The serum and TPA responsive promoter and intron-exon structure of *EGR2*, a human early growth response gene encoding a zinc finger protein

Vivek M.Rangnekar[1], Alfred C.Aplin[2] and Vikas P.Sukhatme[1,2]*
[1]Department of Medicine and [2]Molecular Genetics and Cell Biology, Howard Hughes Medical Institute, The University of Chicago, Chicago, IL 60637, USA

Received November 27, 1989; Revised and Accepted March 13, 1990

ABSTRACT

***EGR2* is a human zinc finger encoding gene whose expression is induced with *fos*-like kinetics by diverse mitogens in several cell types. Since its cDNA sequence predicts a protein which contains zinc finger motifs, *EGR2* may play a transcriptional regulatory role in cellular proliferation. The present study was undertaken to: 1) examine the genomic organization and 5′ flanking sequence of *EGR2* so as to identify upstream regulatory elements; 2) test whether these elements are functional in gel shift assays and by transient expression; and 3) examine whether pathways other than protein kinase C lead to serum induction of *EGR2*, and if they do, ask whether the different pathways converge on a serum response element. The *EGR2* gene spans 4.3 kb and has one intron. The translation initiation site is located within the first exon. The transcription start site of *EGR2* was determined by S1 nuclease and primer extension analysis and a TATA box was identified 28 bp upstream. Two putative serum response elements, designated CArG-1 and CArG-2 were identified in the 5′ flanking sequence. By deletion analyses and mutagenesis, serum and PMA responsiveness of the cloned *EGR2* promoter region was traced to the CArG-1 region in transient CAT assays performed in NIH 3T3 cells. Both protein kinase C dependent and independent pathways were found to converge on the CArG-1 box to induce the expression of *EGR2*.**

INTRODUCTION

To understand the mechanisms by which growth factors modulate cell growth, it is necessary to define the complex series of events that ensue following mitogen-receptor binding. Amongst these events are the rapid generation of second messengers in the cytosol and plasma membrane. In turn, a set of 'immediate-early' genes are activated, whose induction does not require new protein synthesis (1, 2). An important subset of these genes encode transcriptional factors. These include: *c-fos* (3), *fra-1* (4), and *fos B* (5), 3 distinct genes (members of the *EGR* family) designated *Egr-1* (6, 7, 8, 9), *EGR2* (10, 11), and *EGR3* (manuscript in preparation), all of which encode proteins with Cys_2-His_2 zinc finger motifs; a member of the steroid and thyroid hormone receptor family (12, 13); and *c-jun* (14, 15, 16, 17) and its closely related genes (18). The importance of these genes is that by virtue of their structure and induction kinetics, they are likely to play broad roles as 'third messengers' by coupling early biochemical processes to long-term changes in gene expression required to modulate cell growth. Furthermore, these gene products also participate as parts of a regulatory cascade in other cellular processes such as in differentiation and cellular depolarization (3, 9). A major challenge ahead is to define the mechanisms by which specific intracellular second messengers affect expression of one or more members of this subgroup of 'immediate-early' genes and to identify target genes and/or proteins with which products of these genes interact.

Our laboratory is focusing on these questions in the context of the *EGR* family of genes. *Egr-1* displays *c-fos* like induction kinetics following mitogen stimulation in a broad range of cell types including B (19) and T cells (unpublished data), epithelial cells of renal and liver origin and in fibroblasts (20). *Egr-1* expression is also modulated during neuronal and cardiac differentiation and following cellular depolarization (9), as well as during EC cell differentiation (unpublished data), following renal hypertrophy, as well as in renal ischemia and liver regeneration (2, unpublished data). We (21) and others (6) have recently published approximately 1 kb of 5′ flanking sequence of *Egr-1*. This region has a variety of putative elements which might confer promoter responsiveness to cAMP, TPA, and serum. We have recently found that CArG elements (sequences of the form $CC(A/T)_6GG$) present in this region are crucial in serum induction, and its subsequent down-regulation (Cao et al. and Gius et al., submitted). In an attempt to identify other early growth response genes encoding zinc finger structures, low stringency hybridization with an *Egr-1* finger region probe was used to screen a serum plus cycloheximide stimulated fibroblast cDNA library. To date, two novel clones have been identified (designated *EGR2* and *EGR3*). All have zinc fingers very similar to those of *Egr-1*, but dissimilarity exists elsewhere in the deduced amino acid sequence (7, 11, unpublished data). Furthermore, the

* To whom correspondence should be addressed

levels of *EGR2* mRNA are elevated following growth stimulation in fibroblasts as well as in lymphocytes (11), suggesting that, like *Egr-1*, the expression of *EGR2* is not restricted to one cell type. However, unlike *Egr-1*, *EGR2* expression is not induced by nerve growth factor in the rat PC12 pheochromocytoma cell line (11). It also shows a strikingly different developmental profile (22). This paper: 1) presents the intron-exon structure and nucleotide sequence of the flanking region of *EGR2* and the identification of two putative serum response elements designated CArG-1 and CArG-2; 2) provides evidence that CArG-1 confers serum as well as PMA responsiveness to the cloned promoter of *EGR2*; 3) suggests that the CArG-1 and CArG-2 boxes are modulated by the same serum response factor that regulates the CArG/SRE box of *c-fos*; and 4) demonstrates that in addition to a protein kinase C (PKC) dependent pathway, non-protein kinase C pathways also mediate the serum responsiveness of *EGR2* through the CArG-1 box.

MATERIALS AND METHODS

Genomic libraries and isolation of human EGR2 genomic clones. To obtain *EGR2* genomic clones, we screened a WI-38 human fibroblast genomic λFix library (Stratagene, La Jolla, California) and a human placental EMBL-3 library kindly provided by Dr. C. Westbrook, University of Chicago. Approximately 2.4×10^5 clones from the WI-38 library and 1.4×10^5 clones from the placental library, in *Escherichia coli* strain P2392 (F^-, *hsdR514, rk⁻, mk⁺, supE44, supF58, ΔlacIZY6, metB1, trp55*) were hybridized with a ^{32}P-labeled 2 kb Zap2 *EGR2* cDNA fragment (11) containing the three zinc finger motifs and flanking 5′ and 3′ sequences (Fig. 1). Membranes (*GeneScreenPlus*, NEN-DuPont) were hybridized for 16 hours at 55°C in 1% SDS, 10% dextran sulfate and 1 M NaCl. The filters were washed to a final stringency of 55°C in 0.4×SSC (1×SSC = 0.15 M sodium chloride/0.015 M sodium citrate pH 7). Autoradiographs were prepared by exposing the filters for 18 hours at −70°C with an intensifying screen.

Southern blotting and hybridization. All blots used *GeneScreenPlus* membranes (NEN-DuPont) and were hybridized at 65°C in 1% SDS, 10% dextran sulfate and 1 M NaCl. Filters were washed to a final stringency of 65°C in 0.2×SSC.

DNA sequencing. Various DNA fragments were subcloned into the polylinker region of either pUC13 or pUC18 and sequenced using Klenow or *E. coli* T7 polymerase (Pharmacia) according to the dideoxynucleotide chain-termination method (23).

