JOURNAL OF VIROLOGY, Nov. 1992, p. 6338–6352
0022-538X/92/116338-15$02.00/0

Vol. 66, No. 11

Characterization of Nuclear Proteins That Bind the EFII Enhancer Sequence in the Rous Sarcoma Virus Long Terminal Repeat

ROSALIE C. SEARS[1] AND LINDA SEALY[1,2]*

Department of Cell Biology[1] and Department of Molecular Physiology and Biophysics,[2] Vanderbilt University School of Medicine, Nashville, Tennessee 37232

Received 22 April 1992/Accepted 23 July 1992

The EFII *cis* element is a 38-bp sequence at the 5′ end of the Rous sarcoma virus long terminal repeat, extending from nucleotides −229 to −192 (with respect to the viral transcription start site), which is recognized by sequence-specific DNA-binding proteins in avian fibroblast nuclear extracts (L. Sealy and R. Chalkley, Mol. Cell. Biol. 7:787–798, 1987). We demonstrate that multiple copies of the EFII *cis* element strongly activate transcription of a reporter gene in vivo. We correlate the region of the EFII *cis* element which activates transcription in vivo with the in vitro binding site for three nuclear factors, EFIIa, EFIIb, and EFIIc. The sequence motif recognized by EFIIa, -b, and -c is also found in consensus binding sites for members of a rapidly growing family of transcription factors related to the CCAAT/enhancer-binding protein (C/EBP). EFIIa, -b, and -c are present in fibroblast and epithelial cell lines from various species but are much less abundant in differentiated rat liver and kidney cells. The EFIIa binding activity is particularly abundant in an avian B-cell lymphoma line. As judged from molecular weight analysis, cell type distribution, and sequence recognition properties, the EFII factors under study appear to differ from most of the previously described C/EBP-related factors and thus may expand the diversity of the C/EBP family.

The information accumulated about the regulation of gene expression at the level of transcription initiation in eukaryotic cells has revealed an impressive degree of complexity. Among the *cis*-acting DNA sequences that regulate transcription, enhancers are now known to be mosaics of multiple sequence motifs. Each of these sequence motifs represents a discrete binding site for one or more *trans*-acting factors. All of these *cis*-acting elements and regulatory protein factors working in concert specify appropriate levels of transcription (31, 44, 49, 57). Many viruses have evolved to efficiently utilize these host cell mechanisms to promote high-level expression of their own genomes. For example, the avian retrovirus Rous sarcoma virus (RSV) contains strong *cis*-acting enhancer sequences in the long terminal repeat (LTR) regions flanking either end of the proviral DNA (13, 14, 23, 25, 38, 42, 52, 69, 71). The potent transcriptional activity of the RSV LTR enhancer makes it an attractive model system for studying the molecular events involved in the regulation of transcription initiation. Thus, we have been identifying and characterizing nuclear factors which bind in a sequence-specific manner to the RSV LTR enhancer. Although the typical target cells for transformation by RSV and its relative, Rous-associated virus-1, are avian fibroblasts and lymphoid cells, respectively, the RSV LTR enhancer is active in a wide variety of vertebrate cell types (23, 38, 71). This suggests that many of the *trans*-acting proteins which mediate the LTR enhancer activity are common host cell transcription factors.

The major transcriptional control regions for RSV have been localized by deletion mutagenesis (13, 25, 38, 42, 52) and enhancer trap experiments (71) to the U3 region of the RSV LTR, extending from the 5′ end of the LTR at position −229 (relative to the transcription start site) to position −54. We have so far described three nuclear factors, enhancer factors I, II, and III (EFI, EFII, and EFIII, respectively), which recognize specific nucleotide sequences within the RSV LTR enhancer diagrammed in Fig. 1 (4, 5, 17, 64). Further analysis of the EFI protein factor has demonstrated that it specifically recognizes two inverted CCAAT motifs (17). The EFIII factor recognizes two sites (4, 5) which contain a common sequence motif known as the CArG box (48, 66). By its sequence-specific DNA-binding properties and antibody recognition, EFIII has been shown to represent the avian homolog to the serum response factor (4, 5). We now report the further characterization of a number of protein factors present in chicken embryo fibroblast (CEF) nuclei which bind to the EFII region of the RSV LTR enhancer.

Other investigators have also observed nuclear proteins which bind specifically in vitro to the EFII region of the RSV LTR (22, 61) and to related but not identical sequences in the Rous-associated virus-2 LTR (59, 60). However, when this work was begun, the transcriptional relevance of the EFII *cis* element (RSV LTR sequences from −229 to −192) had not been directly addressed. Mutagenesis experiments had demonstrated that removal of the EFII sequences as part of larger deletions crippled enhancer function (13, 42, 52), and smaller deletions demonstrated that at least eight nucleotides (−201 to −208) within the EFII *cis* element are required for maximal enhancer activity (25). We present evidence here that an isolated EFII *cis* element, when multimerized, strongly enhances transcription in vivo. This finding is in agreement with recent experiments performed in GC rat pituitary tumor cells, which show that the EFII *cis* element (RSV LTR sequences −231 to −193), linked to a minimal heterologous promoter, is transcriptionally active, and its activity is increased by expression of a constitutive, calcium/calmodulin-independent mutant of type II calcium/calmodulin-dependent protein kinase (32). We further demonstrate

* Corresponding author.

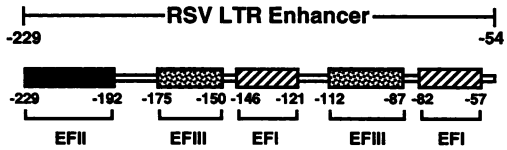

FIG. 1. Schematic representation of the RSV LTR enhancer. Binding sites for three nuclear factors, designated EFI, EFII, and EFIII, are shown. The EFI factor has been found to specifically recognize inverted CCAAT motifs present within each of its binding sites (17, 64). The EFII binding site has been mapped by DNase I footprinting at the 5′ end of the RSV LTR (64). The EFIII factor is suggested to be the avian homolog to the serum response factor, and it has been found to specifically recognize the CArG motif present in both its binding sites (4, 5).

that the upstream most of two near-direct repeat sequences present in the EFII *cis* element is responsible for mediating strong transcriptional activity in vivo. At least three heat-stable protein complexes, referred to as EFIIa, EFIIb, and EFIIc, are identified that specifically bind the EFII DNA sequence in vitro. These three complexes bind selectively to the upstream repeat with high affinity. Thus, the transcriptionally active sequences in the EFII *cis* element in vivo correspond to the EFIIa, EFIIb, and EFIIc binding site in vitro.

All three EFII-binding factors appear to recognize the same nucleotides within the upstream repeat of the EFII DNA sequence. It is now well documented that many transcription factors occur in families which share DNA sequence recognition properties (30). The nucleotide sequence recognized by EFIIa, -b, and -c is also found in consensus binding sites for members of a rapidly growing family of transcription factors related to the previously described CCAAT/enhancer-binding protein (C/EBPα) (1, 7, 10, 15, 27, 29, 34, 39, 56, 58, 61). C/EBP-related family members belong to the basic region-leucine zipper (bZIP) class of transcription factors. bZIP transcription factors are characterized by a conserved DNA-binding domain, containing clusters of basic amino acids, immediately adjacent to a conserved dimerization domain containing a leucine residue at nearly every seventh position, known as the leucine zipper (36, 40, 41, 62). Three reported members of the C/EBP family of transcription factors, C/EBPα, IBF, and Ig/EBP-1 (C/EBPγ), have been shown to be able to bind the EFII region of the RSV LTR in vitro (33, 58, 61). We demonstrate, based on different sequence-specific DNA recognition properties and apparent-molecular-weight analyses, that EFIIa, -b, and -c are probably distinct from C/EBPα and IBF. Moreover, expression of most of the identified C/EBP-related nuclear factors is particularly abundant in differentiated liver, lung, and adipose cells (1, 3, 7, 10, 11, 15, 18, 56, 73), while EFIIa, -b, and -c are much less abundant in differentiated rat liver and kidney cells than in numerous fibroblast and epithelial cell lines from various species. However, Ig/EBP-1, EFIIa, and EFIIb have very similar sequence recognition characteristics, as judged from their ability to bind both the EFII sequence and the E site in the murine immunoglobulin heavy-chain enhancer with high affinity (58). We also note that there is a large excess of the EFIIa factor in an avian B-cell lymphoma line. Thus, it is possible that the EFIIa protein plays a role in B-cell-specific gene expression, and further characterization of the three EFII-binding factors may expand the diversity of the C/EBP-related family of transcription factors.

MATERIALS AND METHODS

Cell culture. CEF were prepared from 10-day-old embryos (SPAFAS Inc., Preston, Conn.) and maintained in medium 199 supplemented with 10% tryptose phosphate broth, 5% fetal calf serum, and 1% chicken serum. The following were also added to the medium: 25 U of penicillin G sodium per ml, 25 μg of streptomycin sulfate per ml, and 0.14% sodium bicarbonate. Dulbecco's modified Eagle's medium (DMEM) (high glucose) supplemented with 10% calf serum was used for all other cell lines, including A431 epidermal carcinoma cells, BALB/c-3T3 fibroblasts, NIH 3T3 fibroblasts, baby hamster kidney (BHK) cells, Rat-1 fibroblasts, avian B-cell lymphoma line Bk3A, and Cos cells, with the following exceptions. BHK cells were maintained in a 1:1 mixture of DMEM and Ham's F12 medium; DMEM for Rat-1 cells was supplemented with 10% fetal calf serum; DMEM for Bk3A cells was supplemented with 10% tryptose phosphate broth, 5% calf serum, and 1% heat-inactivated (60°C, 30 min) chicken serum. BHK cells were a gift from Roger Chalkley (Vanderbilt University), BALB/c-3T3 and NIH 3T3 cells were a gift from Jack Pledger (Vanderbilt University), Rat-1 cells were a gift from Mike Bishop (University of California at San Francisco), Cos and Bk3A cells were a gift from Steve Hann (Vanderbilt University), and A431 cells were a gift from Graham Carpenter (Vanderbilt University).

Plasmid construction. pSRA-CAT was generated in this laboratory as described before (4). pe− CAT was constructed as follows. Viral sequences in pRSV-LTR (64) from the *Eco*RI site at −54 to the *Bam*HI site at +524 (numbering with reference to the start site of transcription in the viral genome) were cloned into the *Pvu*II site of pBR322 to create pRSV-LTRe−. pRSV-LTRe− was digested with *Bst*EII at +103 in the viral sequence, treated with DNA polymerase I at 4°C to create blunt ends as described before (64), and then digested with *Bam*HI at position +524. pSV2-CAT (24) was digested with *Hin*dIII, treated as above to create blunt ends, and then digested with *Bam*HI. The *Hin*dIII-*Bam*HI fragment containing the chloramphenicol acetyltransferase (CAT) gene was then inserted into pRSV-LTRe−, replacing the 423-bp *Bst*EII-*Bam*HI fragment to generate pe− CAT. The *Hin*dIII site from the CAT gene, now at position +103 in pe− CAT (with respect to the start site for transcription), was not regenerated.

The p(EFII)*n*CAT series of reporter constructs were created as follows. The following double-stranded 44-bp oligonucleotide containing the 38-bp EFII *cis* element (−229 to −192 in the RSV LTR are shown in capital letters):

5′ccgagAATGTAGTCTTATGCAATACTCTTGTAGTCTTGCAACAc
3′cTTACATCAGAATACGTTATGAGAACATCAGAACGTTGTgggct5′

with *Ava*I ends, was synthesized and prepared as described below. This 44-bp EFII oligonucleotide was treated with polynucleotide kinase and 1 mM ATP and ligated with T4 DNA ligase (according to the manufacturer's specifications) to create head-to-tail multimers. EFII multimers were treated as above to create blunt ends. pe− CAT was digested with *Acc*I at a unique site 235 nucleotides upstream from the start of transcription and treated as above to create blunt ends. EFII multimers were then inserted at the *Acc*I site to create the p(EFII)*n*CAT constructs. p(EFII3′/5′)2+CAT was created as follows. The double-stranded 50-bp EFII 3′/5′ mutant oligonucleotide (see Fig. 3A for sequence), synthesized and prepared as described below, was treated with polynucleotide kinase and 1 mM ATP and purified by polyacrylamide gel electrophoresis, electroelution, and Qia-

gen chromatography (Qiagen Inc., Chatsworth, Calif.) prior to incubation with T4 DNA ligase at 16°C. The ligated EFII 3′/5′ mutant oligonucleotide was then inserted into the vector pe− CAT, which had been previously digested with *Acc*I and treated as above to generate blunt ends.

