MOLECULAR AND CELLULAR BIOLOGY, July 1992, p. 3070–3077
0270-7306/92/073070-08$02.00/0

Vol. 12, No. 7

Transcriptional Repression by a Novel Member of the bZIP Family of Transcription Factors

IAN G. COWELL,* ANN SKINNER, AND HELEN C. HURST

Oncology Group, Gene Transcription Laboratory, Imperial Cancer Research Fund, Hammersmith Hospital, Du Cane Road, London W12 OHS, England

Received 18 February 1992/Accepted 24 April 1992

We describe here a novel member of the bZIP family of DNA-binding proteins, designated E4BP4, that displays an unusual DNA-binding specificity which overlaps that of the activating transcription factor family of factors. When expressed in a transient transfection assay with a suitable reporter plasmid, E4BP4 strongly repressed transcription in a DNA-binding-site-dependent manner. Examination of a series of deletion mutants revealed that sequences responsible for the repressing potential of E4BP4 lie within the carboxyl-terminal region of the protein. No similarity was found between this region and the repressing domains of other known eukaryotic transcriptional repressors.

Transcription initiation depends to a large extent upon the activating capacity of specific protein factors which bind to gene regulatory regions, usually upstream of the transcription start site (33, 38, 43). However, regulation of transcription also involves inhibitory processes, and a growing number of DNA-binding factors which repress transcription in a gene-specific manner have recently been described elsewhere (3, 5, 11, 24, 26, 34, 50). Various mechanisms as to how transcriptional inhibition is effected may be envisaged. One such mechanism is steric hindrance of or direct competition for DNA binding with positively acting transcription factors. An example of this type of transcriptional repression is the ability of the transcription factor AP1 to inhibit basal and retinoic acid-induced expression of the osteocalcin gene by competition for binding to a common sequence element which is also recognized by the retinoic acid receptor (49). Repression of another kind again involves inhibition of DNA binding of a transcriptional activator, but in this case through protein-protein interactions. Examples of this type of repression include the mutual inhibition of DNA binding of AP1 proteins and the glucocorticoid receptor (48, 55) and negative regulation by the formation of non-DNA-binding heterodimers between helix-loop-helix proteins and the Id protein (5). A third type of transcriptional repressor does not suppress transcription through inhibition of the DNA binding of transcriptional activators. Repressors of this type are DNA-binding proteins whose action appears to be a direct down regulation of transcription. Such repressors will here be collectively termed active transcriptional repressors. Examples include the *Drosophila* krüppel (34) and engrailed proteins (17, 24), the human krüppel-related factor YY1 (50), and the Wilms' tumor gene product (36).

Transcription factors are conveniently divided into a number of structurally related families. The bZIP factors constitute one such family and share an amphipathic α-helical dimerization domain characterized by a heptad repeat of leucine residues, the leucine zipper (30), and a conserved DNA-binding domain which is rich in basic amino acids. The bZIP family contains some well-known transcriptional activators, such as c-*jun* (7), *fos* (1), and CREB (54), and may be subdivided into groups with similar DNA-binding specificities. The CREB-activating transcription factor (ATF) family, for example, all bind a critical core sequence, CGTCA, which is implicated in positive transcriptional control by cyclic AMP (9), by transactivator proteins from the bovine leukemia virus (28) and human T-cell lymphotrophic virus type I (13, 53), and by the adenovirus E1a gene product (6, 20, 25, 31, 32). The functional diversity inherent in the CREB-ATF core binding sequence probably reflects the number of DNA binding factors that interact with it. To date, more than 10 distinct mammalian cDNA clones which encode members of this family have been reported (12, 14, 16, 19, 23, 27, 37, 56). Each of these cloned factors bind as homodimers to the CREB-ATF core sequence, and some combinations can heterodimerize. In addition, heterodimerization between members of the CREB-ATF family and the *fos-jun* family can occur (4, 15).

We have cloned a DNA-binding factor of the bZIP class whose binding activity overlaps with that of the CREB-ATF family of transcription factors but which is a transcriptional repressor.

MATERIALS AND METHODS

Isolation of cDNA clones. A lambda gt11 cDNA library prepared from placental mRNA was obtained from Clontech. Approximately 5×10^5 plaques were screened for binding to an adenovirus E4 promoter ATF site probe, essentially as described by Singh et al. (51). The probe consisted of a double-stranded oligonucleotide, AAGCTTC TAAAAAATGACGTAACGGAAGCTT, corresponding to the proximal ATF site of the E4 promoter. The oligonucleotide was labelled with T4 polynucleotide kinase in the presence of [γ-^{32}P]ATP and concatenated by overnight ligation with T4 DNA ligase. Hybridization analysis of polymerase chain reaction (PCR)-amplified phage insert DNA was used to determine the relationship of these clones to each other and to CREB and ATF-2 cDNAs.

Characterization of cDNA clones. Lambda lysogens were prepared from clones which failed to hybridize to CREB or ATF-2 and were tested for their ability to bind to a monomeric E4 ATF site probe in a gel retardation assay. Phage DNA was prepared, and the cloned cDNA was subcloned into Bluescript SK (Stratagene) and thence into the pGEM-2-derived vector pT7βSRF (40). The sequence of the cDNA

* Corresponding author.

insert of lambda-P4 was determined by the dideoxy chain termination method by using the Sequenase enzyme (USB). A cDNA clone overlapping the insert contained in lambda-P4 was obtained by hybridization screening of the original placental cDNA library with the 0.2-kb *Eco*RI fragment of the lambda-P4 insert as a probe. The resulting lambda clone, lambda-P4/0.2, contained an extra 800 bp of sequence 3′ to the sequence contained in lambda-P4.

Plasmids. The plasmid pT7βP4 was generated by substituting the cDNA insert of lambda-P4 for the SRF insert in the pGEM-2-derived plasmid pT7βSRF (40). A full-length cDNA was generated as follows. The cDNA insert of lambda-P4/0.2 was cloned into the *Eco*RV site of Bluescript SK and a *Bgl*II-*Sma*I fragment was isolated by utilizing a *Bgl*II site within the cDNA sequence and a *Sma*I site 3′ of the cDNA in the Bluescript polylinker sequence. The 1.45-kb fragment was cloned into *Bgl*II-*Sma*I-digested pT7βP4, again making use of the *Bgl*II site internal to the cDNA sequence and a *Sma*I site in the polylinker of the plasmid, 3′ to the cDNA insert. The resulting plasmid, pT7βP4L-2, contained the full-length cDNA, including 5′- and 3′-untranslated regions, cloned downstream of the 5′-untranslated region and initiation codon of the human β-globin gene (40). To remove the 5′-untranslated region of the cloned cDNA, the cDNA insert was amplified by PCR with Vent polymerase (Biolabs) by using an SP6 promoter primer and a 5′ PCR primer ATCCATGGGATTTTAAAC CAGAGTTT; this procedure introduced an *Nco*I site at the initiation codon (and also changed the second codon to glutamine rather than glutamate). The product was sequenced to ensure that no alterations had been introduced during the PCR reaction. After digestion with *Nco*I, the PCR product was recloned into the *Nco*I and *Sma*I sites of pT7βSRF. The coding sequence of the cDNA in the resulting plasmid, p4RS2, was thus fused in frame to the initiation codon and 5′-untranslated region of the β-globin gene contained in the plasmid. The plasmid pP4Ea was generated by cloning the 554-bp *Eco*RI fragment from p4RS2 into *Stu*I- and *Xba*I-digested pT7βStu (40), after all 3′-recessed ends had been filled in with the Klenow fragment of *Escherichia coli* DNA polymerase. The *Nco*I-*Sma*I fragment containing the entire cDNA insert of p4RS2 minus the 5′-untranslated region was cloned into the *Xba*I site downstream of the simian virus 40 promoter of pSVK3 (Pharmacia), following the filling in of the 3′-recessed ends of each fragment. The resulting plasmid, pSVKP4, was digested with *Apa*I or *Pst*I, and the fragments were recircularized to generate pSVKP4delta-Apa and pSVKP4delta-Pst, respectively.