S1 nuclease assay. From plasmid pVA2 (Fig. 3), an 800 bp PvuII-PvuII fragment was isolated, treated with alkaline phosphatase, and end-labeled. It was then annealed at 85−90°C for 10 minutes to 10 μg of total cellular RNA prepared from unstimulated human 303 fibroblast cells (provided by Dr. J. R. Smith, Baylor College of Medicine) and from serum and cycloheximide stimulated 303 cells. Subsequently, hybridization was done for 18 hours in a reaction volume of 30 μl at 50°C in the presence of 60% formamide, 0.4 M NaCl, 0.05 M PIPES and 0.83 mM EDTA. Hybrids were treated with 1 unit of S1 nuclease for 1.5 hours at 37°C in a final volume of 300 μl. The reaction products were extracted with phenol/chloroform, precipitated with 0.3 M sodium acetate pH 4.8, and 2.5 volumes of ethanol, and electrophoresed next to sequencing reaction products used as markers on a 6% polyacrylamide gel containing 8 M urea.

Primer extension. For primer extensions two synthetic 5′ end labeled 29mer oligonucleotides complementary to nucleotides +201 to +229 (PE2) and +243 to +271 (PE1) were hybridized to 10 μg of serum plus cycloheximide treated 303 human fibroblast total RNA for 16 hours at 40°C. The hybridization mixture was extended in 20 μl reverse transcriptase buffer [10 mM $MgCl_2$, 50 mM Tris, 60 mM KCl, 1 mM DTT, 1 mM each dNTP, 50 μg/ml actinomycin D, 2 units RNAsin] with 50 units of AMV reverse transcriptase at 37°C for 2 hours. Products were run on 7.5% acrylamide-8 M urea gels and exposed to Kodak XAR5 film at −80°C with an intensifying screen for 64 hours.

Cell culture and transfection. NIH 3T3 fibroblast cells were grown in Dulbecco's Modified Eagle medium (DMEM) containing 10% fetal calf serum (FCS), gentamicin (20 μg/ml) and 5% CO_2. The cells were transfected with 25 μg of plasmid DNA (5 μg of test construct and 20 μg of pUC18 carrier DNA) per 100 mm dish utilizing 125 mM calcium chloride and HEPES buffered solution pH 7.1 containing 140 mM NaCl, 0.75 mM $Na_2PO_4\cdot12H_20$, and 25 mM HEPES. Precipitates were incubated for 8 hours, after which the cells were washed twice with PBS and maintained for 40 hours in DMEM supplemented with 0.5% FCS. The influence of serum stimulation on promoter activity was examined by maintaining the transfected cells for 40 hours in 0.5% FCS followed by 3 hours of growth in DMEM supplemented with 10% FCS. The effect of PMA on promoter activity was examined by addition of 50 nM PMA for 5 hours to the serum-starved cells. In experiments aimed at studying the effect of serum and PMA on PKC deficient cells, transfected cells were cultivated for 40 hours in DMEM/0.5% FCS in the presence of 1 μM PMA, and depletion of PKC was confirmed by the subsequent loss of PMA inducibility. Cell extracts were prepared and CAT assays were performed as described (24).

Deletion constructs, and site directed mutagenesis. The *EGR2*-CAT construct (plasmid pCAT16) was made by deleting about 280 bp of sequence upstream from the PvuII site of the 1.3 kb PvuII-PvuII fragment (VA4) and placing the resulting fragment upstream of the CAT coding sequences as follows. Plasmid pVA4, which contained the *EGR2* genomic 1.3 kb PvuII-PvuII fragment cloned into the SmaI site of plasmid pUC13, was linearized by digestion with SacI (a unique site in pUC13) and 295 bp from the SacI site were deleted using Bal 31 nuclease. The two ends of the resulting fragment were flushed using Klenow and subsequently digested with XbaI (site in pUC13). The fragment was then cloned into the XbaI-SmaI sites of a CAT-poly vector constructed by ligating a polylinker sequence into the BglII site of CAT3M (24). Plasmid pCAT20 was constructed from pCAT16 by deleting the region between −832 and −201 by cutting with XbaI and SmaI, flushing with Klenow, and ligating.

In order to examine the role of the CArG-1 box in serum and PMA inducibility of the cloned *EGR2* promoter, two constructs pCAT40, containing an unaltered CArG-1 box, and pCAT50 containing a 'G' to 'T' transversion in the CArG-1 box, were utilized. Plasmid pCAT40 was constructed by using PCR to synthesize a 189 bp fragment (Fig. 2 nt. −73 to +116 containing the CArG-1 box sequence of the *EGR2* promoter) by polymerase

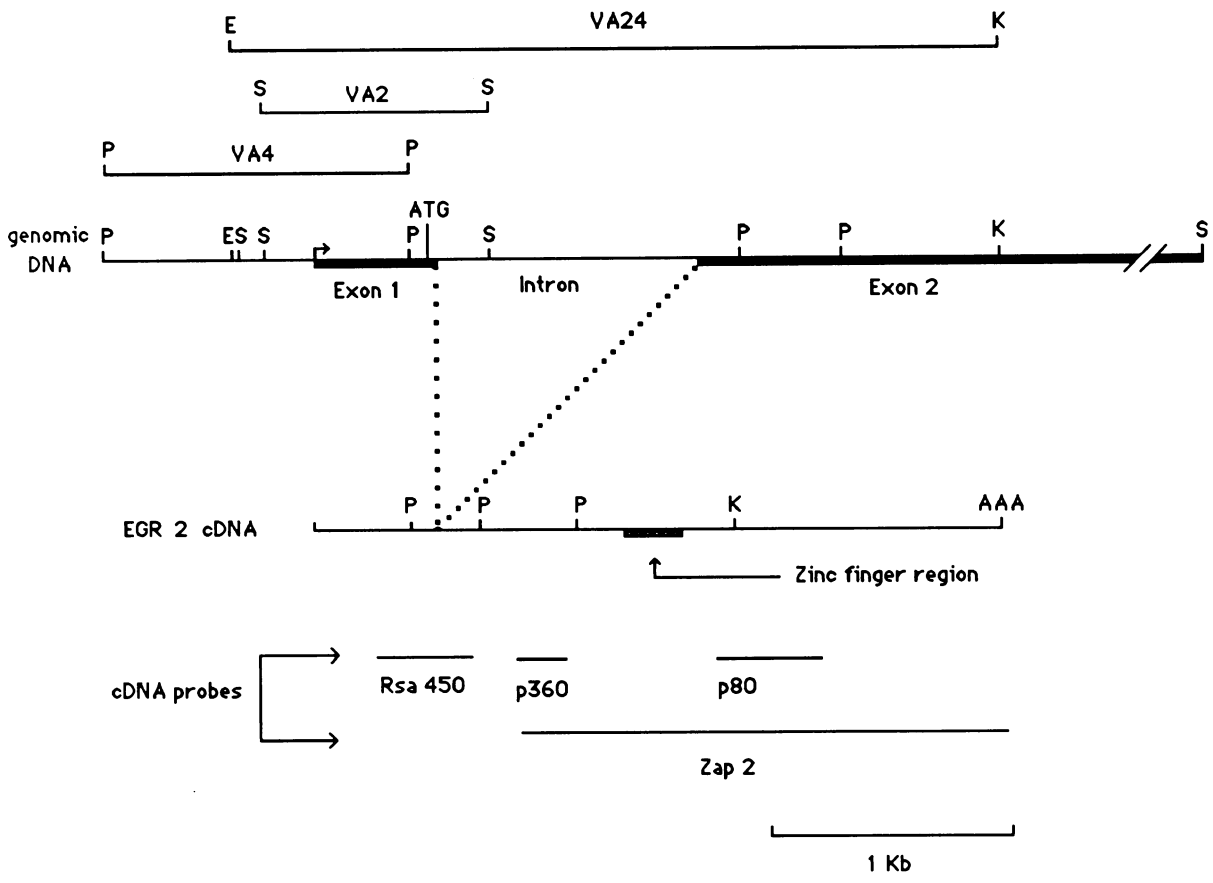

Fig. 1. Composite restriction map and organization of the human *EGR2* gene and comparison with human *EGR2* cDNA (11). The location of restriction sites for EcoRI (E), KpnI (K), PvuII (P), and SmaI (S), the transcription start site (↱), and translation initiation codon (ATG) are indicated.

chain reaction. This fragment was subcloned into the SmaI site of the CAT-poly vector in the appropriate orientation. The strategy for construction of pCAT50 was similar to that for pCAT40, except that the synthetic 5′ primer used for PCR contained CArG-1 box with the 'G' at position −61 altered to a 'T'.