The reporter plasmids p(EFII5′)4−CAT4 and p(EFII3′)2−CAT4 were created as follows. The following double-stranded 20-bp and 18-bp EFII oligonucleotides, containing the 5′ EFII sequences from −226 to −213 (20 bp) or the 3′ sequences from −205 to −194 (18 bp) in the RSV LTR (capital letters):

20 bp: 5′ccgagGTAGTCTTATGCAAc3′
3′cCATCAGAATACGTTgggct5′
18 bp: 5′ccgagGTAGTCTTGCAAc3′
3′cCATCAGAACGTTgggct5′

with *Ava*I ends, were synthesized and prepared as described below. The 20-bp and 18-bp oligonucleotides were then treated with polynucleotide kinase and 1 mM ATP prior to incubation with T4 DNA ligase (according to the manufacturer's specifications). The ligated 20-bp oligonucleotide was then inserted into the *Bam*HI site in the polylinker of a modified Bluescript M13+ (Stratagene, La Jolla, Calif.) plasmid created in this laboratory which lacks both *Xho*I and *Sma*I sites, after treatment of both the ligated oligonucleotide and the modified Bluescript vector as described above to create blunt ends. A clone containing four copies of the 20-bp oligonucleotide in the modified Bluescript vector, designated pII20x4, was digested with *Bam*HI and *Hin*dIII to generate a fragment of 120 bp containing the four ligated 20-bp oligonucleotides. This fragment was then inserted between the *Bam*HI and *Hin*dIII sites of the polylinker at position −105, immediately upstream of the thymidine kinase promoter in the plasmid pBLCAT4 (6), to create p(EFII5′)4−CAT4. The ligated 18-bp oligonucleotide was treated to create blunt ends and inserted into the plasmid pBLCAT4, which had been previously digested with *Bam*HI (position −105) and treated to create blunt ends, to produce the reporter plasmid p(EFII3′)2−CAT4.

The plasmid pMLS-II, used to make the end-labeled probe for footprinting, was generated as follows. The 44-bp EFII oligonucleotide described above was blunt-ended and inserted into the unique *Sma*I site in the polylinker of a modified Bluescript+ which had been previously digested with *Xho*I, blunt-ended as described above, and religated to destroy the *Xho*I site. The *Ava*I site on either end of the insert was regenerated.

The inserts in all of the above plasmids were sequenced (45) to confirm insert number and orientation and the absence of mutations. We noted that the sequence of both copies of the EFII 3′/5′ mutant oligonucleotide in the p(EFII3′/5′)2+CAT clone differed from the sequence given in Fig. 3A in that both are missing the terminal G of the *Bam*HI linker sequence inserted to disrupt the 3′ repeat.

Transfection and enzymatic assays. Transfections were performed by the calcium phosphate coprecipitation technique as described by Graham and Van der Eb (26), with 1 μg of test plasmid DNA, 2.5 μg of pSV40-β-gal DNA (43), and 1.5 μg of pUC19 DNA per 10 ml of medium in 100-mm tissue culture dishes. Cells were plated the day before transfection and transfected when they were approximately 80 to 90% confluent. One hour prior to transfection, CEF were refed with fresh medium 199 (supplemented as described above) containing 10 mM HEPES (*N*-2-hydroxyethylpiperazine-*N*′-2-ethanesulfonic acid, pH 7.3). Cells were exposed to the $CaPO_4$ DNA precipitate for 6 to 8 h, after which they were washed once with Tris-glucose buffer containing 137 mM NaCl, 5 mM KCl, 5.6 mM dextrose, 25 mM Tris (pH 7.4), 25 U of penicillin G sodium per ml, 25 μg of streptomycin sulfate per ml, and 4.5 μg of phenol red sodium salt per ml and then refed with fresh supplemented medium 199. Transfected cells were allowed to grow for an additional 48 h. Cells were then washed twice in cold (4°C) phosphate-buffered saline (PBS) and incubated for 5 min at 37°C in buffer containing 40 mM Tris (pH 7.5), 1 mM EDTA (pH 8), and 150 mM NaCl to loosen cells from the tissue culture dish.

All subsequent steps involving transfected cells were carried out at 4°C. Cells were collected by centrifugation (5 min at 3,700 × *g*) and lysed by sonication, and cell extracts were prepared as described by Boulden and Sealy (4). CAT assays with these cell extracts were performed by the procedure of Gorman et al. (24). CAT activity was calculated for each test plasmid by cutting out the acetylated and unreacted forms of [^{14}C]chloramphenicol, quantitating the radioactivity in each by liquid scintillation counting, and determining the percent acetylation. Units of β-galactosidase activity in the same cell extracts were assayed and calculated as described by Norton and Coffin (51). CAT activity was normalized by dividing percent acetylation by units of β-galactosidase activity for each test plasmid.

Nuclear extracts. Nuclear extracts (0.5 M NaCl) were prepared from 14-day-old chicken embryos (CE) (SPAFAS) as previously described by Sealy and Chalkley (64), with the modifications described by Boulden and Sealy (4). The following changes were also made. The phosphatase inhibitors β-glycerolphosphate (10 mM), sodium vanadate (100 μM), and sodium molybdate (10 mM) were added to all buffers. The protease inhibitor benzamidine (10 mM) replaced leupeptin and pepstatin in buffer G. Homogenized embryos were filtered through cheesecloth prior to passage through Miracloth (Calbiochem Corp., La Jolla, Calif.). Nuclei were prepared and sequentially extracted by one wash with buffer A (5 ml per embryo), one extraction with buffer B (1 ml per embryo), and one extraction with buffer B containing 0.63 M NaCl (3 ml per embryo). This yielded an approximately 0.5 M NaCl extract because of the volume of the nuclear pellet. Following this extraction, the supernatant was collected by ultracentrifugation in an SW27 rotor at 25 krpm for 60 min. The supernatant was dialyzed against 50 volumes of dialysis buffer containing 10 mM HEPES (pH 8), 1 mM EDTA (pH 8), 50 mM NaCl, 50% glycerol, 7 mM β-mercaptoethanol, 1 mM phenylmethylsulfonyl fluoride, 10 mM benzamidine, and the phosphatase inhibitors mentioned above.

Nuclear extracts (0.5 M NaCl) from cells grown in monolayers, including CEF, A431, BALB/c-3T3, NIH 3T3, BHK, Cos, and Rat-1 cells, were prepared as described by Sealy and Chalkley (64), with the following modifications. The phosphatase inhibitors mentioned above were added to all buffers. Nuclei were prepared by only one wash in buffer A. The crude nuclei were then resuspended directly in buffer B containing 0.53 M NaCl (yielding an approximately 0.5 M NaCl extract because of the volume of the nuclear pellet). The supernatant was collected at 16,000 × *g* and dialyzed as described above. Nuclear extracts (0.5 M NaCl) from Bk3A cells were prepared as described for the above monolayer cells except that these suspension cells were pelleted at 750 × *g* for 10 min, washed twice with cold (4°C) PBS, and resuspended in buffer A to prepare nuclei. The 0.5 M NaCl extraction was performed with 1 ml per 10^9 cells. Heated nuclear extracts were prepared as described for the mono-

layer cells with the following modifications. After the 30-min 0.5 M NaCl extraction, the mixture was heated for 5 min at 100°C, placed on ice for 10 min, and subjected to ultracentrifugation in an SW56 rotor at 30 krpm for 30 min. The supernatant was dialyzed as described above. Nuclear extracts (0.5 M NaCl) prepared from the following cells were gifts from the indicated individuals and were prepared as described above: H4, Tony Ip (Vanderbilt University); adult rat liver and kidney, Michelle Cissell (Vanderbilt University). Protein concentrations in extracts were determined by the method of Schaffner and Weissman (63), with bovine serum albumin as the standard.

Radiolabeled and competitor DNAs. Oligonucleotides were synthesized (Diabetes Research Training Center DNA core, Vanderbilt University) and purified as described by Boulden and Sealy (4). Equal amounts of each oligonucleotide, measured by the A_{260}, and its complement were annealed in buffer containing 10 mM Tris (pH 8), 1 mM EDTA (pH 8), and from 200 to 400 mM NaCl by placing the DNAs in a 2-liter boiling-water bath, which was allowed to cool slowly to room temperature overnight. Coding-strand nucleotide sequences of the double-stranded oligonucleotides used for the DNA-binding studies are displayed in Fig. 3A and Table 1, except for the oligonucleotides labeled Core, DEI, IBF DNA, and IgH-E, which are shown below:

Core: 5'agcttGGGCTGTGGAAAGGAGGGc3'
3'aCCCGACACCTTTCCTCCCgagct5'
DEI: 5'tcgacTATGATTTTGTAATGGGGctcga3'
3'agctgATACTAAAACATTACCCCgagct5'
IBF DNA: 5'CTGTTGGCTGCAATTGCGCCACCGCCACAG3'
3'GACAACCGACGTTAACGCGGTGGCGGTGTC5'
IgH-E: 5'TTAAGTTTAAAATATTTTTAAATGAATTGAGCAATGTTGAGTTGAGTCAA3'
3'AATTCAAATTTTATAAAAATTTACTTAACTCGTTACAACTCAACTCAGTT5'

The Core oligonucleotide represents sequences from −48 to −65 from the collagen III promoter, which contains the simian virus 40 (SV40) core enhancer motif (50); linker sequences are indicated by lowercase letters. The DEI oligonucleotide represents sequences from −90 to −107 from the rat albumin promoter (9). IBF DNA represents sequences from the *gag* gene internal enhancer of RSV, designated OpvuII3' by Karnitz et al. (34). IgH-E represents sequences from position 301 to position 350 in the murine immunoglobulin heavy-chain enhancer as defined by Ephrussi et al. (16). Since the 44-bp EFII oligonucleotide described above was unable to bind EFII DNA-binding proteins when radiolabeled and used in an electrophoretic mobility shift assay (EMSA) (unpublished results), a 50-bp EFII oligonucleotide containing RSV LTR sequences from −229 to −192 with *Xho*I and *Bgl*II linkers on the 5' end and a *Sma*I linker on the 3' end was synthesized for use in EMSAs (see Fig. 3A).

Radiolabeled probes for EMSAs were prepared as follows. Double-stranded oligonucleotides (200 ng) were radiolabeled with [γ-^{32}P]ATP and T4 polynucleotide kinase according to the manufacturer's specifications. Full-length radiolabeled oligonucleotides were then purified by electrophoresis on native 12% polyacrylamide gels containing 1× TBE (89 mM Tris base, 89 mM boric acid, 2.5 mM EDTA [pH 8.5]), followed by electroelution in 1× TBE and ethanol precipitation. Nonradiolabeled competitor oligonucleotides were also purified by polyacrylamide gel electrophoresis (PAGE) in 1× TBE, and the amount of DNA recovered was quantitated by A_{260} measurements. All oligonucleotides were resuspended in 10 mM Tris–100 mM NaCl–1 mM EDTA (pH 8).

Radiolabeled probes for footprinting were prepared as follows. The plasmid pMLS-II (see above) was digested with either *Bam*HI (to label the coding strand) or *Hin*dIII (to label the noncoding strand) and end-labeled with [α-^{32}P]dATP, [α-^{32}P]dTTP, [α-^{32}P]dGTP, [α-^{32}P]dCTP, and DNA polymerase I at 4°C according to the manufacturer's specifications. End-labeled, linearized pMLS-II was then digested with either *Hin*dIII (if end-labeled at a *Bam*HI site) or *Bam*HI (if end-labeled at a *Hin*dIII site). The resulting 76-bp end-labeled restriction fragments were purified by PAGE as described above.

EMSAs. Samples of 0.5 M NaCl nuclear extract, containing between 1 and 10 μg of protein, depending on the cell type analyzed, were mixed with 0.5 ng of radiolabeled double-stranded oligonucleotide and 1.3 μg of poly(dI):poly(dC), prepared as described previously (64), in the presence or absence of nonradiolabeled, gel-purified competitor DNAs. Binding reaction mixes were incubated at room temperature for 30 min in a final volume of 20 μl, containing 7.5 mM HEPES (pH 8), 2.5 mM Tris (pH 8), 0.8 mM EDTA (pH 8), 100 mM NaCl, 3.3 mM $MgCl_2$, 5 mM β-mercaptoethanol, and 20% glycerol. Samples were analyzed on native 4% polyacrylamide gels containing 1× TGE (25 mM Tris, 190 mM glycine, and 1 mM EDTA [pH 8.5]) in the absence of tracking dye. Gels were subjected to electrophoresis in 1× TGE for 50 min at 200 V and subsequently dried for autoradiography. When appropriate, the protein-DNA complexes were excised from the dried gel, and the radioactivity in the complex was quantitated by liquid scintillation counting. Background counts were subtracted by counting an equivalent section of the dried gel outside the sample lanes. The relative binding affinity of protein-DNA complexes for specific nonradiolabeled competitor oligonucleotides was determined by calculating the percentage of radiolabeled DNA bound in a specific complex in the presence of competitor DNA relative to that observed in the absence of competitor DNA, plotted versus log (competitor DNA). The amount of competitor DNA required to reduce the formation of a specific complex by 50% was calculated from this plot and divided by the amount of wild-type competitor DNA (representing the sequence of the radiolabeled DNA used in the EMSA) needed to reduce the amount of the same complex by 50% to obtain relative binding affinities.

OP/Cu footprinting. Samples of 0.5 M NaCl nuclear extract containing either 12 or 4 μg of protein from CEF or Bk3A cells, respectively, were mixed with 100 to 200 kcpm (approximately 2 ng) of end-labeled probe (see above) in the presence of 1.3 μg of poly(dI):poly(dC) and subjected to EMSA as described above with the following modification: 20-μl binding reaction mixes were analyzed on 6% polyacrylamide gels containing 1× TG (25 mM Tris base, 190 mM glycine [pH 8.5]). The gels were subjected to electrophoresis in 1× TG for 1.5 h at 200 V. Following electrophoresis, gels were washed once with 50 mM Tris (pH 8) for 15 min at room temperature, placed in fresh 50 mM Tris, and treated in situ with orthophenanthroline-copper (OP/Cu) endonuclease as described by Kuwabara and Sigman (37) for 6 min at room temperature. Free and bound complexes, visualized by autoradiography of the wet gel at 4°C, were excised from the gel and electroeluted in 1× TBE. Cleaved end-labeled DNAs were then ethanol precipitated with 5 μg of carrier tRNA. Samples containing equal counts (defined by Cerenkov radiation), and the end-labeled probes chemically cleaved at purine residues, were separated on a 20% polyacrylamide sequencing gel as described by Maxam and Gilbert (45).

