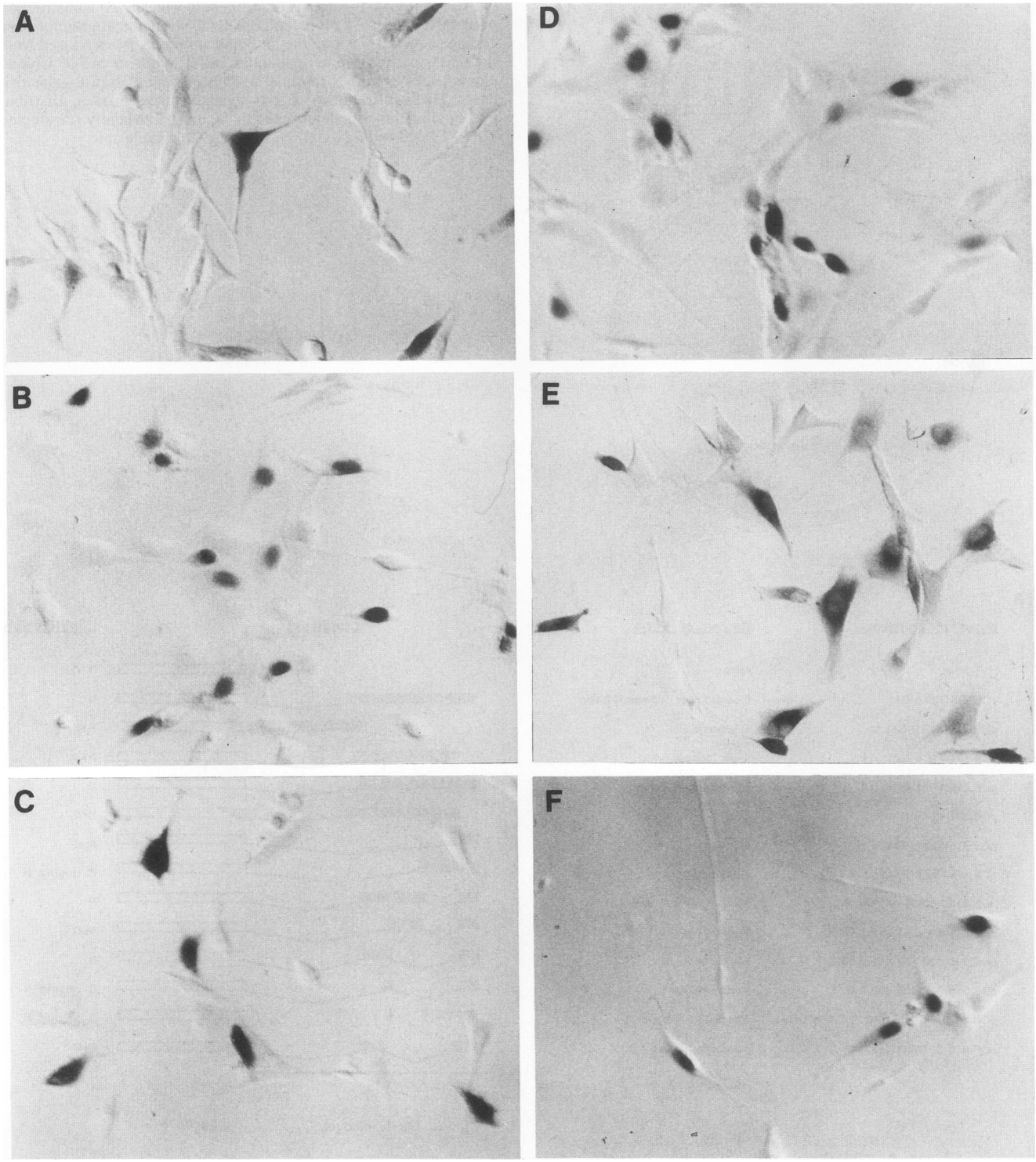


FIG. 10. Localization of Egr-1–β-galactosidase fusions in transfected NIH 3T3 cells. At 40 to 48 h after transfection, cells were fixed with 4% paraformaldehyde and incubated with X-Gal substrate, typically for 1 to 2 h. (A) pCB.β-GAL; (B) pCB.315-429.β-GAL; (C) pCB.331-419.β-GAL; (D) pCB.315-419.β-GAL; (E) pCB.315-367.β-GAL; (F) pCB.315-330/389-419.β-GAL.