Gel mobility shift assays. Nuclear extracts were prepared as described (Cao et al, submitted). A 21 bp probe containing the CArG-1 region of *EGR2* was constructed by annealing synthetic oligonucleotides 5′-AGTCCATATATGGGCAGC-GAC-3′ and 5′-GTCGCTGCCCATATATGGACT-3′ and end-labeling. The unlabeled 21 bp CArG-1 fragment was also used in competition binding experiments. A 36 bp Sau 3A-EcoRI fragment (positions −377 to −342) containing the CArG-2 box was used in competition binding assays. Plasmid p310-XB5 which contains a 25 bp fragment of the *c-fos* SRE (CATGAGGATGTCCATA-TTAGGACATCTG) and p310-MU which contains a mutant SRE (CTGACAGGATGTGGATATTACCACATCTG) were gifts from B. H. Cochran. Both plasmids were cut with PvuII and the 300 bp fragments which contained the normal or mutant SREs were isolated and also used as unlabeled competitors.

Binding reaction mixtures contained 3 μg of pUC18 DNA which had been digested with Sau 3A, 2 μg of poly (dI-dC), 5 to 20 μg nuclear extract, and 1 ng of labeled probe. The binding buffer consisted of 20 mM HEPES, 20% glycerol, 100 mM KCl, 0.2 mM EDTA, 1 mM phenylmethylsulfonyl fluoride, 0.5 mM dithiothreitol and 300 μg/ml bovine serum albumin. The reaction mixtures were incubated at room temperature for 30 minutes and electrophoresed on 4% polyacrylamide gels, which were dried and autoradiographed.

Northern blot analyses. RNA was prepared by the acid-guanidinium thiocyanate-phenol-chloroform extraction technique (25), electrophoresed on formaldehyde-agarose gel, and blotted onto nitrocellulose (Schleicher and Schuell) in the presence of 20×SSC. Prehybridization and hybridization was performed at 50°C using a buffer consisting of 1× Denhardt's solution, 0.1% SDS, 10% dextran sulfate, 4×SSC and 5 mM Tris-HCl. Blots were washed at 50°C with 2×SSC, and autoradiographed for 1 week with an intensifying screen at −70°C.

RESULTS

The EGR2 gene has one intron and two exons. Three genomic clones, WI-38.V1, WI-38.V2, WI-38.V3, from the WI-38 library, and one clone, CWP1 from the placental library hybridized to the *EGR2* cDNA Zap 2 (11). To further confirm the authenticity of these clones, Southern blots of various restriction digests were probed with several *EGR2* cDNA fragments (Rsa450, p360, and p80), located as shown (Fig. 1). The autoradiographs suggested that WI-38.V2 and WI-38.V3, were identical clones but distinct from WI-38.V1 and CWP1 (data not shown). The different *EGR2* clones contained colinear

TABLE 1. Human *EGR2* exon-intron junction sequence.

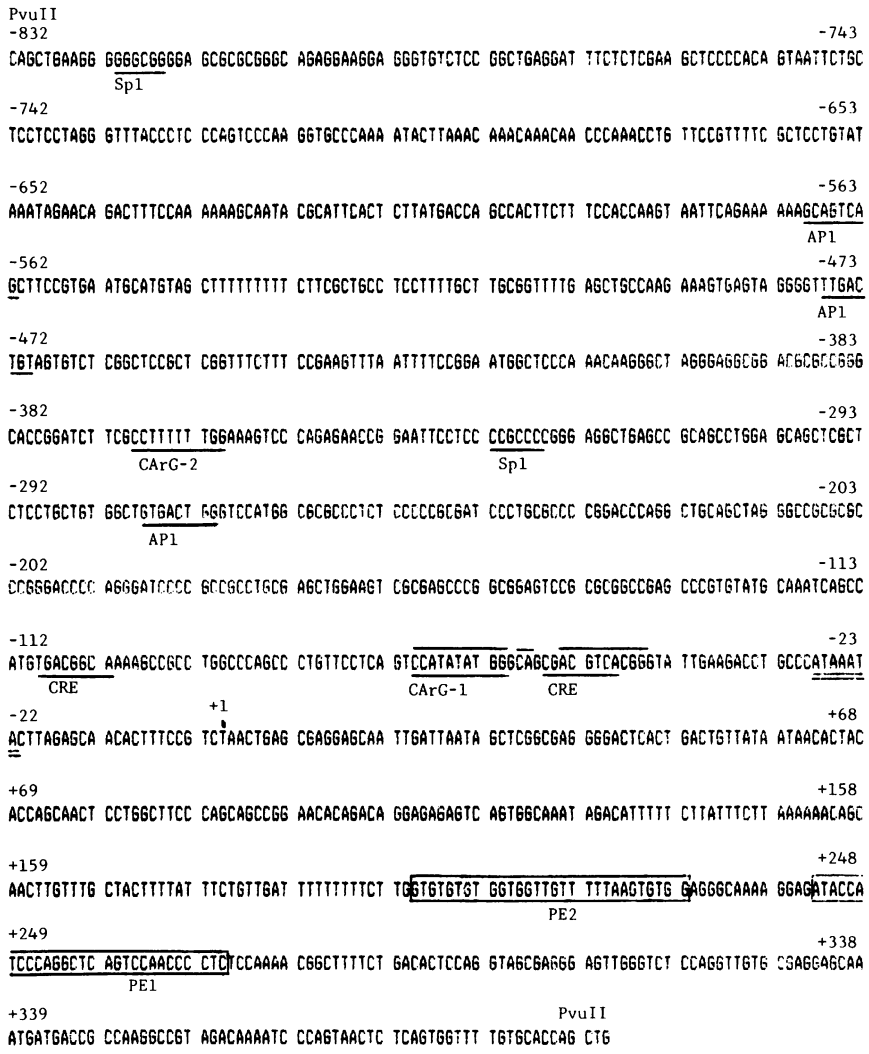

Fig. 2. The 5′ upstream nucleotide sequence of the human *EGR2* gene. Nucleotides are numbered from the transcription start site (+1). Putative regulatory elements: serum response elements (CArG-1 and CArG-2); AP1 binding site (AP1); Sp1 binding site (Sp1); cyclic AMP response element (CRE) are underlined. A possible TATA box is double underlined. The sequence conserved in *Egr-1* and *EGR2* is overlined. Oligomers PE1 and PE2 used for primer extension are boxed.

restriction enzyme fragments, and a composite restriction map (Fig. 1) of the *EGR2* gene is shown. Various DNA fragments including the 1.3 kb PvuII-PvuII (VA4), 0.9 kb SmaI-SmaI (VA2), and 3.2 kb EcoRI-KpnI (VA24) were cloned from WI-38.V3 into the SmaI site of pUC13 and the EcoRI-KpnI sites of pUC18 respectively, and used for mapping and sequence analysis.