Apparent-molecular-mass analysis. Samples of 0.5 M NaCl nuclear extract containing 100 or 30 μg of protein from 14-day CE or Bk3A cells, respectively, were subjected to fractionation on a sodium dodecyl sulfate (SDS)–12% polyacrylamide gel, transferred to an Immobilon-P membrane (Millipore Corp., Bedford, Mass.), and eluted, as described by Faber and Sealy (17), except that the Immobilon-P membrane corresponding to the separating gel was cut into 17 slices, each approximately 3-mm wide, and the proteins in each slice were eluted into 40 μl of cold (4°C) elution solution. Samples (10 μl) of each SDS-sized fraction were analyzed by EMSA. To determine the range of apparent molecular masses of the proteins in each fraction, individual protein standards were separated on the same SDS–12% polyacrylamide gel, transferred to an Immobilon-P membrane, and visualized by amido black staining as described by Schaffner and Weissman (63).

RESULTS

Specific sequences within the EFII *cis* element activate transcription in vivo. We have previously demonstrated by DNase I footprinting that a factor(s) (termed EFII) in nuclear extracts from avian fibroblasts specifically binds to a 38-bp sequence at the 5′ end of the LTR of the Schmidt Ruppin-A strain of RSV, extending from −229 to −192 nucleotides upstream of the start site for transcription (64). To assess the functional significance of this region of the LTR enhancer for transcriptional activation, we inserted one or more copies of the EFII binding site into the reporter plasmid pe− CAT. As shown in Fig. 2A, pe− CAT is similar to the plasmid pSRA-CAT, which was described previously (4), except that sequences in U3 responsible for LTR enhancer activity have been deleted, leaving just 54 nucleotides upstream of the start site for transcription. These 54 nucleotides, which include a TATA box at approximately position −30, are sufficient to direct accurate but very low levels of transcription of the CAT gene in vivo (20). Up to six copies of the EFII oligonucleotide were inserted into pe− CAT at a position (−235) comparable to that found in the wild-type RSV LTR. The resulting constructs, designated p(EFII)*n*CAT, were then tested for their ability to activate transcription in transient-transfection assays in CEF. As shown in Fig. 2B, multiple copies of the EFII *cis* element, in either the sense or antisense orientation, were able to strongly activate transcription of the CAT gene compared with the enhancer-minus construct (pe− CAT). Six copies of the EFII *cis* element in the sense orientation [p(EFII)6+CAT] were able to activate transcription approximately 40-fold.

Inspection of the nucleotide sequence of the EFII *cis* element (Fig. 3A) reveals that a nearly direct repeat of 15 nucleotides is present in the LTR sequence from −227 to −213 and from −206 to −194, with the 3′ repeat lacking an AT present in the 5′ repeat. To determine whether the presence of these repeated sequences is important for transcriptional activation, we created a mutant EFII *cis* element in which both repeats were disrupted by 10-bp *Bam*HI linker sequences (EFII 3′/5′ mutant; see Fig. 3A for sequence). Two copies of this mutant oligonucleotide were also inserted into the pe− CAT vector at −235. When the resulting construct, p(EFII3′/5′)2+CAT, was tested in transient-transfection assays, the mutant EFII *cis* element failed to activate transcription above the level observed with pe− CAT (Fig. 2C). Thus, it appears that one or both of the repeated sequences in the EFII *cis* element are necessary for transcriptional activation, presumably by acting as a binding site(s) for one or more *trans*-acting factors.

In order to directly test the transcriptional activity of each near-direct repeat, we analyzed two abbreviated EFII *cis* elements in which either only the upstream 5′ repeat, from −226 to −213, or only the downstream 3′ repeat, from −205 to −194, was reiterated in a reporter gene construct. Specifically, four copies of the 5′ repeat or two copies of the 3′ repeat (each separated by a 6-bp *Ava*I linker sequence) were inserted immediately upstream of the thymidine kinase promoter and CAT gene in the vector pBLCAT4 (6), creating p(EFII5′)4−CAT4 and p(EFII3′)2−CAT4, respectively. Interestingly, in transient-transfection assays, p(EFII5′)4−CAT4 activated transcription 36-fold relative to pBLCAT4, while p(EFII3′)2−CAT4 was unable to activate transcription above the basal activity of pBLCAT4 (Fig. 2C). Although we have not compared equal copy numbers because such constructs were not available, it is unlikely that more copies of the 3′ repeat would lead to transactivation equivalent to that with the 5′ repeat when two copies are inactive. In general, transactivation by multimerized *cis* elements increases incrementally with increasing copy number, as seen for the complete EFII *cis* element (Fig. 2B). Our data suggest that transcriptional activation by the EFII *cis* element is mediated primarily through the 5′ near-direct repeat sequence. Thus, we wished to further characterize the protein factors which specifically bind this element in vitro and, in particular, determine whether they recognize the upstream repeat.

Nuclear proteins which bind to the EFII *cis* element in vitro specifically recognize the upstream repeat. In our previous study, we identified two EFII DNA-binding factors present in 0.3 to 0.5 M NaCl extracts of quail fibroblast nuclei which specifically bound the EFII *cis* element contained in a 286-bp restriction fragment encompassing the entire RSV LTR enhancer (64). In order to facilitate further analysis of proteins recognizing the EFII *cis* element, we repeated this experiment with a double-stranded 50-bp oligonucleotide, referred to as EFII DNA (see Fig. 3A for coding-strand sequence information), containing the 38-bp EFII *cis* element with linker sequences on both ends. As shown in Fig. 3D, lane 1, two major protein-DNA complexes, which we refer to as EFIIa and EFIIb, are present in an EMSA of radiolabeled EFII DNA following incubation with heated (100°C for 5 min) 0.5 M NaCl nuclear extract prepared from 14-day-old CE. A third minor protein-DNA complex of slower electrophoretic mobility, designated EFIIc, was also detected. The two faster-migrating protein-DNA complexes likely arise from some abundant, nonspecific DNA-binding proteins in the high-NaCl extract, since they were not inhibited by the addition of a variety of nonradiolabeled competitor DNAs (data not shown). Results identical to those presented for the 14-day-old CE nuclear extract were observed with both heated and unheated 0.5 M NaCl nuclear extracts from CEF (data not shown).

To assess whether binding of EFIIa, EFIIb, or EFIIc in vitro would correlate with transcriptional activation in vivo, we investigated the role of the two repeated sequences in the EFII *cis* element in the formation of the three complexes. In addition to the EFII 3′/5′ mutant described above, we obtained two EFII mutant oligonucleotides in which either the 5′ repeat (EFII 5′ mutant) or the 3′ repeat (EFII 3′ mutant) was disrupted by a 10-bp *Bam*HI linker sequence (Fig. 3A). Representative competition experiments with these mutant oligonucleotides and the wild-type EFII DNA are shown in Fig. 3B. The decrease in relative binding affinities of the EFIIa, -b, and -c factors for each of the

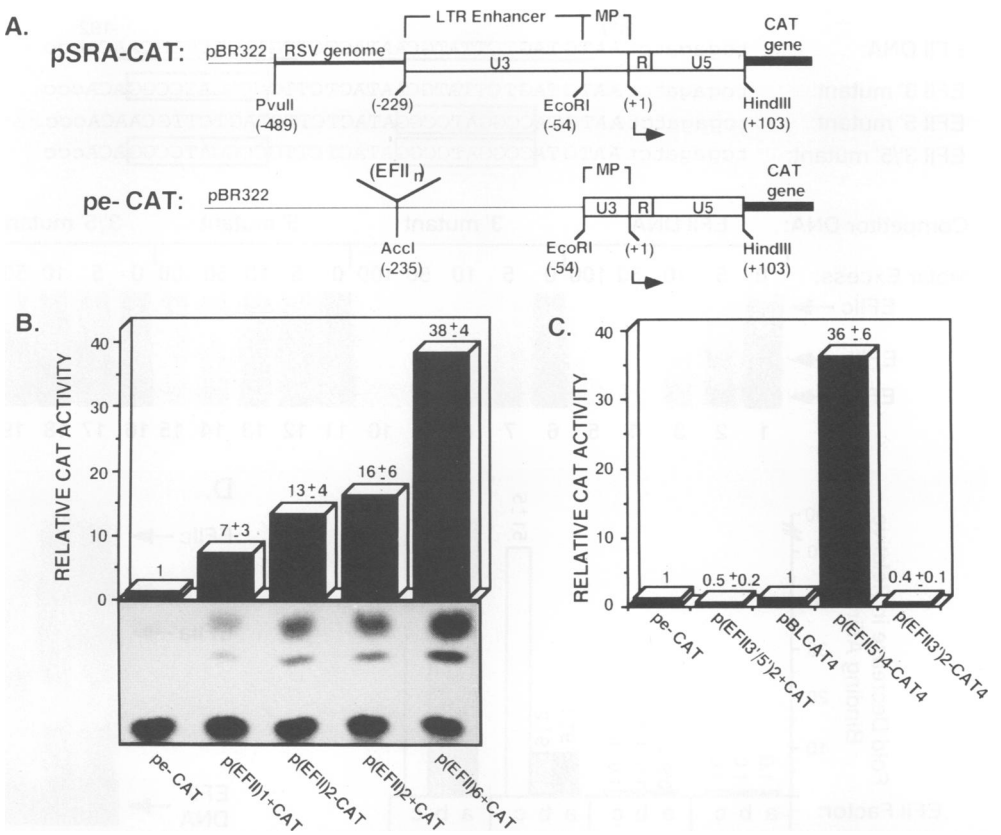

FIG. 2. Specific sequences within the EFII *cis* element activate transcription in vivo. (A) pSRA-CAT is a chimeric construct containing a CAT reporter gene linked to the 330-bp RSV LTR and 260 bp of upstream sequence normally found at the 3′ end of the RSV provirus genome. The RSV LTR enhancer extends from −229 to −54 in pSRA-CAT. pe− CAT is an enhancer-minus construct which lacks the RSV genome sequences from the *Pvu*II site at −489 to the *Eco*RI site at −54. The sequences from −54 to +1 in both plasmids represent a minimal promoter (MP) containing a TATA box at approximately −30. (B) Up to six copies of the EFII *cis* element were inserted into the plasmid pe− CAT at the indicated *Acc*I site, creating the p(EFII)*n*CAT series, in which 1, 2, or 6 copies of the EFII *cis* element in either the sense (+) or antisense (−) orientation were inserted. Samples (1 μg) of either pSRA-CAT, pe− CAT, or p(EFII)*n*CAT were transiently transfected into CEF along with 2.5 μg of pSV40-β-gal, encoding β-galactosidase, as an internal control. A representative CAT assay is shown for the indicated reporter plasmids. CAT activity was calculated and normalized as described in Materials and Methods. Relative CAT activity, indicating the fold induction of normalized CAT activity for each reporter plasmid above that attained with the reference plasmid pe− CAT, is graphed. Data represent the averages of at least three separate experiments ± standard deviation (SD). (C) The additional reporter constructs p(EFII3′/5′)2+CAT, pBLCAT4, p(EFII5′)4−CAT4, and p(EFII3′)2−CAT4 were also tested in transient transfections as described for panel B. p(EFII3′/5′)2+CAT contains two copies of the EFII 3′/5′ mutant oligonucleotide (see Fig. 3A for sequence) in the sense orientation inserted at the *Acc*I site of pe− CAT. p(EFII5′)4−CAT4 contains four copies of the 5′ repeat (EFII sequences from −226 to −213; see Materials and Methods), and p(EFII3′)2-CAT4 contains two copies of the 3′ repeat (EFII sequences −205 to −194; see Materials and Methods); both inserts are in the antisense orientation, and both were placed immediately upstream of the thymidine kinase promoter and CAT gene in the reporter plasmid pBLCAT4. Relative CAT activity, calculated as described for panel B, with pe− CAT as a reference for p(EFII3′/5′)2+CAT and pBLCAT4 as a reference for p(EFII5′)4−CAT4 and p(EFII3′)2−CAT4, is graphed. Data represent the averages of at least three separate experiments ± SD.

mutant oligonucleotides compared with their affinities for the wild-type EFII DNA sequence are presented in Fig. 3C. We observed that the EFIIa, -b, and -c complexes could bind to the EFII 3′ mutant, which retains the upstream repeat, with affinity equal to that seen with the EFII DNA. In contrast, all three EFII DNA-binding complexes demonstrated a marked reduction in relative binding affinity to the EFII 5′ mutant; EFIIa and EFIIb had approximately 9-fold lower relative binding affinity, and EFIIc binding affinity was reduced to approximately 50-fold below that seen with the EFII DNA. When the EFII 3′/5′ mutant was used as the competitor oligonucleotide, all three EFII factors demonstrated greater than 100-fold-lower relative DNA-binding affinity. Thus, at least one intact repeat must be present for EFIIa, -b, and -c complex formation, and all three factors have a higher affinity for the 5′ repeat containing the additional AT nucleotides.

