MOLECULAR AND CELLULAR BIOLOGY, June 1993, p. 3324–3339
0270-7306/93/063324-16$02.00/0

Vol. 13, No. 6

Cloning and Characterization of Subunits of the T-Cell Receptor and Murine Leukemia Virus Enhancer Core-Binding Factor

SHUWEN WANG,† QING WANG, BARBARA E. CRUTE, IRENA N. MELNIKOVA,‡ SUSANNA R. KELLER, AND NANCY A. SPECK*

Department of Biochemistry, Dartmouth Medical School, Hanover, New Hampshire 03755

Received 11 January 1993/Returned for modification 10 February 1993/Accepted 2 March 1993

Moloney murine leukemia virus causes thymic leukemias when injected into newborn mice. A major determinant of the thymic disease specificity of Moloney virus genetically maps to the conserved viral core motif in the Moloney virus enhancer. Point mutations introduced into the core site significantly shifted the disease specificity of the Moloney virus from thymic leukemia to erythroid leukemia (N. A. Speck, B. Renjifo, E. Golemis, T. N. Fredrickson, J. W. Hartley, and N. Hopkins, Genes Dev. 4:233–242, 1990). We previously reported the purification of core-binding factors (CBF) from calf thymus nuclei (S. Wang and N. A. Speck, Mol. Cell. Biol. 12:89–102, 1992). CBF binds to core sites in murine leukemia virus and T-cell receptor enhancers. Affinity-purified CBF contains multiple polypeptides. In this study, we sequenced five tryptic peptides from two of the bovine CBF proteins and isolated three cDNA clones from a mouse thymus cDNA library encoding three of the tryptic peptides from the bovine proteins. The cDNA clones, which we call CBF β p22.0, CBF β p21.5, and CBF β p17.6, encode three highly related but distinct proteins with deduced molecular sizes of 22.0, 21.5, and 17.6 kDa that appear to be translated from multiply spliced mRNAs transcribed from the same gene. CBF β p22.0, CBF β p21.5, and CBF β p17.6 do not by themselves bind the core site. However, CBF β p22.0 and CBF β p21.5 form a complex with DNA-binding CBF α subunits and as a result decrease the rate of dissociation of the CBF protein-DNA complex. Association of the CBF β subunits does not extend the phosphate contacts in the binding site. We propose that CBF β is a non-DNA-binding subunit of CBF and does not contact DNA directly.

Core-binding factor (CBF), otherwise known as SL3-3 enhancer factor 1 (56), polyomavirus enhancer-binding factor 2 (PEBP2) (29, 49), and SL3 and AKV core-binding factor (7), binds the conserved core site in mammalian type C retroviral enhancers and the polyomavirus enhancer. CBF also binds sites in the enhancers of several genes specifically expressed in T cells, including the T-cell receptor (TCR) chain genes (TCRγ, -δ, and -β [22, 42, 44, 60]) and the immunoglobulin μ-chain gene (60). Consensus CBF binding sites have also been noted in the TCR CD3-ε and CD3-δ genes and in the human tumor necrosis factor α- and β-chain genes (54, 57).

Experiments from several laboratories identified the core site as a T-cell-specific *cis*-acting sequence. It was first found that mutations in the core site in Moloney and SL3-3 murine leukemia virus (MLV) enhancers affected transcription from these enhancers specifically in T cells (7, 53, 56). These core site mutations also altered the pathogenic phenotypes of Moloney and SL3-3 MLVs. A 2-bp mutation introduced into both core sites in the Moloney MLV enhancer altered the disease specificity of Moloney MLV from thymic leukemia to erythroid leukemia (52), whereas a 3-bp mutation in all four core sites in the SL3-3 MLV enhancer rendered SL3 MLV nonleukemogenic (20). Subsequently, several laboratories studying the TCR enhancers demonstrated that mutation of core sites in the TCRγ and -δ enhancers causes a large decrease in transcription from these enhancers (5- to 20-fold), thus identifying the core site as a critical *cis*-acting element for transcription of the TCRγ and -δ genes (22, 45). Together, these studies suggest that the core-binding factors contribute to the transcription of genes specifically expressed in T cells and that the thymotropic mammalian type C retroviruses have coopted these proteins for the purpose of viral gene expression.

We and others purified core-binding factors (CBF and PEBP2) from calf thymus and c-Harvey-*ras*-transformed mouse fibroblasts (29, 60). Affinity-purified CBF from both sources consists of multiple polypeptides. Here we describe the isolation of cDNA clones encoding several of the CBF polypeptides from a mouse thymus cDNA library. These cDNA clones encode non-DNA-binding subunits of CBF, which we have named CBF β subunits. CBF β subunits participate in the formation of a protein-DNA complex with DNA-binding subunits of CBF which we call CBF α subunits. The association of CBF β decreases the rate of dissociation of the CBF protein-DNA complex but does not extend the phosphate contacts made by CBF α on the DNA. We also present evidence suggesting that one of the CBF α subunits is the product of a newly identified proto-oncogene called the acute myeloid leukemia 1 (*AML1*) gene (39).

MATERIALS AND METHODS

Electrophoretic mobility shift assays. Complementary 18-base oligonucleotides containing a high-affinity core site deduced by Thornell et al. (57) (HA; 5′-GGATATTTGCG GTTAGCA-3′), wild-type core site from the Moloney MLV enhancer (WT; 5′-GGATATCTGTGGTAAGCA-3′), and

* Corresponding author.
† Present address: Department of Physiology, University of California School of Medicine, San Francisco, CA 94143.
‡ Present address: Institute of Biotechnology, University of Texas Health Science Center, San Antonio, TX 78245.

mutant core site (MUT; 5′-GGATATCTGCCGTAAGCA-3′) were synthesized, purified, and end labeled with [γ-^{32}P]ATP as described previously (60).

Competitor oligonucleotides were prepared by annealing equal amounts of unlabeled complementary oligonucleotides (27). The annealed oligonucleotides were used directly as competitors in binding reactions. Binding reaction mixtures contained 10,000 to 20,000 cpm (2 to 5 fmol) of ^{32}P-end-labeled HA probe, binding buffer (10 mM Tris [pH 7.4], 1 mM β-mercaptoethanol, 1 mM EDTA, 50 mM NaCl, 4% glycerol) and 1 to 10 μl of protein sample in a total volume of 15 μl. To demonstrate the sequence specificity of the protein-DNA complex, an excess of unlabeled WT, HA, or MUT oligonucleotide was included in some of the binding reactions. After 15 min of incubation at room temperature (RT) (unless otherwise indicated), the reaction mixtures were fractionated by electrophoresis through 6% native polyacrylamide gels in 0.5× TBE buffer (22.5 mM Tris-HCl, 22.5 mM boric acid, 0.5 mM EDTA). Electrophoresis was carried out at 130 V for 2 h at RT. Radioactivity was detected by autoradiography of the dried gel.

Purification of CBF. CBF was purified from calf thymus nuclear extracts as described previously (60), but the amount of starting material and the column sizes were increased. All buffers and reagents used in the scaled-up purification were the same as those used in previous purifications (60) except that the concentration of Nonidet P-40 in buffer Z was reduced from 0.1% to 0.02%. Approximately 130 μg of CBF was purified from 10 g of crude nuclear extract by selective pH denaturation, followed by chromatography on heparin-Sepharose columns (2.5 by 16 cm, 80 ml of resin), nonspecific double-stranded DNA-cellulose columns (2.5 by 7 cm, 34 ml of resin), and core oligonucleotide-coupled affinity columns (1.5 by 2.3 cm, 4 ml of the HA resin). Fifty-milliliter aliquots of fractions from the DNA-cellulose column with core-binding activity were passed four times successively over the HA column, and the fractions eluted with 1 M NaCl were collected and pooled. The total volume obtained after the fourth passage through the HA affinity column was approximately 35 ml.

Amino acid sequence determination. Five milliliters of purified CBF was precipitated with 9.1% trichloroacetic acid in the presence of 0.14 mg of deoxycholate per ml as carrier. After incubation on ice for 45 min, the samples were pelleted by microcentrifugation at 16,000 × *g* for 15 min at 4°C. The pellet was resuspended in 100 μl of sample buffer (20 mM dithiothreitol [DTT], 92.5 mM Tris-HCl [pH 6.8], 10% glycerol, 2% sodium dodecyl sulfate [SDS], 2.5 mM EDTA, 0.004% bromophenol blue), the pH of the sample was adjusted to neutrality with 2 M Tris base, and the sample was boiled and electrophoresed through an SDS–15% polyacrylamide gel (15 by 15 cm, 1.0 mm thick) (33) along with low-molecular-weight protein standards (Bio-Rad). Electrophoresis was conducted at constant current of 25 mA for about 5 to 6 h.

Proteins were electrophoretically transferred from the gel to a nitrocellulose membrane (0.45-mm pore size; Schleicher & Schuell) in Tris-glycine transfer buffer (190 mM glycine, 25 mM Tris) at 70 to 75 V (0.8 to 1.4 A) at 4°C for 70 to 120 min. Following transfer, proteins on the membrane were stained with Ponceau S, using a modification of the method described by Salinovich and Montelaro (47). The nitrocellulose membrane was immersed for 1 to 2 min in a solution containing 0.2% Ponceau S dye (FisherBiotech) in 1% aqueous acetic acid. Excess stain was removed from the membrane by gentle agitation for 60 to 90 s in 1% aqueous acetic acid. Protein bands detected by Ponceau S staining were excised and transferred to Eppendorf tubes. The membrane-bound proteins were briefly destained with 0.2 M NaOH for 1 to 2 min and washed with distilled water. Two membrane-bound CBF polypeptides of 24 and 19 kDa were used for amino acid analysis. In situ tryptic digests and amino acid sequence determination were performed by William Lane at the Harvard Microchemistry Facility.

PCR amplification of a 50-bp sequence encoding peptide 3/5. A DNA fragment encoding the central portion of tryptic peptide 3/5 from p24 and p19 CBF (see Fig. 3 for peptide sequences) was amplified by polymerase chain reaction (PCR) from mouse thymus cDNA (11, 35). Primer SW1 (Operon Technology) was designed from the amino acid sequence near the N terminus of peptide 3/5, Glu-Ile-Ala-Phe-Val. The base corresponding to the third position in the codon for the fifth amino acid (Val) was omitted to reduce the degeneracy of the primer by twofold. An additional eight bases which included the recognition site for *Bam*HI were added to the 5′ end, giving rise to primer SW1, containing 22 nucleotides with a degeneracy of 48-fold. Sequences of the primers are as follows:

SW1 5′-CGGGATCCGARATHGCNTTYGT-3′ (*Bam*HI) degeneracy, 48; T_m = 61.6°C
SW2A 5′-CGGAATTC[G]TGYTCNCCYTGCCA-3′ (*Eco*RI) degeneracy, 16; T_m = 60.8°C
SW2B 5′-CGGAATTC(Q)TGYTCYTGYTGCCA-3′ (*Eco*RI) degeneracy, 8; T_m = 58.8°C
SW3A 5′-CARTTYTTYCCNGCNTCNTGGCA-3′ degeneracy, 512; T_m = 64
SW3B 5′-CARTTYTTYCCNGCNAGYTGGCA-3′ degeneracy, 256; T_m = 64

The antisense primer (primer SW2A or SW2B) was designed from five amino acids at the 3′ end of peptide 3/5 (Trp-Gln-Gly/Gln-Glu-Gln). Two separate primers were designed because of the ambiguity at amino acid 25 (Gly and Gln). SW2A is 16-fold degenerate and encodes the amino acid Gly at position 25. SW2B is eightfold degenerate, and encodes Gln at position 25. Both the SW2A and SW2B primers contain a restriction site for *Eco*RI at the 5′ end. Again, the base corresponding to the third position in the codon for the fifth amino acid (Gln) was omitted to decrease the degeneracy of the primers by twofold.

Twenty-three-base oligonucleotides were designed from the amino acid sequence in the middle of peptide 3/5, Gln-Phe-Phe-Pro-Ala-Ser-Trp-Gln, to use as probes in colony hybridizations. Because of the high degeneracy of Ser, two separate pools of degenerate oligonucleotides were designed to reduce the total degeneracy to 768. SW3A was designed from four possible codons of Ser, and SW3B was designed from the remaining Ser codons. SW3A is twofold more degenerate than SW3B; therefore, SW3A and SW3B were combined in a 2:1 molar ratio and used as probes to screen for specific PCR products.

Lambda phage DNA isolated from a liquid phage lysate (34) of a mouse thymus λZAP cDNA library (Stratagene) was used as the template for PCR amplification. Two rounds of PCR were conducted to amplify a unique nucleotide sequence encoding the middle portion of peptide 3/5. The PCR mixture contained 1 μg of λZAP cDNA, 250 pmol each of SW1 plus SW2A or SW1 plus SW2B, and 0.2 mM each deoxynucleoside triphosphate, in reaction buffer (10 mM Tris-HCl, 50 mM KCl, 1.5 mM $MgCl_2$, 0.001% [wt/vol] gelatin), and 1.25 U of *Taq* DNA polymerase, in a total volume of 50 μl. All the reagents used for PCR were supplied in the GeneAmp DNA amplification reagent kit (Perkin-Elmer/Cetus). After the initial denaturation at 94°C for 4 min, the reaction mixtures were incubated at 94°C for 1 min

and 37°C for 2 min for 5 cycles, at 94°C for 1 min and T_m − 15°C for 2 min for 25 cycles, and then for 7 min at 72°C. T_m − 15°C is 46 and 43°C for reactions containing SW2A and SW2B, respectively.