To determine the genomic organization of *EGR2*, specific

oligonucleotides (17-mers at positions 99, 319, 517, 696, 732, 878, 1136, 1431, 1530, 1652, 1750, 1900, 2062, 2254, 2433, and 2579 of the cDNA sequences (11)) were used as primers for double stranded sequencing of CWP1 genomic DNA and plasmids pVA2, pVA4, and pVA24. A comparison of the *EGR2* genomic sequence with the *EGR2* cDNA sequence showed that the *EGR2* gene consists of two exons, and a single intron, 1200 bp in length, located between nucleotides 222 and 223 as numbered in the cDNA sequence (11). The 5′ and 3′ splice junction sequences conform well to the consensus boundary sequence (Table 1; 26).

5′ flanking sequence of the human EGR2 gene shows several consensus regulatory elements. Plasmid pVA4 was employed to obtain the flanking sequence as presented in Fig. 2. The transcription start site was mapped by S1 nuclease analysis as follows. Plasmid pVA2 (900 bp SmaI-SmaI fragment of *EGR2* cloned in pUC13) was digested with PvuII, and an 800 bp PvuII-PvuII fragment was obtained by utilizing a PvuII site of pUC13 and the PvuII site in the *EGR2* genomic portion of the clone (see Fig. 3). This fragment was end-labeled and hybridized to RNA prepared from serum plus cycloheximide treated human 303 fibroblasts cells (Fig. 3, left panel, lane A) or from untreated 303 cells (Fig. 3, left panel, lane B). The size of the S1 nuclease resistant DNA (Fig. 3, left panel, lane A, shown by thick arrow) indicated that the transcription start site was located approximately 400 bp upstream of the PvuII site. To precisely define the start site location, primer extension analysis was utilized (Fig. 3, right panel). Using two primers spaced 42 bp apart (PE1 and PE2 shown in Fig. 2) extension products on total RNA from serum plus cycloheximide stimulated cells were obtained. These fragment sizes (271 and 229 bases) predicted a start site in agreement with the S1 analysis data. A second shorter product (Fig. 3, right panel, arrow) noted with the PE1 primer was not present when hybridization was carried out at a higher temperature (data not shown).

A number of regulatory elements were identified in the 5′ upstream region of the VA4 sequence. A putative TATA box (ATAAATA) was identified 28 bp upstream of the transcription start site (nt + 1 in Fig. 2). Two putative CArG boxes, designated CArG-1 (positions −70 to −61) and CArG-2 (positions −369 to −360) were identified (Fig. 2). Other putative sequences including two Sp1 binding sites (GGGCGG, or in the opposite orientation CCGCCC), three AP1 binding sites ([C/G]TGACT[C/A]A) (27, 28), and two cAMP responsive elements ([T/G][T/A]-CGTCA) (29) were also noted in the 5′ flanking sequence (Fig. 2).

Serum responsive activity in the 5′ flanking region. To determine whether the *EGR2* genomic sequences located 5′ of the transcription start site contained promoter activity that was serum inducible, plasmid pCAT16 was tested in a transient expression assay. Plasmid pCAT16 was constructed by placing the *EGR2* 5′ sequence (nt −832 to +116), containing CArG-1 and CArG-2 boxes as well as other putative regulatory elements, upstream of the bacterial chloramphenicol acetyl transferase gene. CAT activity was induced in pCAT16 transfected cells following stimulation with 10% fetal calf serum (Fig. 4, lane B), whereas cells maintained in 0.5% serum following pCAT16 transfection showed relatively low CAT (Fig. 4, lane A) activity. The percent conversion was 12.8 with extracts from serum induced cells and 3.2 with extracts from serum starved cells. These data indicate

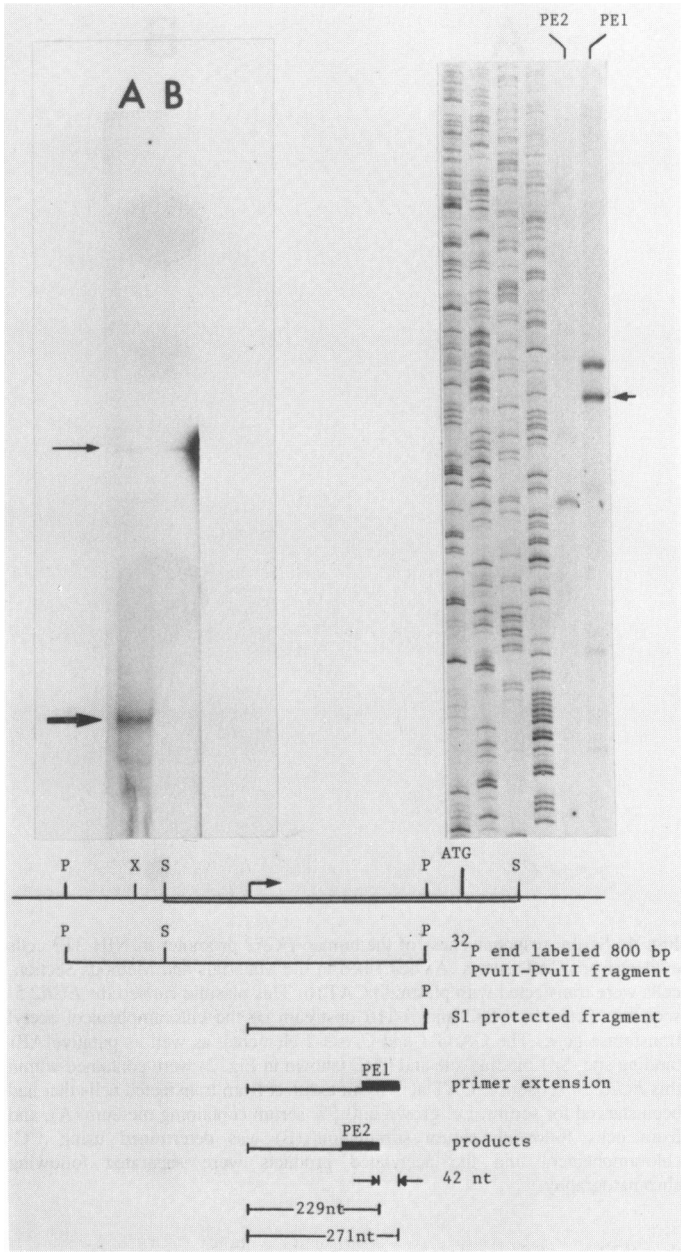

Fig. 3. Mapping of the EGR2 transcription start site. S1 nuclease analysis is shown in the left panel. Plasmid pVA2 was constructed by ligating the 900 bp SmaI-SmaI fragment (VA2) into the unique SmaI site of pUC13. An 800 bp PvuII-PvuII fragment, from pVA2 containing about 600-bp of the human *EGR2* gene and about 200 bp of pUC13 was used as a probe for S1 nuclease analysis. The probe was hybridized to total RNA prepared from uninduced human 303 fibroblast cells (lane B) and from 303 cells that had been stimulated with 10% serum in presence of 10 μg/ml cycloheximide for 3 hours (lane A). A major ~400 bp protected fragment (thick arrow) was seen only in lane A. An 800 bp fragment resulting from reannealing of the two strands of the probe (thin arrow) was seen in both lanes A and B. The restriction sites PvuII (P), SmaI (S), and XbaI (X), transcription start site (↱), and translation initiation codon (ATG) are indicated. Primer extension products are shown in the right panel. Primers (PE1 and PE2 shown in Fig. 2) were hybridized at 40°C to RNA from serum plus cycloheximide stimulated 303 cells. Extension products were analyzed on a 7.5% polyacrylamide-8 M urea gel. The arrow in the right panel is a shorter extension product which was not detected when hybridization was carried out at a higher temperature (data not shown). Sequencing reactions from a known template were co-run to size the extended fragments.

