MOLECULAR AND CELLULAR BIOLOGY, Aug. 1991, p. 3835–3841 Vol. 11, No. 8
0270-7306/91/083835-07$02.00/0

The Early Response Gene NGFI-C Encodes a Zinc Finger Transcriptional Activator and Is a Member of the GCGGGGGCG (GSG) Element-Binding Protein Family

SETH D. CROSBY, JOHN J. PUETZ, KELLI S. SIMBURGER, TIMOTHY J. FAHRNER, AND JEFFREY MILBRANDT*

Departments of Pathology and Internal Medicine, Division of Laboratory Medicine, Box 8118, Washington University School of Medicine, 660 South Euclid Avenue, St. Louis, Missouri 63110

Received 13 March 1991/Accepted 24 April 1991

We have cloned NGFI-C, a nerve growth factor-induced early-response gene which encodes a Cys_2/His_2 zinc finger protein. RNA blot analysis demonstrates that NGFI-C mRNA is induced within minutes of stimulation of PC12 cells by nerve growth factor and is similarly activated in brain after a Metrazol-induced seizure. The cDNA sequence predicts a protein that contains three zinc fingers which show striking homology to the DNA-binding regions of three previously reported zinc finger proteins, NGFI-A, Krox-20, and the Wilms' tumor gene product. NGFI-C binds to the previously described DNA-binding site of these three proteins, which is GCGGGGGCG. Cotransfection experiments revealed that NGFI-C strongly activates transcription from this site in mammalian cells. The isolation of another early-response gene that encodes a member of the G(C/G)G or GSG element-binding family should provide an opportunity to investigate the relative contributions of a family of transcription factors to the cell's response to changes in its environment.

Factors which promote long-term cellular responses such as growth and differentiation induce an orchestrated change in the pattern of gene expression. Initially, the expression of a small subset of genes, termed early-response genes (ERGs) or immediate-early genes, is altered. The induction of ERG expression is mediated by a very rapid but transient transcriptional activation that is independent of de novo protein synthesis. The diversity of signals to which ERGs respond and the realization that many of them encode transcription factors, such as c-*fos* (13), NGFI-A (27) (also called *egr1* [46] and *zif1268* [11]), NGFI-B (28)/*nur77* (19), and c-*jun* (37), suggest that their products may constitute the nuclear arm of the signal transduction process. Because of their possible role in coupling changes in the extracellular environment to long-term cellular responses, the expression of ERGs in the nervous system is being intensely studied. In the central nervous system, ERG expression is modulated by seizure induction (30), sensory stimulation (48), induction of long-term potentiation (12), and shifts in the circadian rhythm (36). In PC12 cells, a frequently used model of neuronal differentiation, a variety of stimuli, such as exposure to nerve growth factor (NGF) (27) and neurotransmitters (18) or membrane depolarization (2), induce the expression of many ERGs.

A number of ERGs expressed in the nervous system encode zinc finger proteins. Two such proteins are NGFI-A and Krox-20 (8), both of which contain very similar, sequence-specific DNA-binding domains composed of three tandemly linked zinc fingers. The zinc finger is a highly conserved protein domain of 28 to 30 amino acid residues containing several invariant residues (29). The most notable of these are pairs of cysteine and histidine residues that stabilize the regional conformation of the finger via coordinate binding to a zinc(II) ion. The high level of conservation within this motif and in the linker region which joins adjacent zinc fingers has enabled the isolation of a number of genes encoding zinc finger proteins by low-stringency hybridization methods (6). More recently, the polymerase chain reaction (PCR) has been exploited to isolate additional zinc finger-encoding sequences (33).

We report here the use of PCR to isolate another ERG cDNA, NGFI-C, that encodes a zinc finger protein which is rapidly induced in PC12 cells by NGF and in brain by seizure activity. Nucleotide sequence analysis revealed that NGFI-C contains three tandemly linked zinc fingers very similar to those present in Sp1 (21), NGFI-A, Krox-20, and the Wilms' tumor gene product (5). We also demonstrate that NGFI-C specifically binds to the nonamer sequence, GCGGGGGCG, recognized by NGFI-A, Krox-20, and the Wilms' tumor gene product (34). Finally, we show that NGFI-C activates transcription in mammalian cells from luciferase reporter plasmids bearing this recognition sequence, thereby establishing NGFI-C as another member of this transcription factor family.

MATERIALS AND METHODS

Isolation of NGFI-C cDNA. RNA from PC12 cells treated with NGF and cycloheximide (CHX) was isolated and reverse transcribed to first-strand cDNA by using an oligo(dT) primer. This pool of cDNA was then amplified by PCR with a forward primer ACIGG(G/C)GAGAAGCC(G/C)T(T/A)(C/T)G(A/C)ITG (where I is inosine) derived from the consensus sequence of the linker region (the well-conserved residues upstream of the first Cys of the zinc finger) and the oligo(dT) reverse primer. The PCR was performed with an annealing temperature of 42°C and an extension time of 2 min at 72°C. The resulting PCR products were cloned into the *Sma*I site of pBS(KS) (Stratagene, La Jolla, Calif.). A ^{32}P-labeled probe (17) was prepared from one of these

---

* Corresponding author.

clones, called NGFI-C, and used to screen a cDNA library prepared from RNA isolated from PC12 cells cultured in the presence of NGF and CHX for 3 h (27). A clone with a 2.1-kb cDNA insert was isolated, and the cDNA fragment was subcloned into the *Eco*RI site of pBS(KS). Nucleotide sequencing was performed by the chain termination method (38), using Sequenase (United States Biochemical) as specified by the supplier.