PCR products from the first amplification were electrophoresed through a 4% low-melting-point agarose gel, and a gel slice corresponding to the expected size of 88 bp was excised from the gel. PCR products recovered from the gel slice were used as templates for a second round of PCR (94°C for 4 min; 94°C for 1 min and T_m − 15°C for 2 min for 30 cycles; 72°C for 7 min). A PCR product of the expected size (88 bp) was easily detectable on an ethidium bromide-stained 4% agarose gel.

The PCR-amplified DNA fragments were digested with *Bam*HI and *Eco*RI, isolated from 4% low-melting-point agarose gels, and subcloned into the pBluescript SK$^+$ vector in the *Bam*HI and *Eco*RI sites. *Escherichia coli* MM294 was transformed with ligation mixtures, and colonies were screened by colony hybridization using the ^{32}P-end-labeled oligonucleotides SW3A and SW3B. DNA-bound nitrocellulose filters were prehybridized at 39°C (T_m − 25°C) for 4 h in 6× SSC (1× SSC is 0.15 M NaCl plus 0.015 M sodium citrate)–5× Denhardt's solution–0.05% sodium pyrophosphate–0.1% SDS–0.1 mg of salmon sperm DNA per ml and then hybridized with 1 × 10^6 to 2 × 10^6 cpm of ^{32}P-end-labeled SW3A and SW3B per ml in the same buffer for 20 h at 39°C. The filters were then washed four times in 2× SSC–0.1% SDS for 15 min each time at RT and once for 15 min at 44°C (T_m − 20°C). Following autoradiography of the washed filters, putative positive clones were sequenced. A 50-bp sequence encoding the amino acids between the primers was identified in several subclones obtained by using primers SW1 and SW2A.

Isolation of CBF β cDNA clones. Two 25-base oligonucleotides, SW11 (5′-GTGGCTACAGGAACCAATCTGTCTC-3′) and SW12 (5′-TCCAGTTTTTTCCGGCCAGCTGGCA-3′) (Macromolecular Resources, Colorado State University), were designed from each end of the 50-bp sequence encoding the middle portion of peptide 3/5 and used to screen the mouse thymus λZAP cDNA library. Duplicate filters (one set of filters for each oligonucleotide probe) were hybridized for 48 h at 42°C in 6× SSC–5× Denhardt's solution–0.1% SDS–0.05% sodium pyrophosphate–0.1 mg of salmon sperm DNA per ml–10^5 cpm of ^{32}P-end-labeled SW11 or SW12 oligonucleotide probe per ml. The filters were washed four times at RT with 6× SSC–0.1% SDS and exposed overnight.

One positive phage was isolated from screening 5.6 × 10^5 phage. In vivo excision of the phagemid containing the insert was performed as instructed by the manufacturer (Stratagene). The 1.4-kb insert from the phagemid was excised with *Eco*RI, labeled by random priming (14, 15), and used as a probe to rescreen the mouse thymus cDNA library. Eleven independent clones were isolated and purified.

The CBF cDNA clones were sequenced by sequential oligonucleotide mapping along the cDNA inserts in both directions, using the dideoxy-chain termination method (48). All reagents were provided in a Sequenase version 2.0 kit (U.S. Biochemical). Sequences of the cDNA clones were read and input into the computer by using the DNA Parrot program and analyzed by DNA Strider and the Genetics Computer Group package of the University of Wisconsin.

Preparation of CBF β-GST fusion proteins. DNA fragments containing the open reading frames of CBF β p22.0, CBF β p21.5, and CBF β p17.6 were subcloned into the glutathione *S*-transferase (GST) fusion vector pGEX2T (Pharmacia) for the purpose of expressing CBF β-GST fusion proteins. Two PCR primers hybridizing to positions 82 to 102 of CBF β p21.5 (5′-AGACGGATCCATGCCGC GCGTCGTCCCGGAC-3′, starting at the ATG translation start codon [underlined] and including a *Bam*HI site at the 5′ end of the primer) and positions 688 to 703 (5′-AGAGATGGGGCACATAAG-3′) were used to amplify a 603-bp fragment. The PCR-amplified fragment was subsequently digested with *Bam*HI and at an internal *Hae*II site (position 201 of CBF β p21.5), yielding a 125-bp fragment which contains sequences common to all three CBF β cDNA clones. This PCR-amplified 125-bp DNA fragment was subcloned along with a *Hae*II-*Pst*I fragment from each of the CBF β clones (corresponding to nucleotides 202 to 777 of CBF β p21.5) into the *Bam*HI and *Pst*I sites in the polylinker of pBluescript SK$^+$. The correct sequence of the PCR-amplified region was confirmed by dideoxy sequencing from the T3 primer. The CBF β inserts were then excised from the pBluescript SK$^+$ phagemid with *Bam*HI and *Eco*RI and subcloned into the *Bam*HI-*Eco*RI sites of pGEX2T. The CBF β-GST fusion proteins were expressed in *E. coli* MM294 and purified on glutathione-agarose beads (51).

Recovery of DNA-binding CBF α subunits from SDS-polyacrylamide gels. DNA-binding CBF α subunits were isolated from affinity-purified bovine CBF (19, 60). One milliliter of purified bovine CBF was trichloroacetic acid precipitated and fractionated through an SDS–15% polyacrylamide gel. The gel was stained with Coomassie brilliant blue, the visible polypeptides were excised from the gel, and the gel slices were chopped into small pieces. Proteins were eluted by incubating the gel slice in 200 to 300 μl of a buffer containing 150 mM NaCl, 20 mM *N*-2-hydroxyethylpiperazine-*N*′-2-ethanesulfonic acid (HEPES; pH 7.5), 5 mM DTT, 0.1 mM EDTA, 0.1% SDS, and 0.1 mg of bovine serum albumin (BSA) per ml overnight at RT. The eluted proteins were precipitated with 4 volumes of cold (−20°C) acetone. Following incubation for 45 min in a dry ice-ethanol bath, proteins were pelleted by microcentrifugation at 16,000 × *g* for 30 min. The pellets were washed with a solution containing 80% acetone, 20% dilution buffer (150 mM NaCl, 20 mM HEPES [pH 7.5], 1 mM DTT, 0.1 mM EDTA, 20% glycerol), dried, dissolved in 5 to 10 μl of dilution buffer supplemented with 6 M guanidine-HCl, and incubated at RT for 30 min. The proteins were renatured by diluting the sample 1:50 with dilution buffer and incubating it for 2 h at RT. The protein samples were then assayed for core-binding activity by electrophoretic mobility shift assay. Protein samples that yielded only one protein-DNA complex and were thus judged to be free from contamination by CBF β subunits were used in subsequent analyses.

DNase I footprinting analysis. Probes from the top and bottom strands of the TCRγ enhancer probe were prepared from plasmid J21-.20RR (54), kindly provided by David Raulet. J21-.20RR contains a 196-bp *Rsa*I-*Rsa*I fragment from the TCRγ enhancer subcloned into the polylinker of plasmid J21 (61). The plasmid was labeled at the *Xho*I or *Cla*I site in the polylinker of J21 as described previously (60). DNase I footprinting was conducted exactly as previously described (60).

Measurement of dissociation rate constants. The rate of dissociation of CBF α and CBF α+β subunits was measured by using electrophoretic mobility shift assays. A fixed amount of the CBF α and CBF α plus CBF β-GST fusion protein (approximately 10^{-9} M CBF α and 10^{-6} M CBF β p21.5-GST) was incubated with a saturating amount of the HA probe (75-μl total volume per binding reaction) at RT for 20 min and then on ice for 15 min. Binding reactions were

also performed with bovine CBF polypeptides renatured from SDS-polyacrylamide gels that contained both CBF α and CBF β subunits (polypeptide bands F to M; see Fig. 5). A 15-μl aliquot of the binding reaction was loaded onto a running native polyacrylamide gel. A 50-fold molar excess of unlabeled HA binding site was then added to the remainder of the binding reaction to trap the dissociated CBF proteins, and 15-μl aliquots were removed at defined intervals (15, 30, 60, and 300 s) and loaded onto the running gel. Electrophoresis was conducted at 4°C. Following autoradiography of the dried gel, the autoradiogram was photographed, scanned, and quantified with the National Institutes of Health Image 1.44 program (36). The rate constant for dissociation was calculated from the slope of the plot of log(% bound at each time − % bound at $t = \infty$) (∞ = 300 s) versus time.

Ethylation interference analysis. The probes used in the ethylation interference analysis were 37-bp oligonucleotides containing a core site from the TCRδ enhancer (δE3, kind gift from Mike Krangel):

5′-CTAGAAGCAATGCATGTGGTTTCCAACCGTTAATGCTAGAG-3′
3′-TTCGTTACGTACACCAAAGGTTGGCAATTACGATCTCCTAG-5′

Double-stranded probes ^{32}P end labeled on either the top or bottom strand were modified on phosphates with *N*-ethyl-*N*-nitrosourea (Sigma) prior to use in binding reactions, as described by Siebenlist and Gilbert (50). The binding reaction (50-μl total volume) contained 100,000 cpm of end-labeled probe and either 20 μl of bovine CBF α subunit or 20 μl of CBF α plus 5 μl (12.5 μg) of CBF β p21.5-GST protein in the same buffer used for electrophoretic mobility shift assays. The binding reaction was electrophoresed through a 6% native polyacrylamide gel. Following overnight exposure of the wet gel, the bands corresponding to the CBF α and CBF α+β protein-DNA complexes and free DNA bands were excised from the gel, electroeluted onto a NA45 membrane (Schleicher & Schuell), and eluted from the membrane in elution buffer (20 mM Tris-HCl [pH 8.0], 0.1 mM EDTA, 1.0 M NaCl) at 68°C for 1 h (5). The DNA was then extracted once each with phenol and chloroform, recovered by ethanol precipitation, and subjected to alkali cleavage (50). Free ^{32}P-labeled single-stranded probe was also subjected to alkali cleavage. Equal amounts of radioactivity for each of the cleavage products were analyzed by electrophoresis through a 15% polyacrylamide–7 M urea sequencing gel.

Northern (RNA) and Southern analyses. RNA from mouse tissues was isolated by the guanidinium isothiocyanate method (10). Poly(A)$^+$ mRNA was purified on oligo(dT) columns (Stratagene) according to the manufacturer's instructions. One microgram of mRNA was electrophoresed through each lane of a 1.5% agarose gel and transferred to a Nitropure membrane (MSI, Westboro, Mass.). Following hybridization with the 1.4-kb insert from a partial CBF β cDNA clone (encompassing nucleotides 139 to 1463 of CBF β p21.5), the blot was washed with 2× SSPE–0.1% SDS at RT and 0.1× SSPE–0.1% SDS at elevated temperatures until background was diminished as monitored with a Geiger counter (approximately 65°C).

Genomic DNA was isolated from mouse liver (55). Twenty micrograms of genomic DNA was digested with restriction endonucleases *Bam*HI, *Eco*RI, *Hin*dIII, *Pst*I, *Pvu*II, *Sca*I, and *Xba*I. The digested DNA was fractionated through a 0.7% agarose gel and transferred to a nylon membrane (Nytran; Schleicher & Schuell). The membrane was prehybridized in 50% formamide–6× SSC–10× Denhardt's solution–1.0% SDS–0.1 mg of salmon sperm DNA per ml and hybridized in 40% formamide–6× SSPE–1.0% SDS–50 μg of salmon sperm DNA per ml–10% dextran sulfate, using the insert from the 1.4-kb partial CBF β cDNA clone (1 × 10^6 to 2 × 10^6 cpm/ml) at 42°C. The filters were washed twice with 2× SSC–0.1% SDS at RT for 15 min, once with 0.5× SSC–0.1% SDS at RT, and once with 0.1× SSC–0.1% SDS at 50°C for 15 min.

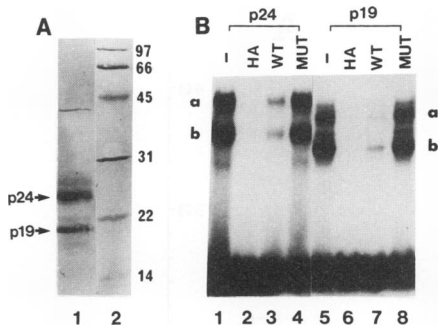


FIG. 1. Purification of bovine CBF polypeptides p24 and p19. (A) Ponceau S-stained nitrocellulose membrane following transfer of affinity purified bovine CBF from an SDS-polyacrylamide gel. Lanes: 1, affinity-purified CBF; 2, low-molecular-weight protein standards. The arrows indicate positions of the two CBF polypeptides (p24 and p19) that were excised and sequenced. Sizes are indicated in kilodaltons. (B) Assay for core-binding activity of renatured CBF p24 and p19 polypeptides recovered from a Coomassie brilliant blue-stained SDS-polyacrylamide gel. Lanes: 1, binding of CBF p24 to the HA probe; 2, binding of p24 to the HA probe in the presence of a 500-fold molar excess of unlabeled HA site; 3, competition of p24 binding to the HA probe by the WT core site; 4, competition with the MUT site; 5, binding of CBF p19 to the HA probe; 6, binding of p19 to the HA probe in the presence of unlabeled HA; 7, competition of p19 by unlabeled WT; 8, competition of p19 with unlabeled MUT core site. The letters a and b indicate two protein-DNA complexes obtained with proteins renatured from both the p24 and p19 polypeptide bands.