We also analyzed the ability of the three EFII complexes to directly bind each of the EFII mutant oligonucleotides (Fig. 3D). We observed that all three complexes formed equally well with the radiolabeled EFII DNA (Fig. 3D, lane 1) or the radiolabeled EFII 3′ mutant (Fig. 3D, lane 2). In contrast, EMSA with the radiolabeled EFII 5′ mutant (Fig. 3D, lane 3) resulted in two less-prominent DNA-binding activities that displayed proportionally slower migration in the gel than EFIIa and EFIIb, and no EFIIc-like complex was present. There is evidence that the same DNA-binding protein can have different mobilities in a gel matrix when it occupies different positions on oligonucleotides of the same length (35, 55). Therefore, it is possible that the two DNA-

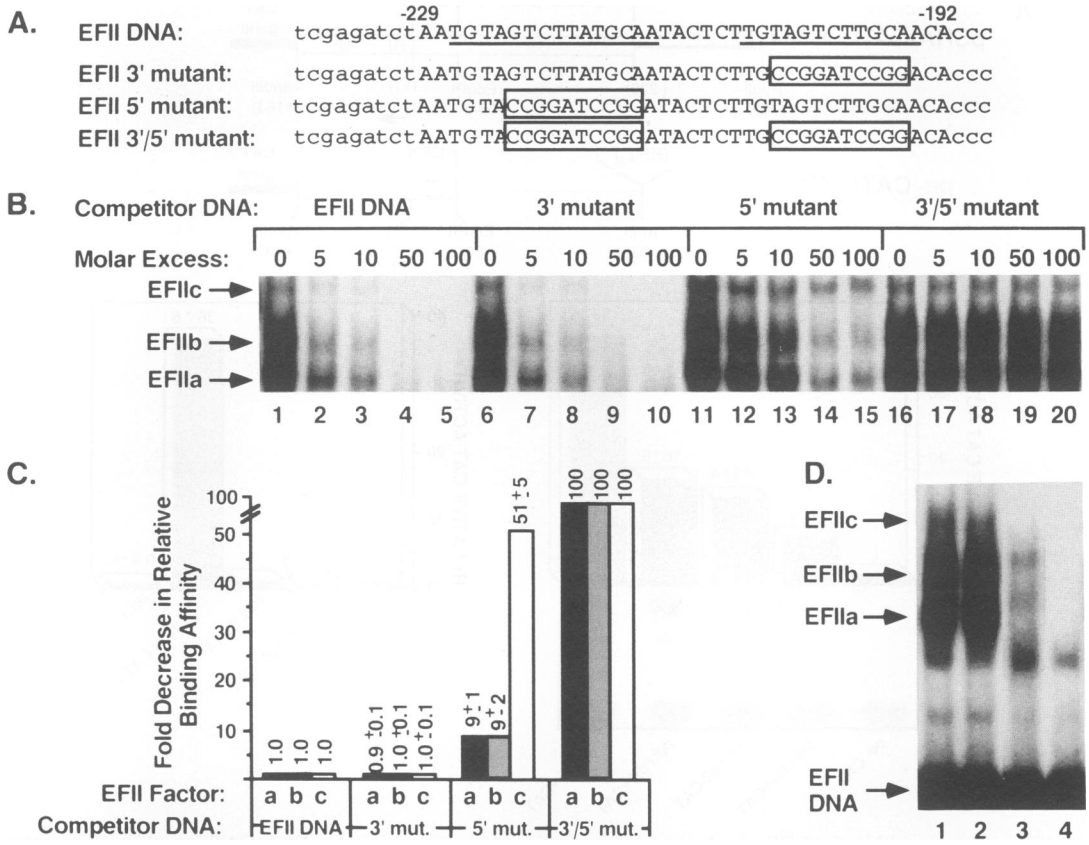

FIG. 3. EFIIa, EFIIb, and EFIIc nuclear factors recognize the upstream repeat in the EFII *cis* element. (A) Coding-strand nucleotide sequences of double-stranded 50-bp oligonucleotides containing either the wild-type EFII *cis* element (EFII DNA) or the EFII *cis* element in which either the 3′ repeat (EFII 3′ mutant), the 5′ repeat (EFII 5′ mutant), or both (EFII 3′/5′ mutant) have been disrupted by 10-bp *Bam*HI linker sequences, indicated by open boxes. The 38-bp EFII *cis* element (representing RSV LTR sequences from −229 to −192) is indicated by capital letters. The near-direct repeat sequences present in the wild-type EFII *cis* element are underlined. (B) EMSAs were performed as described in Materials and Methods. Briefly, 2 μg of heated 0.5 M NaCl nuclear extract from 14-day-old CE were incubated with 0.5 ng of ^{32}P-labeled EFII DNA in the presence of 1.3 μg of poly(dI):poly(dC) and in the absence or presence of a 5- to 100-fold molar excess of nonradiolabeled competitor DNAs. The positions of the retarded EFIIa, EFIIb, and EFIIc protein-DNA complexes are shown. (C) The relative binding affinities of EFIIa, -b, and -c for each competitor oligonucleotide were calculated as described in Materials and Methods. The fold decrease in binding affinity relative to wild-type EFII DNA is graphed, and data represent the averages of at least three separate experiments ± SD. (D) EMSAs were performed with 2 μg of heated 0.5 M NaCl nuclear extract from 14-day-old CE in the presence of 1.3 μg of poly(dI):poly(dC) and 0.5 ng of the following ^{32}P-labeled EFII wild-type and mutant oligonucleotides: EFII DNA (lane 1), EFII 3′ mutant (lane 2), EFII 5′ mutant (lane 3), and EFII 3′/5′ mutant (lane 4). The unbound ^{32}P-labeled EFII DNAs are indicated along with the retarded EFIIa, EFIIb, and EFIIc protein-DNA complexes.

binding activities in Fig. 3D, lane 3, represent EFIIa and EFIIb, since both can bind with lower affinity to the EFII 5′ mutant (Fig. 3B and C). On the other hand, the two complexes in Fig. 3D, lane 3, could represent unrelated proteins, which are normally masked by the presence of the EFIIa, -b, and -c complexes. EMSA with the radiolabeled EFII 3′/5′ mutant (Fig. 3D, lane 4) did not result in any specific EFII DNA-binding activity. The results of both the competition and direct binding studies in Fig. 3 are highly consistent with our analysis of the transactivation properties of the EFII *cis* element in vivo. The upstream repeat in the EFII *cis* element mediates both high-affinity recognition by the EFIIa, -b, and -c factors in vitro and strong transcriptional activation in vivo.

OP/Cu footprinting of EFIIa, EFIIb, and EFIIc. In order to further define the nucleotides within the EFII *cis* element which EFIIa, -b, and -c recognize for sequence-specific DNA binding, OP/Cu nuclease footprinting of each DNA-protein complex was performed. The EFIIa, -b, and -c DNA-protein complexes were resolved by EMSA and treated in situ with the chemical endonuclease OP/Cu. Cleaved DNA was then isolated from each complex and analyzed, along with the cleaved unbound DNA, on a high-resolution sequencing gel (Fig. 4A). Interestingly, all three complexes exhibited essentially identical footprints. Strong protection was observed over 15 nucleotides from −210 to −224 on the noncoding strand, with weaker protection of 9 nucleotides from −210 to −218 on the coding strand. The footprints on both strands were observed to coincide closely but not exactly with the 5′ repeat, as illustrated in Fig. 4B. No protection was observed with any of the complexes over the 3′ repeat on either strand. These findings are similar to observations by Goodwin, who found by footprinting analysis that erythroid nuclear proteins protected sequences from −224 to −215 in the RSV LTR enhancer (22).

Since the EFIIa, -b, and -c complexes footprint the same

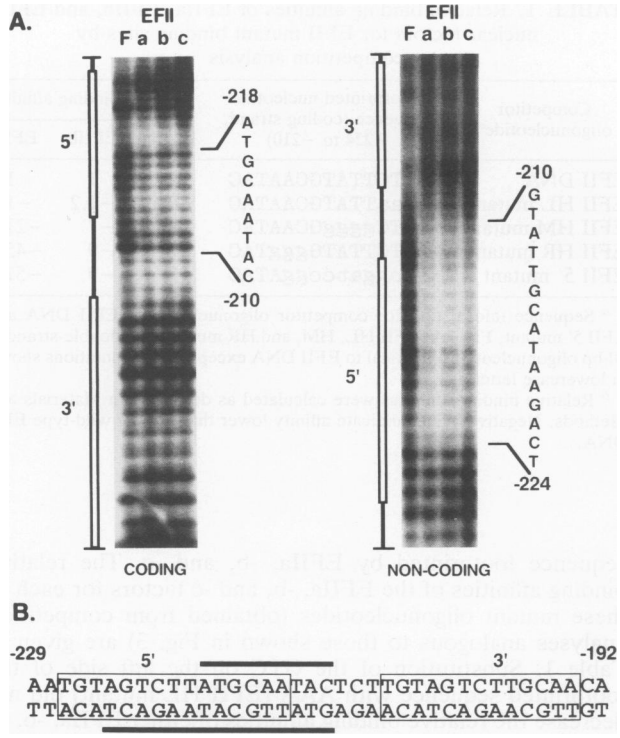

FIG. 4. EFIIa, EFIIb, and EFIIc protect the same nucleotides within the EFII *cis* element. (A) EMSAs were performed with 2 ng of a double-stranded end-labeled 76-bp restriction fragment probe, containing the 38-bp EFII *cis* element, and 12 μg of CEF 0.5 M NaCl nuclear extract. Following EMSA, the gel was treated with OP/Cu endonuclease. The EFIIa, -b, and -c protein-DNA complexes and the free probe were excised from the gel. The cleaved, end-labeled DNA was recovered by electroelution and analyzed on a high-resolution sequencing gel. The nuclease cleavage patterns of the free probe (lane F) and the probe bound by either EFIIa (lane a), EFIIb (lane b), or EFIIc (lane c), for both the coding and noncoding strands of the EFII *cis* element, are shown. A schematic of the EFII *cis* element with the 5′ and 3′ near-direct repeats indicated by open bars is shown beside the lanes. Footprinted nucleotides are indicated for each strand. (B) Schematic representation of the EFII *cis* element, indicating the EFIIa, EFIIb, and EFIIc footprints on the coding (overlined) and noncoding (underlined) strands. 5′ and 3′ near-direct repeat sequences are boxed.

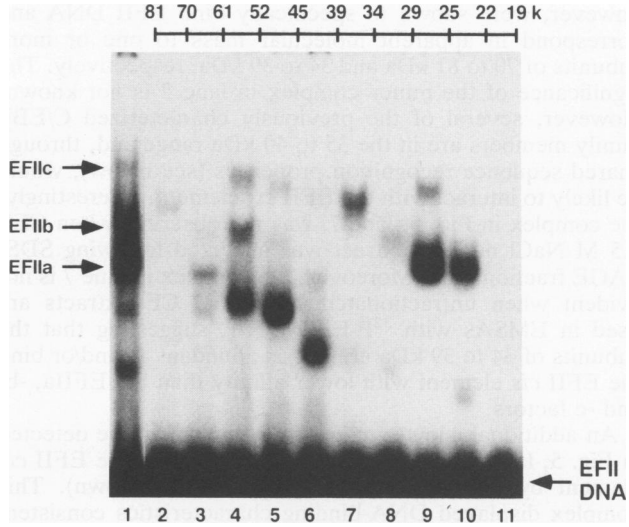

FIG. 5. Apparent-molecular-mass determination. Fourteen-day-old CE 0.5 M NaCl nuclear extract (120 μg) was subjected to SDS-PAGE and transferred to an Immobilon-P membrane, which was cut into 17 3-mm-wide slices. The proteins in each slice were eluted separately and tested in EMSAs. EMSAs were performed with 0.5 ng of ^{32}P-labeled EFII DNA and either 2 μg of 14-day-old CE 0.5 M NaCl nuclear extract in the presence of 1.3 μg of poly(dI):poly(dC) (lane 1) or 10 μl of SDS-PAGE-fractionated 14-day-old CE extract, ranging in size from 81 to 19 kDa, in the presence of 100 ng of poly(dI):poly(dC) (lanes 2 to 12). EFIIa, EFIIb, and EFIIc retarded complexes in the unfractionated extract (lane 1) are indicated. Sizes are shown above the lanes (in kilodaltons).

sequences in the EFII *cis* element, it is possible that they represent the same DNA-binding protein and that the differences in electrophoretic mobility are due to the fact that EFIIa and EFIIb are proteolytic products of EFIIc. However, no increase is observed in the relative amounts of EFIIa and EFIIb versus EFIIc with increasing age of the extract. Alternatively, the same DNA-binding protein could give rise to the slower-migrating EFIIb and EFIIc complexes if it was differentially posttranslationally modified or associated with other transcription factors. On the other hand, the EFIIa, -b, and -c complexes could contain distinct DNA-binding proteins which have very similar, or perhaps identical, sequence recognition properties. This latter possibility is consistent with the growing number of transcription factors that have been found to exist in families which demonstrate nearly identical sequence recognition properties (30).