RNA preparation and analysis. PC12 cells were cultured as previously described (26). When cells were treated with NGF or CHX, the final concentrations were 50 ng/ml and 10 μg/ml, respectively. Seizure was induced in adult Wistar rats by an intraperitoneal injection of pentylenetetrazole (Metrazole) at a dose of 50 mg/kg. Forty-five minutes after seizure induction, animals were sacrificed by decapitation and the brains were rapidly removed and frozen on dry ice. Poly(A)-enriched RNA was isolated with the Fastrack Kit (Invitrogen, La Jolla, Calif.). RNA transfer analysis was performed as previously described, using a ^{32}P-labeled antisense RNA probe generated from the NGFI-C 3′ untranslated region (nucleotides [nt] 2110 to 1751) by using T7 RNA polymerase as instructed by the manufacturer (Promega).

DNA binding analysis. Proteins for DNA binding analysis were produced as fusion proteins by linking them to the bacterial TrpE protein encoded in the pATH-3 expression plasmid (15). Three restriction fragments of the NGFI-C cDNA were cloned into the pATH-3 expression plasmid to create the following constructs: pNCFL, which includes the entire NGFI-C protein (nt 1 to 1625); pNCZF, which contains the carboxy-terminal domain including the three zinc fingers (nt 1122 to 1625); and pNCAM, which contains the amino-terminal domain and lacks the zinc fingers (nt 1 to 800). *Escherichia coli* DH5α bacteria harboring these constructs or nonrecombinant pATH-3 were grown, and the fusion proteins were induced as previously described (16). The NGFI-A cDNA was cloned into the pET3d vector, and full-length NGFI-A was expressed as described previously (44). After induction of the fusion proteins, the bacteria were pelleted and resuspended in denaturing protein loading buffer (2% sodium dodecyl sulfate [SDS], 100 mM Tris [pH 7.5], 280 mM β-mercaptoethanol, 20% glycerol). After the suspension was heated at 100° for 5 min, the proteins were separated by electrophoresis through a 10% SDS-polyacrylamide gel and transferred to nitrocellulose. The protein blot was incubated for 2 h at 25°C in 5% nonfat dry milk in renaturation buffer (100 mM KCl, 25 mM *N*-2-hydroxyethylpiperazine-*N*′-2-ethanesulfonic acid [HEPES; pH 7.8], 10 μM ZnCl, 1 mM dithiothreitol, 20% glycerol), washed briefly three times in renaturation buffer without milk, and incubated in renaturation buffer containing salmon sperm DNA (1 μg/ml) and a 1,000-fold excess of unlabeled oligonucleotides (either specific or unrelated) for 30 min at 25°C with gentle agitation. A double-stranded oligonucleotide containing two copies of the NGFI-A DNA-binding site (GGATCCAGCGGGGGCGAGCGGGGGCGA) was end labeled with [γ-^{32}P]ATP and T4 polynucleotide kinase to a specific activity of 2×10^8 cpm/μg. The labeled probe was added (final concentration, 10^6 cpm/ml), and binding was allowed to proceed at 25°C for 30 min. The protein blot was washed three times in renaturation buffer for 5 min and subjected to autoradiography. In addition, an identical blot was probed in the same manner with a labeled double-stranded oligonucleotide whose sequence (GTTTT AAAAGGTCATGCTGACCTGACCCGTA) is unrelated to the NGFI-A binding site.

Mammalian transfection. The pCMV mammalian expression plasmid (pCB6) was obtained from M. Roth (University of Texas, Southwestern). pCMV-NGFI-C was constructed in two steps. First, the majority of the 3′ untranslated region (nt 1625 to 2103) was deleted by digestion with *Tth*111I and *Eco*RV; the ends were blunted with Klenow fragment and deoxynucleoside triphosphates, and recircularization of the plasmid was done by ligation. The truncated NGFI-C cDNA was excised from pBS with *Bam*HI and *Cla*I and ligated into the *Bgl*II and *Cla*I sites of pCMV. The reporter vector, Pro36-Luc (obtained from S. Adler, Washington University), contains the firefly luciferase coding region downstream of a minimal prolactin promoter (1). Oligonucleotides containing either one or two NGFI-A binding sites were inserted upstream of the prolactin promoter at the *Bam*HI site of Pro36-Luc.

Cos-7 cells were grown in Dulbecco's modified Eagle's medium supplemented with 10% fetal calf serum. Cells were plated 24 h before transfection at a density of 700,000 cells per 10-cm dish. The cells were transfected by $CaPO_4$ precipitation essentially as described previously (9), using a total of 10 μg of plasmid DNA per plate (5 μg of activator plasmid and 5 μg of reporter plasmid). Three days after transfection, the cells were washed twice with ice-cold phosphate-buffered saline and lysed by incubation in 0.7 ml of 50 mM Tris-morpholine ethanesulfonic acid (pH 7.8)–1 mM dithiothreitol–1% Triton X-100 for 5 min at 25°C. The lysate was cleared of cellular debris by centrifugation. Luciferase assays were performed on 50 μl of cell lysate, using a Monolight 2010 luminometer (Analytical Bioluminescence Laboratory, San Diego, Calif.) (14).