Nucleotide sequence accession numbers. The GenBank accession numbers for the CBF β cDNA clones are as follows: CBF β p21.5, L03279; CBF β p22.0, L03306; and CBF β p17.6, L03305. The CBF β p22.0 sequence entered in GenBank begins at the putative ATG translation start codon.

RESULTS

Purification and amino acid sequence analysis of CBF. We scaled up the purification of CBF, following the basic purification scheme described previously (60). The affinity-purified CBF preparation was fractionated through an SDS–15% polyacrylamide gel, and the proteins were transferred onto a nitrocellulose membrane (2) (Fig. 1A). Only two polypeptides migrating with apparent molecular sizes of 24 and 19 kDa (p24 and p19) were present in adequate amounts and transferred to the nitrocellulose membrane with sufficient efficiency for microsequence analysis. We eluted and renatured both proteins from a Coomassie brilliant blue-stained polyacrylamide gel run in parallel and assayed the renatured proteins to confirm that they had core-binding activity (Fig. 1B). Both renatured p24 and p19 bound to the core site, and both yielded two protein-DNA complexes (a and b) in electrophoretic mobility shift assays. We will discuss the significance of these two protein-DNA complexes later.

Both p24 and p19 were excised from the nitrocellulose membrane and digested with trypsin in situ, and the tryptic

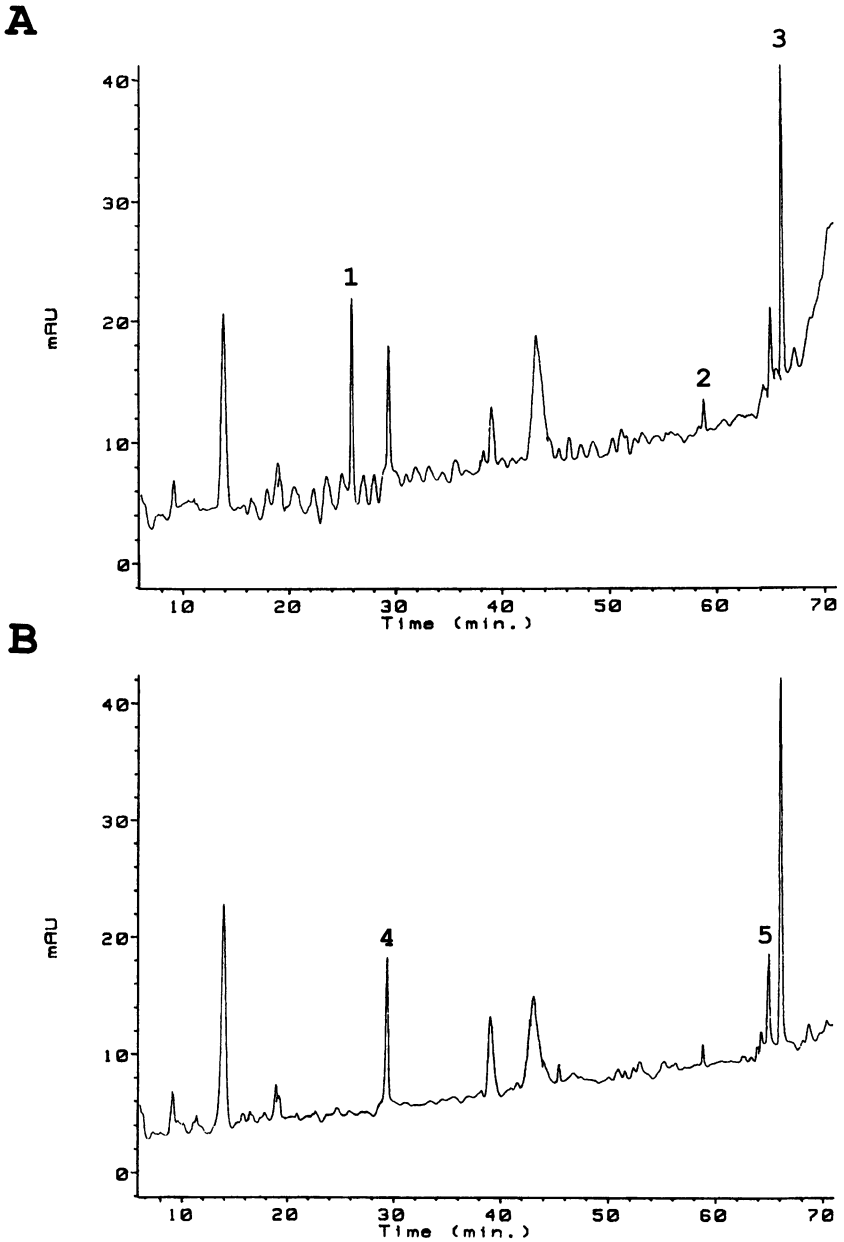


FIG. 2. Reverse-phase HPLC of tryptic peptides from CBF polypeptides p24 and p19. (A) Absorbance profile of tryptic peptides from p24 CBF at 210 nm. Numbered peaks correspond to sequenced peptides in Fig. 3. Note that peak 1 is the difference peak between p24 and p19. (B) Absorbance profile of tryptic peptides from p19 CBF. mAU, milli-absorption units.

peptides were separated by high-pressure liquid chromatography (HPLC) (Fig. 2). The elution profiles of tryptic peptides from p24 (Fig. 2A) and p19 (Fig. 2B) were superimposable, except that p24 contained an extra peptide peak (peak 1; Fig. 2A). The peptides in peaks 1 to 5 in Fig. 2 were sequenced (Fig. 3). Peak 4 contained two peptides and yielded two sequences (4a and 4b). Three peptide sequences from peaks 3 and 5 are combined into one sequence in Fig. 3 (peptide 3/5). Both the similar tryptic peptide elution profiles and the fact that both p24 and p19 contain tryptic peptides (in peaks 3 and 5) having the same amino acid sequence indicate that the two proteins are highly related.

Four tryptic peptides from p24 and p19 show no sequence homology to any proteins in computer data bases. However, peptide 2 from p24 shares 70% sequence identity with the protein encoded by the *Drosophila* segmentation gene *runt* (30) and 100% homology to a protein encoded by the human *AML1* gene (39) (Fig. 3). *runt* is a pair rule gene and plays a key role in regulating the expression of all other pair rule genes (8, 9, 17, 21, 24). *runt* encodes a nuclear protein which by sequence analysis does not contain either a recognizable DNA-binding or dimerization domain (30). The *AML1* gene is located on chromosome 21 in humans, at the t(8;21) translocation breakpoint frequently found in the leukemic cells from patients with the M2 subtype of acute myeloid leukemia (39). The function of the AML1 protein was not reported (39). The proteins encoded by *runt* and *AML1* are homologous in a 118-amino-acid (aa) region that includes a

Peptide	Sequence	p24	p19
1	R Q Q D P S P G S S L G G G D D L K	+	-
2	V V A L G D V P D G T L V T V M A G N D [E] N Y	+	+
3/5	G R S E I A F V A T G T N L S L Q F F P A S W Q [G] E Q R (Q)	+	+
4a	E Y V D F E R	+	+
4b	[R] - (N) P Q E A G K	+	+

Peptide 2	V V A L G D V P D G T L V T V M A G N D E N Y
AML1	V V A L G D V P D G T L V T V M A G N D E N Y
runt	V I A L D D W P D G T L V S I K C G N D E N Y

FIG. 3. Sequences of CBF p24 and p19 tryptic peptides. Peptide numbers correspond to peaks in Fig. 2. The sequences of the three peptides from peaks 3 and 5 are combined. Peak 5 from p19 contained two peptides. One peptide in peak 5 is an incomplete trypsin cleavage product, containing a Gly-Arg in the first two positions. The sequences of the other peptide in peak 5 and the peptide in peak 3 from p24 are identical and start at the third amino acid (Ser). An amino acid in brackets indicates an interpretation with reasonable confidence, and an amino acid in parentheses indicates a low-confidence result. A hyphen indicates that no sequence determination could be made at that position. The sequence of peptide 2 is compared with those of the corresponding amino acid sequences of the human AML1 and *Drosophila* runt proteins below. The sequence of AML1 is from aa 91 to 113 (39), and the runt sequence corresponds to aa 146 to 168 (30). Amino acids that are identical in runt and peptide 2 are underlined.

consensus ATP or GTP binding domain (13). Peptide 2 from p24 is contained within this 118-aa *runt/AML1* homologous region. A tryptic peptide from p19 elutes at the same position as does peptide 2 from p24 (Fig. 2), suggesting that the *runt/AML1* homologous peptide is present in both p24 and p19. However, the amino acid sequence of the analogous peptide from p19 was not determined; therefore, we did not conclusively demonstrate that the *runt/AML1* homologous peptide is present in the p19 polypeptide band.

Molecular cloning and sequence analysis of the CBF β subunits. Two PCR primers deduced from the amino acid sequence of peptide 3/5, which was the longest peptide with the lowest degeneracy, were used to amplify the sequence between the primers from template DNA prepared from a mouse thymus cDNA library. Two 25-mers were designed from the amplified 50-bp sequence and used as probes to screen the mouse thymus cDNA library. One cDNA clone that hybridized with both 25-mers was isolated from the cDNA library and found to contain a 1.4-kb insert. Preliminary DNA sequence analysis indicated that this cDNA clone encoded peptide 3/5 but lacked a complete 5′ end. Therefore, using the 1.4-kb insert as a probe, we isolated 11 additional clones from the mouse thymus cDNA library. The complete nucleotide sequence of one of these cDNA clones, CBF β p21.5, is presented in Fig. 4A. The structures of all three cDNA clones are diagrammed schematically in Fig. 4B, and the deduced amino acid sequences of the three proteins are presented in Fig. 4C.

The CBF β p21.5 cDNA clone contains a 2,742-bp insert and a long open reading frame encoding a polypeptide of 182 aa with a calculated molecular weight (M_r) of 21,500 (Fig. 4A). The calculated molecular weight is in good agreement with the molecular weights of the purified bovine CBF polypeptides, p24 and p19. Coding sequences for two tryptic peptides, 3/5 and 4a, are contained within the long open reading frame. The coding sequence for tryptic peptide 1 from bovine p24, which was the difference peptide not present in p19, is contained in CBF β p21.5 but is in the +1 reading frame relative to the coding sequences for peptides 3/5 and 4a. CBF β p21.5 contains a 1,997-nucleotide 3′ untranslated region, ending in a poly(A) tail. A consensus polyadenylation sequence AATAAA is not found, but a related sequence ATTAAA is located 13 nucleotides from the poly(A) tail.

The nucleotide sequence of the cDNA clone CBF β p22.0 predicts a long open reading frame encoding a polypeptide of 187 aa (M_r = 22,000). The sequence of CBF β p22.0 contains a 1,319-bp 5′ untranslated region that starts with a stretch of thymidines and contains a site for *Eco*RI restriction endonuclease cleavage at nucleotide 584 (data not shown). Since the mouse thymus cDNA library was prepared by using *Eco*RI linkers, we suspect that the insert contains two unrelated cDNAs fused head to head. Sequences between the 5′ end of the insert and the *Eco*RI site at nucleotide 584 contain a 412-nucleotide-long open reading frame (in reverse orientation), supporting the hypothesis that the 5′ end of the insert is derived from an unrelated cDNA. We cannot, therefore, determine the bona fide 5′ end of the CBF β p22.0 cDNA and have numbered the CBF β p22.0 sequence in Fig. 4B starting at the putative ATG translation start codon. The 3′ untranslated sequence of CBF β p22.0 diverges from that of CBF β p21.5 starting at nucleotide 1192 (corresponding to nucleotide 1304 in CBF β p21.5) and contains 331 nucleotides of nonhomologous sequence.

Coding sequences for three tryptic peptides (peptides 1, 3/5, and 4a) are contained within the long open reading frame of CBF β p22.0 (Fig. 4C). A 31-bp deletion near the 3′ end of the long open reading frame of CBF β p22.0 (relative to CBF β p21.5) brings the coding sequences for peptide 1 into the same reading frame as the coding sequences for peptides 3/5 and 4a. This is consistent with the presence of peptide 1 in the larger 24-kDa bovine protein. We identified three other independent cDNA clones that had the same 31-bp deletion and a total of eight that did not (data not shown). Thus, we conclude that the 31-bp difference between CBF β p22.0 and CBF β p21.5 results from alternative splicing and is not a cloning artifact. The coding sequence of CBF β p21.5 at the 5′ endpoint of the 31-bp deletion contains a characteristic donor splice site (40) (Fig. 4B), suggesting that it serves as an alternative splice donor to yield the CBF β p22.0 transcript. Isolation and characterization of the genomic clone of CBF β will be necessary to confirm or disprove this hypothesis.