The reporter plasmid pSS0.2CAT was described by Dixon et al. (10). pπS12(34)CAT was derived from this by ligation of three copies of the double-stranded oligonucleotide AGC TTAAAAAATTATGTAACGGTAAGCT into the *Hin*dIII site at position −94 in the GST-π promoter.

In vitro transcription and translation. p4RS2 or pP4Ea was digested with *Sma*I, and RNA was transcribed from 1 μg of the linearized template in a 25-μl reaction mixture containing 40 mM Tris (pH 7.5), 50 mM NaCl, 12.5 mM Mg_2Cl, 2 mM spermidine, 0.4 mM UTP, 0.4 mM ATP, 0.4 mM TTP, 40 μM GTP, 0.3 mM dithiothreitol, 0.5 mM GpppG, 1 U of RNasin (Promega), and 10 U of T7 RNA polymerase for 30 min at 37°C. In vitro translation was performed in a rabbit reticulocyte lysate (Promega) in a standard reaction volume of 50 μl. The reactions were carried out for 20 min at 30°C. For optimum synthesis of full-length product, potassium was adjusted to a final concentration of 120 mM by the addition of 4 μl of 1 M potassium acetate (pH 7.5) to each standard reaction mix. At the end of the translation reaction, an equal volume of 2× dialysis buffer (40 mM HEPES [*N*-2-hydroxyethylpiperazine-*N*′-2-ethanesulfonic acid], 20% glycerol, 2 mM $MgCl_2$, 1 mM EDTA, 1 mM phenylmethylsulfonyl fluoride, 1 mM dithiothreitol, 50 mM KCl) was added. Yields of in vitro-translated protein were sufficient for dilution of the extract to a level at which endogenous DNA-binding activity was negligible.

Gel retardation analysis. Probes for gel retardation were labelled as described previously (20), and gel retardation assays were carried out as described by Hurst et al. (21). The probe that was used in all of the gel retardation experiments shown here was a double-stranded oligonucleotide containing the proximal ATF site of the adenovirus E4 promoter, as used in concatenated form for expression library screening.

Binding site selection. The optimum DNA-binding sites for E4BP4 were selected from a mixture of random sequence oligonucleotides essentially as described by Pollock and Treisman (41). Binding reactions were performed in GRA buffer (25 mM HEPES [pH 7.9], 1 mM EDTA, 5 mM dithiothreitol, 150 mM NaCl, 20% [vol/vol] glycerol) with in vitro-translated E4BP4. DNA-protein complexes were immunoprecipitated with a rabbit antiserum raised against the 15 C-terminal amino acids of E4BP4. Selected DNA was cloned into Bluescript SK and sequenced with a T7 promoter primer.

Cell culture and transfection assays. HeLa JEG-3 and HepG2 cells were maintained in Dulbecco's modified Eagle medium supplemented with 10% fetal calf serum. The cells were transfected by the calcium phosphate procedure (2), with 3 μg of reporter plasmid and up to 10 μg of expression plasmid in a standard experiment. Cell extracts were prepared by lysis of the cells in situ, and the extracts were assayed for chloramphenicol acetyltransferase (CAT) activity as described by Ausubel et al. (2). In each experiment, cells were cotransfected with pJATLAC^{3}z, which contains the β-galactosidase coding region driven by the rat β-actin promoter. Extracts were routinely assayed for β-galactosidase activity (18) to normalize for transfection efficiency. Each experiment was repeated at least three times.

Nucleotide sequence accession number. The E4BP4 cDNA sequence reported here has been lodged with the EMBL data library and received the accession number X64318.

RESULTS

Isolation and sequence analysis of cDNA clones encoding E4 ATF site binding factors. Of the independent cDNA clones that were isolated following the screening of a human placental cDNA library with a concatenated adenovirus E4 ATF site probe (see Materials and Methods), several corresponded on initial examination by hybridization and sequence analysis to the previously characterized CRE-BP1/ATF-2 factor (13, 16). However, one clone, lambda-P4, represented a novel factor, which when expressed in a lambda lysogen system, efficiently formed protein-DNA complexes with an ATF site oligonucleotide probe (data not shown). The protein factor encoded by lambda-P4 was designated E4BP4 to reflect its affinity for the E4 promoter.