RESULTS

Isolation of the NGFI-C cDNA. To identify genes which encode zinc finger proteins and are rapidly activated by NGF, a reverse transcriptase/PCR method was used. Template RNA was isolated from PC12 cells treated for 3 h with NGF and CHX and reverse transcribed by using an oligo(dT) primer. For the subsequent PCR, the forward primer [AC IGG(G/C)GAGAAGCC(G/C)T(T/A)(C/T)G(A/C)ITG] was a consensus of the highly conserved linker region between zinc fingers (encoding amino acids TGQKPYDC) from a number of zinc finger proteins. The reverse primer was identical to that used to generate the first-strand cDNA. Fragments produced by this PCR reaction were cloned, and the nucleotide sequences of many of the resultant clones were determined. The sequence of one of the cDNA fragments predicted a peptide containing a zinc finger motif that was very similar to that of the ERGs NGFI-A and Krox-20. This cDNA fragment was therefore used to screen a library constructed from RNA prepared from PC12 cells treated with NGF and CHX for 3 h. Several clones were isolated, and the largest, called NGFI-C, was used for further characterization.

NGFI-C is an ERG. To examine the induction time course for the NGFI-C mRNA, poly(A)$^+$ RNA was isolated from PC12 cells treated with NGF for various lengths of time. RNA blot analysis with a probe derived from the NGFI-C 3′ untranslated region (see below) showed that the basal level of transcript was very low, but that within 30 min after NGF administration the level of NGFI-C mRNA was greatly increased (Fig. 1A). NGFI-C mRNA peaked at 60 min and had rapidly declined to its low basal level by 3 h. In contrast, if CHX was present when NGF was administered, NGFI-C mRNA levels remained elevated for an extended

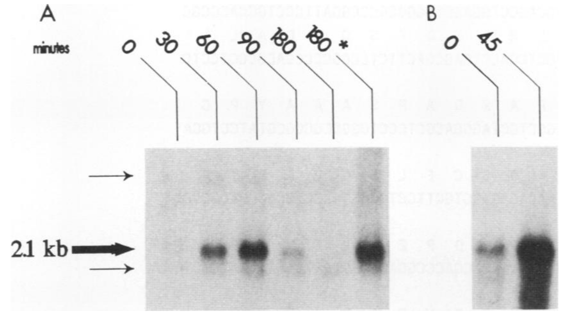

FIG. 1. Evidence that NGFI-C is an ERG. (A) RNA transfer analysis of 5 μg of poly(A)$^+$ RNA isolated from PC12 cells treated with NGF for the indicated time. The lane labeled 180* indicates that CHX (10 μg/ml) was also present during the NGF treatment. (B) Analysis of 15 μg of total RNA isolated from rat brain before or 45 min after a Metrazol-induced seizure. Small arrows indicate the 28S and 18S ribosomal bands. Large arrow indicates the position of the NGFI-C mRNA.

period. The kinetics of induction and disappearance of NGFI-C mRNA and its persistence in the presence of protein synthesis inhibitors are similar to those observed for other ERGs (28).

Many ERGs, including c-*fos*, NGFI-A, NGFI-B, *fra-1*, and *jun-B*, are activated in the brain by seizure (for a review, see reference 41). Because the NGFI-C gene is regulated similarly to these genes in PC12 cells, we investigated whether it would also be activated by seizure. RNA blot analysis demonstrated that the level of NGFI-C mRNA in normal brain was moderately abundant and that seizure activity resulted in a large increase in NGFI-C transcripts within 45 min (Fig. 1B).

Sequence analysis of the NGFI-C cDNA. The nucleotide sequence of the NGFI-C cDNA was determined (Fig. 2). The 5′ end of the NGFI-C mRNA was identified by primer extension and confirmed by S1 analysis with a genomic clone containing the 5′ region (data not shown); both procedures revealed that transcriptional initiation occurs 53 nt upstream from the 5′ end of the cDNA clone. Analysis of the NGFI-C cDNA sequence revealed that the mRNA is 2,093 nt long excluding the poly(A) tail. It contains a 138-nt 5′ untranslated region, followed by an open reading frame beginning at the Met codon at nt 139 and terminating at nt 1573, followed by a 521-nt-long 3′ untranslated region. A typical polyadenylation signal is present 11 nt upstream of the poly(A) tail. The sequence predicts a polypeptide of 478 amino acids with an unmodified molecular mass of 49,667 Da. The most striking feature of the predicted protein is the presence of three zinc fingers of the Cys$_2$-His$_2$ subtype near the carboxy terminus. This highly conserved motif is a well-characterized DNA-binding domain that is found in a number of transcription factors.

Inspection of the NGFI-C zinc finger domain revealed that it is strikingly similar to the DNA-binding domains of NGFI-A and Krox-20. The homologies between this region of NGFI-C and the corresponding regions of NGFI-A and Krox-20 are 85 and 81%, respectively. However, the similarities between the DNA-binding domains of these proteins and the conserved, intragenic differences in each individual finger motif are more apparent when respective zinc fingers of each protein are compared with one another (Fig. 3). For example, in each protein, the first zinc finger contains four residues between the cysteine pair whereas the cysteines of the other two fingers are separated by only two amino acids. We have also noted that the third zinc finger is very atypical in these proteins because it lacks the invariant leucine located 10 residues downstream from the second cysteine. This leucine is thought to interact with other hydrophobic residues within the motif to stabilize the zinc finger structure (3). Recently Nardelli et al. (32) identified two residues within the zinc fingers of Krox-20 and Sp1 that dictate whether these particular finger motifs recognize the sequence GCG or GGG. It should be noted that both NGFI-C and NGFI-A are identical to Krox-20 at these apparently critical positions, thus predicting that NGFI-C should also recognize the GCGGGGGCG nonamer.

Outside of the zinc finger region, the NGFI-C protein bears no similarities to either NGFI-A or Krox-20, just as those two proteins show little homology to each other outside of their respective DNA-binding domains. However, as has been observed for a number of DNA-binding proteins, NGFI-C has an unusually high proline content. The proline composition is especially high (25%) in a region extending from residues 111 through 188. This region is similar to the transcriptional activating domains present in CCAAT transcription factor (25), AP-2 (20), Oct-2 (22), and c-Jun (43) and may play an analogous role in NGFI-C. A further search of GenBank by using sequences outside of the zinc finger domain revealed no significant homology to previously described proteins.