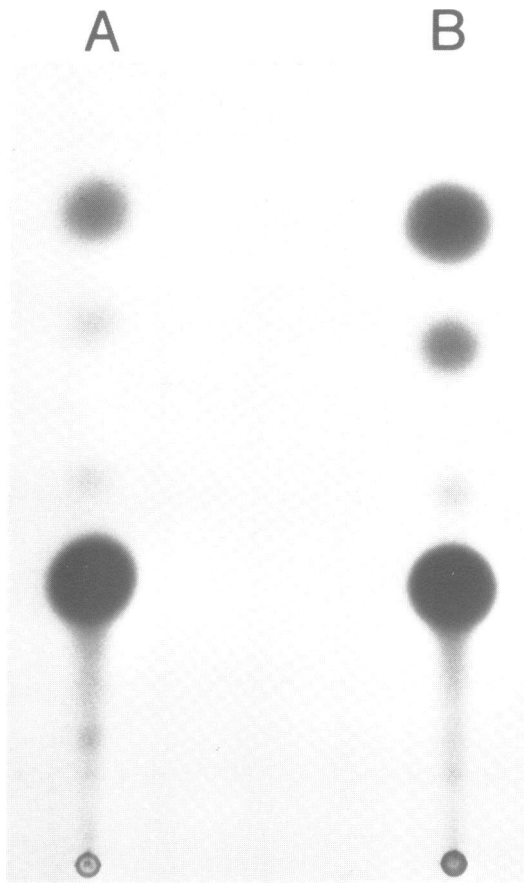

Fig. 4. Serum responsiveness of the human *EGR2* promoter in NIH 3T3 cells evident from CAT assays. As described in the Materials and Methods Section, cells were transfected with plasmid pCAT16. This plasmid carried the *EGR2* 5′ sequence between −832 and +116 upstream of the chloramphenicol acetyl transferase gene. The CArG-1 and CArG-2 elements, as well as putative AP1 binding site, Sp1 binding site and CRE (shown in Fig. 2) were contained within this *EGR2* 5′ sequence. CAT activity in extracts from transfected cells that had been starved for serum, i.e. grown in 0.5% serum containing medium (A), and from cells following serum stimulation (B) was determined using ^{14}C-chloramphenicol and the acetylated products were separated following chromatography.

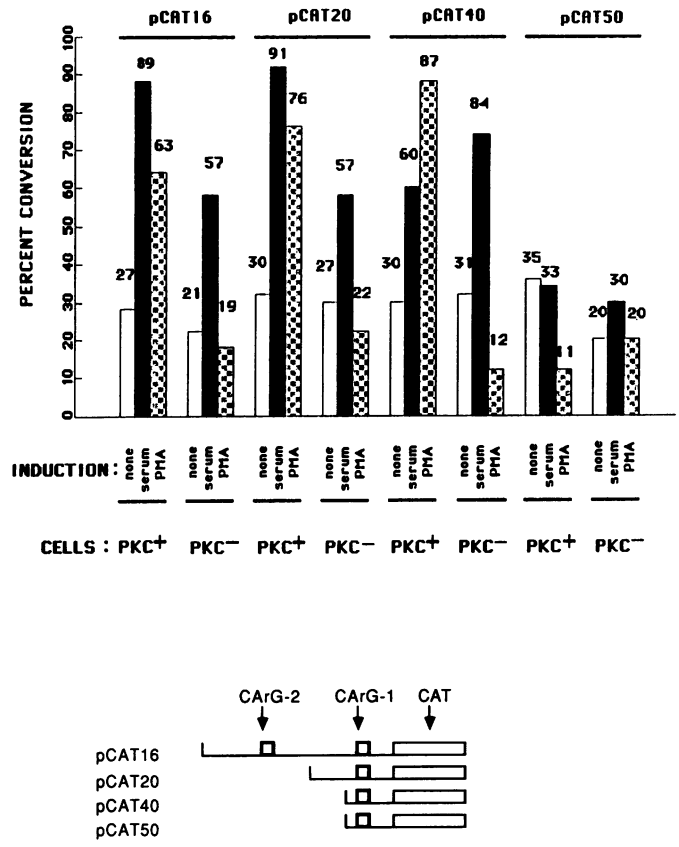

Fig. 5. Effect of deletions and CArG box mutation on serum and PMA responsiveness of the cloned *EGR2* promoter. Cells were transfected with the different constructs (pCAT16, pCAT20, pCAT40, or pCAT50). Plasmid pCAT16 contained both CArG-1 and CArG-2 elements, whereas pCAT20 and pCAT40 contained only CArG-1, and pCAT50 contained mutant CArG-1 sequence. One set of cells was grown in presence of 0.5% serum plus PMA (1 μM) for 40 hours to deplete PKC, while a second set was grown in 0.5% serum without PMA. Cells were then induced with serum, PMA, or left untreated (none), as indicated. CAT activity in the extracts was determined using 14C-chloramphenicol and the acetylated products were separated by chromatography. The acetylation reactions were deliberately incubated for extended periods of time so as to allow detectable conversion with extracts from cells transfected with pCAT50 (PKC^- = protein kinase C depleted cells; PKC^+ = 'normal' cells).

that 832 bp of 5′ upstream sequence is sufficient to confer serum inducibility.

5′ sequence containing CArG-1 is sufficient to confer serum responsiveness to the EGR2 promoter. The 5′ sequence of *EGR2* contains two CArG boxes designated CArG-1 and CArG-2 (see Fig. 2). In order to analyze the functional properties of each CArG box, three constructs, pCAT20 and pCAT40 containing only CArG-1, and pCAT50 containing a mutated CArG-1 element were tested for serum responsiveness in transient expression assays. Data is presented as part of Fig. 5 (see CAT activity data on normal (PKC^+) cells). Deletion of CArG-2 did not decrease the serum responsiveness of pCAT20 compared to that of pCAT16 indicating that CArG-1 was sufficient for serum inducibility. This observation was confirmed by the finding that alteration of CArG-1 in plasmid pCAT50 resulted in complete loss of serum responsiveness.

Serum responsiveness has been attributed to the post-translational modification of a DNA binding protein, serum response factor (SRF), which binds to the SRE region in the 5′ sequence of the *c-fos* promoter and activates transcription (30, 31). The inner core of the SRE is a CArG box and since we have shown that the CArG-1 sequence in *EGR2* is serum responsive, we examined whether this element interacted with nuclear proteins. In gel mobility shift assays, the end-labeled 21 bp CArG-1 fragment migrated as a discrete band in the absence of nuclear extract (Fig. 6A, lane 1; Fig. 6B, lane 1), but in the presence of extract a single shifted band (Fig. 6A, lane 2; Fig. 6B, lanes 2−4) was evident. The specificity of this interaction was tested by addition of a 10-fold (Fig. 6A, lane 3), 60-fold (Fig. 6A, lane 4) and 120-fold (Fig. 6A, lane 5) molar excess of unlabeled CArG-1 fragment. Total abolition of binding to the labeled probe was accomplished only with 60-fold and 120-fold molar excess of unlabeled CArG-1. Furthermore, up to 120-fold molar excess of the unlabeled mutant CArG-1 did not affect formation of the labeled CArG-1-protein complex (compare Fig. 6B, lane 2, i.e. no competition, to lanes 6−8 i.e. with competitor DNA).