Sequence analysis of the lambda-P4 cDNA insert revealed a long open reading frame (ORF) commencing with an ATG codon 214 bp from the 5′ end of the cDNA sequence (Fig. 1). This ATG presumably represents the initiation codon, since a TAA stop codon is positioned 84 bp upstream of this triplet in the cDNA sequence. Rescreening of the placental cDNA library by hybridization with a 3′ fragment of the lambda-P4

```
1    GCCCCTTTCT TTCTCCTCGT CGGCCCGAGA GCAGGAACAC GATAACGAAG GAGGCCCAAC TTCATTCAAT

71   AAGGAGCCTG ACGGATTTAT CCCAGACGGT AGAACAAAAG GAAGAATATT GATGGATTTT AAACCAGAGT

141  TTTTAAAGAG CTTGAGAATA CGGGGAAATT AATTTGTTCT CCTACACACA TAGATAGGGT AAGGTTGTTT

     M  Q  L  R  K  M  Q  T  V  K  K  E  Q  A  S  L  D
211  CTG ATG CAG CTG AGA AAA ATG CAG ACC GTC AAA AAG GAG CAG GCG TCT CTT GAT

     A  S  S  N  V  D  K  M  M  V  L  N  S  A  L  T  E  V
265  GCC AGT AGC AAT GTG GAC AAG ATG ATG GTC CTT AAT TCT GCT TTA ACG GAA GTG

     S  E  D  S  T  T  G  E  D  V  L  L  S  E  G  S  V  G
319  TCA GAA GAC TCC ACA ACA GGT GAG GAC GTG CTT CTC AGT GAA GGA AGT GTG GGG

     K  N  K  S  S  A  C  R  R  K  R  E  F  I  P  D  E  K
373  AAG AAC AAA TCT TCT GCA TGT CGG AGG AAA CGG GAA TTC ATT CCT GAT GAA AAG

     K  D  A  M  Y  W  E  K  R  R  K  N  N  E  A  A  K  R
427  AAA GAT GCT ATG TAT TGG GAA AAA AGG CGG AAA AAT AAT GAA GCT GCC AAA AGA

     S  R  E  K  R  R  L  N  D  L  V  L  E  N  K  L  I  A
481  TCT CGT GAG AAG CGT CGA CTG AAT GAC CTG GTT TTA GAG AAC AAA CTA ATT GCA

     L  G  E  E  N  A  T  L  K  A  E  L  L  S  L  K  L  K
535  CTG GGA GAA GAA AAC GCC ACT TTA AAA GCT GAG CTG CTT TCA CTA AAA TTA AAG

     F  G  L  I  S  S  T  A  Y  A  Q  E  I  Q  K  L  S  N
589  TTT GGT TTA ATT AGC TCC ACA GCA TAT GCT CAA GAG ATT CAG AAA CTC AGT AAT

     S  T  A  V  Y  F  Q  D  Y  Q  T  S  K  S  N  V  S  S
643  TCT ACA GCT GTG TAC TTT CAA GAT TAC CAG ACT TCC AAA TCC AAT GTG AGT TCA

     F  V  D  E  H  E  P  S  M  V  S  S  S  C  I  S  V  I
697  TTT GTG GAC GAG CAC GAA CCC TCG ATG GTG TCA AGT AGT TGT ATT TCT GTC ATT

     K  H  S  P  Q  S  S  L  S  D  V  S  E  V  S  S  V  E
751  AAA CAC TCT CCA CAA AGC TCG CTG TCC GAT GTT TCA GAA GTG TCC TCA GTA GAA

     H  T  Q  E  S  S  V  Q  G  S  C  R  S  P  E  N  K  F
805  CAC ACG CAG GAG AGC TCT GTG CAG GGA AGC TGC AGA AGT CCT GAA AAC AAG TTC

     Q  I  K  Q  E  P  M  E  L  E  S  Y  T  R  E  P  R
859  CAG ATT ATC AAG CAA GAG CCG ATG GAA TTA GAG AGC TAC ACA AGG GAG CCA AGA

     D  D  R  G  S  Y  T  A  S  I  Y  Q  N  Y  M  G  N  S
913  GAT GAC CGA GGC TCT TAC ACA GCG TCC ATC TAT CAA AAC TAT ATG GGG AAT TCT

     F  S  G  Y  S  H  S  P  P  L  L  Q  V  N  R  S  S  S
967  TTC TCT GGG TAC TCA CAC TCT CCC CCA CTA CTG CAA GTC AAC CGA TCC TCC AGC

     N  S  P  R  T  S  E  T  D  D  G  V  V  G  K  S  S  D
1021 AAC TCC CCG AGA ACG TCG GAA ACT GAT GAT GGT GTG GTA GGA AAG TCA TCT GAT

     G  E  D  E  Q  Q  V  P  K  G  P  I  H  S  P  V  E  L
1075 GGA GAA GAC GAG CAA CAG GTC CCC AAG GGC CCC ATC CAT TCT CCA GTT GAA CTC

     K  H  V  H  A  T  V  V  K  V  P  E  V  N  S  S  A  L
1129 AAG CAT GTG CAT GCA ACT GTG GTT AAA GTT CCA GAA GTG AAT TCC TCT GCC TTG

     P  H  K  L  R  I  K  A  K  A  M  Q  I  K  V  E  A  F
1183 CCA CAC AAG CTC CGG ATC AAA GCC AAA GCC ATG CAG ATC AAA GTA GAA GCC TTT

     D  N  E  F  E  A  T  Q  K  L  S  S  P  I  D  M  T  S
1237 GAT AAT GAA TTT GAG GCC ACG CAA AAA CTT TCC TCA CCT ATT GAC ATG ACA TCT

     K  R  H  F  E  L  E  K  H  S  A  P  S  M  V  H  S  S
1291 AAA AGA CAT TTC GAA CTC GAA AAG CAT AGT GCC CCA AGT ATG GTA CAT TCT TCT

     L  T  P  F  S  V  Q  V  T  N  I  Q  D  W  S  L  K  S
1345 CTT ACT CCT TTC TCA GTG CAA GTG ACT AAC ATT CAA GAT TGG TCT CTC AAA TCG

     E  H  W  H  Q  K  E  L  S  G  K  T  Q  N  S  F  K  T
1399 GAG CAC TGG CAT CAA AAA GAA CTG AGT GGC AAA ACT CAG AAT AGT TTC AAA ACT

     G  V  V  E  M  K  D  S  G  Y  K  V  S  D  P  E  N  L
1453 GGA GTT GTT GAA ATG AAA GAC AGT GGC TAC AAA GTT TCT GAC CCA GAG AAC TTG

     Y  L  K  Q  G  I  A  N  L  S  A  E  V  V  S  L  K  R
1507 TAT TTG AAG CAG GGG ATA GCA AAC TTA TCT GCA GAG GTT GTC TCA CTC AAG AGA

     L  I  A  T  Q  P  I  S  A  S  D  S  G  ***
1561 CTT ATA GCC ACA CAA CCA ATC TCT GCT TCA GAC TCT GGG TAA ATTACTACTG AGTAAGAGCT

1623 GGGCATTTAG AAAGATGTCA TTTGCAATAG AGCAGTCCAT TTTGTATTAT GCTGAATTTT CACTGGACCT

1693 GTGATGTCAT TTCACTGTGA TGTGCACATG TTGTCTGTTT GGTGTCTTTT TGTGCACAGA TTATGATGAA

1763 GATTAGATTG TGTTATCACT CTGCCTGTGT ATAGTCAGAT AGTCATATGC GTAAGGCTGT ATATATTAAA

1833 CTTTTATTTT TGTTGTTCTA TTATAAAGTG TGTAAGTTAC CAGTTTCAAT AAAGGATTGG TGACAAACAC

1903 AGAAAAAAAA AAAAAAAAA A
```

FIG. 1. Nucleotide sequence of the cDNA insert of lambda-P4 and the encoded amino acid sequence of E4BP4. Leucine residues of the leucine heptad are underlined.