CBF β p17.6 encodes a 148-aa protein with an M_r of 17,600. CBF β p17.6 contains the same 31-bp deletion near

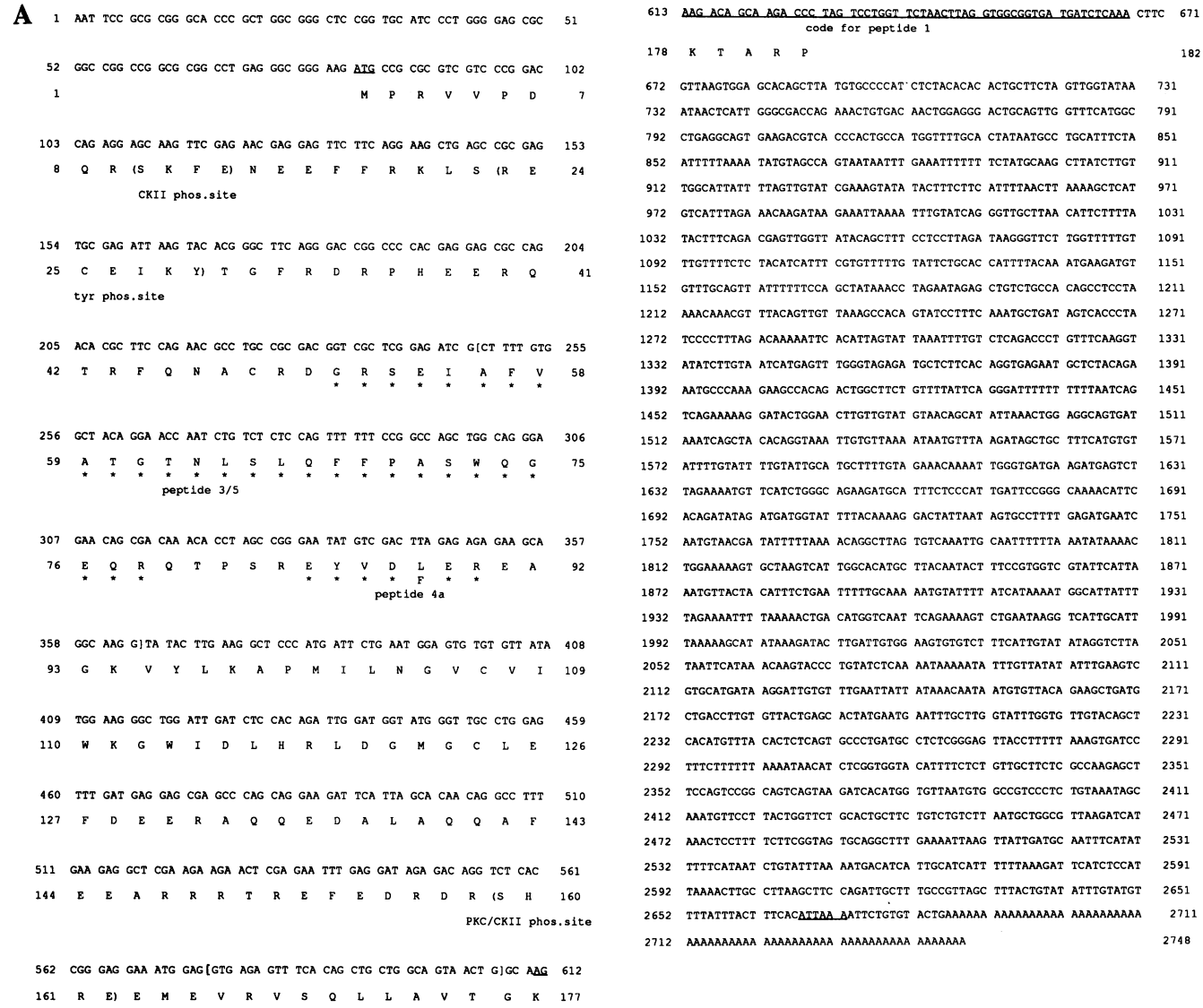
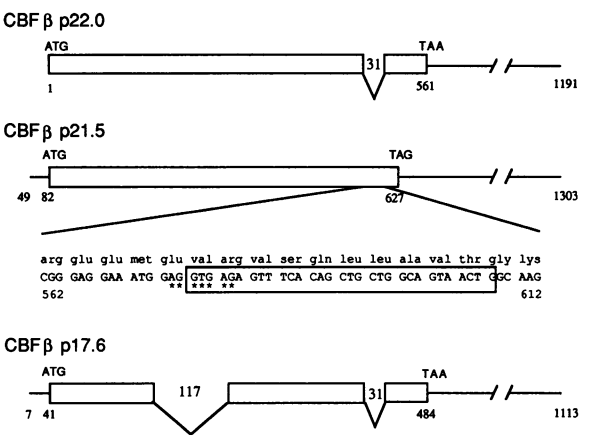


FIG. 4. Nucleotide and amino acid sequences of CBF β subunits. (A) Nucleotide sequence of CBF β p21.5, with the deduced amino acid sequence (single-letter code) in the major long open reading frame. The putative ATG start codon is underlined. The sequences corresponding to bovine peptides 3/5 and 4a are shown beneath the deduced amino acid sequence. Positions of amino acid sequence identity with the bovine peptides are indicated by asterisks; positions of amino acid sequence divergence are indicated by the substituted amino acid. Positions of sequences corresponding to peptide 1 (in the +1 reading frame) are underlined and noted. The putative polyadenylation sequence (ATTAAA) is also underlined. The boundaries of the 117- and 31-bp deletions in CBF β p22.0 and CBF β p17.6 (nucleotides 248 to 364 and 577 to 607) are indicated by brackets in the nucleotide sequence. The partial 1.4-kb cDNA clone that was first isolated contains sequences from nucleotides 139 to 1463 of CBF β p21.5. Putative casein kinase II (CKII), protein kinase C (PKC), and tyrosine (tyr) phosphorylation (phos.) sites are noted. (B) Schematic diagram showing the relationship between the CBF β p21.5, CBF β p22.0, and CBF β p17.6 cDNA clones (GenBank accession numbers L03279, L03306, and L03305, respectively). Only the overlapping regions of the cDNA clones are shown. Numbers correspond to the nucleotide sequence of each cDNA clone as entered into GenBank. Since the 5′ end of CBF β p22.0 appears to include sequences from another unrelated cDNA resulting from head-to-head fusion, we cannot determine the bona fide 5′ end of the CBF β p22.0 cDNA. Therefore, we have numbered the CBF β p22.0 sequences starting at the putative ATG start codon. White rectangles indicate long open reading frames. ATG and TAA or TAG translation start and stop codons are shown. Positions of the 31- and 117-bp deletions in CBF β p22.0 and CBF β p17.6 are indicated by jagged lines. The nucleotide sequence including and immediately adjacent to the site of the 31-bp deletion found in both CBF β p22.0 and CBF β p17.6 is shown below the CBF β p21.5 diagram. The 31 nucleotides that are deleted are boxed. Asterisks indicate the residues at the 5′ deletion endpoint that are conserved in typical splice donor sequences (40). (C) Deduced amino acid sequences of the proteins encoded by the CBF β cDNA clones. Numbers correspond to the CBF β p22.0 amino acid sequence. Amino acids corresponding to the bovine tryptic peptides are underlined.

the 3′ end of the open reading frame found in CBF β p22.0 and an additional 117-bp deletion near the middle of the open reading frame (Fig. 4B). This 117-bp deletion removes coding sequences for most of peptide 3/5 and all of peptide 4a (Fig. 4C). CBF β p17.6 was the only cDNA that we isolated which contains this deletion, and thus the deletion could be an artifact of cloning. The 3′ untranslated region of CBF β p17.6 is identical to that of CBF β p21.5.



C

```
                                    1                                  50
CBF β p22.0    MPRVVPDQRS KFENEEFFRK LSRECEIKYT GFRDRPHEER QTRFQNACRD
CBF β p17.6    MPRVVPDQRS KFENEEFFRK LSRECEIKYT GFRDRPHEER QTRFQNACRD
CBF β p21.5    MPRVVPDQRS KFENEEFFRK LSRECEIKYT GFRDRPHEER QTRFQNACRD

                                   51          peptide 3/5          peptide 4a  100
CBF β p22.0    GRSEIAFVAT GTNLSLQFFP ASWQGEQRQT PSREYVDLER EAGKVYLKAP
CBF β p17.6    GRSEI..... .......... .......... .......... ....VYLKAP
CBF β p21.5    GRSEIAFVAT GTNLSLQFFP ASWQGEQRQT PSREYVDLER EAGKVYLKAP

                                   101                                 150
CBF β p22.0    MILNGVCVIW KGWIDLHRLD GMGCLEFDEE RAQQEDALAQ QAFEEARRRT
CBF β p17.6    MILNGVCVIW KGWIDLHRLD GMGCLEFDEE RAQQEDALAQ QAFEEARRRT
CBF β p21.5    MILNGVCVIW KGWIDLHRLD GMGCLEFDEE RAQQEDALAQ QAFEEARRRT

                                   151                peptide 1     187
CBF β p22.0    REFEDRDRSH REEMEARRQQ DPSPGSNLGG GDDLKLR
CBF β p17.6    REFEDRDRSH REEMEARRQQ NPSPGSNLGG GDDLKLR
CBF β p21.5    REFEDRDRSH REEMEVRVSQ LLAVTGKKTA RP.....

               CBF β p22.0: 187 aa.; 22.0 kDa
               CBF β p17.6: 148 aa.; 17.6 kDa
               CBF β p21.5: 182 aa.; 21.5 kDa
```

FIG. 4—*Continued.*

Although not shown in Fig. 4A, the long open reading frame of CBF β p21.5 (and CBF β p17.6) continues to the 5′ end of the insert. We have tentatively designated the ATG at position 82 in CBF β p21.5 as the initiation codon, based on the sequence of the CBF β p22.0 cDNA clone, in which the open reading frame ends 696 nucleotides 5′ to this ATG (3′ to the internal *Eco*RI site) and does not contain another ATG. The putative ATG initiation codon is in the context of an imperfect Kozak consensus sequence for translational initiation, in that a C at the +4 position is not favorable (AAGATGC) (32).

The proteins encoded by the CBF β cDNA clones do not contain any DNA-binding motifs commonly found in DNA-binding proteins (for a review, see reference 26), nor is there a predominantly basic region in the protein suggestive of a DNA-binding domain. There are also no acidic, glutamine-rich, or proline-rich regions that are characteristic of activation domains. The CBF β proteins do not contain the prototypic nuclear targeting sequence first found in the simian virus 40 large T antigen (28); however, a potential bipartite nuclear targeting sequence is present at aa 19 to 35 (**RKLSRECEIKYTGFRDR**) (46). Both the protein and nucleotide sequences of CBF β p22.0 were used to search computer data bases for other related sequences (3). No significant homology to any other protein was revealed in the search.

The coding sequence for peptide 2, which is the *runt/AML1* homologous peptide, is not contained within any of the CBF β cDNA clones. Therefore, at least two proteins migrated in the SDS-polyacrylamide gel from which the bovine proteins were purified with an apparent molecular size of 24 kDa (and probably also 19 kDa) and were excised from what had appeared to be a single band. The two proteins generated a mixture of tryptic peptides. One protein is encoded by the CBF β gene, and it contributed peptides 1, 3/5, and 4a. We believe that the protein that comigrated with the CBF β-encoded protein is the bovine homolog of human *AML1*, and it contributed peptide 2. We could not find coding sequences for the short peptide 4b either in the mouse CBF β cDNA clones or in the published human *AML1* sequence (39).

The CBF β subunits form a complex with DNA-binding CBF α subunits. We subcloned the open reading frames of CBF β p22.0, CBF β p21.5, and CBF β p17.6 into the GST vector pGEX2T and expressed the CBF β proteins as GST fusion proteins. The GST fusion proteins were purified on glutathione-agarose beads, and fractions of the eluate were analyzed by SDS-polyacrylamide gel electrophoresis (PAGE) (data not shown). The fusion proteins migrated with apparent molecular sizes of 49 kDa (CBF β p22.0-GST), 47 kDa (CBF β p21.5-GST), and 44.5 kDa (CBF β p17.6-GST), consistent with their predicted molecular sizes plus that of the 27-kDa GST domain.

We tested whether the CBF β-GST fusion proteins would bind to the core site but could detect no protein-DNA complexes by electrophoretic mobility shift assay. We then examined whether the CBF β proteins are part of a protein complex that binds to the core site and purified the DNA-binding component of CBF from affinity-purified bovine CBF to test this hypothesis. Bovine CBF was fractionated through an SDS-polyacrylamide gel, the gel was stained with Coomassie brilliant blue, and each visible protein band was excised from the gel (Fig. 5A). Proteins were eluted from each gel slice (A to N), renatured, and tested for core-binding activity (Fig. 5B).