NGFI-C recognizes the sequence element GCGGGGGCG. The nucleotide sequence GCGGGGGCG serves as a recognition site for NGFI-A (10), and an essentially identical element was determined to be a cognate site for Krox-20 (7). Because of the similarities between the zinc fingers of these proteins and NGFI-C, we used Southwestern (DNA-protein) blot analysis (42) to explore whether NGFI-C would also recognize this sequence element. Restriction fragments encoding the entire NGFI-C protein or portions of it corresponding to the amino-terminal (non-zinc finger containing) or carboxy-terminal (containing the zinc fingers) domains were cloned into the pATH vector (15). Bacterial lysates containing the corresponding fusion proteins or NGFI-A were electrophoresed on SDS-polyacrylamide gels and transferred to nitrocellulose. The proteins were renatured and incubated with a ^{32}P-labeled oligonucleotide containing the sequence GCGGGGGCG and a 1,000-fold excess of nonradioactive unrelated oligonucleotide. Binding of the oligonucleotide to fusion proteins containing the NGFI-C or NGFI-A zinc fingers was detected by autoradiography (Fig. 4). The specific nature of this interaction was determined by showing that no binding was detected when the incubation was performed in the presence of a 1,000-fold excess of nonradioactive oligonucleotide containing the GCGGGGGCG sequence (data not shown). In addition, no signal was detected when an identical protein blot was incubated with ^{32}P-labeled oligonucleotide which did not contain the GCGGGGGCG sequence (data not shown). The results clearly demonstrate that NGFI-C specifically recognizes the same nucleotide sequence as NGFI-A, Krox-20, and the Wilms' tumor gene product.