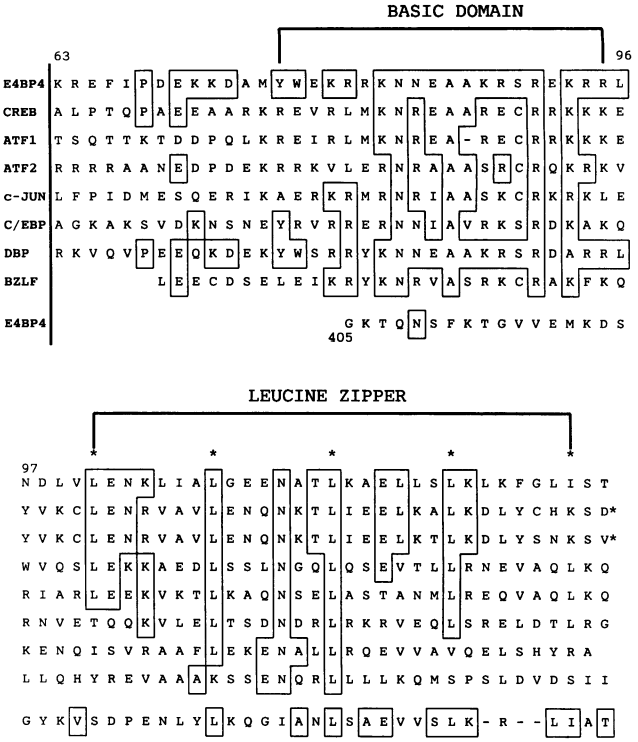

FIG. 2. E4BP4 is a member of the bZIP family of proteins. A region toward the N terminus of E4BP4 was aligned with the basic and leucine zipper domains of the following representative bZIP factors CREB (14), ATF-1 (16), CRE-BP1 (37), c-*jun*, C/EBP (29), DBP (39), and BZLF (35). Residues identical to those of E4BP4 are boxed. Within the basic domain, certain residues, such as Asn-83, Ala-86, Arg-91, and Lys-93, are highly conserved, and while the basic region is most similar to that of DBP, the leucine zipper is most like that of CREB-ATF-1. A second region of homology with the leucine heptad repeat regions of bZIP factors near the C terminus of E4BP4 is shown on the last line of the alignment.

cDNA resulted in the isolation of an overlapping clone extending the length of the 3′ sequence of the cDNA. The resulting ORF terminated at a TAA stop codon and was followed by 300 bp of 3′-untranslated region, including a canonical AATAAA polyadenylation signal. The molecular mass of E4BP4 was calculated to be 51.4 kDa from an ORF containing 462 amino acids.

E4BP4 is a member of the bZIP family of DNA-binding factors. Inspection of the ORF contained in the cDNA insert of lambda-P4 revealed a region of homology with the bZIP family of transcription factors (52). The characteristic features of this family of DNA-binding proteins are a linked basic DNA-binding domain and a leucine zipper structure (30) which is necessary for dimerization. Figure 2 shows an alignment between the basic region and leucine zippers of E4BP4 and a number of other bZIP proteins. E4BP4 appears to be most closely related (26 of 69 residues) to a bZIP factor, DBP, which binds a sequence in the human albumin promoter (39); the next closest relation is CREB (20 of 69 residues), while c-*jun* (14 of 69) and C/EBP (13 of 69 residues) appear more distantly related. No sequence similarity was detected between E4BP4 and any members of the bZIP family outside the DNA binding-dimerization domain. Similarly, no significant similarity was found between E4BP4 and other proteins in the SwissProt protein sequence data base, with the exception of the bZIP factors CREB and DBP.

E4BP4 binds DNA as a dimer. The cDNA insert of lambda-P4 was cloned into a pGEM2-derived plasmid for in vitro transcription and translation. A truncated protein, SmP4, retaining the predicted DNA-binding and dimerization domains of E4BP4, was generated in the same way, by translation of RNA transcribed in vitro from the plasmid pP4Ea (see Materials and Methods). Protein synthesized in vitro was tested for specific DNA binding. As shown in Fig. 3, both the full-length and the truncated forms of E4BP4 formed protein-DNA complexes with the E4 ATF site probe. As expected, the complex generated with the truncated protein exhibited greater mobility in a gel retardation assay than did the full-length protein. When the truncated and full-length proteins were mixed prior to incubation with the

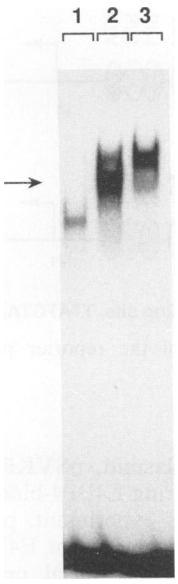

FIG. 3. DNA binding and dimerization of E4BP4. The truncated form, SmP4 (lane 1), and the full-length E4BP4 (lane 3) were translated in vitro and allowed to bind a ^{32}P-labelled E4 ATF site oligonucleotide probe. Protein-DNA complexes were resolved by gel retardation analysis. In lane 2, E4BP4 and SmP4 were mixed prior to incubation with the binding site probe. The arrow indicates the position of migration of the heterodimeric complex formed between E4BP4 and SmP4.

binding site oligonucleotide, a complex of intermediate mobility was also observed (Fig. 3), demonstrating dimerization between the long and short forms of E4BP4. Heterodimerization has been shown to occur between some pairs of bZIP factors; for example, AP1 consists largely of heterodimeric c-*jun*–c-*fos*, and heterodimerization has been reported between ATF-2 and c-*jun* (4, 15), between CREB and ATF-1 (21, 22), and between different members of the C/EBP family. No heterodimerization between E4BP4 and CREB, c-*jun*, ATF-1, or ATF-2 was observed by the type of assay described above (data not shown). However, this study was not exhaustive, and it is therefore still possible that E4BP4 might interact with a still unidentified partner.

DNA-binding of E4BP4: specificity for E4 ATF sites. E4BP4 efficiently forms DNA-protein complexes with the ATF site from the adenovirus E4 promoter. Its affinity for other CREB-ATF sequences and for the related AP1, C/EBP, and USF sites was tested by competition analysis with a series of unlabelled oligonucleotides. As illustrated in Fig. 4, oligonucleotides containing the CREs from the human chorionic gonadotrophin gene (*hcg*), fibronectin gene, and somatostatin gene and the ATF site from the adenovirus E3 gene all failed to compete efficiently for binding, even at a 50-fold molar excess. Oligonucleotides containing AP1-, C/EBP-, and USF-binding sites also failed to compete. Thus, while E4BP4 is able to bind the E4 ATF site, it does not bind the canonical CRE sequence, even though both contain the core sequence CGTCA. To further probe the DNA-binding preferences of E4BP4, a number of E4 ATF site point mutants were tested for their ability to compete with the wild-type ATF site. The results of this analysis and of similar experiments carried out with ATF-1 and CREB are shown in Table 1. While the point mutant PM1 was unable to compete for binding to E4BP4, CREB, or ATF-1, other point mutants