As shown previously, each protein in the molecular size range of 17 to 33 kDa binds to the core site (60). Several proteins (B to E and N) give rise to one protein-DNA complex. Proteins migrating in the molecular size range of 19 to 27 kDa (F to M) yield two protein-DNA complexes. We noted this phenomenon previously and at that time hypothesized that the second protein-DNA complex with lower mobility contains two CBF subunits bound to the probe (60). As we will demonstrate later, we now believe that the fast-mobility protein-DNA complex contains only a DNA-binding subunit of CBF (which we call CBF α subunits) and that the lower-mobility protein-DNA complex contains both a DNA-binding subunit of CBF (CBF α) and a non-DNA-binding subunit of CBF (CBF β). The large molecular size range of the CBF α subunits appears to be contributed by a combination of proteolysis and splicing variants (4, 12, 25, 39). The hypothesis that the CBF β cDNA clones that we isolated encode non-DNA-binding subunits of CBF is consistent with the 19- to 27-kDa molecular size range in which the bovine proteins that yield two protein-DNA complexes migrate, which corresponds to the predicted size of several of the CBF β subunits. A range of 19 to 27 kDa is also reasonably consistent with the size range of the core-binding PEBP2 β subunits (20 to 24 kDa) purified by Kamachi et al. from c-Harvey-*ras*-transformed mouse fibroblasts (29). Therefore, we reasoned that (i) gel slices that give rise to one protein-DNA complex (B to E and N) contain only a

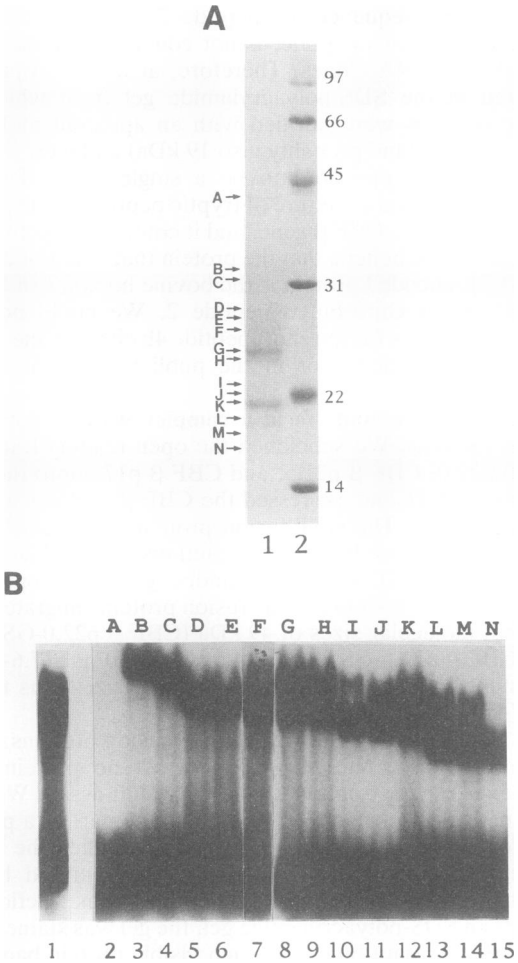


FIG. 5. Recovery of DNA-binding bovine CBF α subunits from SDS-polyacrylamide gels. (A) Fractionation of polypeptides by SDS-PAGE. Shown is the Coomassie brilliant blue-stained SDS-polyacrylamide gel from which the different polypeptides were excised. The CBF preparation is the same as that used to obtain purified bovine p24 and p19. Five hundred microliters of the affinity-purified fraction was trichloroacetic acid precipitated and electrophoresed through an SDS–15% polyacrylamide gel, and the gel was stained with Coomassie brilliant blue. Lanes: 1, affinity-purified CBF; 2, low-molecular-weight protein standards (positions indicated in kilodaltons). Arrows indicate the bands that were excised from the Coomassie brilliant blue-stained SDS-polyacrylamide gel. The CBF polypeptide bands are labeled sequentially by letters from top to bottom (A to N). Polypeptides G and K are p24 and p19, respectively. (B) Electrophoretic mobility shift assay of core-binding activity from renatured polypeptides. The excised polypeptides were eluted from the SDS-gel slices and subjected to a denaturation/renaturation regimen as described previously (19, 60). Ten microliters of each renatured protein was assayed for core-binding activity, using 20,000 cpm of the HA probe. The letter above each lane corresponds to the letter of the polypeptide band in panel A. Lane 1 contains 2 μl of the affinity-purified CBF sample prior to fractionation by SDS-PAGE.

DNA-binding subunit of CBF (CBF α), (ii) gel slices that give rise to two protein-DNA complexes (F to M) contain both a DNA-binding CBF α subunit and a CBF β subunit, and (iii) protein-DNA complexes containing both CBF α and CBF β subunits migrate more slowly on native polyacrylamide gels than do protein-DNA complexes containing CBF α subunits alone.

To test the hypothesis that proteins encoded by the CBF β cDNA clones are the non-DNA-binding subunits of CBF, we tested the ability of the CBF β-GST fusion proteins to shift the mobility of protein-DNA complexes formed with bovine CBF α subunits (polypeptides B to E and N in Fig. 5). CBF β p22.0-GST, CBF β p21.5-GST, and CBF β p17.6-GST alone do not bind to the core site (Fig. 6A, lanes 6, 11, and 15). However, when included in binding reactions with renatured bovine CBF α subunits, both CBF β p22.0-GST and CBF β p21.5-GST decrease the mobility of the protein-DNA complex. An approximately 10:1 molar ratio of CBF β to CBF α drives 50% of the CBF protein-DNA complex into the α+β form (lanes 3 and 8). We found that all of the renatured bovine CBF α polypeptides (B to E and N) associate with the CBF β p22.0-GST and CBF β p21.5-GST subunits (data for all CBF α subunits not shown). Furthermore, we find that when CBF β p22.0-GST and CBF β p21.5-GST are combined with CBF α polypeptides that yield two protein-DNA complexes (F to M), the higher-mobility CBF α protein-DNA complex is driven into a lower-mobility form (data not shown). The CBF β subunits do not, however, associate with an unrelated protein Ets-1 to form a lower-mobility protein-DNA complex (18).

CBF β p17.6-GST cannot participate in the formation of a protein-DNA complex with CBF α subunits (Fig. 6A, lanes 12 to 14). We determined whether CBF β p17.6 can interfere with the association of the CBF β p22.0 and CBF β p21.5 subunits with CBF α subunits. If, for example, the association of CBF β subunits to the CBF α subunit was dependent on CBF β dimerization, then CBF β p17.6 could potentially dimerize with CBF β p22.0 and CBF β p21.5 and inhibit their association to the CBF α subunit. Mixing experiments were performed in which a 1:1 or 10:1 molar ratio of CBF β p17.6-GST to either CBF β p21.5-GST or CBF β p22.0-GST was added to binding reactions containing CBF α subunits (Fig. 6B). CBF β p17.6-GST did not interfere with the association of CBF β p21.5 or CBF β p22.0 to the CBF α protein-DNA complex even when present in a 10-fold molar excess. One caution in interpreting this result, however, is that CBF β dimers may exchange subunits too slowly for significant exchange to occur during the binding reaction.

Association of CBF β subunits stabilizes DNA binding by CBF α subunits. DNase I footprinting analyses were performed to determine whether the association of CBF β subunits qualitatively alters the footprint generated by the CBF α subunits. We found, however, that CBF α by itself does not generate a footprint on the core sites in the TCRγ enhancer (Fig. 7). Increasing the amount of CBF α in the binding reactions resulted in either nonspecific binding or repression of DNase I cleavage (Fig. 7A, lane 4). Upon addition of either CBF β p22.0-GST or CBF β p21.5-GST (but not CBF β p17.6-GST or GST alone), a clear footprint appears over the two core sites in the TCRγ enhancer. Association of the CBF β subunits therefore appears to increase the stability of the CBF protein-DNA complex.

The observation that CBF α binding can be detected by electrophoretic mobility shift but not by DNase I footprinting assays suggests that CBF α dissociates from the DNA during the 1-min incubation with DNase I. To determine whether this is indeed the case and whether the association of CBF β decreases the rate of dissociation, we measured the dissociation rates of the CBF α protein-DNA complex and the CBF α+β protein-DNA complex by electrophoretic mobility shift assays (Fig. 8). The dissociation of both CBF α alone and CBF α+β (CBF β p21.5) from the HA probe was too rapid to measure when both the binding reactions and

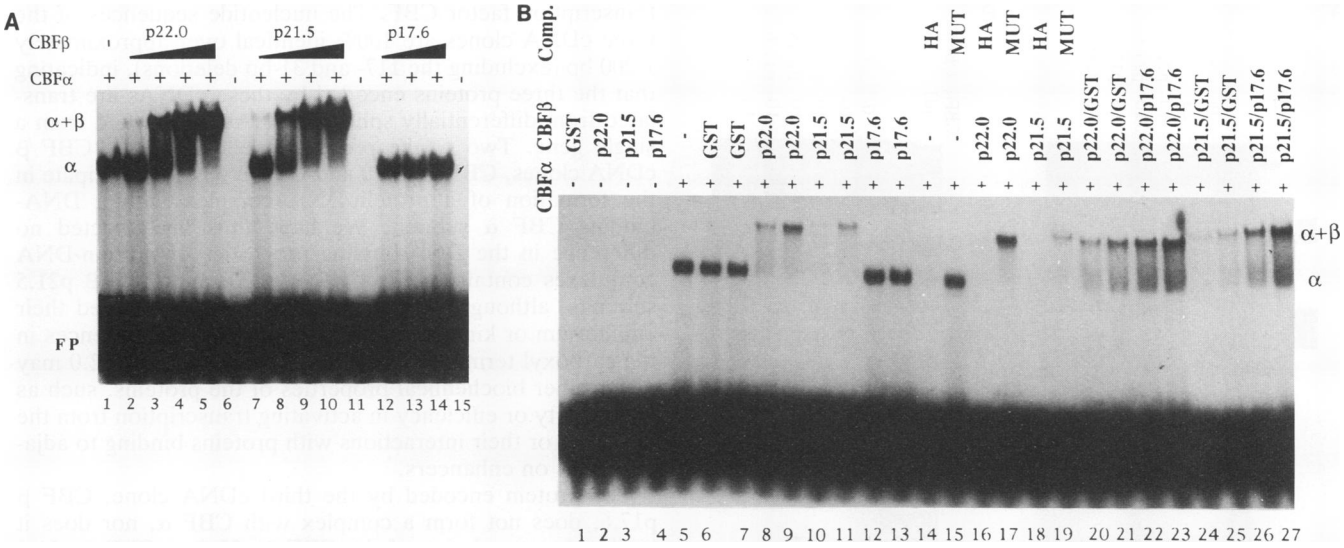


FIG. 6. Association of CBF β-GST fusion proteins with DNA-binding CBF α subunits. (A) Various amounts of CBF β-GST fusion proteins were added to 4 μl (approximately 1 to 5 ng; 2×10^{-9} M to 6×10^{-9} M, final concentration) of the CBF α subunit recovered from the SDS-polyacrylamide gel (polypeptide C from Fig. 5), and binding was measured by electrophoretic mobility shift assay using the HA probe. Lanes: 1, binding of the CBF α subunit alone; 2 to 5, CBF α plus 5 ng, 50 ng, 500 ng, and 2.5 μg of CBF β p22.0-GST; 6, 2.5 μg of CBF β p22.0-GST alone; 7 to 10, CBF α plus 5 ng, 50 ng, 500 ng, and 5 μg of CBF β p21.5-GST; 11, 5 μg of CBF β p21.5-GST alone; 12 to 14, CBF α plus 10 ng, 100 ng, and 1 μg of CBF β p17.6-GST; 15, 1 μg of CBF β p17.6-GST alone. FP shows the position of the free probe; α and α+β indicate positions of the protein-DNA complexes formed with the CBF α subunit in the absence and presence of the CBF β-GST fusion proteins, respectively. (B) CBF β p17.6 does not inhibit association of CBF β p21.5 and CBF β p22.0 to the CBF α protein-DNA complex. Lanes: 1, binding of 500 ng of GST alone; 2, 500 ng of CBF β p22.0-GST; 3, 500 ng of CBF β p21.5-GST; 4, 500 ng of CBF β p17.6-GST; 5, binding of 2.5 μl (estimated to be <1 to 5 ng) of renatured bovine CBF α subunit alone (polypeptide C from Fig. 5); 6 and 7, CBF α plus 50 and 500 ng of GST; 8 and 9, CBF α plus 50 and 500 ng of CBF β p22.0-GST; 10 and 11, CBF α plus 50 and 500 ng of CBF β p21.5-GST; 12 and 13, CBF α plus 50 and 500 ng of CBF β p17.6-GST; 14 and 15, competition of CBF α binding by 20 ng (200-fold molar excess) of the unlabeled HA and MUT sites; 16 and 17, competition of CBF α plus 500 ng of CBF β p22.0-GST with HA and MUT sites; 18 and 19, competition of CBF α plus 500 ng of CBF β p21.5 with HA and MUT sites; 20, CBF α plus 50 ng of CBF β p22.0-GST and 50 ng of GST; 21, CBF α plus 50 ng of CBF β p22.0-GST and 500 ng of GST; 22 and 23, CBF α plus 50 ng of CBF β p22.0-GST and 50 (lane 22) and 500 (lane 23) ng of CBF β p17.6-GST; 24 and 25, CBF α plus 50 ng of CBF β p21.5-GST and 50 (lane 24) and 500 (lane 25) ng of GST; 26 and 27, CBF α plus 50 ng of CBF β p21.5-GST and 50 (lane 26) and 500 (lane 27) ng of CBF β p17.6-GST. α and α+β represent the protein-DNA complexes formed with the CBF α subunit in the absence and presence of the CBF β-GST fusion proteins, respectively.

electrophoresis were conducted at RT (data not shown). When the competition was performed at 0°C and the binding reactions were electrophoresed at 4°C, the CBF α protein-DNA complex still dissociated rapidly but at a rate that could be measured experimentally by electrophoretic mobility shift assay (Fig. 8A, lanes 1 to 5). Addition of CBF β p21.5-GST to the binding reactions significantly decreased the rate of dissociation of the CBF protein-DNA complex (Fig. 8A, lanes 6 to 10). We plotted the data as shown in Fig. 8B and from this calculated a dissociation rate constant for the CBF α protein-DNA complex of $8.8 \times 10^{-2}\ s^{-1}$ and for the CBF α+β protein-DNA complex (β = CBF β p21.5-GST) of $1.25 \times 10^{-2}\ s^{-1}$ at 0°C, corresponding to a sevenfold decrease in the dissociation rate constant upon association of CBF β.