the bacterial protein β-galactosidase to the nucleus. This 5′ basic stretch is conserved in other members of the Egr family (Egr-2 and Egr-3) which have homologous DNA-binding domains but generally diverge outside this region (reviewed in reference 68). Our results agree with early suggestions by Day et al. that the C terminus of Egr-1 is required for nuclear localization (12). Precedents for the incorporation of nuclear targeting signals within a DNA-

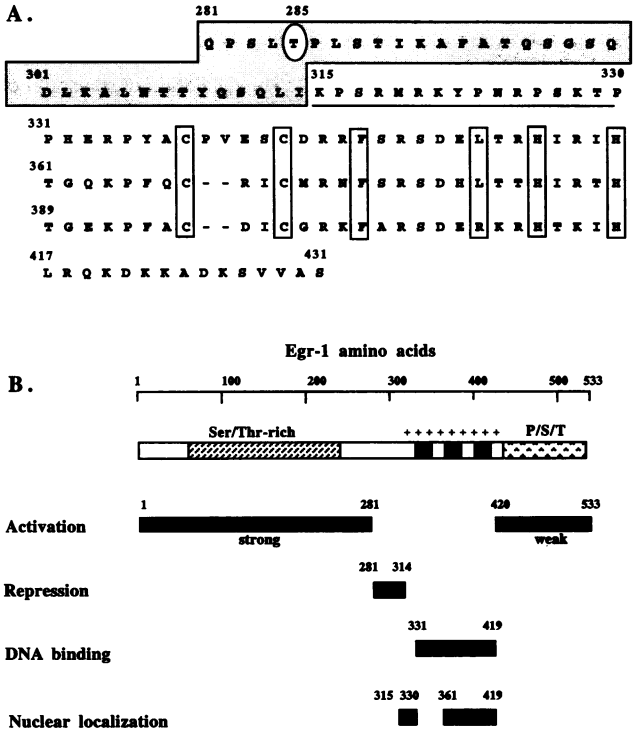


FIG. 11. Summary of Egr-1 domains. (A) Sequence of Egr-1 repression domain and zinc fingers. The repression domain is shaded, and a potential phosphorylation site is circled. The 5′ basic region involved in nuclear localization is underlined. The three zinc fingers of Egr-1 are aligned for comparison, with residues conserved among Cys_2His_2 zinc finger proteins enclosed. (B) Functional domains of Egr-1. The serine/threonine-rich N terminus of Egr-1 is shown. The basic region of Egr-1 is indicated by + symbols. Each zinc finger is designated by a black bar. The proline/serine/threonine-rich C-terminal domain is indicated P/S/T. Residues 1 to 281 activate transcription 100-fold, while the C terminus of Egr-1 (residues 420 to 533) encodes a weaker transactivation function. Amino acids 281 to 314 are sufficient to act as a repressor of transcription when fused to a heterologous DNA-binding domain. The DNA-binding activity of Egr-1 has been mapped to amino acids 331 to 419. The NLS of Egr-1 is bipartite; a basic region (amino acids 315 to 330) and part of the zinc finger domain target the bacterial protein β-galactosidase to the nucleus.

binding domain include Fos (72); the progesterone receptor, in which the second finger but not the first functions as an NLS (19); GAL4 (66); and the homeodomain proteins α2 and Pit-1/GHF-1 (22, 70). If nuclear localization signals and Cys_2His_2 finger domains, both typified by basic residues, have coevolved, NLS sequences may generally be found adjacent to or integrated within zinc finger domains.