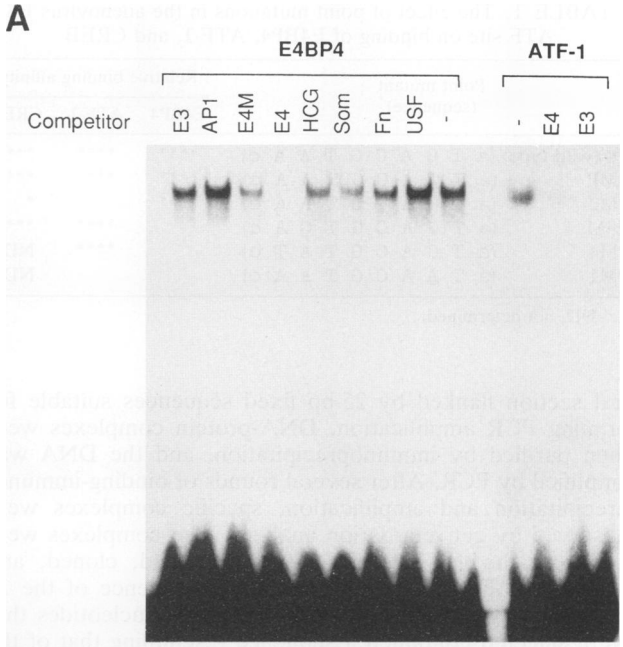

FIG. 4. Competition between CRE-ATF sites for E4BP4 binding. (A) Oligonucleotides containing the sequences indicated in panel B were used as competitors at a 50-fold M excess for binding of either E4BP4 or ATF-1. The probe in each case was a ^{32}P-labelled E4 ATF site oligonucleotide. (B) Sequences of the CREB-ATF sites contained in the competitor oligonucleotides. E4 (WT), adenovirus E4 gene ATF site; E3, adenovirus E3 gene ATF site; AP1, simian virus 40 AP1 site; E4M, E4 ATF site containing a single point mutation to a canonical CRE sequence; HCG, Som, and Fn, CREs from the human chorionic gonadotrophin, somatostatin, and fibronectin genes, respectively; USF, USF (MLTF)-binding site within the adenovirus major late promoter.

clearly discriminated between E4BP4 and ATF-1. For example, PM4 competed efficiently with the wild-type site for binding to ATF-1 but not for binding to E4BP4, while the converse was true of PM2.

Selection of optimum E4BP4-binding sites. Competition experiments of the kind described above suggest that the DNA-binding specificity of E4BP4 differs markedly from those of members of the CREB-ATF family and indeed of other bZIP factors. To determine the true binding-site preferences of E4BP4, oligonucleotides containing E4BP4-binding sites were selected from a pool of random-sequence oligonucleotides by a method described by Pollock and Treisman (41). Briefly, E4BP4 synthesized in vitro was incubated in a DNA-binding reaction containing a pool of oligonucleotides constructed with a random-sequence cen-

TABLE 1. The effect of point mutations in the adenovirus E4 ATF site on binding of E4BP4, ATF-1, and CREB

Point mutant (sequence)		Relative binding affinity		
		E4BP4	ATF-1	CREB
E4 (wild type)	(a T G A C G T A A c)	****	****	****
PMP	(a T T A C G T A A c)	****	***	***
PM2	(a T G A T G T A A c)	****		*
E4M	(a T G A C G T C A c)		****	****
PM4	(a T G A C G T A T c)		****	ND[a]
PM1	(a T A A C G T A A c)			ND

[a] ND, not determined.

tral section flanked by 25-bp fixed sequences suitable for priming PCR amplification. DNA-protein complexes were then purified by immunoprecipitation, and the DNA was amplified by PCR. After several rounds of binding-immunoprecipitation and amplification, specific complexes were observed by gel retardation analysis. The complexes were cut from the gel, and DNA was amplified, cloned, and sequenced. Figure 5 shows the DNA sequence of the 23 clones that were analyzed. All of the oligonucleotides that were selected contained a sequence resembling that of the E4 ATF site, and none of the selected oligonucleotides occurred more than once. Interestingly, only a small number of variations on the E4 site were represented; more than 50% of the clones contained the sequence TTATGTAA, and just more than 25% contained the palindromic sequence TTACGTAA. This suggests that TTATGTAA represents the highest affinity-binding site for E4BP4. From this study, the consensus sequence (G/A)T(G/T)A(C/T)GTAA(C/T) for optimal E4BP4 binding emerged.

E4BP4 has transcriptional repressing activity. By analogy with other leucine zipper factors, such as CREB or c-*jun*, it was anticipated that E4BP4 was a transcriptional activator. To test this hypothesis, cells were cotransfected with an

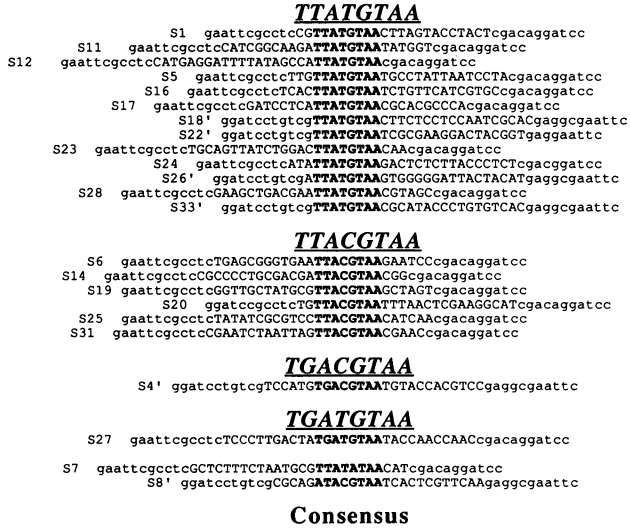

FIG. 5. E4BP4-binding site selection analysis. High-affinity binding sites were recovered from a pool of random-sequence oligonucleotides. The sequences of cloned, selected oligonucleotides were aligned around the ATF-like sequence that each contained. The consensus sequence that emerged is shown at the bottom.

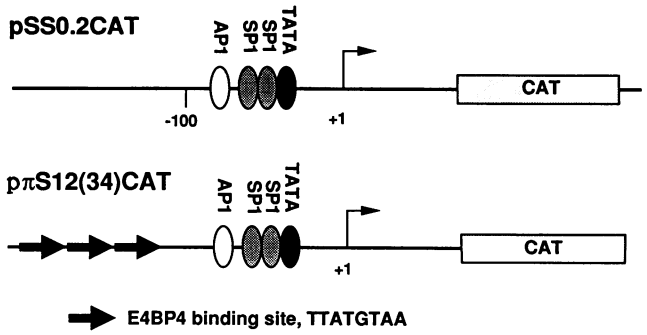

FIG. 6. Structures of the reporter plasmids pSS0.2CAT and pπS12(34)CAT.