We also determined the dissociation rate constants for the two protein-DNA complexes generated from the bovine CBF proteins eluted from gel slices F to M in Fig. 5 (Fig. 8A, lanes 11 to 15; Fig. 8C). The faster-migrating complex which we hypothesized contained only CBF α subunits dissociated at a rate comparable to that of CBF α alone ($7.0 \times 10^{-2}\ s^{-1}$). The protein-DNA complex that migrates more slowly and presumably contains both CBF α and CBF β subunits dissociates more slowly, with a rate constant of 1.6×10^{-2} s^{-1}, supporting the hypothesis that it contains both a DNA-binding α and non-DNA-binding β subunit.

Association of the CBF β subunit does not extend the phosphate contacts made by CBF in the core site. To determine whether the association of CBF β changes the contacts made by the CBF α subunit, we compared the phosphate contacts made by CBF α in the absence and presence of the CBF β subunit. The ethylation interference analysis was performed by using the δE3 core site from the TCRδ enhancer (45). CBF α contacts six phosphates on the top strand and four phosphates on the bottom strand of the δE3 core site (lanes 3 to 5 in Fig. 9A and B). This finding is consistent with the pattern of phosphate contacts made by affinity-purified bovine CBF on the core site in the Moloney MLV enhancer (37). This pattern of phosphate contacts is not extended upon association of the CBF β (CBF β p21.5-GST) subunit (lanes 6 to 8 in Fig. 9A and B); on the contrary, there is some suggestion that the effect of ethylation at some phosphates on the bottom strand is slightly less severe upon association of the CBF β subunit (Fig. 9C). This result suggests that the dissociation rate constant of the CBF protein-DNA complex is not decreased through the establishment of additional contacts between the protein and the DNA upon association of CBF β.

Southern and Northern analyses of the gene for the CBF β subunits. The insert from the partial CBF β p21.5 cDNA clone (nucleotides 139 to 1463 of CBF β p21.5) was used as a probe in Southern blot analysis under stringent hybridization conditions (Fig. 10A). Multiple bands were observed with use of a variety of restriction endonucleases, suggesting

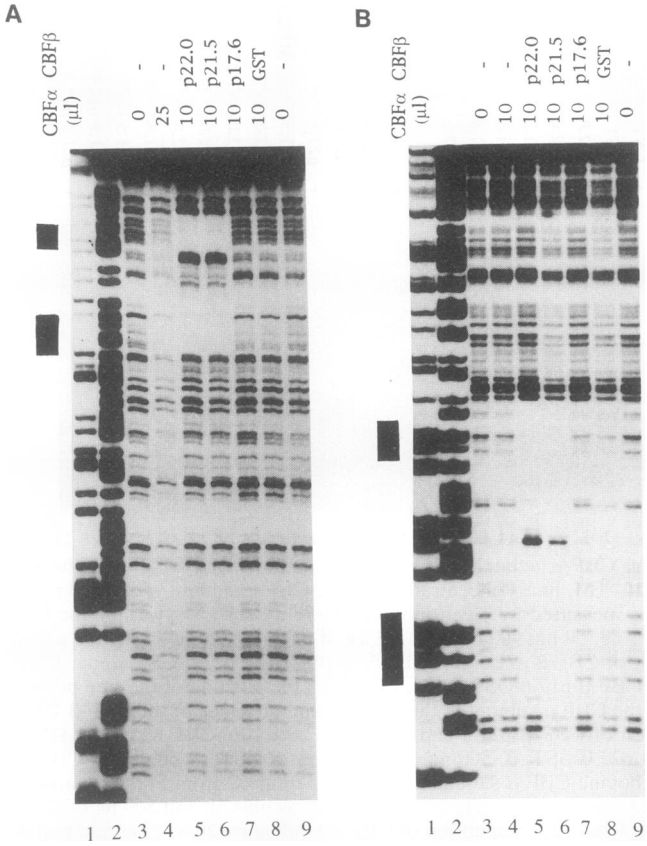


FIG. 7. DNase I footprint analysis of CBF α subunit binding to the enhancer from the TCRγ gene in the presence of CBF β-GST fusion proteins. (A) Binding to the core sites on the top strand of the TCRγ enhancer. Lanes: 1 and 2, G and G+A chemical sequencing ladders; 3, BSA; 4, 25 μl of CBF α (polypeptide D from Fig. 5); 5, 10 μl of CBF α plus 2.5 μg of CBF β p22.0-GST; 6, 10 μl of CBF α plus 2.5 μg of CBF β p21.5-GST; 7, 10 μl of CBF α plus 2.5 μg of CBF β p17.6-GST; 8, 10 μl of CBF α plus 2.5 μg of GST; 9, BSA. Boxes at the left indicate positions of the two core motifs (CTGTG GTTT and CTGTGGTCT) in the TCRγ enhancer. (B) Binding of CBF to the core sites on the bottom strand of the TCRγ enhancer. Lanes are the same as those in panel A except that lane 4 contains 10 μl of CBF α alone. CBF α polypeptide 5C was used in footprinting reactions.

either that the CBF β gene is relatively large or that it is a member of a family of related genes.

We examined the expression of CBF β mRNA in various mouse tissues. CBF β mRNA is present at some level in all tissues tested, although the mRNA is more abundant in thymus than in other tissues (Fig. 10B). The predominant mRNA species in most tissues is 3.1 kb, a less abundant mRNA species in thymus migrates at 2.5 kb, and muscle contains a 2.8-kb CBF β mRNA species not found in other tissues. Differences in the level of CBF β mRNA levels between thymus and other tissues range from 2.6-fold (thymus/muscle) to 29-fold (thymus/testis), with an average difference of 11-fold. When normalized by densitometry to the actin signal, the differences in mRNA levels range from 0.9-fold (thymus/liver) to 6.6-fold (thymus/spleen).

DISCUSSION

We have isolated and sequenced three cDNA clones from mouse thymus encoding non-DNA-binding β subunits of the transcription factor CBF. The nucleotide sequences of the three cDNA clones are 100% identical over approximately 1,200 bp (excluding the 117- and 31-bp deletions), indicating that the three proteins encoded by these cDNAs are translated from differentially spliced mRNAs transcribed from a single gene. Two of the proteins encoded by the CBF β cDNA clones, CBF β p22.0 and CBF β p21.5, participate in the formation of a protein-DNA complex with a DNA-binding CBF α subunit. We have thus far detected no difference in the DNA-binding properties of protein-DNA complexes containing the CBF β p22.0 and CBF β p21.5 subunits, although we have not rigorously compared their equilibrium or kinetic binding constants. The differences in the carboxyl termini of CBF β p21.5 and CBF β p22.0 may affect other biochemical properties of the proteins, such as their ability or efficiency in activating transcription from the promoter or their interactions with proteins binding to adjacent sites on enhancers.

The protein encoded by the third cDNA clone, CBF β p17.6, does not form a complex with CBF α, nor does it inhibit the association of the CBF β p22.0 or CBF β p21.5 subunits to the CBF α protein-DNA complex. CBF β p17.6 may, however, interfere with other biochemical processes, perhaps through squelching of essential components of the transcriptional apparatus.

The association of CBF β p21.5 decreases the rate constant of dissociation of the CBF protein-DNA complex. This decrease in the dissociation rate constant is not accompanied by an extension of the phosphate contacts in the binding site, as evidenced by the similarity of the phosphate interference patterns generated by the CBF α subunit alone and the CBF α+β complex. Two models are compatible with the observations that CBF β stabilizes the protein-DNA complex yet does not extend the phosphate contacts on the DNA. One model is that CBF β stabilizes a conformation assumed by CBF α that is favorable for binding but that CBF β does not itself contact DNA directly. This model presupposes that CBF α binds as a monomer, and therefore that the association of CBF β does not involve exchange of a DNA-binding CBF α subunit. An alternative model is that CBF β does displace a subunit of CBF α and contacts the DNA directly but that the CBF α+β heterodimer establishes essentially the same contacts as CBF α alone.

We predict that CBF α binds as a monomer, that the association of CBF β to CBF α does not involve a subunit exchange, and therefore that the first model (CBF β stabilizes a conformation assumed by CBF α) is correct. The reasons for our prediction are as follows. (i) Association of CBF β to the protein-DNA complex decreases the mobility of the protein-DNA complex relative to CBF α alone, suggesting that rather than displacing a subunit, CBF β piggybacks onto the complex. For example, in Fig. 5, when proteins from individual polypeptide bands excised from an SDS-polyacrylamide gel which contain both CBF α and CBF β subunits of identical molecular weights are assayed for binding, the association of CBF β decreases the mobility of the protein-DNA complex. If CBF β were instead displacing one subunit of a homodimeric CBF α complex, the mobility of the protein-DNA complex should not change significantly. (ii) Results from selected and amplified binding sequence footprinting (6) indicate that the binding site selected by the CBF α subunit alone is asymmetrical (PyGPyGGTPy), suggesting that it binds as a monomer (37, 38). Furthermore, the distributions of binding sites selected by CBF α and CBF α+β subunits are indistinguishable; thus, association of the CBF β subunit does not appear to alter sequence specificity

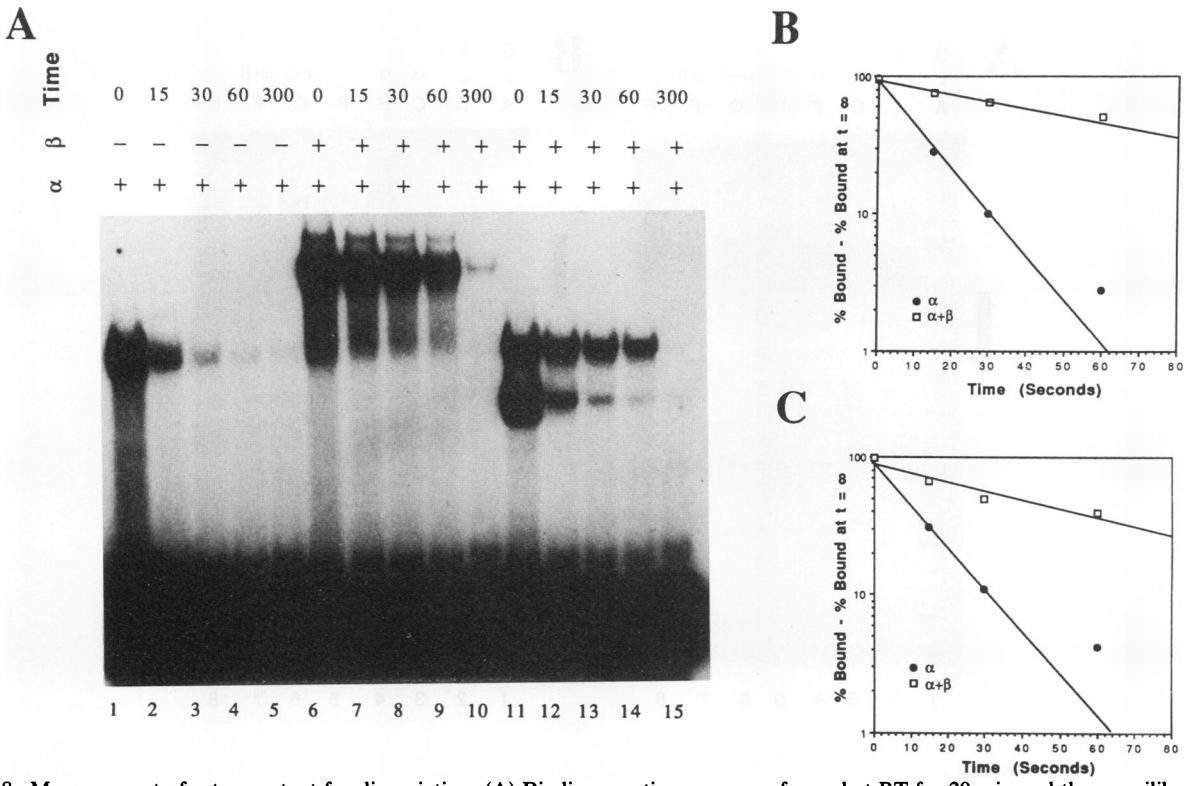


FIG. 8. Measurement of rate constant for dissociation. (A) Binding reactions were performed at RT for 20 min and then equilibrated on ice for 15 min. A 50-fold molar excess of HA competitor was added to trap the dissociated protein, and 15-μl aliquots were removed at the indicated time points and loaded onto a running gel (equilibrated at 4°C). Lanes: 1 to 5, CBF α (polypeptide B) at t = 0, 15, 30, 60, and 300 s; 6 to 10, CBF α plus CBF β p21.5-GST at t = 0, 15, 30, 60, and 300 s; 11 to 15, bovine CBF α+β (polypeptide band J) at t = 0, 15, 30, 60, and 300 s. The mobility of the CBF α+β protein-DNA complex in lanes 6 to 10 is lower than that in lanes 11 to 15, presumably because of the added molecular weight of the GST domain in the CBF β-GST fusion protein. (B and C) Plots of log(% bound at each time [relative to t = 0] − % bound at t = ∞) versus time (seconds) for the bovine CBF α protein-DNA complex (data from lanes 1 to 5) and bovine CBF α plus CBF β p21.5-GST protein-DNA complex (data from lanes 6 to 10) (B) and for bovine CBF α and bovine CBF α+β protein-DNA complexes from polypeptide band J (data from lanes 11 to 15) (C).