Other bipartite NLSs have been characterized in the polymerase basic protein 1 of influenza virus (49); *Xenopus* protein N1 (33); adenovirus DNA-binding protein (47); and the yeast repressor α2, which has two nonhomologous signals, a basic NLS found at the N terminus as well as a signal located in the homeodomain (22, 23). Because each α2 signal gives a different phenotype individually, Hall et al. (22) suggest that these nonhomologous signals mediate separate steps in nuclear accumulation. The peripheral nuclear staining seen with α2 mutants with only the N-terminal NLS intact may reveal a signal for binding to but not translocation across the nuclear pore (22). This may be the case for several Egr-1–β-galactosidase mutants containing the 5′ basic stretch (but neither finger 2 nor 3 intact) which ring the nucleus and may be accumulating at nuclear pores (Fig. 10E). It will be interesting to determine whether these mutants act in a dominant manner to block import of nuclear localization-competent β-galactosidase fusions.

Each of the assays used in these studies measured the equilibrium nuclear/cytoplasmic distribution of protein. Future kinetic analyses may reveal additional sequences required for prompt nuclear localization, as demonstrated by Rihs and Peters for the SV40 T antigen (56). In addition, it is possible that nuclear localization of Egr-1 is regulated as is the case for the serum-dependent nuclear import of Fos (58). Since each of the assays used here monitored the subcellular distribution in exponentially growing cells (maintained in 10% calf serum), more work is needed to address this issue.

Conclusion. Many lines of evidence have intimated that Egr-1 is a factor of central importance. Like the proto-oncogene c-*fos*, *Egr-1* is induced within 30 min by extracellular growth signals in the absence of protein synthesis; this early growth response is common to all cell types. In addition, Egr-1 is strikingly induced when some cell lineages are cued to differentiate. The high degree of cross-species sequence conservation further suggests the importance of this protein (69). However, establishing an in vivo role for Egr-1 as essential for cell growth, contributing to abnormal proliferation, or involved in specialized cell functions has proved more difficult. The recent results of Nguyen et al. demonstrate a requirement for Egr-1 in macrophage differentiation of HL60 cells and further show that Egr-1 functions to restrict the response of HL60 cells by preventing granulocyte differentiation (50). Our functional dissection of Egr-1 provides reagents and insight into the role of this protein that will complement such phenotypic studies. Further work elucidating target genes of Egr-1 transcriptional regulation and potentially other phenotypes by over- or underexpression studies will contribute to our understanding of this multifunctional transcription factor.

ACKNOWLEDGMENTS

We are especially grateful to Lenore Pereira for the gift of monoclonal antibody CH28-2. We thank Colleen Brewer for vector pCB6, Ivan Sadowski and Mark Ptashne for the GAL4(1-147) expression plasmid pSG424, Jim Lillie for the $GAL4_5$E1bCAT reporter, Tom Shenk for pGAL4$_5$tkCAT, and Malcolm Casadaban for the β-galactosidase-encoding plasmid pMC1871. Harinder Singh was invaluable for his comments while this work was in progress. We are grateful to Kevin Morris and Richard Roller, who supplied essential technical advice, and to Nancy Zeleznik-Le, Iain Drummond, David Mack, and Terry Martin for helpful discussions.

This work was supported by NIH grant CA40046-06A1 to V.P.S. and the Howard Hughes Medical Institute. A.L.G. was supported by NCI grant CA09594 and NIH grants DK07199 and DK44921.