E4BP4 expression plasmid, pSVKP4, together with a reporter plasmid harboring E4BP4-binding sites. The reporter plasmid used for this experiment, pπS12(34)CAT (Fig. 6), contained three binding sites for E4BP4 (TTATGTAA) positioned upstream of the basal promoter of the human glutathione transferase π gene (see Materials and Methods). The E4BP4-binding sequence, which is in effect a double-point-mutated E4 ATF site, corresponds to a high-affinity E4BP4-binding site (as shown in Fig. 5) and was chosen since it is not bound by members of the CREB-ATF family or other factors in HeLa, HepG2, or JEG-3 nuclear extracts (unpublished results). When transfected alone into HeLa cells, pπS12(34)CAT and the pSS0.2CAT control plasmid, which lacked E4BP4-binding sites, expressed similar levels of CAT activity. However, when increasing amounts of pSVKP4 were cotransfected with the reporter plasmids, CAT activity was quantitatively reduced from pπS12(34) CAT, while CAT activity in cells transfected with pSS0.2CAT was essentially unaffected (Fig. 7). This effect was observed with as little as 100 ng of pSVKP4 plasmid per plate and was maximal by 2 μg per plate. Thus, at least in this assay, E4BP4 acts as a binding site-dependent transcrip-

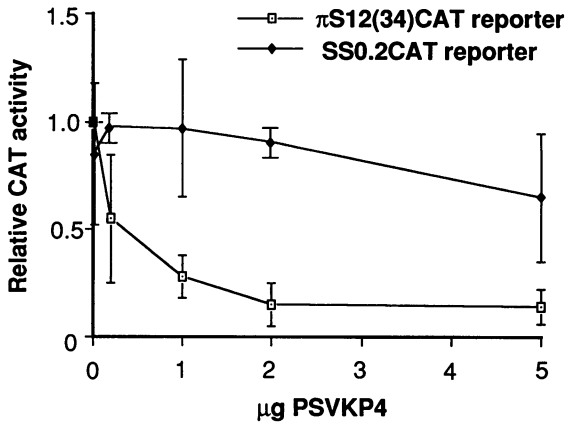

FIG. 7. E4BP4 is a transcriptional repressor. HeLa cells were cotransfected with up to 5 μg of the E4BP4 expression plasmid pSVKP4 and 3 μg of either pSS0.2CAT or pπS12(34)CAT reporter. Each experiment was repeated at least three times, and transfection efficiency was corrected for by comparison with β-galactosidase internal control. Data are the mean relative CAT activities plus or minus the standard errors of the means. All of the data are expressed relative to the CAT activity obtained following transfection of pπS12(34)CAT alone.

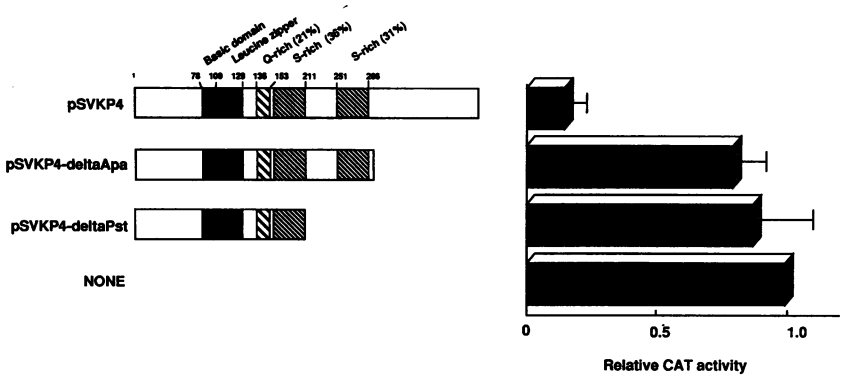

FIG. 8. Localization of transcriptional repressing activity to the C-terminal domain of E4BP4. HepG2 cells were cotransfected with pπS12(34)CAT and 2 μg of either pSVKP4 or one of two E4BP4 deletion constructs, pSVKP4-deltaApa or pSVKP4-deltaPst. The resulting CAT activities were normalized against that obtained with reporter alone. The data are the mean results plus or minus the standard errors of the means obtained from at least three experiments.

tional repressor. Essentially identical results were obtained when these experiments were conducted with either HepG2 or JEG cells (data not shown).

The C-terminal region of E4BP4 is required for transcriptional repressing activity. In order to map the domain or domains of E4BP4 that are involved in its transcriptional repressor activity, a series of C-terminal deletions was constructed (Fig. 8). DNA-binding activity corresponding to that of either wild-type E4BP4 or the truncated proteins could be detected in nuclear extracts prepared from appropriately transfected cells (data not shown), demonstrating that the truncated proteins are stable, active in DNA binding, and not defective in nuclear localization. However, neither of the forms of E4BP4 truncated at the C terminus was able to repress the pπS12(34)CAT reporter (Fig. 8). This strongly suggests either that the transcriptional repressing activity of E4BP4 lies in its C-terminal region or at least that C-terminal amino acids are necessary for this activity.

DISCUSSION

In this paper, we describe the expression cloning of a novel bZIP factor, which we have called E4BP4. Surprisingly, this factor acts as a repressor of transcription. The repressing activity of E4BP4 depends on binding to the test promoter and is therefore not due to the "squelching" (42) of positively acting factors. In experiments with the pπS12(34) CAT reporter, it is unlikely that E4BP4 is acting by displacing positively acting factors, because in the absence of exogenous E4BP4, the test promoter that contains E4BP4-binding sites has the same activity as the parent basal promoter (pSS0.2CAT) that lacks such sites. Furthermore, forms of E4BP4 truncated at the C terminus retain DNA-binding activity but lose the ability to repress transcription. Consequently, E4BP4 would appear to be an active repressor of transcription. However, we cannot rule out the interpretation that E4BP4 could behave as an activator under some circumstances. Such characteristics have recently been described for the factor YY1 (50), which in the absence of the adenovirus E1a product can behave as a transcriptional repressor but which activates transcription when E1a is present.

Relatively little is known about the mechanisms of active transcriptional repression. By analogy to what is understood of the mechanism of transcriptional activation, active repressors could destabilize or inhibit the formation of preinitiation complexes (8), perhaps by interfering with the interactions between positively acting transcription factors and the general transcription machinery. The *Drosophila* transcriptional repressors krüppel, engrailed, and even-skipped each contain alanine-rich regions, which, in the case of krüppel, have been mapped to the region responsible for its repressing activity (34). Inspection of its amino acid sequence revealed no similar alanine-rich regions in E4BP4. The transcriptional repressing domain of the Wilms' tumor gene product has been mapped and found to contain a predominance of proline and glutamine residues (36). However, the C-terminal region of E4BP4, which is essential for transcriptional repression, again bears no significant similarity to this sequence. Interestingly, the C-terminal domain of E4BP4 does contain a second heptad repeat of leucine residues (Fig. 2). In this region (residues 432 to 448), three leucine residues are spaced at seven amino acid intervals in a short region of probable α-helical structure (Chou and Fasman) bounded by helix-breaking proline residues. Finer deletion analysis should determine the importance of this motif for the transcriptional repressing activity of E4BP4. In addition to the leucine zipper(s) and basic domain, E4BP4 contains a small glutamine-rich region just C terminal of the leucine zipper (4 of 18 residues) and two regions rich in serine (20 of 55 and 11 of 36 residues) (Fig. 8), but any functions of these regions have yet to be elucidated, since they fall outside the C-terminal transcriptional repressing domain.

Another protein that is known to bind the sites in the E4 promoter that are also bound by E4BP4 is the cellular factor E4F (45, 46). Interestingly, the DNA-binding specificities of E4F (47) and E4BP4 appear to be very similar. Both factors bind poorly to CRE sequences and to the PM1 mutant of the E4 ATF site, while they both bind efficiently to the wild-type ATF site and the PM2 mutant (Table 1). However, we believe these two proteins to be distinct, since an antiserum raised against the 15 C-terminal amino acids of E4BP4 failed to interact with purified samples of E4F (44).