NGFI-C is a transcriptional activator. NGFI-A and Krox-20 both activate transcription from the GCGGGGGCG sequence element in *Drosophila* Schneider cells (7, 24). To test whether NGFI-C could activate transcription when bound to this sequence, a Cos cell cotransfection transactivation assay was used. For these experiments, oligonucleotides containing one or two copies of the GCGGGGGCG recog-

```
   1  CGAGCCAGGGCGCCGCGGCCCCCCAACCCTCGTCCCCGAGCCAAGCGCCGGGAGCCTGGAGCTGGGGCGCCCGGATTCCCTGCCAGCCGC
   1                                                  M  L  H  L  S  D  F  S  G  P  D  A  L  L
  91  GCGCGCCCCAGCCCAGCGAGGCCCCGGGCGCCCCGCCGCCACCACGCCATGCTCCACCTGAGCGACTTCTCCGGCCCCGACGCGCTCCTC

  15  S  K  P  T  E  G  C  A  H  T  S  P  E  L  P  R  L  P  A  R  D  A  P  S  A  A  A  Y  P  G
 181  TCCAAGCCCACCGAAGGCTGCGCCCACACCAGCCCCGAGCTGCCCCGGCTGCCTGCTAGGGACGCTCCCTCGGCCGCCGCGTATCCTGGA

  45  G  D  F  L  S  W  A  L  S  T  C  G  A  G  G  D  L  T  D  S  C  F  L  E  G  P  A  P  T  P
 271  GGCGACTTCTTGAGCTGGGCTCTGAGCACCTGCGGCGCCGGGGGGGACTTAACAGACTCCTGCTTCCTGGAGGGCCCTGCACCCACGCCC

  75  P  S  G  L  S  Y  S  G  S  F  F  I  Q  A  V  P  E  H  P  H  D  P  E  A  L  F  N  L  M  S
 361  CCTTCGGGCCTCAGCTACAGCGGCAGCTTCTTCATCCAGGCGGTTCCCGAACACCCGCACGACCCGGAGGCCCTCTTCAACCTCATGTCT

 105  G  I  L  G  L  A  P  F  P  S  P  E  A  A  A  S  R  S  P  L  D  V  P  F  P  A  G  P  D  A
 451  GGCATCTTGGGCTTGGCACCCTTCCCTAGCCCCGAGGCGGCAGCGTCTCGGTCCCCCCTGGATGTCCCTTTCCCCGCGGGTCCCGATGCC

 135  L  L  P  D  L  Y  S  P  D  L  S  S  A  A  F  P  E  A  F  W  E  A  A  P  S  A  G  A  P  S
 541  TTGCTGCCGGACCTTTACTCCCCGGATCTGAGTTCGGCCGCCTTCCCGGAGGCGTTTTGGGAGGCCGCGCCTTCGGCGGGCGCTCCCTCG

 165  Q  C  L  F  E  P  Q  L  S  P  P  D  V  K  P  G  L  R  A  P  P  A  S  P  A  L  D  A  A  A
 631  CAGTGCCTGTTCGAGCCCCAGCTCTCCCCGCCCGACGTCAAGCCCGGGCTGAGGGCGCCTCCCGCTTCGCCAGCGCTGGACGCTGCTGCT

 195  S  A  F  K  G  P  Y  A  P  W  E  L  L  S  A  G  A  P  G  N  C  G  S  Q  G  S  F  Q  T  T
 721  TCGGCCTTCAAAGGGCCCTACGCCCCCTGGGAGCTGCTGTCGGCCGGGGCTCCGGGGAACTGTGGGTCGCAGGGAAGCTTCCAGACCACC

 225  P  E  A  R  F  S  A  V  G  T  K  V  E  D  L  L  S  I  S  C  P  A  E  L  P  G  P  A  S  R
 811  CCGGAGGCACGCTTTTCCGCCGTGGGGACCAAGGTCGAGGACCTGCTGTCCATCAGCTGCCCCGCCGAGCTGCCCGGTCCGGCTAGCAGA

 255  L  Y  P  P  G  A  Y  D  A  F  S  L  A  P  G  D  L  G  E  G  T  E  G  L  P  A  L  L  T  P
 901  CTCTACCCGCCAGGGGCCTACGATGCCTTCTCGCTGGCCCCAGGTGACTTAGGGGAGGGGACCGAGGGCCTCCCGGCGCTGCTGACCCCT

 285  P  G  G  E  G  G  S  G  G  E  G  G  E  F  L  A  V  P  Q  A  Q  L  S  P  L  G  L  R  G  A
 991  CCGGGCGGGGAGGGAGGGAGCGGCGGCGAAGGCGGAGAGTTCCTGGCCGTCCCTCAAGCGCAGCTGTCCCCGCTGGGCCTGCGCGGCGCC

 315  A  T  A  D  F  S  K  A  L  V  A  D  L  P  G  G  S  G  V  A  A  P  S  S  P  A  T  S  F  P
1081  GCCACGGCAGACTTCTCCAAAGCCCTGGTGGCCGACCTCCCGGGGGGCAGCGGAGTGGCGGCGCCTTCATCCCCCGCCACCTCCTTCCCC

 345  A  A  K  A  R  R  K  G  R  R  G  G  K  C  S  A  R  C  F  C  P  R  P  H  V  K  A  F  A  C
1171  GCGGCCAAAGCCCGGCGCAAGGGACGCCGGGGCGGCAAGTGCAGCGCGCGCTGCTTCTGCCCGCGGCCGCACGTCAAGGCCTTCGCCTGC

 375  P  V  E  S  C  V  R  T  F  A  R  S  D  E  L  N  R  H  L  R  I  H  T  G  H  K  P  F  Q  C
1261  CCCGTGGAGAGCTGCGTGCGGACGTTCGCGCGCTCCGACGAGCTCAACCGCCACCTGCGCATCCACACGGGCCACAAGCCCTTCCAGTGC

 405  R  I  C  L  R  N  F  S  R  S  D  H  L  T  T  H  V  R  T  H  T  G  E  K  P  F  A  C  D  V
1351  CGCATCTGCCTGCGCAACTTCAGCCGCAGCGACCACCTCACCACGCACGTGCGCACCCACACCGGCGAGAAGCCCTTCGCCTGCGACGTG

 435  C  G  R  R  F  A  R  S  D  E  K  K  R  H  S  K  V  H  L  K  Q  K  A  R  A  E  E  R  L  K
1441  TGCGGCCGCCGCTTCGCGCGCAGCGACGAGAAGAAGCGACACAGCAAGGTGCACCTCAAGCAGAAGGCGCGCGCCGAGGAGCGCCTCAAG

 465  G  L  G  F  Y  S  L  G  L  S  F  A  A  L  *
1531  GGCCTGGGCTTCTACTCGCTGGGCCTCTCTTTCGCCGCGCTGTAGCCGGAGCTGGCTCCGTAGGTCGGCGCCGCCGGCCGTCGGCGCACG
1621  CGACACGGTCCTGCCGCTCCCTCGTCCCTGCTGCCTTCCCTGCCTCTTCCACGCACGTCCGGGGCCACCCGCAGCCCAGCTCCAGTTCCC
1711  CCGAAGCGCCCGCCGCTCACGCCCTTCAGCACGGGCTCCGCGGACAGCGCCCGCTGTTTTCGGAGCCGCCTTCCTCTAGCCACCCGCTCT
1801  GGGGACTGTCCTCTCGGTCCACCCACAGAGCAGGCGATACCTTAGGACTGAAGAGAGTTTTTGTAACTGGCGTACGCCCCACGCCTTCCT
1891  CTTTATCCCTTCCCAGAGTCAAGCTGGGGATGTACCGAGCCGGTCTCTCAAGAACTTTGTACAGCAAGTCCAGCAAGCCTTTGGATGTGA
1981  TGTCTTTGCTTTGGGGTTATTTCCTTTTTGTTGTCGTTCATTTTTTGTAAAGCAGACGCTACTCTCAAGCATTTGACAAAACTGTTTATT
2071  TTTGCAATTAAAATTATTGTGCTAAAAAAAAAAAAAAAGG
```

FIG. 2. Nucleotide and deduced amino acid sequences of the NGFI-C cDNA. Numbers on the left refer to the nucleotide sequence (upper) and the amino sequence (lower). The lines are drawn over region corresponding to the zinc finger domain.

nition element, but not containing the closely related Spl binding sites, were cloned upstream of the basally inactive prolactin promoter present in a luciferase reporter vector (1). When these luciferase reporter plasmids were transfected into Cos cells along with the nonrecombinant pCMV expression vector, very low expression was observed. However, when they were cotransfected with the NGFI-C expression vector (pCMV-NGFI-C), luciferase activity was increased 20-fold from a reporter plasmid carrying one copy and 40-fold from a reporter plasmid containing two copies of the GCGGGGGCG sequence (Fig. 5). Experiments performed in parallel with an NGFI-A expression vector demonstrated a similar level of activation, thereby demonstrating that NGFI-C and NGFI-A function equivalently as transcriptional activators in Cos cells.