REFERENCES

1. **Abate, C., D. Luk, and T. Curran.** 1991. Transcriptional regulation by Fos and Jun in vitro: interaction among multiple activator and regulatory domains. Mol. Cell. Biol. **11**:3624–3632.
2. **Adams, A. E. M., and W. R. Pringle.** 1984. Relationship of actin and tubulin distribution to bud growth in wild-type and morphogenic mutant *Saccharomyces cerevisiae*. J. Cell Biol. **98**:934–945.
3. **Biggin, M. D., and R. Tjian.** 1992. A purified Drosophila homeodomain protein represses transcription in vitro. Cell **58**:433–440.
4. **Bohmann, D., and R. Tjian.** 1989. Biochemical analysis of transcriptional activation by Jun: differential activity of c- and

v-Jun. Cell **59**:709–717.
5. **Bonventre, J. V., V. P. Sukhatme, M. Bamberger, A. J. Ouellette, and D. Brown.** 1991. Localization of the protein product of the immediate early growth response gene, Egr-1, in the kidney after ischemia and reperfusion. Cell Regul. **2**:251–260.
6. **Cao, X., R. A. Koski, A. Gashler, M. McKiernan, C. F. Morris, R. Gaffney, R. V. Hay, and V. P. Sukhatme.** 1990. Identification and characterization of the Egr-1 gene product, a DNA-binding protein induced by differentiation and growth signals. Mol. Cell. Biol. **10**:1931–1939.
7. **Christy, B. A., L. F. Lau, and D. Nathans.** 1988. A gene activated in mouse 3T3 cells by serum growth factors encodes a protein with zinc finger sequences. Proc. Natl. Acad. Sci. USA **85**:7857–7861.
8. **Christy, B. A., and D. Nathans.** 1989. DNA binding site of the growth factor-inducible protein Zif268. Proc. Natl. Acad. Sci. USA **86**:8737–8741.
9. **Cole, A. J., D. W. Saffen, J. M. Baraban, and P. F. Worley.** 1989. Rapid increase of an immediate early gene messenger RNA in hippocampal neurons by synaptic NMDA receptor activation. Nature (London) **340**:474–475.
10. **Corden, J. L.** 1990. Tails of RNA polymerase II. Trends Biochem. Sci. **15**:383–387.
11. **Courey, A. J., and R. Tjian.** 1988. Analysis of Sp1 in vivo reveals multiple transcriptional domains, including a novel glutamine-rich activation motif. Cell **55**:887–898.
12. **Day, M. L., T. J. Fahrner, S. Aykent, and J. Milbrandt.** 1990. The zinc finger protein NGFI-A exists in both nuclear and cytoplasmic forms in nerve growth factor-stimulated PC12 cells. J. Biol. Chem. **25**:15253–15260.

12a. **Drummond, I. A., P. Rohwer-Nutter, and V. P. Sukhatme.** The zebrafish EGR1 gene encodes a highly conserved zinc finger transcriptional regulator. Submitted for publication.