We further examined whether CArG-2 and the *c-fos* SRE interfered with binding of nuclear factors to CArG-1, and found that binding reactions containing unlabeled CArG-2 (Fig. 6A,

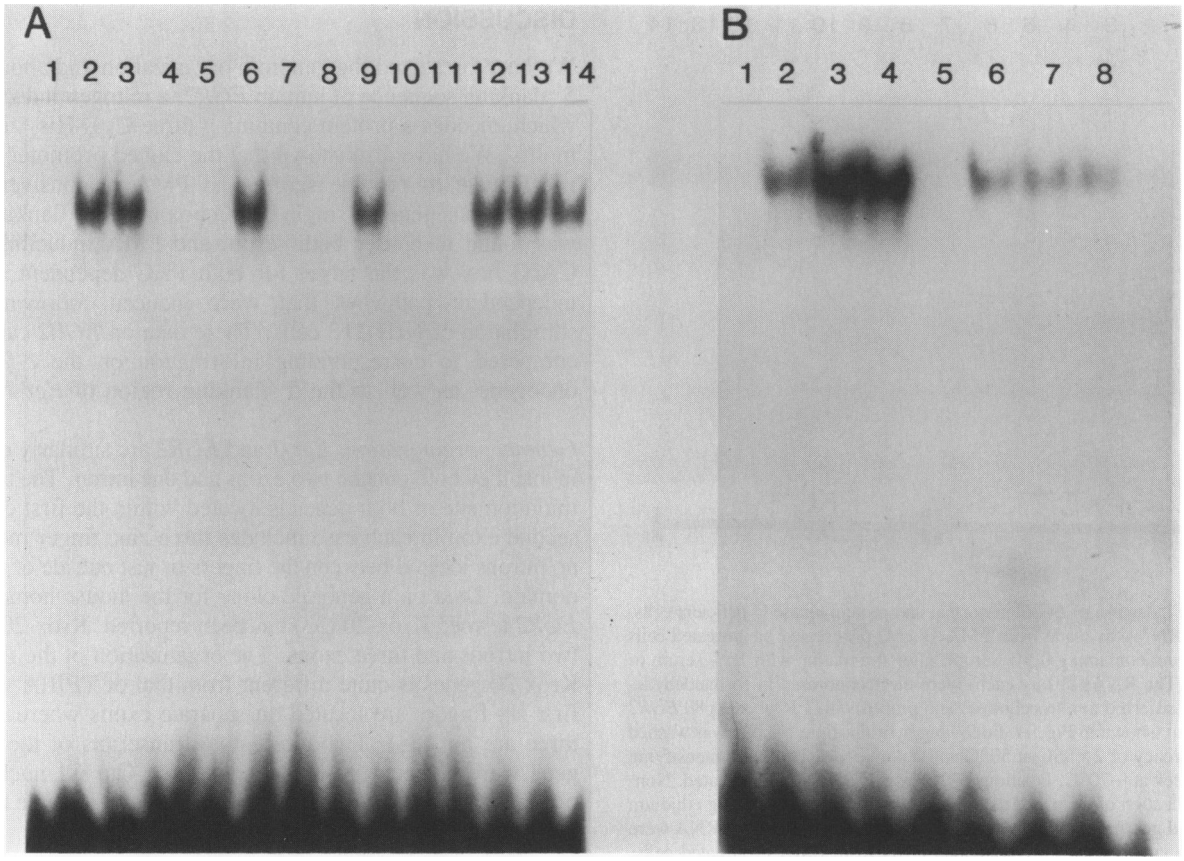

Fig. 6. Gel mobility shift of CArG-1 due to specific binding of nuclear factors. Binding reactions containing 1 ng of ^{32}P end-labeled CArG-1 probe and either no extract or 10 μg (unless otherwise specified) of extract were electrophoresed on 4% acrylamide gels. For examining competition binding, molar excess of unlabeled CArG fragments were included in the reaction mixtures. Dried gels were exposed overnight at −70°C with an intensifying screen. A. lane 1: probe alone; 2: probe plus extract and no competition; 3−5: competition with CArG-1; 6−8: competition with *c-fos* SRE; 9−11: competition with CArG-2; 12−14: competition with *c-fos* SRE mutant. [Unlabeled competitor DNA was in 10-fold molar excess (A. lanes 3; 6; 9; and 12), 60-fold molar excess (A. lanes 4; 7; 10; and 13), or 120-fold molar excess (A. lanes 5; 8; 11; and 14) over probe DNA.] B. lane 1: probe alone; 2: probe plus 4 μg extract and no competition; 3: probe plus 10 μg extract and no competition; 4: probe plus 20 μg extract and no competition; 5: probe and 4 μg extract and competition with 60-fold molar excess of CArG-1; 6−8: probe plus 4 μg extract and competition with CArG-1 mutant at 10, 60, and 120-fold molar excess respectively.

lanes 9−11) or the *c-fos* SRE (Fig. 6A, lanes 6−8) abolished labeled CArG-1-protein interactions at 60-fold and 120-fold molar excess (Fig. 6A, lanes 10, 11 and 7, 8). The mutant SRE of *c-fos*, on the other hand, did not compete with binding of labeled CArG-1 (Fig. 6A, lanes 12−14). These data suggest that the same nuclear factor, most likely the SRF, binds predominantly to CArG-1, CArG-2, and the SRE of *c-fos*.

The CArG-1 element confers both serum and PMA responsiveness. Besides serum, other mitogenic stimuli such as PMA can induce *EGR2* in human fibroblasts and lymphocytes (11). PMA activates the protein kinase C pathway, and affects transcription of genes containing a PMA responsive element, i.e. the AP-1 binding site. Three putative AP-1 binding sites were identified in the 5′ region of *EGR2*. We, therefore, asked whether these elements mediate PMA responsiveness of the cloned *EGR2* promoter. Transient expression studies indicated that CAT activity was inducible in pCAT16, which contained the three putative AP-1 binding elements. However, pCAT20 in which the three AP-1 binding elements were deleted retained PMA responsiveness (Fig. 5) suggesting that the AP-1 binding elements were not essential for PMA inducible transcription. Plasmid pCAT40 contains one putative cAMP binding site and the CArG-1 box. Of these two elements, CArG-1 was specifically examined for PMA responsiveness because recent studies (32) in which the SRE of *c-fos* was mutated, had revealed that it contributes responsiveness to both serum and PMA. Our construct, pCAT50, retains sequences in the 5′ region of *EGR2* intact except for CArG-1, which was altered by site directed mutagenesis without affecting the distance between the response elements and the cap site. This plasmid had lost the ability to confer PMA responsiveness (Fig. 5). These data imply that the CArG-1 sequence of *EGR2* is crucial for responsiveness to both serum and PMA.