EFIIa, EFIIb, and EFIIc are composed of multiple DNA-binding proteins. To address whether the same or different DNA-binding proteins are present in the EFIIa, -b, and -c complexes, 14-day-old CE nuclear extract was subjected to SDS-PAGE and blotted to Immobilon paper. The paper was cut into small slices, and the proteins in each slice were eluted and assayed for EFII DNA-binding activity. As shown in Fig. 5, three protein-DNA complexes (lanes 4 and 9) migrated to the positions corresponding to EFIIa, -b, and -c observed in unfractionated nuclear extracts (lane 1). The dominant shifted band in lane 9 (also found in lane 10) is equivalent in mobility to EFIIa and represents a protein complex composed of one or more subunits of 22 to 29 kDa. The two slower-mobility complexes in lane 4 migrate to positions corresponding to EFIIb and EFIIc and represent complexes made up of subunits of 52 to 61 kDa. These three renatured EFIIa, EFIIb, and EFIIc complexes specifically recognized the EFII *cis* element in competition assays with the wild-type and mutant EFII DNAs (data not shown).

In addition to the renatured EFIIa, -b, and -c complexes, a number of other protein-DNA complexes were observed, although most did not correspond in mobility to EFIIa, EFIIb, or EFIIc. The faster-migrating complexes in Fig. 5, lanes 4 to 6, were found to be nonspecific in competition assays (data not shown). Although the complex in lane 3 migrates to a position similar to that of EFIIa, it did not demonstrate high affinity for the EFII *cis* element in competition assays (data not shown). For this analysis, a minimal amount of nonspecific competitor DNA was used in the EMSA in order to maximize recovery. The addition of more nonspecific poly(dI):poly(dC) to the EMSA eliminated the nonspecific complexes mentioned above but had no effect on the EFIIa, -b, and -c complexes in Fig. 5, lanes 4 and 9 (data not shown). The complexes in Fig. 5, lanes 2 and 7,

however, were shown to specifically bind EFII DNA and correspond in apparent molecular mass to one or more subunits of 70 to 81 kDa and 34 to 39 kDa, respectively. The significance of the minor complex in lane 2 is not known. However, several of the previously characterized C/EBP family members are in the 35 to 40 kDa range and, through shared sequence recognition properties (see below), would be likely to interact with the EFII *cis* element. Interestingly, the complex in Fig. 5, lane 7, was not observed when CEF 0.5 M NaCl nuclear extract was analyzed following SDS-PAGE fractionation. Moreover, the complex in lane 7 is not evident when unfractionated 14-day-old CE extracts are used in EMSAs with ^{32}P-EFII DNA, suggesting that the subunits of 34 to 39 kDa are not as abundant as and/or bind the EFII *cis* element with lower affinity than the EFIIa, -b, and -c factors.

An additional, slower-migrating complex can be detected in Fig. 5, lane 9, which specifically recognizes the EFII *cis* element by competition analysis (data not shown). This complex displayed DNA-binding characteristics consistent with the addition of a second protein complex to a repeated binding site. For example, we observed that the additional band in lane 9 was only present at high concentrations of the 25- to 29-kDa fraction, and in competition experiments with unlabeled EFII DNA, it was seen to disappear prior to any diminution in the renatured EFIIa band. Moreover, the additional complex in Fig. 5, lane 9, was absent when the 25- to 29-kDa fraction was subjected to EMSA with the radiolabeled EFII 3′ mutant, which lacks the lower-affinity downstream repeat (data not shown). Thus, we interpret the additional shifted complex seen in lane 9 to be a second EFIIa complex, binding to the 3′ repeat under conditions of sufficiently high factor concentration. This interpretation is further substantiated by the footprinting experiments presented in Fig. 7 (see below).

Since these molecular mass determinations were performed with the factors present in a crude nuclear extract, the potential for proteolysis of the factors is a concern. However, we performed a similar analysis of the EFII factors after lysing intact CEF in SDS-containing buffer and immediately boiling the total-cell lysate. This approach should minimize proteolysis as much as possible. Nonetheless, results identical to those presented in Fig. 5 were obtained (data not shown). It is entirely possible that proteins in addition to the polypeptides detected here after SDS electrophoretic separation are components of the EFII DNA-binding activities in 14-day-old CE nuclear extracts. A polypeptide whose DNA-binding activity cannot be renatured after SDS exposure will not be detected, nor will a complex which requires the association of heterologous subunits for DNA binding, unless all polypeptides in the native complex have similar molecular masses and are renatured in the same fraction. Despite these limitations, this analysis suggests that, at minimum, two polypeptides of 22 to 29 kDa and 52 to 61 kDa are present in the EFIIa and in the EFIIb and EFIIc complexes, respectively.

EFIIa, EFIIb, and EFIIc recognize the same nucleotide sequence. Although EFIIa, -b, and -c protected the same nucleotides within the EFII *cis* element from OP/Cu cleavage (Fig. 4), these complexes appear to be composed of at least two polypeptides of very different molecular masses which could be recognizing different sequence motifs within the 15-nucleotide footprint. To attempt to discern any differences in DNA sequence recognition among these polypeptides, we obtained a series of EFII oligonucleotides harboring small substitution mutations across the 15-nucleotide sequence footprinted by EFIIa, -b, and -c. The relative binding affinities of the EFIIa, -b, and -c factors for each of these mutant oligonucleotides (obtained from competition analyses analogous to those shown in Fig. 3) are given in Table 1. Substitution of the GTC on the left side of the footprinted sequence with AAA (EFII HL mutant) did not decrease the relative binding affinities of either EFIIa, -b, or -c. Thus, we can narrow down the sequence important for the DNA-binding activities of EFIIa, -b, and -c to the 11 nucleotides TTATGCAATAC (coding strand). Substitution of either TTAT with GGGG (EFII HM mutant) or CAA with GGG (EFII HR mutant) reduced the binding affinities of both EFIIa and EFIIb 6- to 9-fold and the binding affinity of EFIIc 21- and 45-fold, respectively, relative to wild-type EFII DNA. The decrease in the relative binding affinities of all three complexes for the EFII HM and HR mutants is very similar to that observed with the EFII 5′ mutant (Table 1 and Fig. 3C), in which most of the footprinted sequence (10 of 15 nucleotides) has been disrupted. The only exception is that EFIIc binds to the EFII HM mutant with a slightly higher relative affinity than to either the EFII 5′ mutant or the EFII HR mutant, possibly because of a partial ability to recognize the remaining GCAATAC portion of the 11-nucleotide binding site. The low-affinity binding of EFIIa and EFIIb to oligonucleotides harboring mutations within the 11-nucleotide binding site is likely due to the ability of these factors to recognize similar sequences in the 3′ repeat.

TABLE 1. Relative binding affinities of EFIIa, EFIIb, and EFIIc nuclear factors for EFII mutant binding sites by competition analysis

Competitor oligonucleotide[a]	Footprinted nucleotide sequence (coding strand, −224 to −210)	Relative binding affinity[b]		
		EFIIa	EFIIb	EFIIc
EFII DNA	AGTCTTATGCAATAC	1	1	1
EFII HL mutant	AaaaTTATGCAATAC	−1.2	−1.2	−1.3
EFII HM mutant	AGTCggggGCAATAC	−6	−7	−21
EFII HR mutant	AGTCTTATGgggTAC	−8	−9	−45
EFII 5′ mutant	AccggatccggATAC	−9	−9	−51

[a] Sequence information for competitor oligonucleotides: EFII DNA and EFII 5′ mutant, Fig. 3A; EFII HL, HM, and HR mutants are double-stranded 50-bp oligonucleotides identical to EFII DNA except for the mutations shown in lowercase letters.

[b] Relative binding affinities were calculated as described in Materials and Methods. Negative values indicate affinity lower than that of wild-type EFII DNA.

At this level of resolution, the EFIIa and EFIIb factors appear to have identical sequence-specific DNA-binding properties, while EFIIc has a more restricted sequence requirement for the 5′ repeat. It is interesting that the EFII sequences important for high-affinity binding by all three EFII DNA-binding factors contain the sequence motif TTNNGCAAT. This motif is found in a number of high-affinity binding sites for the previously described transcription factor C/EBPα (27, 61, 65, 70) as well as for a growing number of C/EBP-related family members (1, 7, 10, 15, 34, 56, 58). In fact, C/EBPα has been shown to bind the EFII *cis* element in the RSV LTR in vitro; Ryden and Beemon demonstrated by DNase I footprinting experiments that a heat-stable factor present in chicken liver nuclear extracts and purified C/EBPα both bound the EFII region of the RSV LTR (61). Interestingly, unlike the EFII-binding proteins described here, footprints appeared concomitantly over both repeats in the EFII *cis* element with increasing concentrations of either

TABLE 2. Relative binding affinities of EFIIa, EFIIb, and EFIIc nuclear factors for various binding sites by competition analysis

Competitor oligonucleotide[a]	Nucleotide sequence	Relative binding affinity[b]		
		EFIIa	EFIIb	EFIIc
EFII DNA	TTATGCAATAC	1	1	1
IgH-E	TTGAGCAATGT	−1.6	−1.5	−4
IBF DNA	GGCTGCAATTG	−3	−3	−4
DE I	TTTTGTAATGG	−11	−9	−14
Core	TGTGGAAAGGA	>100	>100	>100

[a] Sequence information for competitor oligonucleotides: EFII DNA, Fig. 3A; IgH-E, IBF DNA, DE I, and Core, Materials and Methods.
[b] See Table 1, footnote *b*.

the liver nuclear factor or purified C/EBPα (61). The murine C/EBP-related factor immunoglobulin enhancer-binding protein (Ig/EBP-1) has also been shown by Roman et al. to bind the EFII *cis* element in the RSV LTR in vitro as a β-galactosidase fusion protein encoded in part by the murine Ig/EBP-1 cDNA, and the role of Ig/EBP-1 in regulating RSV transcription in murine cells was discussed by these investigators (58). In addition, a putative C/EBP-related factor, the internal binding factor (IBF), which was characterized and purified from BHK cells, was reported by Karnitz et al. to bind the EFII *cis* element in competition analysis (33). Given that the EFII factors recognize a sequence motif which fits the consensus binding site for C/EBP-related transcription factors, we wished to compare the sequence recognition requirements of EFIIa, -b, and -c in avian fibroblasts with those of the members of the C/EBP family which have been reported to bind the EFII *cis* element in other species and cell types.

Comparison of the DNA-binding properties of EFIIa, EFIIb, and EFIIc with those of C/EBPα, IBF, and Ig/EBP-1. Initially, a series of competition experiments were performed to ascertain whether the EFIIa, -b, and -c factors could also bind to high-affinity recognition sites used to characterize C/EBPα, IBF, and Ig/EBP-1. As shown in Table 2, the EFII factors showed negligible to low recognition of two high-affinity C/EBPα binding sites: the SV40 core enhancer sequence found in the collagen III promoter (9, 29, 61) and the DE I site in the rat albumin promotor (18, 50, 73). In contrast, oligonucleotide DNAs containing binding sites for either Ig/EBP-1 or IBF competed relatively well for EFIIa, -b, and -c binding activities. EFIIa and EFIIb bound IgH-E, an oligonucleotide which encompasses the E site from the murine immunoglobulin heavy-chain enhancer, with only slightly lower affinity than they did EFII DNA. The IgH-E nucleotide sequence was used to identify the Ig/EBP-1 expression clone from murine L-cell and B-cell cDNA libraries (58). Sequence analysis of Ig/EBP-1 revealed that it has high homology to C/EBPα throughout the DNA-binding domain and leucine zipper region (58). EFIIa and EFIIb also bound to IBF DNA, an oligonucleotide which contains sequences from an internal enhancer located within the *gag* gene of RSV (2, 33), with threefold lower affinity than to EFII DNA. The IBF nucleotide sequence was used to characterize and purify the IBF DNA-binding activity from BHK cell nuclear extracts (33, 34). IBF has been reported to be related to C/EBPα from its overlapping DNA sequence recognition properties and heat stability (8, 34, 39). EFIIc recognized both IgH-E and IBF DNA with fourfold lower affinity than the EFII sequence. These differences in the affinities of EFIIa, -b, and -c for the EFII, IBF, and IgH-E DNAs, although relatively small, were reproducibly observed.

Comparison of the cell type distribution of EFIIa, -b, and -c with IBF, Ig/EBP-1, and C/EBPα. Although the binding sites for IBF and Ig/EBP-1 were recognized by the three EFII factors present in CE nuclear extracts with moderate to high affinity, similar DNA-binding specificity could be exhibited by different proteins that may be selectively present in different cell types. Thus, we next characterized the EFII DNA-binding activity in cell types in which IBF and Ig/EBP-1 expression has been documented. As mentioned above, IBF was purified from BHK cell nuclear extracts, and Ig/EBP-1 expression, monitored by Northern (RNA blot) analysis, is observed in all murine cell types tested but is particularly strong in murine B lymphocytes. We prepared nuclear extracts from BHK cells and from an avian B-cell lymphoma line, Bk3A, to test for EFII DNA-binding activity. As shown in Fig. 6A, lane 1, and Fig. 6C, lane 5, typical EFIIa, -b, and -c complexes are observed when radiolabeled EFII DNA is incubated with a BHK cell nuclear extract. To confirm that the BHK extract contained active IBF, we incubated the extract with radiolabeled IBF DNA. A striking difference between the radiolabeled EFII and IBF DNA-binding patterns was observed (Fig. 6A, lanes 1 and 2). The IBF DNA-binding pattern more closely matched that reported by Karnitz et al. (34). We performed competition analyses with unlabeled EFII DNA in an EMSA with radiolabeled IBF DNA and BHK extract. As shown in Fig. 6A, lanes 2 to 7, by our analysis, IBF has negligible affinity for the EFII DNA sequence; an 80-fold molar excess of EFII DNA was required to compete for 50% of the IBF binding activity. Thus, we conclude that the IBF factor characterized in BHK cells is unlikely to be an integral part of either the EFIIa, -b, or -c DNA-binding complex. Moreover, this IBF factor does not appear to be very abundant in nuclear extracts from 14-day-old CE since we do not observe its formation in EMSAs with radiolabeled IBF DNA and 14-day-old CE extract (data not shown).