The distinct amino acid sequence and DNA-binding specificity of E4BP4 place it in a separate class within the bZIP family of factors distinct from the CREB-ATF family, whose DNA-binding specificities it overlaps. From binding-site selection studies, the optimum binding site for E4BP4 emerged as (G/A)TTATGTAA(C/T). The sequence TTACG TAA was selected less frequently than the asymmetric TTATGTAA motif, suggesting that the sequence TTAT− is a better half site than TTAC−; however, the corresponding

palindrome, TTATATAA, occurred only once in the 23 sequenced clones and presumably represents a lower-affinity-binding site. Thus, although E4BP4 is a dimeric DNA-binding factor, its optimal DNA-binding site consists of one TTAC− and one TTAT− half site. A search of the GenBank DNA sequence data base for consensus sequence binding sites revealed E4BP4-binding sites in a number of cellular and viral promoters. Unfortunately, these give little clue to the role that E4BP4 may play. Northern (RNA) analysis of a number of cell lines and human fetal tissues has shown that the E4BP4 gene is ubiquitously expressed. However, gel retardation and Western blot (immunoblot) analysis have shown that the cell lines contain little or no E4BP4 protein (unpublished results). It is possible, therefore, that normal expression of E4BP4 is controlled at the level of translation, as has been suggested for the bZIP factor DBP (39). We are currently further investigating the expression of E4BP4 and are studying in more detail its mechanism of transcriptional repression and its mode of interaction with other cellular components. One interesting possibility is that E4BP4 may have a still unidentified cellular dimerization partner which may profoundly affect its overall activity.

ACKNOWLEDGMENTS

We thank Nic Jones, Steve Goodbourn, and Mike Fried for critical reading of the manuscript. We also thank Iain Goldsmith for the synthesis of oligonucleotides.

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. **Ausubel, F. M., R. Brent, R. E. Kingston, D. D. Moore, J. G. Seidman, J. A. Smith, and K. Struhl.** 1987. Current protocols in molecular biology. John Wiley & Sons, New York.
3. **Auwerx, J., and P. Sassone-Corsi.** 1991. IP-1: a dominant inhibitor of fos/jun whose activity is modulated by phosphorylation. Cell **64:**983–993.
4. **Benbrook, D. M., and N. C. Jones.** 1990. Heterodimer formation between CREB and JUN proteins. Oncogene **5:**295–302.
5. **Benezra, R., R. L. Davis, D. Lockshon, D. L. Turner, and H. Weintraub.** 1990. The protein id: a negative regulator of helix-loop-helix DNA-binding proteins. Cell **61:**49–59.
6. **Berk, A. J.** 1986. Adenovirus promoters and E1a transactivation. Annu. Rev. Genet. **20:**45–79.
7. **Bohmann, D., and R. Tjian.** 1989. Biochemical analysis of transcriptional activation by jun: differential activity of c- and v-jun. Cell **59:**709–717.
8. **Buratowski, S., S. Hahn, L. Guarante, and P. A. Sharp.** 1989. Five intermediate complexes in transcription initiation by RNA polymerase II. Cell **56:**549–561.
9. **Deutsch, P. J., J. P. Hoeffler, L. J. Jameson, J. C. Lin, and J. F. Habener.** 1988. Structural determinants of transcriptional activation by cyclic AMP responsive DNA elements. J. Biol. Chem. **263:**18466–18472.
10. **Dixon, K. H., I. G. Cowell, C. L. Xia, S. E. Pemble, B. Ketterer, and J. B. Taylor.** 1989. Control of expression of the human glutathione S-transferase π gene differs from its rat homologue. Biochem. Biophys. Res. Commun. **163:**815–822.
11. **Dostatni, N., P. F. Lambert, R. Sousa, J. Ham, P. M. Howley, and M. Yaniv.** 1991. The functional bpv-1 e2 trans-activating protein can act as a repressor by preventing formation of the initiation complex. Genes Dev. **5:**1657–1671.
12. **Foulkes, N. S., E. Borrelli, and P. Sassone Corsi.** 1991. CREM gene: use of alternative DNA-binding domains generates multiple antagonists of cAMP-induced transcription. Cell **64:**739–749.
13. **Giam, C.-Z., and Y.-L. Xu.** 1989. HTLV-1 tax gene product activates transcription via pre-existing cellular factors and c-AMP response element. J. Biol. Chem. **264:**15236–15241.
14. **Gonzalez, G. A., K. K. Yamamoto, W. H. Fischer, D. Karr, P. Manzel, W. Biggs III, W. W. Vale, and M. R. Montminy.** 1989. A cluster of phosphorylation sites on the c-AMP regulated nuclear factor CREB predicted from its sequence. Nature (London) **337:**749–752.
15. **Hai, T., and T. Curran.** 1991. Cross-family dimerization of transcription factors Fos/Jun and ATF/CREB alters DNA binding specificity. Proc. Natl. Acad. Sci. USA **88:**3720–3724.
16. **Hai, T., F. Liu, W. J. Coukos, and M. R. Green.** 1989. Transcription factor ATF cDNA clones: an extensive family of leucine zipper proteins able to selectively form DNA-binding heterodimers. Genes Dev. **3:**2083–2090.
17. **Han, K., M. S. Levine, and J. L. Manley.** 1989. Synergistic activation and repression of transcription by Drosophila homeobox proteins. Cell **56:**573–583.
18. **Harshman, K. D., W. S. Moye-Rowley, and C. S. Parker.** 1988. Transcriptional activation by the AP1 recognition element in yeast is mediated by a factor similar to AP1 that is distinct from GCN4. Cell **53:**321–330.
19. **Hoeffler, J. P., T. E. Meyer, Y. Yun, J. L. Jameson, and J. F. Habner.** 1988. Cyclic AMP-responsive DNA-binding protein: structure based on a cloned placental cDNA. Science **242:**1430–1433.
20. **Hurst, H. C., and N. C. Jones.** 1987. Identification of factors that interact with the E1a-inducible adenovirus E3 promoter. Genes Dev. **1:**1132–1146.
21. **Hurst, H. C., N. Masson, N. C. Jones, and K. A. W. Lee.** 1990. The cellular transcription factor CREB corresponds to activating transcription factor 47 (ATF-47) and forms complexes with a group of polypeptides related to ATF-43. Mol. Cell. Biol. **10:**6192–6203.
22. **Hurst, H. C., N. F. Totty, and N. C. Jones.** 1991. Identification and functional-characterization of the cellular activating transcription factor-43 (ATF-43) protein. Nucleic Acids Res. **19:**4601–4609.
23. **Ivashkiv, L. B., H.-C. Liou, C. J. Kara, W. W. Lamph, I. M. Verma, and L. H. Glimcher.** 1990. mXBP/CRE-BP2 and c-Jun form a complex which binds to the cyclic AMP, but not to the 12-*O*-tetradecanoylphorbol-13-acetate, response element. Mol. Cell. Biol. **10:**1609–1621.
24. **Jaynes, J. B., and P. H. O'Farrell.** 1991. Active repression of transcription by the engrailed homeodomain protein. EMBO J. **10:**1427–1433.
25. **Jones, C., and K. A. W. Lee.** 1991. E1a-mediated activation of the adenovirus E4 promoter can occur independently of the cellular transcription factor E4F. Mol. Cell. Biol. **11:**4297–4305.
26. **Kageyama, R., and I. Pastan.** 1989. Molecular cloning and characterization of a human DNA binding factor that represses transcription. Cell **59:**815–825.
27. **Kara, C. J., H. C. Liou, L. B. Ivashkiv, and L. H. Glimcher.** 1990. A cDNA for a human cyclic AMP response element-binding protein which is distinct from CREB and expressed preferentially in brain. Mol. Cell. Biol. **10:**1347–1357.
28. **Katoh, I., Y. Yoshinaka, and Y. Ikawa.** 1989. Bovine leukemia virus trans-activator P38 tax activates heterologous promoters with a common sequence known as the cAMP-responsive element or the binding site of a cellular transcription factor ATF. EMBO J. **8:**497–503.
29. **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.
30. **Landschulz, W. H., P. F. Johnson, and S. L. McKnight.** 1988. The leucine zipper: a hypothetical structure common to a new class of DNA binding proteins. Science **240:**1759–1764.
31. **Lee, K. A. W., J. S. Fink, R. H. Goodman, and M. R. Green.** 1989. Distinguishable promoter elements are involved in transcriptional activation by E1a and cyclic AMP. Mol. Cell. Biol. **9:**4390–4397.
32. **Lee, K. A. W., and M. R. Green.** 1987. A cellular transcription factor E4F1 interacts with an E1a-inducible enhancer and mediates constitutive enhancer function in vitro. Proc. Natl. Acad. Sci. USA **84:**8355–8359.
33. **Lewin, B.** 1990. Commitment and activation at DNA pol II