DISCUSSION

We have identified and characterized NGFI-C, a new member of a family of zinc finger proteins that contain very similar DNA-binding domains. In addition to NGFI-C, this family presently includes NGFI-A, Krox-20, Spl, and the

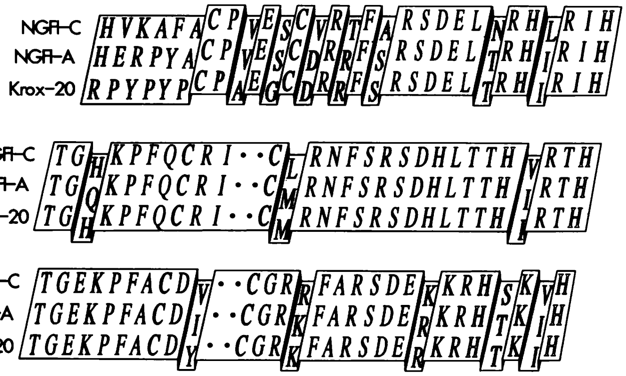

FIG. 3. Alignment of the NGFI-C zinc finger domains with those present in NGFI-A and Krox-20. The recessed, shaded regions denote areas of nonidentity.

Wilms' tumor gene product. Like NGFI-A and Krox-20, NGFI-C is an ERG. It is rapidly activated by NGF in PC12 cells independent of de novo protein synthesis and is induced by seizure in brain.

Many families of transcription factors that bind to identical or highly similar nucleotide elements have been described, including the CREB protein family, whose members recognize the cyclic AMP-responsive element (4), the c-Fos and c-Jun families, whose members interact with and recognize the AP-1 element (35), the octamer-binding proteins Oct-1 (45), Oct-2 (39), Oct-3-10 (40), C/EBP (23), and DBP (31), which recognize the D box of the albumin promoter, and the NF-E1 homologs which recognize the consensus motif WGATAR (49). The similarities between the zinc fingers present in NGFI-A, Krox-20, the Wilms' tumor gene product, and NGFI-C prompted us to test whether NGFI-C recognizes the same nucleotide sequence. Experiments in this study demonstrated that NGFI-C recognizes the nonamer GCGGGGGCG and activates transcription of a cotransfected reporter gene bearing this recognition site. These data provide evidence that these proteins comprise a family of transcription factors defined by their nucleotide recognition site.

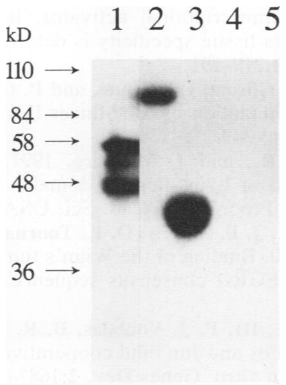

FIG. 4. Specific binding of NGFI-C to the sequence GCGG GGGCG. A ^{32}P-labeled oligonucleotide containing the GCGGGGG CG sequence was used to probe a protein blot containing bacterial lysates containing the following constructs: lane 1, NGFI-A in the pET3d vector (44); Lane 2, pNCFL, which expresses a TrpE fusion protein containing the entire NGFI-C protein; lane 3, pNCZF, which expresses a TrpE fusion protein containing the NGFI-C zinc finger region; lane 4, pNCAM, which expresses a TrpE fusion protein containing the amino-terminal portion of NGFI-C which does not include the zinc fingers; and lane 5, pATH-3, which expresses TrpE. Positions of molecular weight standards are indicated.

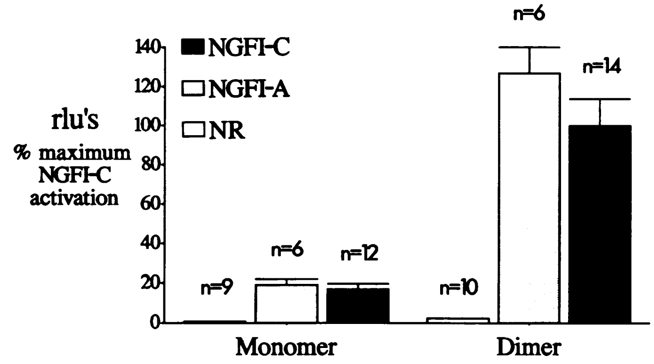

FIG. 5. Evidence that NGFI-C is a transcriptional activator in mammalian cells. Shown are relative luciferase units (rlu's) obtained with the indicated expression constructs (NR, nonrecombinant pCMV) and luciferase reporter plasmids containing one copy (monomer) or two copies (dimer) of the GCGGGGGCG sequence. To normalize data between experiments, values are expressed as the percentage of the relative luciferase units obtained relative to the NGFI-C/dimer combination. The number of plates assayed for each combination (n) is indicated along with the error bar. All values represent the results of at least three different experiments, performed with two different preparations of DNA.

Why are multiple proteins, each capable of activating transcription from the GCGGGGGCG nonamer, activated by the same extracellular stimuli? Several possibilities exist. First, although these proteins each recognize an identical nonamer, differences in their optimal binding affinities for closely related recognition sequences may allow them to regulate genes differently in vivo. Second, the local nucleotide environment or chromosomal structure into which the GSG motif is embedded may alter the relative affinity of these proteins for the site. Third, these proteins may each interact with different accessory factors that control their activities in a cellular or developmentally regulated fashion. Finally, the extent, and possibly type, of posttranslational modification of ERG products is dependent on the inducing stimulus (16) and may play a role in determining the inherent activities of these proteins.