In contrast, radiolabeled IgH-E and EFII DNAs yielded nearly identical binding patterns in the Bk3A lymphoma extract except that no EFIIc complex was observed to form with radiolabeled IgH-E (Fig. 6B). Also, radiolabeled IgH-E bound EFIIa and EFIIb in EMSAs with 14-day-old CE extract (data not shown). Thus, it is possible that either of the avian EFIIa and EFIIb factors is highly related to murine Ig/EBP-1. A great excess of the EFIIa factor, far exceeding the level of EFIIb (or EFIIc), was detected with either radiolabeled EFII or IgH-E DNA in the Bk3A extract (Fig. 6B). Although the proportions of EFIIa, -b, and -c are different from those observed in CEF, 14-day-old CE, and BHK extracts, all three complexes in the Bk3A extract are heat stable and demonstrated characteristic DNA-binding specificities in competition analysis (Fig. 6C, lane 9, and data not shown). We also observed that an increase in the amount of Bk3A nuclear extract resulted in the disappearance of the EFIIb and EFIIc shifts and the formation of a new complex (Fig. 6C, lane 10), termed ABF (additional binding factor). The ABF complex appeared only with the Bk3A extract; titration of all other extracts led to a uniform increase in all three factors (data not shown). As will be discussed later (Fig. 7), the ABF complex appears to represent the addition of a second EFIIa factor to the downstream near-direct repeat in the EFII *cis* element, caused by the high concentration of EFIIa in the B-cell extract.

In addition to BHK and Bk3A nuclear extracts, we tested a number of fibroblast and epithelial cell lines for the EFIIa,

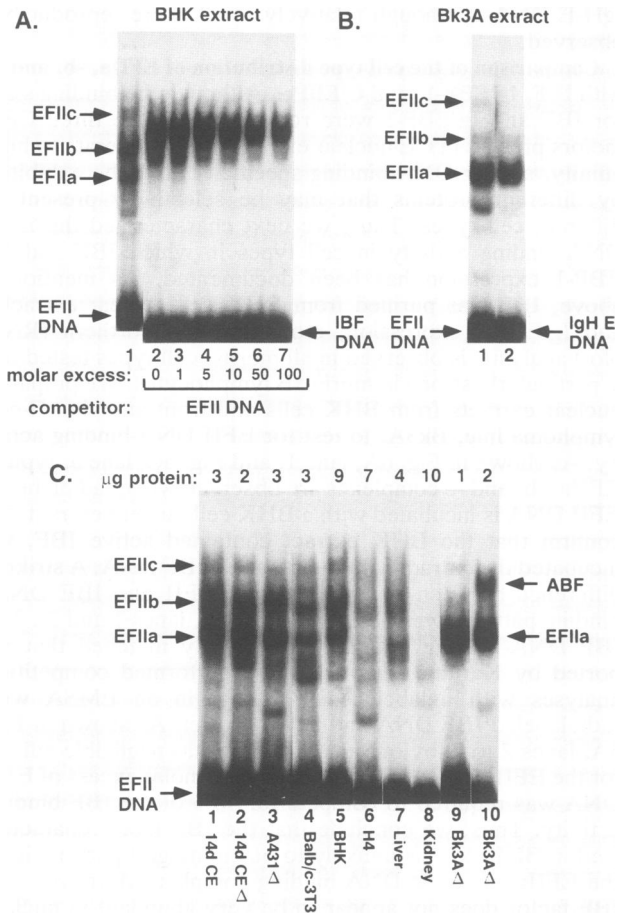

FIG. 6. Comparing the cell type distribution of EFIIa, -b, and -c with that of IBF, Ig/EBP-1, and C/EBP. (A) EMSAs were performed with 9 μg of BHK 0.5 M NaCl nuclear extract and 0.5 ng of ^{32}P-labeled EFII DNA (lane 1) or IBF DNA (lanes 2 to 7) in the presence of 1.3 μg of poly(dI):poly(dC) and in the absence (lanes 1 and 2) or presence (lanes 3 to 7) of a 1- to 100-fold molar excess of nonradiolabeled competitor EFII DNA. (B) EMSAs were performed with 1 μg of heated Bk3A 0.5 M NaCl nuclear extract and 0.5 ng of ^{32}P-labeled EFII DNA (lane 1) or IgH-E DNA (lane 2) in the presence of 1.3 μg of poly(dI):poly(dC). (C) EMSAs were performed with samples of 0.5 M NaCl nuclear extracts from the indicated cell types and 0.5 ng of ^{32}P-labeled EFII DNA in the presence of 1.3 μg of poly(dI):poly(dC). Lanes: 1 and 2, 14-day-old (14d) CE; 3, A431 epidermal carcinoma cells; 4, BALB/c-3T3 fibroblasts; 5, BHK cells; 6, rat H4 hepatoma cells; 7, adult rat liver; 8, adult rat kidney; 9 and 10, avian B-cell lymphoma line Bk3A. The amount of protein included in each EMSA is listed above the corresponding lane. Δ, heated extract. EFIIa, EFIIb, EFIIc, and ABF retarded protein-DNA complexes are indicated.

-b, and -c DNA-binding pattern (Fig. 6C). We observed a pattern of binding activity similar to that of EFIIa, -b, and -c in nuclear extracts from human epidermal carcinoma A431 cells (Fig. 6C, lane 3) and murine fibroblasts (BALB/c-3T3; Fig. 6C, lane 4) as well as several other fibroblast lines (Rat-1, NIH 3T3, and Cos cells; data not shown). The EFII DNA-binding activities in these extracts, as in our standard CE extract, were heat stable (Fig. 6C, compare lanes 1 and 2, also lane 3; data not shown). We also obtained nuclear extracts from a differentiated rat liver heptoma cell line, H4, and adult rat liver and kidney tissue. Reproducibly less (hepatoma and liver; Fig. 6C, lanes 6 and 7, respectively) or negligible (kidney; Fig. 6C, lane 8) EFII DNA-binding activity was detected in these extracts. Furthermore, retarded complexes corresponding in mobility to EFIIa were not observed. Although we cannot rule out the possibility that poor extraction or selective proteolysis of the EFII factors occurred consistently in the H4, liver, and kidney extracts, all extracts tested were prepared by the same method, and analysis of proteins that bind a hypersensitive site 4,800 bp upstream of the phosphoenolpyruvate carboxykinase gene demonstrated that the H4, liver, and kidney extracts we analyzed were not significantly proteolyzed (12). Although the tissues and cell types tested were from various species, it appears from this preliminary survey that the EFIIa, -b, and -c binding activities are not uniformly present in all cell types.

EFIIa binds to the 3′ repeat sequence in the EFII *cis* element at high concentrations. We were intrigued by the unusually large amount of the EFIIa factor present in nuclear extracts from Bk3A cells (Fig. 6B and C), as well as the presence of the ABF complex when higher concentrations of the extract were used in the EMSA (Fig. 6C, lane 9). We analyzed Bk3A nuclear extract by SDS gel electrophoresis and Immobilon transfer and obtained results identical to those shown for the 14-day-old CE extract in Fig. 5 except that in the 25- to 29-kDa fraction, greater amounts of both the EFIIa-like factor and the slower-mobility complex were seen. As shown in Fig. 7A, when we compared the mobility shift pattern seen in the Bk3A 25- to 29-kDa fraction (lane 2) with the EFII DNA-binding pattern of unfractionated Bk3A extract, in which the ABF complex is observed (lane 1), we found that the ABF factor was identical in mobility to the slower complex in the 25- to 29-kDa fraction. This slower complex was previously interpreted to represent an additional EFIIa complex binding to the 3′ repeat in the EFII *cis* element. We therefore suggest that the ABF complex also represents the binding of a second EFIIa factor to the 3′ repeat in the radiolabeled EFII DNA, because of the abundance of EFIIa in the Bk3A cells. In support of this, we also noted that the ABF complex did not form in an EMSA with the radiolabeled EFII 3′ mutant, which contains only one intact repeat, and that ABF demonstrated the binding kinetics expected for the addition of a second factor to a repeated binding site by competition analysis (data not shown) and extract titration (Fig. 6C, lanes 9 and 10).

In order to directly analyze the possible presence of an additional protein complex bound to the 3′ repeat of the EFII *cis* element in the ABF complex, we performed OP/Cu nuclease footprinting of both the EFIIa protein-DNA complex and the ABF protein-DNA complex present in the Bk3A nuclear extract. Comparison of the EFIIa footprints with the ABF footprints (Fig. 7B) revealed that proteins in the ABF complex protected the same sequences in the 5′ repeat on the coding and noncoding strands as the proteins in the EFIIa complex, although protection was enhanced in these regions. However, only proteins in the ABF complex protected similar nucleotides in the 3′ repeat. The additional ABF footprints in the 3′ repeat protected 12 nucleotides from −203 to −192 on the noncoding strand and 10 nucleotides from −201 to −192 on the coding strand. A diagram of the ABF footprint along with the EFIIa footprint is presented in Fig. 7C.

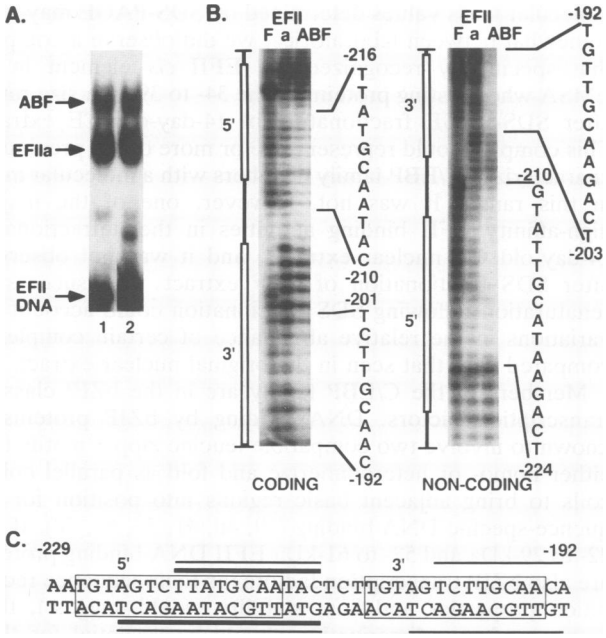

FIG. 7. Binding of EFIIa to the 3′ repeat in the EFII *cis* element. (A) EMSAs were performed with 0.5 ng of ^{32}P-labeled EFII DNA and either 2 μg of heated Bk3A 0.5 M NaCl nuclear extract in the presence of 1.3 μg of poly(dI):poly(dC) (lane 1) or 10 μl of SDS-PAGE-fractionated Bk3A 0.5 M NaCl nuclear extract from the 25- to 29-kDa fraction in the presence of 100 ng of poly(dI):poly(dC) (lane 2). EFIIa and ABF retarded protein-DNA complexes in the unfractionated Bk3A extract are indicated. (B) EMSAs were performed with 2 ng of an end-labeled 76-bp restriction fragment containing the 38-bp EFII *cis* element and 4 μg of heated Bk3A 0.5 M NaCl nuclear extract. The EFIIa and ABF protein-DNA complexes and the free probe were treated in situ with the OP/Cu endonuclease as described in the legend to Fig. 4A. The nuclease cleavage patterns of the free probe (lane F) and the probe bound by EFIIa (lane a) or ABF (lane ABF) for both the coding and noncoding strands are shown. A schematic of the EFII *cis* element, with the 5′ and 3′ near-direct repeats indicated by open bars, is shown beside the lanes. Footprinted nucleotides are indicated for each strand. (C) Schematic representation of the EFII *cis* element, indicating nucleotides protected by EFIIa on the coding (overlined in bold) and noncoding (underlined in bold) strands, and nucleotides protected by ABF on the coding (overlined) and noncoding (underlined) strands. The 5′ and 3′ near-direct repeat sequences are boxed.

DISCUSSION

We have continued our characterization of nuclear proteins which specifically bind the EFII region of the RSV LTR enhancer (64). In this report, we initially demonstrated that multiple copies of the EFII *cis* element strongly activate transcription of a reporter gene in vivo. Furthermore, the transcription-activating potential of this segment of the LTR enhancer lies primarily within the upstream of two near-direct repeat sequences present in the EFII *cis* element. We found that the 5′ repeat sequence alone, when reiterated, could strongly activate transcription, while two copies of the 3′ repeat showed no transcriptional activity. If the 5′ and 3′ near-direct repeats were equivalent in their transactivation potential, we would expect two copies of the 3′ repeat to exhibit at least modest transactivation. Multimerization of either the 5′ repeat sequence or the entire EFII *cis* element did not, however, lead to the level of transcriptional activation exhibited by the complete LTR enhancer (which activates transcription over 600-fold relative to pe− CAT; data not shown). It is likely that the extremely high level of transcriptional activity mediated by the entire RSV LTR enhancer results from the synergistic interaction of multiple *trans*-acting factors binding to a number of different *cis* elements within the enhancer.