(38). We therefore favor the hypothesis that CBF α binds as a monomer and that CBF β does not contact DNA directly.

CBF is a new member of a relatively small group of transcription factor complexes shown to be composed of a DNA-binding subunit(s) associated with a non-DNA-binding subunit, in which the non-DNA-binding subunit increases the affinity of the protein complex for the binding site. Another member of this group is the transcription factor AP-1, which consists of a DNA-binding subunit, Jun, in a heteromeric complex with a Fos subunit (Fos-Fos homodimers cannot bind DNA). DNA-binding Jun-Jun homodimers dissociate from the DNA at a faster rate than do Fos-Jun heterodimers (43). The DNA-binding domain of the Fos-Jun heterodimer is contributed by both Fos and Jun, and both proteins have been demonstrated to contact the DNA directly (1). Jun-Jun homodimers and Jun-Fos heterodimers make the same contacts on the DNA (as assayed by DNase I footprinting) (1), although phasing analysis demonstrated that Jun-Jun and Fos-Jun bend DNA in opposite directions (31). We predict that the mode of DNA binding by CBF is significantly different from that of Fos and Jun because we believe that the association of CBF β to CBF α does not involve a subunit exchange.

Another example of a protein complex consisting of DNA-binding and non-DNA-binding subunits is the yeast transcription factor complex HAP2/3/4. HAP2/3/4 consists of three nonidentical subunits, all of which are required for DNA binding and transcriptional activity (16, 41). The DNA-binding component of the complex is believed to be contributed by the HAP2 and HAP3 proteins, and the transactivation function is provided predominantly by HAP4 (41). However, CBF α can bind DNA by itself, whereas no detectable DNA binding of the DNA-binding subunits HAP2/3 occurs in the absence of HAP4 (16).

A third example of a transcription factor associated with a non-DNA-binding subunit is the COUP (chicken ovalbumin upstream promoter) protein, which is a member of the steroid hormone receptor superfamily. The binding of COUP to its recognition site is stabilized by the association of a non-DNA-binding factor S300-II (58). Ethylation and methylation interference analysis detected no alterations in the contacts made by COUP and COUP plus S300-II on the DNA (58). S300-II was recently cloned and found to be identical to the basal transcription factor TFIIB (23), raising the possibility that CBF β could also be a general transcription factor.

We sequenced five tryptic peptides from the bovine CBF p24 and p19 proteins, only three of which are encoded by the CBF β cDNA clones. One of the remaining peptides (peptide 2) is from the bovine homolog of a protein encoded by the human *AML1* proto-oncogene. We believe that the *AML1* gene encodes a DNA-binding α subunit of CBF. This hypothesis stems from several observations. (i) The bovine proteins renatured from gel slices had core-binding activity

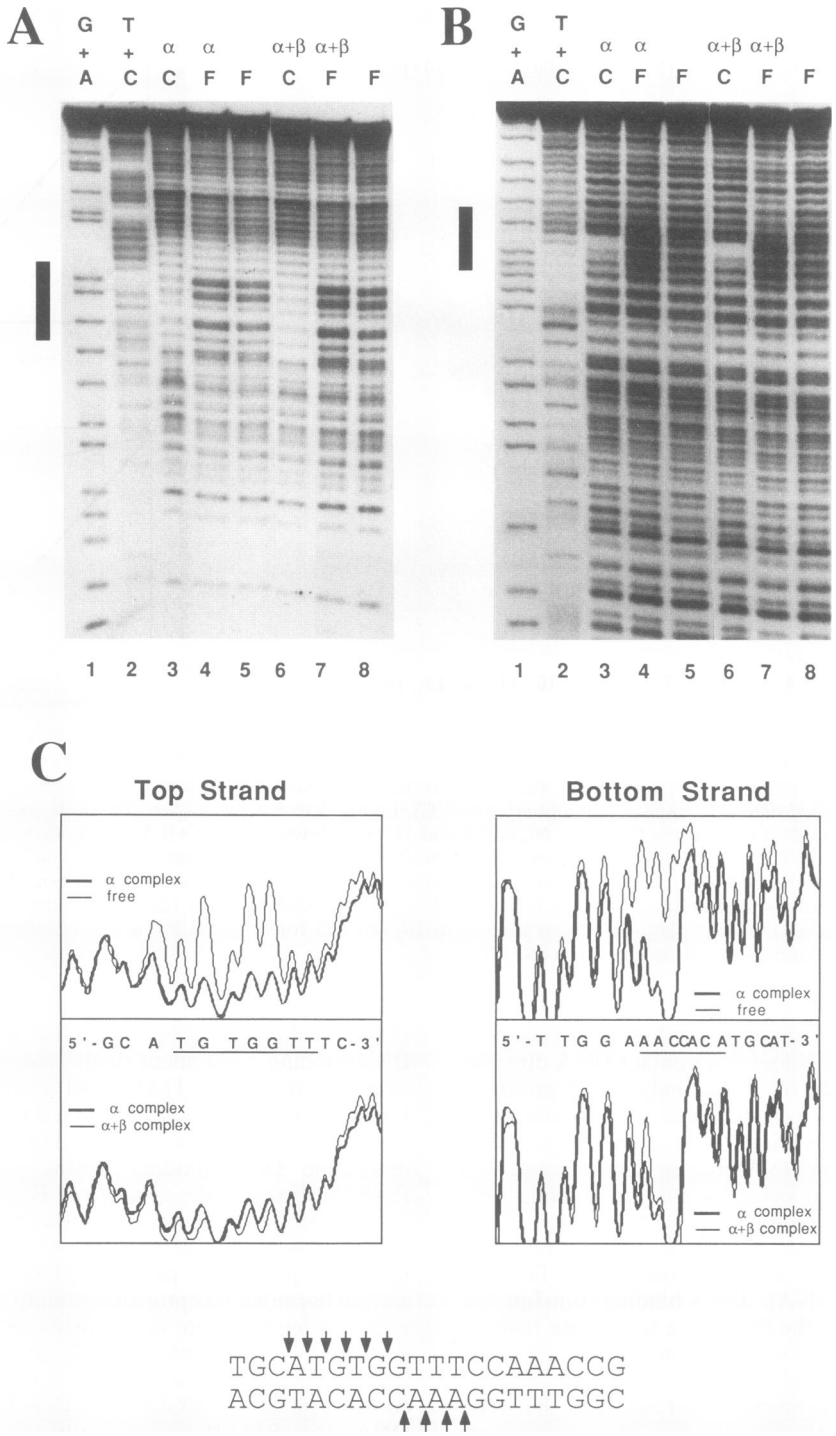


FIG. 9. Ethylation interference. (A) Ethylation interference of CBF on the top strand of the δE3 site (45) from the TCRδ enhancer. Lanes: 1, G+A chemical sequencing tract; 2, T+C; 3, CBF α protein-DNA complex from the mobility shift gel; 4, free DNA band from the CBF α lane on the mobility shift gel; 5, free DNA; 6, CBF α plus CBF β p21.5-GST protein-DNA complex from the mobility shift gel; 7, free DNA band from the CBF α-plus-CBF β p21.5-GST lane on the mobility shift gel; 8, free DNA. Vertical rectangles indicate the location of the core site (TGTGGTT). (B) Ethylation interference of CBF on the bottom strand of the δE3 site from the TCRδ enhancer. Lanes correspond to those in panel A. (C) Analysis of ethylation interference results. On the top are densitometric analyses of autoradiographs. The graphs show traces of the autoradiographic densities of the region immediately surrounding the δE3 site from the top and bottom strands of the probe, as indicated. The top graphs compare the tracing of the CBF α protein-DNA complex (lanes 3 from panels A and B) with the tracing from the free probe (lanes 5 from panels A and B). The bottom graphs compare the tracing from the CBF α protein-DNA complex (lanes 3 from panels A and B) with the tracing from the CBF α–CBF β p21.5-GST protein-DNA complex (lanes 6 from panels A and B). The *x* coordinate is named by the nucleotide whose 5′ phosphate is ethylated. A summary of the ethylation interference results is illustrated on the bottom. The phosphate contacts are shown by arrows above and below the sequence.

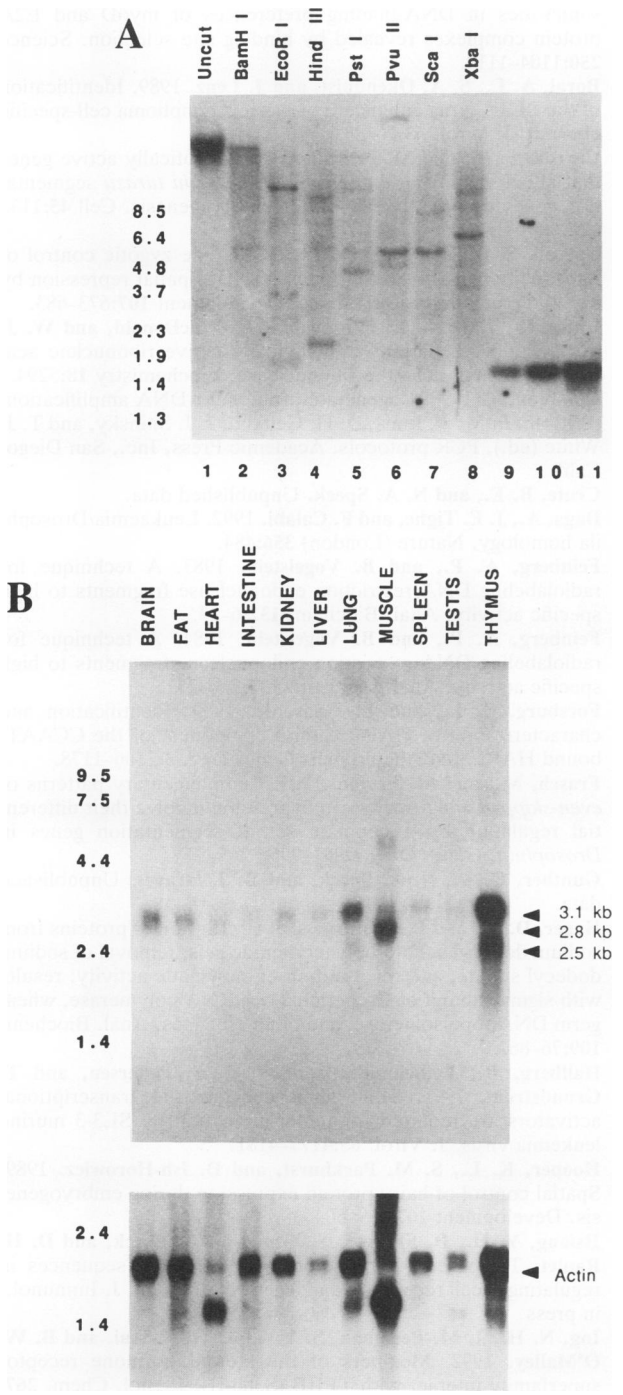


FIG. 10. Southern and Northern analyses. (A) Southern analysis. Lanes: 1, uncut DNA; 2, *Bam*HI-digested DNA; 3, *Eco*RI-digested DNA; 4, *Hin*dIII-digested DNA; 5, *Pst*I-digested DNA; 6, *Pvu*II-digested DNA; 7, *Sca*I-digested DNA; 8, *Xba*I-digested DNA; 9 to 11, 2, 5, and 10 copies of the 1.4-kb partial CBF β p21.5 cDNA clone (encompassing nucleotides 139 to 1463 of CBF β p21.5). Molecular weight markers are shown on the left in kilobases. (B) Northern analysis of Poly(A)$^+$ RNA hybridized with the 1.4-kb insert from the partial CBF β p21.5 cDNA clone. The tissue source of the mRNA is indicated above each lane. Molecular weight markers are shown on the left in kilobases. Positions of the 3.1-, 2.8-, and 2.5-kb bands are noted. The gel at the bottom shows a smaller region of the same blot probed for actin.