A combination of these factors and possibly others is likely to determine the overall activity of these DNA-binding proteins in vivo. The proteins Oct-1 and Oct-2 are perhaps the most intensively studied examples of transcription factors which recognize the same nucleotide sequence yet have different activities. There is precisely controlled cell-type-specific expression of these proteins, but also dramatic promoter context-specific differences in transcriptional activation (47). The latter is thought to be secondary to differences in the intrinsic abilities of these proteins to interact with different classes of transcriptional initiation complexes (47). Because NGFI-A, Krox-20, and NGFI-C bear little resemblance to each other outside of the DNA-binding domain, it is likely that their functions are influenced by differences in their interactions with proteins of the transcriptional machinery. It will therefore be of interest to study not only cell-type-specific expression of these proteins in vivo but also the peptide determinants that direct the protein-DNA and protein-protein interactions of these factors. The cloning of a third ERG encoding a GSG element-binding protein provides an additional opportunity to examine the relative contributions of members of a family of transcription factors to the cellular response to extracellular stimuli.

ACKNOWLEDGMENTS

We thank Xi He and G. Rosenfeld for providing the protocol for Southwestern blotting.

This work was supported by grants from the National Institute of Neurological and Communicative Disorders and Stroke and the Multiple Sclerosis Foundation. S.D.C. was supported by National Research Service award 1F32NS09005-01 from the National Institute of Neurological and Communicative Disorders and Stroke.

REFERENCES

1. **Adler, S., M. L. Waterman, X. He, and M. G. Rosenfeld.** 1988. Steroid receptor-mediated inhibition of rat prolactin gene expression does not require the receptor DNA-binding domain. Cell **52**:685–695.
2. **Bartel, D. P., M. Sheng, L. F. Lau, and M. E. Greenberg.** 1989. Growth factors and membrane depolarization activate distinct programs of early response gene expression: dissociation of fos and jun induction. Genes Dev. **3**:304–313.
3. **Berg, J. M.** 1988. Proposed structure for the zinc-binding domains from transcription factor IIIA and related proteins. Proc. Natl. Acad. Sci. USA **85**:99–102.
4. **Berkowitz, L. A., and M. Z. Gilman.** 1990. Two distinct forms of active transcription factor CREB (cAMP response element binding protein). Proc. Natl. Acad. Sci. USA **87**:5258–5262.
5. **Call, K., T. Glaser, C. Y. Ito, A. J. Buckler, J. Pelletier, D. A. Haber, E. A. Rose, A. Kral, H. Yeger, W. H. Lewis, C. Jones, and D. E. Housman.** 1990. Isolation and characterization of a zinc finger polypeptide gene at the human chromosome 11 Wilm's tumor locus. Cell **60**:509–520.
6. **Chavrier, P., P. Lemaire, O. Revelant, R. Bravo, and P. Charnay.** 1988. Characterization of a mouse multigene family that encodes zinc finger structures. Mol. Cell. Biol. **8**:1319–1326.
7. **Chavrier, P., C. Vesque, B. Galliot, M. Vigneron, P. Dolle, D. Duboule, and P. Charnay.** 1990. The segment specific gene krox 20 encodes a transcription factor with binding sites in the promoter region of the Hox 1.4 gene. EMBO J. **9**:1209–1218.
8. **Chavrier, P., M. Zerial, P. Lemaire, J. Almendral, R. Bravo, and P. Charnay.** 1988. A gene encoding a protein with zinc fingers is activated during G0/G1 transition in cultured cells. EMBO J. **7**:29–35.
9. **Chen, C., and H. Okayama.** 1987. High-efficiency transformation of mammalian cells by plasmid DNA. Mol. Cell. Biol. **7**:2745–2752.
10. **Christy, B., and D. Nathans.** 1989. DNA binding site of the growth factor-inducible protein zif/268. Proc. Natl. Acad. Sci. USA **86**:8737–8741.
11. **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.
12. **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 snyaptic NMDA receptor activation. Nature (London) **340**:474–476.
13. **Curran, T., and J. I. Morgan.** 1985. Superinduction of c-fos by nerve growth factor in the presence of peripherally active benzodiazepines. Science **229**:1265–1268.
14. **de Wet, J. R., K. V. Wood, M. DeLuca, D. R. Helinski, and S. Subramani.** 1987. Firefly luciferase gene: structure and expression in mammalian cells. Mol. Cell. Biol. **7**:725–737.
15. **Dieckmann, C., and A. Tzagoloff.** 1985. Assembly of the mitochondrial membrane system. J. Biol. Chem. **260**:1513–1520.
16. **Fahrner, T. J., S. L. Carroll, and J. Milbrandt.** 1990. The NGFI-B protein, an inducible member of the thyroid/steroid receptor family, is rapidly modified posttranslationally. Mol. Cell. Biol. **10**:6454–6459.
17. **Feinberg, A., and B. Vogelstein.** 1984. A technique for radiolabeling DNA restriction endonuclease fragments to high specific activity. Anal. Biochem. **67**:15–28.
18. **Greenberg, M. E., E. B. Ziff, and L. A. Greene.** 1986. Stimulation of neuronal acetylcholine receptors induces rapid gene transcription. Science **234**:80–83.
19. **Hazel, T. G., D. Nathans, and L. F. Lau.** 1988. A gene inducible by serum growth factors encodes a member of the steroid and thyroid hormone receptor superfamily. Proc. Natl. Acad. Sci. USA **85**:8444–8448.
20. **Imagawa, M., R. Chiu, and M. Karin.** 1987. Transcription factor AP-2 mediates induction by two different signal-transduction pathways: protein kinase C and cAMP. Cell **51**:251–260.
21. **Kadonaga, J. T., K. R. Carner, F. R. Masiarz, and R. Tjian.** 1987. Isolation of a cDNA encoding transcription factor Sp1 and functional analysis of the DNA binding domain. Cell **51**:1079–1090.
22. **Ko, H. S., P. Fast, W. McBride, and L. M. Staudt.** 1988. A human protein specific for the immunoglobulin octamer DNA motif contains a functional homeobox domain. Cell **55**:135–144.
23. **Landschulz, W. H., P. F. Johnson, E. Y. Adashi, B. J. Graves, and S. L. McKnight.** 1988. Isolation of a recombinant copy of the gene encoding C/EBP. Genes Dev. **2**:786–800.
24. **Lemaire, P., C. Vesque, J. Schmitt, H. Stunnenberg, R. Frank, and P. Charney.** 1990. The serum-inducible mouse gene Krox-24 encodes a sequence specific transcriptional activator. Mol. Cell. Biol. **10**:3456–3467.
25. **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.
26. **Milbrandt, J.** 1986. Nerve growth factor rapidly induces c-fos mRNA in PC12 rat pheochromocytoma cells. Proc. Natl. Acad. Sci. USA **83**:4789–4793.
27. **Milbrandt, J.** 1987. A nerve growth factor-induced gene encodes a possible transcriptional regulatory factor. Science **238**: 797–799.
28. **Milbrandt, J.** 1988. Nerve growth factor induces a gene homologous to the glucocorticoid receptor gene. Neuron **1**:183–188.
29. **Miller, J., A. D. McLachlan, and A. Klug.** 1985. Repetitive zinc-binding domains in the protein transcription factor IIIA from Xenopus oocytes. EMBO J. **4**:1609–1614.
30. **Morgan, J. I., D. R. Cohen, J. L. Hempstead, and T. Curran.** 1987. Mapping patterns of c-fos expression in the central nervous system after seizure. Science **237**:192–197.
31. **Mueller, C. R., P. Maire, and U. Schibler.** 1990. DBP, a liver-enriched transcriptional activator, is expressed late in ontogeny and its tissue specificity is determined posttranscriptionally. Cell **61**:279–291.
32. **Nardelli, J., T. Gibson, C. Vesque, and P. Charnay.** 1991. Base sequence discrimination by zinc-finger DNA-binding domains. Nature (London) **349**:175–178.
33. **Pellegrino, G. R., and J. M. Berg.** 1991. Identification and characterization of "zinc-finger" domains by the polymerase chain reaction. Proc. Natl. Acad. Sci. USA **88**:671–675.
34. **Rauscher, F. J., J. F. Morris, O. E. Tournay, D. M. Cook, and T. Curran.** 1990. Binding of the Wilm's tumor locus zinc finger protein to the EGR-1 consensus sequence. Science **250**:1259–1262.
35. **Rauscher, F. J., III, P. J. Voulalas, B. R. Franza, Jr., and T. Curran.** 1988. Fos and Jun bind cooperatively to the AP-1 site: reconstitution in vitro. Genes Dev. **2**:1687–1699.
36. **Rusak, B., H. A. Robertson, W. Wisden, and S. P. Hunt.** 1990. Light pulses that shift rhythms induce gene expression in the suprachiasmatic nucleus. Science **24**:1237–1240.
37. **Ryseck, R. P., S. I. Hirai, M. Yaniv, and R. Bravo.** 1988. Transcriptional activation of c-jun during the G0/G1 transition in mouse fibroblasts. Nature (London) **334**:535–539.
38. **Sanger, F., S. Nicklen, and A. R. Coulson.** 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA **74**:5463–5467.
39. **Scheidereit, C., J. A. Cromlish, T. Gerster, K. Kawakami, C. Balmaceda, R. A. Currie, and R. G. Roeder.** 1988. A human lymphoid-specific transcription factor that activates immuno-