13. **Friedman, A. D., and S. L. McKnight.** 1990. Identification of two polypeptide segments of CCAAT/enhancer-binding protein required for transcriptional activation of the serum albumin gene. Genes Dev. **4**:416–426.
14. **Garcia-Bustos, J., J. Heitman, and M. N. Hall.** 1991. Nuclear protein localization. Biochim. Biophys. Acta **1071**:83–101.
15. **Gerster, T., C.-G. Balmaceda, and R. G. Roeder.** 1990. The cell type specific octamer transcription factor OTF-2 has two domains required for the activation of transcription. EMBO J. **9**:1635–1643.
16. **Gilman, M. Z., R. N. Wilson, and R. A. Weinberg.** 1986. Multiple protein-binding sites in the 5′-flanking region regulates c-Fos expression. Mol. Cell. Biol. **6**:4305–4316.
17. **Gorman, C. M., L. F. Moffat, and B. H. Howard.** 1982. Recombinant genomes which express chloramphenicol acetyltransferase in mammalian cells. Mol. Cell. Biol. **2**:1044–1051.
18. **Greenberg, M. E., and E. B. Ziff.** 1984. Stimulation of 3T3 cells induces transcription of the c-*fos* proto-oncogene. Nature (London) **311**:433–438.
19. **Guiochon-Mantel, A., P. Lescop, S. Christin-Maitre, H. Loosfelt, M. Perrot-Applanat, and E. Milgrom.** 1991. Nucleocytoplasmic shuttling of the progesterone receptor. EMBO J. **10**:3851–3859.
20. **Hahn, S.** 1992. The yin and yang of mammalian transcription. Curr. Biol. **2**:152–154.
21. **Hall, F. R., and P. R. Vulliet.** 1991. Proline-directed protein phosphorylation and cell cycle regulation. Curr. Biol. **3**:176–184.
22. **Hall, M. N., C. Craik, and Y. Hiraoka.** 1990. Homeodomain of yeast repressor α2 contains a nuclear localization signal. Proc. Natl. Acad. Sci. USA **87**:6954–6958.
23. **Hall, M. N., L. Hereford, and I. Herskowitz.** 1984. Targeting of *E. coli* β-galactosidase to the nucleus in yeast. Cell **36**:1057–1065.
24. **Han, K., M. S. Levine, and J. L. Manley.** 1989. Synergistic activation and repression of transcription by Drosophila homeobox proteins. Cell **56**:573–583.
25. **Hariharan, N., D. E. Kelley, and R. P. Perry.** 1991. δ, a transcription factor that binds to downstream elements in several polymerase II promoters, is a functionally versatile zinc finger protein. Proc. Natl. Acad. Sci. USA **88**:9799–9803.
26. **Hope, I. A., and K. Struhl.** 1986. Functional dissection of a eukaryotic transcriptional activator protein, GCN4 of yeast. Cell **46**:885–894.
27. **Hunter, T., and M. Karin.** 1992. The regulation of transcription by phosphorylation. Cell **70**:375–387.
28. **Inostroza, J. A., F. H. Mermelstein, I. Ha, W. S. Lane, and D. Reinberg.** 1992. Dr1, a TATA-binding protein-associated phosphoprotein and inhibitor of class II gene transcription. Cell **70**:477–489.
29. **Jaynes, J. B., and P. H. O'Farrell.** 1991. Active repression of transcription by the Engrailed homeodomain protein. EMBO J. **10**:1427–1433.
30. **Kadonaga, J. T., A. J. Courey, J. Ladika, and R. Tjian.** 1988. Distinct regions of Sp1 modulate DNA binding and transcriptional activation. Science **242**:1566–1569.
31. **Kalderon, D., B. L. Roberts, W. D. Richardson, and A. E. Smith.** 1984. A short amino acid sequence able to specify nuclear localization. Cell **39**:499–509.