Serum induces EGR2 in protein kinase C deficient cells. Serum contains a mixture of growth factors, which utilize both PKC dependent and independent second messenger systems for signal transduction. We therefore examined whether serum could activate *EGR2* by non-protein kinase C pathways. NIH 3T3 cells were grown for 40 hours in the presence of 0.5% serum or high concentrations of PMA (1 μM) and 0.5% serum, and following induction with PMA (50 nM) or 10% serum for different periods of time, total RNA was extracted from the cells. The RNA was electrophoresed, transferred to nitrocellulose and probed with ^{32}P-labeled p80 fragment of *EGR2* cDNA (Fig. 1). Uninduced PKC+ cells (Fig. 7, lane 8) expressed low levels of *EGR2*. However, following treatment with PMA (50 nM) or 10% serum,

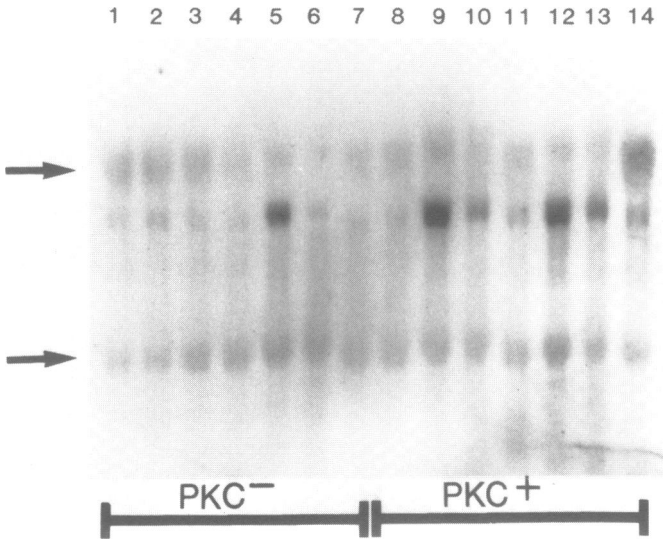

Fig. 7. Serum induction of *EGR2* in normal and protein kinase C deficient cells. Total cellular RNA was made from PMA (1 μM) pre-treated or untreated cells grown in medium containing 0.5% serum, after stimulation with 10% serum or 50 nM PMA. The RNAs (10 μg each) were electrophoresed in formaldehyde-agarose gel, transferred to nitrocellulose, and probed with ^{32}P-labeled p80 *EGR2* cDNA fragment (shown in Fig. 1). Following hybridization, the blot was washed at a final stringency of 2×SSC at 50°C and autoradiographed with intensifying screen for 1 week at −70°C. Positions of 28S and 18S RNA are indicated. Non-specific hybridization of the probe to 18S RNA, and the corresponding ethidium bromide stained gel (not shown) indicated that equivalent amounts of RNA were loaded in each lane. Lanes 1−7 contained RNA from PMA pre-treated cells. Lanes 8−14 contained RNA from cells that were not pre-treated with PMA. Lanes 1, and 8 are 0 hour samples, contain RNA from PMA stimulated cells.Stimulation with PMA (lanes 2−4, and 9−11) or serum (lanes 5−7, and 12−14) was done for 1 hour (lanes 2, 5, 9, and 12); 3 hours (lanes 3, 6, 10, and 13); or 5 hours (lanes 4, 7, 11, and 14). PKC$^+$='normal' cells; PKC$^-$=protein kinase C depleted cells.

transient induction of *EGR2* was seen at 1 hr (Fig. 7, lanes 9 and 12) and the message levels then dropped at 3 hours (lanes 10 and 13) and 5 hours (Fig. 7, lanes 11 and 14). Inactivation of the protein kinase C pathway in these cells was confirmed by the inability of PMA, that was subsequently added, to induce *EGR2* (Fig. 7, lanes 2, 3 and 4). The protein kinase C deficient cells, however, retained the ability to induce *EGR2* expression following serum stimulation (Fig. 7, lane 5), implying that non-protein kinase C dependent pathways are also involved in *EGR2* induction.

The EGR2 CArG-1 element responds to PKC dependent and independent pathways. We have demonstrated earlier that, in transient expression assays, the CArG-1 sequence of *EGR2* is essential for PMA induction, a mechanism which is known to involve activation of the protein kinase C pathway. When cells transfected with pCAT16, pCAT20, or pCAT40 were made protein kinase C deficient and then treated with serum, they retained serum responsiveness (Fig. 5). In contrast, pCAT50 transfected cells, when rendered protein kinase C deficient by 40 hours of PMA treatment, lost serum responsiveness in the transient assays. Taken together, these data imply that protein kinase C-dependent as well as protein kinase C-independent pathways operate through the CArG-1 box in order to effect *EGR2* transcription.

DISCUSSION

We have presented the genomic organization and about 1 kb of 5′ flanking sequence of human *EGR2*, a mitogen inducible gene which encodes a protein containing three Cys_2-His_2 zinc finger motifs. We have also shown that the cloned promoter of *EGR2* is adequate to provide serum and PMA responsiveness to a heterologous gene. A single CArG box in the 5′ flanking region was found to confer both serum and PMA inducibility. This CArG box was the target for both PKC dependent and PKC independent pathways that were induced following serum stimulation of NIH 3T3 cells. These data on *EGR2* can now be compared to corresponding information on the *c-fos* proto-oncogene, as well to the 5′ flanking region of *Egr-1*.

Genomic organization. Egr-1 and *EGR2* are similarly organized in that they both contain two exons and one intron. The translation initiation site of both genes is located within the first exon. The second exon in each case includes three zinc finger motifs with no introns located between the fingers or just outside of the finger domain. Data on a genomic clone for the mouse homologue of *EGR2* termed Krox-20 (33) has been reported. Krox-20 contains two introns and three exons. The organization of the *EGR2* and Krox-20 genes is quite different from that of TFIIIA where the first six fingers are located on separate exons whereas the last three are not (34). Two alternative transcripts of the Krox-20 gene were detected in mouse cells (33). Our S1 nuclease and primer extension data indicate the presence of a single transcript for *EGR2*.

5′ flanking sequence. Egr-1, EGR2, and *c-fos* are all mitogen inducible genes in a variety of cell types. In the case of *c-fos*, this effect is mediated by 22 bp element termed serum response element (SRE) (30). The inner core of 10 nucleotides (CC followed by six A + T's followed by GG) has been referred to as a CArG element and has been found in the *X. laevis* cytoskeletal actin (35) and human cardiac actin promoters (36). The *Egr-1* upstream sequence contains six such elements (6, 21). (In addition to the five indicated in our paper (21), there is another one at position −83). It is of interest that *EGR2* mRNA, which is induced by serum to peak levels comparable to those of *c-fos* but five-to-ten fold less than those seen for *Egr-1*, has only two such elements. The difference in the number of elements is reflected in our *EGR2* CAT expression data when compared with that of *Egr-1* (Cao et al., submitted). The CArG elements as well as the other putative regulatory elements underscored in Fig. 2 are conserved in the mouse Krox-20 sequence. Deletion analysis of the 5′ flanking region of *EGR2* revealed that the CArG-1 element was sufficient for serum responsiveness. Mutagenesis of this CArG-1 element completely abolished serum responsiveness of the cloned *EGR2* promoter, thereby indicating that this element is critical for serum responsiveness.

EGR2 expression is also inducible by PMA in human peripheral blood lymphocytes (11) and in mouse NIH 3T3 cells (present study). Previous studies have suggested that the effect of PMA is mediated by binding of a fos-jun heterodimer to a *cis* element of the form [C/G]TGACT[C/A]A (27, 28, 37−41). However, the deletion construct pCAT20, in which the region containing all the putative AP-1 binding sites was eliminated, still responded to PMA, indicating that the AP-1 binding elements were unessential for PMA responsiveness. We then asked whether the CArG element contained within the sequence of pCAT20

conferred PMA responsiveness, because the SRE element of *c-fos* was recently shown to confer PMA responsiveness as well. Indeed, construct pCAT50, in which the CArG-1 box was mutagenized, failed to respond to PMA, indicating that the CArG-1 box of *EGR2* dictates PMA responsiveness, in addition to serum responsiveness. Furthermore, our data indicates that extended dyad symmetry (present in the *c-fos* SRE but not surrounding the CArG-1 box of *EGR2*) is not a prerequisite for serum and PMA responsiveness.