Our in vitro analysis identified three heat-stable protein complexes, EFIIa, EFIIb, and EFIIc, in nuclear extracts from CEF and 14-day-old CE which specifically bound the EFII *cis* element. Mutational analysis as well as footprinting experiments demonstrated that all three EFII DNA-binding factors preferentially recognized the 5′ repeat. The DNA-binding activity of all three EFII factors was completely abolished by a double mutation in both repeats (3′/5′ mutant). This EFII 3′/5′ mutant oligonucleotide was also incapable of activating transcription of a reporter gene in vivo. The correspondence of both transcription-activating potential and factor binding with the 5′ repeat suggests that transcriptional activation by the EFII *cis* element could be mediated by one or more of the proteins present in the EFIIa, -b, and/or -c complexes.

EFIIa, EFIIb, and EFIIc are composed of multiple DNA-binding proteins which recognize the same consensus sequences as members of the C/EBP-related family of transcription factors. Although EFIIa, -b, and -c footprinted the same nucleotides in the 5′ repeat, molecular mass analysis revealed that multiple protein factors are likely involved. The EFIIa binding activity represents a protein complex composed of one or more subunits of approximately 22 to 29 kDa, and the EFIIb and EFIIc factors appear to be composed of at least one or more polypeptides of 52 to 61 kDa. In performing more detailed mutagenesis across the 5′ repeat, we defined an 11-nucleotide sequence important for high-affinity binding by these proteins. This 11-nucleotide binding site was essential for DNA binding by the EFIIc factor. However, both EFIIa and EFIIb could compensate for mutations in this region by recognizing similar sequences in the 3′ repeat with lower affinity. The nucleotide sequence recognized by EFIIa, -b, and -c with high affinity conforms to the consensus recognition sequence for the previously characterized transcription factor C/EBPα (27, 29, 39, 61, 65, 70) as well as a growing number of C/EBP-related family members, including NF-IL6 (nuclear factor for interleukin-6 [IL-6] expression) (1), which is identical to LAP, IL-6DBP, AGP/EBP, C/EBPβ, and CRP2 (7, 10, 15, 56, 73); C/EBPδ, which is identical to CRP3 (7, 73); CRP1 (73); Ig/EBP-1 (C/EBPγ) (58), which may be identical to μEBP-E (54, 55); and IBF, a putative C/EBP family member (33, 34).

Although C/EBPα and IBF have been shown to bind to the EFII *cis* element in vitro (33, 61), our analysis indicates that they are unlikely to be components of any of the EFII binding activities characterized here. EFIIa, -b, and -c demonstrated negligible or low affinity for two different high-affinity C/EBPα binding sites, the SV40 core enhancer motif and the DE I site from the albumin promoter. Apparently, C/EBPα is more promiscuous in its sequence recognition requirements than EFIIa, -b, and -c. The cell type distribution of the EFIIa, -b, and -c factors also differs from what would be expected for C/EBPα and most related family members. Expression of C/EBPα as well as of LAP (C/EBPβ, CRP2), C/EBPδ (CRP3), and CRP1 proteins is abundant in differentiated liver, lung, and adipose cells (3, 7, 15, 73). In contrast, the EFIIa, -b, and -c binding pattern, assayed by EMSA, was detected in a number of fibroblast and epithelial cell lines from various species, while much lower levels were observed in rat liver and in a differentiated

rat hepatoma cell line, and negligible activity was found in rat kidney. Interestingly, EFIIa, -b, and -c binding activity was present in BHK cells, in which IBF was characterized, and all three EFII factors could recognize the IBF binding site with moderately high affinity. However, the IBF factor present in BHK cells had negligible affinity for the EFII *cis* element in our analysis, suggesting that IBF is not a component of the EFIIa, -b, or -c complex and that it may have a more restricted sequence recognition requirement than the EFII factors.

Our comparison of the DNA-binding properties of the EFII factors with those of Ig/EBP-1, which was also demonstrated to be able to bind the EFII *cis* element in vitro (58), revealed that both EFIIa and EFIIb in the 14-day-old CE nuclear extract could bind to either the EFII *cis* element or the Ig/EBP-1 binding site present in the murine immunoglobulin heavy-chain enhancer (denoted IgH-E) with nearly identical affinity. In addition, the DNA-binding activities seen in the avian B-cell line Bk3A and the 14-day-old CE extract in an EMSA with radiolabeled IgH-E DNA were quantitatively and qualitatively identical to those of EFIIa and EFIIb bound to EFII DNA in these extracts. Ig/EBP-1 and μEBP-E are expressed in murine fibroblast and epithelial cells (55, 58), with particularly high levels of Ig/EBP-1 observed in murine B cells (58). It has been reported that mutations within the murine IgH enhancer E site reduce enhancer activity to 36% of wild-type levels (67). Thus, μEBP-E and/or Ig/EBP-1 is likely to play an important role in the regulation of murine immunoglobulin heavy-chain gene transcription. It is possible that the high-level expression of EFIIa observed in avian B cells reflects its normal cellular function in binding the immunoglobulin heavy-chain enhancer E site in these cells, although since the chicken immunoglobulin heavy-chain gene has not been cloned to date, it is not known whether a homologous E site is present in the avian enhancer.

From our analysis of the relative binding affinities of EFIIa, EFIIb, and EFIIc for the EFII 5′ and 3′ near-direct repeat sequences and for the IBF and Ig/EBP-1 (IgH-E) binding sites, we derived the following consensus sequence, TTNTGCAATA, that was critical for EFIIc DNA binding and required for high-affinity DNA binding by EFIIa and EFIIb. A slightly looser consensus sequence, T(T/C)NNG CAA(T/C), was also recognized by EFIIa and EFIIb with lower affinity. This looser consensus sequence for EFII factor binding is very similar to the consensus sequence derived by Ryden and Beemon for C/EBPα binding (61). Our comparison of the DNA-binding properties of EFIIa, EFIIb, EFIIc, C/EBPα, IBF, and Ig/EBP-1 suggests that, although transcription factors can be placed into a family by their similar consensus binding sites, they can also be distinguished within a family by varying affinity for a specific *cis* element.

EFIIa, -b, and -c are distinguished from other members of the C/EBP family by their apparent molecular mass. Molecular mass analysis of the EFIIa, -b, and -c nuclear factors revealed that not only do they appear to be composed of at least two different polypeptides, but the apparent molecular masses of the 22- to 29-kDa and 52- to 61-kDa polypeptides distinguish them from most of the previously reported members of the C/EBP family of transcription factors. C/EBPα, NF-IL6 (LAP, IL-6DBP, AGP/EBP, C/EBPβ, CRP2), C/EBPδ (CRP3), CRP1, μEBP-E, and IBF (1, 7, 10, 15, 34, 55, 56, 73) have apparent molecular masses in the range of 32 to 45 kDa. The molecular mass of Ig/EBP-1 is unknown, since a full-length clone has not been reported (58). Although molecular mass values determined by SDS-PAGE may vary somewhat between laboratories, we did observe a complex that specifically recognized the EFII *cis* element in an EMSA when testing proteins in the 34- to 39-kDa size range after SDS-PAGE fractionation of 14-day-old CE extract. This complex could represent one or more of the previously characterized C/EBP family members with a molecular mass in this range. It was not, however, one of the major, high-affinity EFII binding activities in the unfractionated 14-day-old CE nuclear extracts, and it was not observed after SDS fractionation of CEF extract. The success of renaturation following SDS fractionation could account for variations in the relative abundance of certain complexes compared with that seen in the original nuclear extract.

Members of the C/EBP family are in the bZIP class of transcription factors. DNA binding by bZIP proteins is known to involve two compatible leucine zipper motifs that either homo- or heterodimerize and fold as parallel coiled coils to bring adjacent basic regions into position for sequence-specific DNA binding (19, 40, 41, 53, 68, 70). If the 22- to 29-kDa and 52- to 61-kDa EFII DNA-binding proteins are also bZIP transcription factors (as their sequence recognition properties and heat stability would suggest), then homo- or heterodimerization should be essential for their DNA-binding properties. Our ability to recover DNA-binding activity subsequent to SDS size fractionation of nuclear extracts suggests that each of these factors is capable of homodimerizing or forming a heterodimer of similarly sized subunits. However, we cannot rule out the possibility that heterodimerization between different-size subunits may be an important component of EFII DNA-binding activity in unfractionated nuclear extracts or in vivo.

Many transcription factor families have been shown to contain inducible *trans*-activating proteins, allowing even more complex control of gene transcription (21, 28, 30, 49, 74). A number of the C/EBP-related family members are inducible. For example, the C/EBPβ- and C/EBPδ-encoding genes are directly activated by adipogenic hormones (7). NF-IL6 mRNA is induced by stimulation with either lipopolysaccharide, IL-1, or IL-6 (1). Recently, NF-IL6 (referred to as rNFIL-6 by these investigators) was shown to bind the c-*fos*, serum response element (46), and its activity was demonstrated to be induced by cyclic AMP. The cyclic AMP-induced binding of NF-IL6 to the serum response element was accompanied by increased phosphorylation of the factor and *trans*-location to the nucleus (47).

The EFII sequences in the RSV LTR have also been reported to function as inducible *cis*-acting elements. Kapiloff et al. recently found that calcium/calmodulin-independent mutants of the brain-specific type II calcium/calmodulin-dependent protein kinase specifically activated the RSV LTR through the EFII *cis* element in GC pituitary cells (32). In addition, while the manuscript of this article was under review, Wegner et al. demonstrated that C/EBPβ binds to the EFII *cis* element in GC pituitary cells in vitro and is a target for calcium/calmodulin-dependent protein kinase phosphorylation. We have observed that the EFII DNA-binding activities in CEF are rapidly induced by serum stimulation of quiescent cell populations; furthermore, transformation of CEF mediated by v-*src* or v-*myc* results in substantially increased levels of the EFII DNA-binding activities (data not shown). Therefore, further characterization of the EFII DNA-binding factors seems warranted, not only because these factors are likely to expand the diversity of the C/EBP-related family of transcription factors, but also

because their role in modulating transcription in response to various regulatory signals needs to be elucidated.

ACKNOWLEDGMENTS

We thank our colleagues in the Sealy and Chalkley laboratories for their support, gifts of extracts and oligonucleotides, and helpful suggestions and comments about this work. We especially thank Denise Malone for timely construction and sequencing of the p(EFII3')2-CAT4 construct, Kathy Matix for assistance in construction of the p(EFII)*n*CAT constructs, and Mary Faber for the preparation of the pII20x4 and pMLS-II plasmids.

This work was supported by Public Health Service grant GM39826 (to L.S.). R.C.S. was a predoctoral fellow of the Multidisciplinary Basic Research Training in Cancer grant CA09592, and L.S. was a Scholar of the Leukemia Society of America.