32. **Kemp, B. E., and R. B. Pearson.** 1990. Protein kinase recognition sequence motifs. Trends Biochem. Sci. **15**:342–346.
33. **Kleinschmidt, J. A., and A. Seiter.** 1988. Identification of domains involved in nuclear uptake and histone binding of protein N1 of Xenopus laevis. EMBO J. **7**:1605–1614.
34. **Kruijer, W., J. A. Cooper, T. Hunter, and I. M. Verma.** 1984. Platelet-derived growth factor induced rapid but transient expression of the c-fos gene and protein. Nature (London) **312**:711–716.
35. **Lemaire, P., O. Relevant, R. Bravo, and P. Charnay.** 1988. Two mouse genes encoding potential transcription factors with identical DNA-binding domains are activated by growth factors in cultured cells. Proc. Natl. Acad. Sci. USA **85**:4691–4695.
36. **Lemaire, P., C. Vesque, J. Schmitt, H. Stunnenberg, R. Frank, and P. Charnay.** 1990. The serum-inducible mouse gene Krox-24 encodes a sequence-specific transcriptional activator. Mol. Cell. Biol. **10**:3456–3467.
37. **Levine, M., and J. L. Manley.** 1989. Transcriptional repression of eukaryotic promoters. Cell **59**:405–408.
38. **Licht, J. D., M. J. Grossel, J. Figge, and U. M. Hansen.** 1990. *Drosophila* Krüppel protein is a transcriptional repressor. Nature (London) **346**:76–79.
39. **Lillie, J. W., and M. R. Green.** 1989. Transcriptional activation by the adenovirus E1a protein. Nature (London) **338**:39–44.
40. **Lim, R. W., B. C. Varnum, and H. R. Herschman.** 1987. Cloning of tetradecanoyl phorbol ester-induced "primary response" sequences and their expression in density-arrested Swiss 3T3 cells and a TPA nonproliferative variant. Oncogene **1**:263–270.
41. **Ma, J., and M. Ptashne.** 1987. Deletion analysis of GAL4 defines two transcriptional activating segments. Cell **48**:847–853.
42. **Madden, S. L., D. M. Cook, J. F. Morris, A. Gashler, V. P. Sukhatme, and F. J. Raucher III.** 1991. Transcriptional repression mediated by the WT1 Wilms tumor gene product. Science **253**:1550–1553.
43. **Martin, D. I. K., and S. H. Orkin.** 1990. Transcriptional activation and DNA binding by the erythroid factor GF-1/NF-E1/Eryf 1. Genes Dev. **4**:1886–1898.
44. **Mermod, N., E. A. O'Neill, T. J. Kelly, and R. Tjian.** 1989. The proline-rich transcriptional activator of CTF/NF-1 is distinct from the replication and DNA binding domain. Cell **58**:741–753.
45. **Milbrandt, J.** 1988. A nerve growth factor-induced gene encodes a possible transcriptional regulatory factor. Science **238**:797–799.
46. **Moniku, E. S., R. Kuhn, G. Weinmaster, B. D. Trapp, and G. Lemke.** 1990. Expression and activity of the POU transcription factor SCIP. Science **249**:1300–1303.
47. **Morin, N., C. Delsert, and D. F. Klessig.** 1989. Nuclear localization of the adenovirus DNA-binding protein: requirement for two signals and complementation during viral infection. Mol. Cell. Biol. **9**:4372–4380.
48. **Müller-Immerglück, M. M., W. Schaffner, and P. Mattias.** 1990. Transcription factor Oct-2A contains functionally redundant activating domains and works selectively from a promoter but not a remote enhancer position in non-lymphoid (HeLa) cells.