promoters: a tail of protein-protein interactions. Cell **61:**1161–1164.

34. **Licht, J. D., M. J. Grossel, J. Figge, and U. M. Hansen.** 1990. Drosophila Kruppel protein is a transcriptional repressor. Nature (London) **346:**76–79.
35. **Lieberman, P. M., J. M. Hardwick, J. Sample, G. S. Hayward, and S. D. Hayward.** 1990. The zta transactivator involved in induction of lytic cycle gene expression in Epstein-Barr virus-infected lymphocytes binds to both AP-1 and ZRE sites in target promoter and enhancer regions. J. Virol. **64:**1143–1155.
36. **Madden, S. L., D. M. Cook, J. F. Morris, A. Gashler, V. P. Sukhatme, and F. J. Rausher.** 1991. Transcriptional repression by the WT1 Wilms tumour gene product. Science **253:**1550–1553.
37. **Maekawa, T., H. Sakura, C. Kanei Ishii, T. Sudo, T. Yoshimura, J. Fujisawa, M. Yoshida, and S. Ishii.** 1989. Leucine zipper structure of the protein CRE-BP1 binding to the cyclic AMP response element in brain. EMBO J. **8:**2023–2028.
38. **Maniatis, T., S. Goodbourn, and J. A. Fischer.** 1987. Regulation of inducible and tissue-specific gene expression. Science **236:**1237–1244.
39. **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.
40. **Norman, C., M. Runswick, R. Pollock, and R. Treisman.** 1988. Isolation and properties of cDNA clones encoding SRF, a transcription factor that binds to the c-Fos serum response element. Cell **55:**989–1003.
41. **Pollock, R., and R. Treisman.** 1990. A sensitive method for the determination of protein-DNA binding specificities. Nucleic Acids Res. **18:**6197–6204.
42. **Ptashne, M.** 1988. How transcriptional activators work. Nature (London) **335:**683–689.
43. **Ptashne, M., and A. A. F. Gann.** 1990. Activators and targets. Nature (London) **346:**329–331.
44. **Ray, S. K., and P. S. Raychaudhuri.** 1991. Personal communication.
45. **Raychaudhuri, P., S. Bagchi, and J. R. Nevins.** 1989. DNA binding activity of the adenovirus-induced E4F transcription factor is regulated by phosphorylation. Genes Dev. **3:**620–627.
46. **Raychaudhuri, P., R. Rooney, and J. R. Nevins.** 1987. Identification of an E1a-inducible factor that interacts with regulatory sequences within the adenovirus E4 promoter. EMBO J. **6:**4073–4081.
47. **Rooney, R. J., P. Raychaudhuri, and J. R. Nevins.** 1990. E4F and ATF, two transcription factors that recognize the same site, can be distinguished both physically and functionally: a role for E4F in E1a *trans* activation. Mol. Cell. Biol. **10:**5138–5149.
48. **Schüle, R., P. Rangarajan, S. Kliewer, L. J. Ransone, J. Bolado, N. Yang, I. M. Verma, and R. M. Evans.** 1990. Functional antagonism between oncoprotein c-jun and the glucocorticoid receptor. Cell **62:**1217–1226.
49. **Schüle, R., K. Umesono, D. J. Mangelsdorf, J. Bolado, J. W. Pike, and R. M. Evans.** 1990. Jun-fos and receptors for vitamins A and D recognize a common response element in the osteocalcin gene. Cell **61:**497–504.
50. **Shi, Y., E. Seto, L.-S. Chang, and T. Shenk.** 1991. Transcriptional repression by YY1, a human GL1-Kruppel-related protein, and relief of repression by adenovirus E1a protein. Cell **67:**377–388.
51. **Singh, H., J. H. LeBowitz, A. S. Baldwin, and P. A. Sharp.** 1988. Molecular cloning of an enhancer binding protein: isolation by screening of an expression library with a recognition site DNA. Cell **52:**415–423.
52. **Vinson, C. R., P. B. Sigler, and S. L. McKnight.** 1989. Scissors-grip model for DNA recognition by a family of leucine zipper proteins. Science **246:**911–916.
53. **Xu, Y. L., N. Adya, E. Siores, Q. Gao, and C. Z. Giam.** 1990. Cellular factors involved in transcription and Tax-mediated trans-activation directed by the TGACGT motifs in human T-cell leukemia virus type I promoter. J. Biol. Chem. **265:**20285–20292.
54. **Yamamoto, K. K., G. A. Gonzalez, P. Menzel, J. Rivier, and M. R. Montminy.** 1990. Characterization of a bipartite activator domain in transcription factor CREB. Cell **60:**611–617.
55. **Yang Yen, H. F., J. C. Chambard, Y. L. Sun, T. Smeal, T. J. Schmidt, J. Drouin, and M. Karin.** 1990. Transcriptional interference between c-Jun and the glucocorticoid receptor: mutual inhibition of DNA binding due to direct protein-protein interaction. Cell **62:**1205–1215.
56. **Yoshimura, T., J. Fujisawa, and M. Yoshida.** 1990. Multiple cDNA clones encoding nuclear proteins that bind to the tax-dependent enhancer of HTLV-1: all contain a leucine zipper structure and basic amino acid domain. EMBO J. **9:**2537–2542.