Multiple pathways converge on CArG-1 box. Serum is a complex mixture of growth factors that activate several distinct intracellular signalling pathways. Through mechanisms yet to be defined, it activates a protein termed serum response factor (SRF) which subsequently binds to the SRE of *c-fos* to induce expression of the gene. The gel mobility shift data from the present study suggested that the SRF might also be binding the CArG-1 box of *EGR2*. PMA which mainly activates a PKC pathway was found to activate *EGR2* expression through the CArG-1 box. We, therefore, asked whether non-PKC pathways also induce *EGR2* expression, and if so, whether they operate through the CArG-1 box or other sequences. Northern analysis of PKC depleted cells revealed that serum could still induce *EGR2* expression. Functional studies of the cloned promoter indicated that mutagenesis of the CArG-1 box completely abolished *EGR2* inducibility with serum as well as PMA in PKC+ and PKC depleted cells. Taken together, these data indicate that the *EGR2* promoter is activated by both PKC^+ and PKC^- pathways, all converging on the CArG-1 element.

ACKNOWLEDGMENTS

We would like to thank D. Gius for the CAT-poly construct, and C. Deshich for tissue culture work. Thanks are also due to L. Joseph, X. Cao and C. -H. Tsai-Morris for helpful suggestions.

REFERENCES

1. Lau,L.F., and Nathans,D. (1985) EMBO J. **4**,3145-3151.
2. Lau,L.F., and Nathans,D. (1987) Proc. Natl. Acad. Sci. USA **84**,1182-1186.
3. Verma,I.M., and Graham,W.R. (1987) Adv. Cancer Res. **49**,29-52.
4. Cohen,D.R., and Curran,T. (1988) Mol. Cell. Biol. **8**,2063-2069.
5. Zerial,M., Toschi,L., Ryseck,R.P., Schuermann,M., Muller,R., and Bravo,R. (1989) EMBO J. **8**,805-813.
6. Christy,B.A., Lau,L.F., and Nathans,D. (1988) Proc. Natl. Acad. Sci. USA **85**,7857-7861.
7. Lemaire,P., Revelant,O., Bravo,R., and Charnay,P. (1988) Proc. Natl. Acad. Sci. USA **85**,4691-4695.
8. Milbrandt,J. (1987) Science **238**,797-799.
9. Sukhatme,V.P., Cao,X., Chang,L.C., Tsai-Morris,C.-H., Stamenkovich,D., Ferreira,P.C.P., Cohen,D.R., Edwards,S.A., Shows,T.B., Curran,T., Le Beau,M.M., and Adamson,E.D. (1988) Cell **53**,37-43.
10. Chavrier,P., Zerial,M., Lemaire,P., Almendral,J., Bravo,R., and Charnay,P. (1988) EMBO J. **7**,29-36.
11. Joseph,L.J., Le Beau,M.M., Jamieson Jr.,G.A., Acharya,S., Shows,T.B., Rowley,J.D., and Sukhatme,V.P. (1988) Proc. Natl. Acad. Sci. USA **85**,7164-7168.
12. Hazel,T.G., Nathans,D., and Lau,L.F. (1988) Proc. Natl. Acad. Sci. USA **85**,8444-8448.
13. Milbrandt,J. (1988) Neuron **1**,183-188.
14. Bohmann,D., Bos,T.J., Admon,A., Nishimura,T., Vogt,P.K., and Tjian,R. (1987) Science **238**,1386-1392.
15. Lamph,W.W., Wamsley,P., Sassone-Corsi,P., and Verma,I.M. (1988) Nature (London) **334**,629-631.
16. Ryder,K., and Nathans,D. (1988) Proc. Natl. Acad. Sci. USA **85**,8454-8467.
17. Ryseck,R.-P., Harai,S.I., Yaniv,M., and Bravo,R. (1988) Nature (London) **334**,535-537.
18. Nakabeppu,Y., Ryder,K., and Nathans,D. (1988) Cell **55**,907-915.
19. Seyfert,V.L., Sukhatme,V.P., and Monroe,J.G. (1989) Mol. Cell. Biol. **9**,2083-2088.
20. Sukhatme,V.P., Kartha,S., Toback,F.G., Taub,R., Hoover,R.G., and Tsai-Morris,C.-H. (1987) Oncogene Res. **1**,343-355.
21. Tsai-Morris,C.-H., Cao,X., and Sukhatme,V.P. (1988) Nucleic Acids Res. **16**,8835-8846.
22. Wilkinson,D.G., Bhatt,S., Chavrier,P., Bravo,R., and Charnay,P. (1989) Nature **337**,461-464.
23. Sanger,F., Nicklen,S., and Coulson,A.R. (1977) Proc. Natl. Acad. Sci. USA **74**,5463-5467.
24. Gorman,C.M., Moffat,L.F., and Howard,B.H. (1982) Mol. Cell. Biol. **2**,1044-1051.
25. Chomczynski,P., and Sacchi,N. (1987) Anal. Biochem. **162**,156-159.
26. Mount,S.M. (1982) Nucleic Acids Res. **10**,459-472.
27. Angel,P., Imagawa,M., Chiu,R., Stein,B., Imbra,R.J., Rahmsdorf,H.J., Jonat,C., Herrlich,P., and Karin,M. (1987) Cell **49**,729-739.
28. Lee,W., Mitchell,P., and Tjian,R. (1987) Cell **49**,741-752.
29. Montminy,M.R., and Bilezsikjian,L.M. (1987) Nature (London) **328**,175-178.
30. Treisman,R. (1986) Cell **46**,567-574.
31. Triesman,R. (1987) EMBO J. **6**,2711-2717.
32. Gilman,M.Z. (1988) Genes Dev. **2**,394-402.
33. Chavrier,P., Janssen-Timmen,U. Mattei,M.G., Zerial,M., Bravo,R., and Charnay,P. (1989) Mol. Cell. Biol. **9**,787-797.
34. Tso,J.Y., Van Den Berg,D.J., and Korn,L.J. (1986) Nucleic Acids Res. **14**,2187-2200.
35. Mohun,T., Garrett,N., and Treisman,R. (1987) EMBO J. **6**,667-673.
36. Gustafson,T.A., Miwa,T., Boxer,L.M., and Kedes,L. (1988) Mol. Cell. Biol. **8**,4110-4119.
37. Chiu,R., Boyle,W.J., Meek,J., Smeal,T., Hunter,T., and Karin,M. (1988) Cell **54**,553-564.
38. Halazonetis,T.D., Georgopoulos,K., Greenberg,M.E., and Leder,P. (1988) Cell **55**,917-924.
39. Kouzarides,T., and Ziff,E. (1988) Nature **336**,646-651.
40. Rauscher (III),F.J., Voulalas,P.J., Franza, Jr.,B. R., and Curran,T. (1988) Genes Dev. **2**,1687-1699.
41. Sassone-Corsi,P., Ransone,L.J., Lamph,W.W., and Verma,I.M. (1988) Nature **336**,692-695.