EMBO J. 9:1625–1634.
49. **Nath, S. T., and D. P. Nayak.** 1990. Function of two discrete regions is required for nuclear localization of polymerase basic protein 1 of A/WSN/33 influenza virus (H1 N1). Mol. Cell. Biol. **10**:4139–4145.
50. **Nguyen, H., B. Hoffman-Liebermann, and D. Liebermann.** 1992. The zinc finger transcription factor Egr-1 is essential for and restricts differentiation along the macrophage lineage. Cell **72**: 1–20.
51. **Ouellette, A. J., R. A. Malt, V. P. Sukhatme, and J. V. Bonventre.** 1990. Expression of two "immediate early genes," Egr-1 and c-fos, in response to renal ischemia and during compensatory renal hypertrophy in mice. J. Clin. Invest. **85**: 766–771.
52. **Park, K., and M. L. Atchison.** 1991. Isolation of a candidate repressor/activator, NF-E1 (YY-1,δ), that binds to the immunoglobulin κ 3′ enhancer and the immunoglobulin heavy-chain μE1 site. Proc. Natl. Acad. Sci. USA **88**:9804–9808.
53. **Patwardhan, S., A. Gashler, M. G. Siegel, L. C. Chang, L. J. Joseph, T. B. Shows, M. M. Le Beau, and V. P. Sukhatme.** 1991. EGR3, a novel member of the Egr family of genes encoding immediate-early transcription factors. Oncogene **6**:917–928.
54. **Pavletich, N. P., and C. O. Pabo.** 1991. Zinc finger-DNA recognition: crystal structure of a Zif268-DNA complex at 2.1 Å. Science **252**:809–817.
55. **Ptashne, M.** 1988. How eukaryotic transcriptional activators work. Nature (London) **335**:683–689.
56. **Rihs, H.-P., and R. Peters.** 1989. Nuclear transport kinetics depend on phosphorylation-site-containing sequences flanking the karyophilic signal of the simian virus 40 T-antigen. EMBO J. **8**:1479–1484.
57. **Rosenthal, N.** 1987. Identification of regulatory elements of cloned genes with functional assays. Methods Enzymol. **152**: 704–720.
58. **Roux, P., J.-M. Blanchard, A. Fernandez, N. Lamb, P. Jeanteur, and M. Piechaczyk.** 1990. Nuclear localization of c-Fos, but not v-Fos proteins, is controlled by extracellular signals. Cell **63**: 341–351.
59. **Ryseck, R.-P., S. I. Harai, M. Yaniv, and R. Bravo.** 1988. Transcriptional activation of c-jun during the G_0G_1 transition in mouse fibroblasts. Nature (London) **334**:535–537.
60. **Sadowski, I., J. Ma, S. Triezenberg, and M. Ptashne.** 1988. GAL4-VP16 is an unusually potent transcriptional activator. Nature (London) **335**:563–564.
61. **Sanes, J. R., J. L. R. Rubenstein, and J.-F. Nicolas.** 1986. Use of a recombinant retrovirus to study post-implantation cell lineage in mouse embryos. EMBO J. **5**:3133–3142.
62. **Sanger, F., S. Nicklen, and A. R. Coulson.** 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA **74**:5463–5467.
63. **Shastry, B. S.** 1991. *Xenopus* transcription factor IIIA (XTFIIIA): after a decade of research. Biophys. Mol. Biol. **56**:135–144.
64. **Shi, Y., E. Seto, L.-S. Chang, and T. Shenk.** 1991. Transcriptional repression by YY1, a human GLI-Krüppel-related protein, and relief of repression by adenovirus E1A protein. Cell **67**:377–388.
65. **Silver, P. A.** 1991. How proteins enter the nucleus. Cell **64**:489–497.
66. **Silver, P. A., L. P. Keegan, and M. Ptashne.** 1984. Amino terminus of the yeast GAL4 gene product is sufficient for nuclear localization. Proc. Natl. Acad. Sci. USA **81**:5951–5955.
67. **Spaete, R. R., and E. S. Mocarski.** 1985. Regulation of cytomegalovirus gene expression: α and β promoters are *trans* activated by viral functions in permissive human fibroblasts. J. Virol. **56**:135–143.
68. **Sukhatme, V. P.** 1990. Early transcriptional events in cell growth: the Egr family. J. Am. Soc. Nephrol. **1**:859–966.
69. **Sukhatme, V. P., X. Cao, L. C. Chang, C.-H. Tsai-Morris, D. Stamenkovich, P. C. P. Ferreira, D. R. Cohen, S. A. Edwards, T. B. Shows, T. Curran, M. M. Le Beau, and E. D. Adamson.** 1988. A zinc finger-encoding gene coregulated with c-fos during growth and differentiation, and after cellular depolarization. Cell **53**:37–43.
70. **Theill, L. E., J.-L. Castrillo, D. Wu, and M. Karin.** 1989. Dissection of functional domains of the pituitary-specific transcription factor GHF-1. Nature (London) **342**:945–948.
71. **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. USA **76**:4350–4354.
72. **Tratner, I., and I. M. Verma.** 1991. Identification of a nuclear targeting sequence in the Fos protein. Oncogene **6**:2049–2053.
73. **Waters, C. M., D. C. Hancock, and G. I. Evan.** 1990. Identification and characterisation of the *egr-1* gene product as an inducible, short-lived, nuclear phosphoprotein. Oncogene **5**:669–674.
74. **Zeleznik-Le, N., and V. P. Sukhatme.** Unpublished data.
75. **Zuo, P., D. Stanojevic, J. Colgan, K. Han, M. Levine, and J. L. Manley.** 1991. Activation and repression of transcription by the gap proteins *hunchback* and *Krüppel* in cultured *Drosophila* cells. Genes Dev. **5**:254–264.