globulin genes is a homeobox protein. Nature (London) **336:** 551–556.

40. **Scholer, H. R., A. K. Hatzopoulos, R. Balling, N. Suzuki, and P. Gruss.** 1989. A family of octamer-specific proteins present during mouse embryogenesis: evidence for germline-specific expression of an Oct factor. EMBO J. **8:**2543–2550.
41. **Sheng, M., and M. E. Greenberg.** 1990. The regulation and function of c-fos and other immediate early genes in the nervous system. Neuron **4:**477–485.
42. **Singh, S. P., and M. F. Lavin.** 1990. DNA-binding protein activated by gamma radiation in human cells. Mol. Cell. Biol. **10:**5279–5285.
43. **Struhl, K.** 1988. The JUN oncoprotein, a vertebrate transcription factor, activates transcription in yeast. Nature (London) **332:**649–650.
44. **Studier, F. W., A. H. Rosenberg, J. J. Dunn, and J. W. Dubendorff.** 1990. Use of T7 polymerase to direct the expression of cloned genes. Methods Enzymol. **185:**60–89.
45. **Sturm, R. A., G. Das, and W. Herr.** 1988. The ubiquitous octamer-binding protein Oct-1 contains a POU domain with a homeo box subdomain. Genes Dev. **2:**1582–1599.
46. **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. 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.
47. **Tanaka, M., and W. Herr.** 1990. Differential transcriptional activation by Oct-1 and Oct-2: interdependent activation domains induce Oct-2 phosphorylation. Cell **60:**375–386.
48. **Wisden, W., M. L. Errington, S. Williams, S. B. Dunnett, C. Waters, D. Hitchcock, G. Evan, T. V. P. Bliss, and S. P. Hunt.** 1990. Differential expression of immediate early genes in the hippocampus and spinal cord. Neuron **4:**603–614.
49. **Yamamoto, M., L. Ko, M. W. Leonard, H. Beug, S. H. Orkin, and J. D. Engel.** 1990. Activity and tissue-specific expression of the transcription factor NF-E1 multigene family. Genes Dev. **4:**1650–1662.