REFERENCES

1. **Akira, S., H. Isshiki, T. Sugita, O. Tanabe, S. Kinoshita, Y. Nishio, T. Nakajima, T. Hirano, and T. Kishimoto.** 1990. A nuclear factor for IL-6 expression (NF-IL6) is a member of a C/EBP family. EMBO J. **9**:1897–1906.
2. **Arrigo, S., M. Yun, and K. Beemon.** 1987. *cis*-Acting regulatory elements within *gag* genes of avian retroviruses. Mol. Cell. Biol. **7**:388–397.
3. **Birkenmeier, E. H., B. Gwynn, S. Howard, J. Jerry, J. I. Gordon, W. H. Landschulz, and S. L. McKnight.** 1989. Tissue-specific expression, developmental regulation, and genetic mapping of the gene encoding CCAAT/enhancer binding protein. Genes Dev. **3**:1146–1156.
4. **Boulden, A., and L. Sealy.** 1990. Identification of a third protein factor which binds to the Rous sarcoma virus LTR enhancer: possible homology with the serum response factor. Virology **174**:204–216.
5. **Boulden, A., and L. Sealy.** 1992. Maximal serum stimulation of the c-*fos* serum response element (SRE) requires both the serum response factor and a novel binding factor, SRE-binding protein (SRE-BP). Mol. Cell. Biol. **12**:4769–4783.
6. **Buscher, M., H. J. Rahmsdorf, M. Litfin, M. Karin, and P. Herrlich.** 1988. Activation of the c-*fos* gene by UV and phorbol ester: different signal pathways converge to the same enhancer element. Oncogene **3**:301–311.
7. **Cao, Z., R. M. Umek, and S. L. McKnight.** 1991. Regulated expression of three C/EBP isoforms during adipose conversion of 3T3-L1 cells. Genes Dev. **5**:1538–1552.
8. **Carlberg, K., T. A. Ryden, and K. Beemon.** 1988. Localization and footprinting of an enhancer within the avian sarcoma virus *gag* gene. J. Virol. **62**:1617–1624.
9. **Cereghini, S., M. Raymondjean, A. G. Carranca, P. Herbomel, and M. Yaniv.** 1987. Factors involved in control of tissue-specific expression of albumin gene. Cell **50**:627–638.
10. **Chang, C.-J., T.-T. Chen, H.-Y. Lei, S.-S. Chen, and S.-C. Lee.** 1990. Molecular cloning of a transcription factor, AGP/EBP, that belongs to members of the C/EBP family. Mol. Cell. Biol. **10**:6642–6653.
11. **Christy, R. J., V. W. Yang, J. M. Ntambi, D. E. Geimen, W. H. Landschulz, A. D. Friedman, Y. Nakabeppu, T. J. Kelly, and M. D. Lane.** 1989. Differentiation-induced gene expression in 3T3-L1 preadipocytes: CCAAT/enhancer binding protein interacts with and activates the promoters of two adipocyte-specific genes. Genes Dev. **3**:1323–1335.
12. **Cissell, M., and R. Chalkley.** 1991. Personal communication.
13. **Cullen, B. R., K. Raymond, and G. Ju.** 1985. Functional analysis of the transcription control region located within the avian retroviral long terminal repeat. Mol. Cell. Biol. **5**:438–447.
14. **Cullen, B. R., K. Raymond, and G. Ju.** 1985. Transcriptional activity of avian retroviral long terminal repeats directly correlates with enhancer activity. J. Virol. **53**:515–521.
15. **Descombes, P., M. Chojkier, S. Lichtsteiner, E. Falvey, and U. Schibler.** 1990. LAP, a novel member of the C/EBP gene family, encodes a liver-enriched transcriptional activator protein. Genes Dev. **4**:1541–1551.
16. **Ephrussi, A., G. M. Church, S. Tonegawa, and W. Gilbert.** 1985. B lineage-specific interactions of an immunoglobulin enhancer with cellular factors *in vivo*. Science **227**:134–140.
17. **Faber, M., and L. Sealy.** 1990. Rous sarcoma virus enhancer factor I is a ubiquitous CCAAT transcription factor highly related to CBF and NF-Y. J. Biol. Chem. **265**:22243–22254.
18. **Friedman, A. D., W. H. Landschulz, and S. L. McKnight.** 1989. CCAAT/enhancer binding protein activates the promoter of the serum albumin gene in cultured hepatoma cells. Genes Dev. **3**:1314–1322.
19. **Gentz, R., F. J. Rauscher III, C. Abate, and T. Curran.** 1989. Parallel association of fos and jun leucine zippers juxtaposes DNA binding domains. Science **243**:1695–1699.
20. **Gilmartin, G. M., and J. T. Parsons.** 1983. Identification of transcriptional elements within the long terminal repeat of Rous sarcoma virus. Mol. Cell. Biol. **3**:1834–1845.
21. **Gonzalez, G. A., K. K. Yamamoto, W. H. Fischer, D. Karr, P. Menzel, W. Biggs III, W. W. Vale, and M. R. Montminy.** 1989. A cluster of phosphorylation sites on the cyclic AMP-regulated nuclear factor CREB predicted by its sequence. Nature (London) **337**:749–752.
22. **Goodwin, G. H.** 1988. Identification of three sequence-specific DNA-binding proteins which interact with the Rous sarcoma virus enhancer and upstream promoter elements. J. Virol. **62**:2186–2190.
23. **Gorman, C. M., G. T. Merlino, M. C. Willingham, I. Pastan, and B. H. Howard.** 1982. The Rous sarcoma virus long terminal repeat is a strong promoter when introduced into a variety of eukaryotic cells by DNA-mediated transfection. Proc. Natl. Acad. Sci. USA **79**:6777–6781.
24. **Gorman, C. M., L. F. Moffat, and B. H. Howard.** 1982. Recombinant genomes which express chloramphenicol acetyltransferase in mammalian cells. Mol. Cell. Biol. **2**:1044–1051.
25. **Gowda, S., A. S. Rao, Y. W. Kim, and R. V. Guntaka.** 1988. Identification of sequences in the long terminal repeat of avian sarcoma virus required for efficient transcription. Virology **162**:243–247.
26. **Graham, F. L., and A. J. Van der Eb.** 1973. A new technique for the assay of infectivity of human adenovirus 5 DNA. Virology **52**:456–467.
27. **Graves, B. J., P. F. Johnson, and S. L. McKnight.** 1986. Homologous recognition of a promoter domain common to the MSV LTR and the HSV *tk* gene. Cell **44**:565–576.
28. **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.
29. **Johnson, P. F., W. H. Landschulz, B. J. Graves, and S. L. McKnight.** 1987. Identification of a rat liver nuclear protein that binds to the enhancer core element of three animal viruses. Genes Dev. **1**:133–146.
30. **Johnson, P. F., and S. L. McKnight.** 1989. Eukaryotic transcriptional regulatory proteins. Annu. Rev. Biochem. **58**:799–839.
31. **Jones, N. C., P. W. J. Rigby, and E. B. Ziff.** 1988. *trans*-Acting protein factors and the regulation of eukaryotic transcription: lessons from studies on DNA tumor viruses. Genes Dev. **2**:267–281.
32. **Kapiloff, M. S., J. M. Mathis, C. A. Nelson, C. R. Lin, and M. G. Rosenfeld.** 1991. Calcium/calmodulin-dependent protein kinase mediates a pathway for transcriptional regulation. Proc. Natl. Acad. Sci. USA **88**:3710–3714.
33. **Karnitz, L., S. Faber, and R. Chalkley.** 1987. Specific nuclear proteins interact with the Rous sarcoma virus internal enhancer and share a common element with the enhancer located in the long terminal repeat of the virus. Nucleic Acids Res. **15**:9841–9859.
34. **Karnitz, L., D. Poon, P. A. Weil, and R. Chalkley.** 1989. Purification and properties of the Rous sarcoma virus internal enhancer binding factor. Mol. Cell. Biol. **9**:1929–1939.
35. **Kerppola, T. K., and T. Curran.** 1991. DNA bending by fos and jun: the flexible hinge model. Science **254**:1210–1214.
36. **Kouzarides, T., and E. Ziff.** 1988. The role of the leucine zipper in the fos-jun interaction. Nature (London) **336**:646–651.
37. **Kuwabara, M., and D. Sigman.** 1987. Footprinting DNA-protein

complexes in situ following gel retardation assays using 1,10-phenanthroline-copper ion: Escherichia coli RNA polymerase-lac promoter complexes. Biochemistry **26:**7234–7238.
38. **Laimins, L. A., P. Tsichlis, and G. Khoury.** 1984. Multiple enhancer domains in the 3′ terminus of the Prague strain of Rous sarcoma virus. Nucleic Acids Res. **12:**6427–6442.
39. **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.
40. **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.
41. **Landschulz, W. H., P. F. Johnson, and S. L. McKnight.** 1989. The DNA binding domain of the rat liver nuclear protein C/EBP is bipartite. Science **243:**1681–1688.
42. **Luciw, P. A., J. M. Bishop, H. E. Varmus, and M. R. Capecchi.** 1983. Location and function of retroviral and SV40 sequences that enhance biochemical transformation after microinjection of DNA. Cell **33:**705–716.
43. **MacGregor, G. R., and C. T. Caskey.** 1989. Construction of plasmids that express *E. coli* β-galactosidase in mammalian cells. Nucleic Acids Res. **17:**2365.
44. **Maniatis, T., S. Goodbourn, and J. A. Fischer.** 1987. Regulation of inducible and tissue-specific gene expression. Science **236:** 1237–1245.
45. **Maxam, A., and W. Gilbert.** 1980. Sequencing end-labeled DNA with base-specific chemical cleavages. Methods Enzymol. **65:** 499–560.
46. **Metz, R., and E. Ziff.** 1991. The helix-loop-helix protein rE12 and the C/EBP-related factor rNFIL-6 bind to neighboring sites within the c-*fos* serum response element. Oncogene **6:**2165–2178.
47. **Metz, R., and E. Ziff.** 1991. cAMP stimulates the C/EBP-related transcription factor rNFIL-6 to *trans*-locate to the nucleus and induce c-*fos* transcription. Genes Dev. **5:**1754–1766.
48. **Minty, A., and L. Kedes.** 1986. Upstream regions of the human cardiac actin gene that modulate its transcription in muscle cells: presence of an evolutionarily conserved repeated motif. Mol. Cell. Biol. **6:**2125–2136.
49. **Mitchell, P. J., and R. Tjian.** 1989. Transcriptional regulation in mammalian cells by sequence-specific DNA binding proteins. Science **245:**371–378.
50. **Mudryj, M., and B. de Crombrugghe.** 1988. Deletion analysis of the mouse alpha1 (III) collagen promoter. Nucleic Acids Res. **16:**7513–7526.
51. **Norton, P. A., and J. M. Coffin.** 1985. Bacterial β-galactosidase as a marker of Rous sarcoma virus gene expression and replication. Mol. Cell. Biol. **5:**281–290.
52. **Norton, P. A., and J. M. Coffin.** 1987. Characterization of Rous sarcoma virus sequences essential for viral gene expression. J. Virol. **61:**1171–1179.
53. **O'Shea, E. K., R. Rutkowski, and P. S. Kim.** 1989. Evidence that the leucine zipper is a coiled coil. Science **243:**538–542.
54. **Peterson, C. L., and K. L. Calame.** 1987. Complex protein binding within the mouse immunoglobulin heavy-chain enhancer. Mol. Cell. Biol. **7:**4194–4203.
55. **Peterson, C. L., S. Eaton, and K. Calame.** 1988. Purified μEBP-E binds to immunoglobulin enhancers and promoters. Mol. Cell. Biol. **8:**4972–4980.
56. **Poli, V., F. P. Mancini, and R. Cortese.** 1990. IL-6DBP, a nuclear protein involved in interleukin-6 signal transduction, defines a new family of leucine zipper proteins related to C/EBP. Cell **63:**643–653.
57. **Ptashne, M.** 1988. How eukaryotic transcriptional activators work. Nature (London) **335:**683–689.
58. **Roman, C., J. S. Platero, J. Shuman, and K. Calame.** 1990. Ig/EBP-1: a ubiquitously expressed immunoglobulin enhancer binding protein that is similar to C/EBP and heterodimerizes with C/EBP. Genes Dev. **4:**14041415.
59. **Ruddell, A., M. L. Linial, and M. Groudine.** 1989. Tissue-specific lability and expression of avian leukosis virus long terminal repeat enhancer-binding proteins. Mol. Cell. Biol. **9:**5660–5668.
60. **Ruddell, A., M. Linial, W. Schubach, and M. Groudine.** 1988. Lability of leukosis virus enhancer-binding proteins in avian hematopoietic cells. J. Virol. **62:**2728–2735.
61. **Ryden, T. A., and K. Beemon.** 1989. Avian retroviral long terminal repeats bind CCAAT/enhancer-binding protein. Mol. Cell. Biol. **9:**1155–1164.
62. **Sassone-Corsi, P., L. J. Ransone, W. W. Lamph, and I. M. Verma.** 1988. Direct interaction between fos and jun nuclear onco-proteins: role of the "leucine zipper" domain. Nature (London) **336:**692–695.
63. **Schaffner, W., and C. Weissman.** 1973. A rapid, sensitive and specific method for the determination of protein in dilute solutions. Anal. Biochem. **56:**502–514.
64. **Sealy, L., and R. Chalkley.** 1987. At least two nuclear proteins bind specifically to the Rous sarcoma virus long terminal repeat enhancer. Mol. Cell. Biol. **7:**787–798.
65. **Shuman, J. D., C. R. Vinson, and S. L. McKnight.** 1990. Evidence of changes in protease sensitivity and subunit exchange rate on DNA binding by C/EBP. Science **249:**771–774.
66. **Treisman, R.** 1986. Identification of a protein-binding site that mediates transcriptional response of the c-*fos* gene to serum factors. Cell **46:**567–574.
67. **Tsao, B. T., X.-F. Wang, C. L. Peterson, and K. Calame.** 1988. *In vivo* analysis of *in vitro* protein binding sites in the immunoglobulin heavy chain enhancer. Nucleic Acids Res. **16:**3239–3253.
68. **Turner, R., and R. Tjian.** 1989. Leucine repeats and an adjacent DNA binding domain mediate the formation of functional cfos-cjun heterodimers. Science **243:**1689–1694.
69. **Varmus, H. E., and R. Swanstrom.** 1982. Replication of retroviruses, p. 369–512. *In* R. Weiss, N. Teich, J. Coffin, and H. E. Varmus (ed.), The molecular biology of RNA tumor viruses. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
70. **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.
71. **Weber, F., and W. Schaffner.** 1985. Enhancer activity correlates with the oncogenic potential of avian retroviruses. EMBO J. **4:**949–956.
72. **Wegner, M., Z. Cao, and M. G. Rosenfeld.** 1992. Calcium-regulated phosphorylation within the leucine zipper of C/EBPβ. Science **256:**370–373.
73. **Williams, S. C., C. A. Cantwell, and P. F. Johnson.** 1991. A family of C/EBP-related proteins capable of forming covalently linked leucine zipper dimers *in vitro*. Genes Dev. **5:**1553–1567.
74. **Ziff, E. B.** 1990. Transcription factors: a new family gathers at the cAMP response site. Trends Genet. **6:**69–72.