and thus contained a DNA-binding α subunit. The DNA-binding subunit must be present in the p24 and p19 bands in a stoichiometric ratio of at least 1:10 relative to CBF β, since the proteins eluted from the p24 and p19 bands yielded an equal distribution of α and α+β protein-DNA complexes, and we know from other experiments that an approximately 10-fold molar excess of CBF β will drive 50% of the α protein-DNA complexes into the α+β form. The molar ratio of the *AML1/runt* homologous peptide 2 relative to peptides 1 and 3 from p24 was approximately 1:5, reasonably consistent with the expected stoichiometric ratio. A tryptic peptide from the mixture of proteins that migrated at 19 kDa had the same retention time on the HPLC column as did peptide 2 from p24 and is thus also likely to be the same *AML1/runt* homologous peptide. The *AML1* gene gives rise to multiply spliced transcripts which could encode proteins of different sizes (4, 12, 39). This could explain why the peptide encoded by the *AML1* gene might be present in proteins migrating at two different molecular sizes (24 and 19 kDa). The observation that the p24 and p19 CBF β proteins were the predominant proteins in the affinity-purified CBF preparation could result from relatively few CBF β subunits being associated with a much larger number of CBF α subunits, each of which is present in the affinity-purified preparation in smaller amounts. (ii) AML1 and CBF β remained associated through multiple purification steps, including a core site affinity column. This finding suggests that AML1 and the CBF β subunits are tightly associated, that the copurification of CBF β and AML1 was not merely coincidental, and that both proteins are part of the protein complex that binds the core site. (iii) AML1 is 92% homologous to another core-binding α subunit, the PEBP2 α subunit, through the 118-amino-acid region that is also conserved between the human *AML1*- and the *Drosophila runt*-encoded proteins and contains the peptide 2 sequence (25). PEBP2 was purified from c-Harvey-*ras*-transformed mouse fibroblasts on the basis of its ability to bind two core sequences in the polyomavirus enhancer (TGTGGTTT and TGCGGTC) (29). As with CBF, purified preparations of PEBP2 were shown to contain multiple polypeptides consisting of DNA-binding α subunits and non-DNA-binding β subunits (25, 29). The PEBP2 β subunits are identical to CBF β (25).

Since the *AML1/runt* homologous peptide 2 is from a bovine protein and the *AML1* and PEBP2 α cDNA clones were isolated from human and mouse cells, respectively, there was some question of whether peptide 2 is derived from the bovine homolog of *AML1* or PEBP2 α. We PCR amplified and sequenced a 220-bp sequence from mouse thymus cDNA, using degenerate primers chosen from within the 118-aa region that is conserved in the *Drosophila* runt and human AML1 proteins (data not shown). The PCR-amplified sequence from mouse cDNA encodes 75 aa that are 100% identical to the human AML1 protein and are therefore 100% homologous to the bovine peptide 2 sequence. Mouse PEBP2 α, on the other hand, contains two conservative amino acid substitutions within the 23-aa peptide 2 sequence (25). Since we were able to identify a mouse *AML1* homolog that is 100% identical to human *AML1* in the region encompassed by peptide 2, we believe that peptide 2 is derived from the bovine homolog of *AML1*. While this report was under review, Bae et al. (4) published the nucleotide sequence of a cDNA derived from the murine homolog of human *AML1* and claim that it is encoded by a gene distinct from murine PEBP2α. We predict that at least one of the proteins encoded by the *AML1* gene will bind the

core site and will therefore represent one of the DNA-binding α subunits of CBF.

CBF β mRNA was detected in a broad range of mouse tissues, although levels were somewhat higher in thymus. In contrast, only nuclear extracts prepared from thymus and spleen contained core-binding activity (56, 59). Lung nuclear extracts, for example, which contain relatively high levels of CBF β mRNA, had no demonstrable core-binding activity. The tissue distribution of core-binding activity, however, is presumably determined by the expression of both the CBF α and β subunits. PEBP2 α is highly expressed in thymus (25). The tissue distribution of *AML1* was reported to be widely expressed (4, 39). The fact that CBF β is expressed in tissues other than thymus and spleen also raises the possibility that CBF β has other interaction partners besides PEBP2 α and *AML1*.

ACKNOWLEDGMENTS

We thank Michael Krangel for the δE3 oligonucleotides, Paul Gardner, Gus Lienhard, Marvin Adams, and Stan Froehner for technical advice, and Gus Lienhard for help in analyzing the kinetic binding data. We thank Mike Krangel and Gus Lienhard for critically reading this manuscript and Yoshiaki Ito for communicating results prior to publication. We particularly acknowledge and thank Bill Lane for excellent advice and analysis.

This work was supported by Public Health Service grant CA51065-01A1 from the National Cancer Institute, by a Cancer Center CORE grant (CA23018) from the National Cancer Institute to the Norris Cotton Cancer Center, and by the Leukemia Research Foundation. N.A.S. is a recipient of a Junior Faculty Research Award from the American Cancer Society, and S.W. was an awardee of a predoctoral fellowship from the Norris Cotton Cancer Center at the Mary Hitchcock Hospital.

ADDENDUM IN PROOF

While this article was being revised, we learned that the work cited in reference 25 will be described elsewhere (E. Ogawa, M. Inuzuka, M. Maruyama, M. Satake, M. Naito-Fujimoto, Y. Ito, and K. Shigesada, Virology, in press; E. Ogawa, M. Maruyama, H. Kagoshima, M. Inuzuka, J. Lu, M. Satake, K. Shigesada, and Y. Ito, Proc. Natl. Acad. Sci. USA, in press).

REFERENCES

1. **Abate, C., C. Luk, R. Gentz, F. J. Rauscher III, and T. Curran.** 1990. Expression and purification of the leucine zipper and DNA-binding domains of Fos and Jun: both Fos and Jun contact DNA directly. Proc. Natl. Acad. Sci. USA **87:**1032–1036.
2. **Aebersold, R. H., J. Leavitt, R. A. Saavedra, and L. E. Hood.** 1987. Internal amino acid sequence analysis of proteins separated by one- or two-dimensional gel electrophoresis after in situ protease digestion on nitrocellulose. Proc. Natl. Acad. Sci. USA **84:**6970–6974.
3. **Altschul, S. F., W. Gish, W. Miller, E. W. Myers, and D. J. Lipman.** 1990. Basic local alignment search tool. J. Mol. Biol. **215:**403–410.
4. **Bae, S. C., Y. Yamaguchi-Iwai, E. Ogawa, M. Maruyama, M. Inuzuka, H. Kagoshima, K. Shigesada, M. Satake, and Y. Ito.** 1993. Isolation of PEBP2αB cDNA representing the mouse homolog of human acute myeloid leukemia gene, *AML1*. Oncogene **8:**809–814.
5. **Baldwin, A.** 1989. Methylation interference assay for analysis of DNA-protein interactions, p. 12.3.2–12.3.3. *In* F. M. Ausubel, R. Brent, R. E. Kingston, D. D. Moore, J. G. Seidman, J. A. Smith, and K. Struhl (ed.), Current protocols in molecular biology. Greene Publishing Associates and Wiley-Interscience, New York.
6. **Blackwell, T. K., and H. Weintraub.** 1990. Differences and similarities in DNA-binding preferences of myoD and E2A protein complexes revealed by binding site selection. Science **250:**1104–1110.
7. **Boral, A. L., S. A. Okenquist, and J. Lenz.** 1989. Identification of the SL3-3 virus enhancer core as a T lymphoma cell-specific element. J. Virol. **63:**76–84.
8. **Carroll, S. B., and M. P. Scott.** 1986. Zygotically active genes that affect the spatial expression of the *fushi tarazu* segmentation gene during early Drosophila embryogenesis. Cell **45:**113–126.
9. **Carroll, S. B., and S. H. Vavra.** 1989. The zygotic control of Drosophila pair-rule gene expression. II. Spatial repression by gap and pair-rule gene products. Development **107:**673–683.
10. **Chirgwin, J. J., A. E. Przybyla, R. J. MacDonald, and W. J. Rutter.** 1979. Isolation of biologically active ribonucleic acid from sources enriched in ribonuclease. Biochemistry **18:**5294.
11. **Compton, T.** 1990. Degenerate primers for DNA amplification, p. 39–45. *In* M. A. Innis, D. H. Gelfand, J. J. Sninsky, and T. J. White (ed.), PCR protocols. Academic Press, Inc., San Diego, Calif.
12. **Crute, B. E., and N. A. Speck.** Unpublished data.
13. **Daga, A., J. E. Tighe, and F. Calabi.** 1992. Leukaemia/Drosophila homology. Nature (London) **356:**484.
14. **Feinberg, A. P., and B. Vogelstein.** 1983. A technique for radiolabeling DNA restriction endonuclease fragments to high specific activity. Anal. Biochem. **132:**6–13.
15. **Feinberg, A. P., and B. Vogelstein.** 1984. A technique for radiolabeling DNA restriction endonuclease fragments to high specific activity. Anal. Biochem. **137:**266–277.
16. **Forsburg, S. L., and L. Guarente.** 1989. Identification and characterization of HAP4: a third component of the CCAAT-bound HAP2/HAP3 heteromer. Genes Dev. **3:**1166–1178.
17. **Frasch, M., and M. Levine.** 1987. Complementary patterns of *even-skipped* and *fushi tarazu* expression involve their differential regulation by a common set of segmentation genes in *Drosophila*. Genes Dev. **1:**981–995.
18. **Gunther, C. V., N. A. Speck, and B. J. Graves.** Unpublished data.
19. **Hager, D. A., and R. R. Burgess.** 1980. Elution of proteins from sodium dodecyl sulfate-polyacrylamide gels, removal of sodium dodecyl sulfate, and renaturation of enzymatic activity: results with sigma subunit of Escherichia coli RNA polymerase, wheat germ DNA topoisomerase, and other enzymes. Anal. Biochem. **109:**76–86.
20. **Hallberg, B., J. Schmidt, A. Luz, F. S. Pedersen, and T. Grundstrom.** 1991. SL3-3 enhancer factor 1 transcriptional activators are required for tumor formation by SL3-3 murine leukemia virus. J. Virol. **65:**4177–4181.
21. **Hooper, K. L., S. M. Parkhurst, and D. Ish-Horowicz.** 1989. Spatial control of hairy protein expression during embryogenesis. Development **107:**489–504.
22. **Hsiang, Y. H., D. Spencer, S. Wang, N. A. Speck, and D. H. Raulet.** The role of viral "core" motif-related sequences in regulating T cell receptor γ and δ gene expression. J. Immunol., in press.
23. **Ing, N. H., J. M. Beekman, S. Y. Tsai, M.-J. Tsai, and B. W. O'Malley.** 1992. Members of the steroid hormone receptor superfamily interact with TFIIB (S300-II). J. Biol. Chem. **267:**17617–17623.
24. **Ingham, P., and P. Gergen.** 1988. Interactions between the pair-rule genes *runt*, *hairy*, *even-skipped* and *fushi tarazu* and the establishment of periodic pattern in the Drosophila embryo. Development **104**(Suppl.):51–60.
25. **Ito, Y. (Kyoto University).** Personal communication.
26. **Johnson, P. F., and S. L. McKnight.** 1989. Eukaryotic transcriptional regulatory proteins. Annu. Rev. Biochem. **58:**799–839.
27. **Kadonaga, J. T., and R. Tjian.** 1986. Affinity purification of sequence-specific DNA binding proteins. Proc. Natl. Acad. Sci. USA **83:**5889–5893.
28. **Kalderon, D., W. D. Richardson, A. T. Markham, and A. E. Smith.** 1984. Sequence requirements for nuclear location of simian virus 40 large T antigen. Nature (London) **311:**33–38.

29. **Kamachi, Y., E. Ogawa, M. Asano, S. Ishida, Y. Murakami, M. Satake, Y. Ito, and K. Shigesada.** 1990. Purification of a mouse nuclear factor that binds to both the A and B cores of the polyomavirus enhancer. J. Virol. **64**:4808–4819.
30. **Kania, M. A., A. S. Bonner, J. B. Duffy, and J. P. Gergen.** 1991. The *Drosophila* segmentation gene *runt* encodes a novel nuclear regulatory protein that is also expressed in the developing nervous system. Genes Dev. **4**:1701–1713.
31. **Kerppola, T. K., and T. Curran.** 1991. Fos-jun heterodimers and jun homodimers bend DNA in opposite orientations: implications for transcription factor cooperativity. Cell **66**:317–326.
32. **Kozak, M.** 1987. At least six nucleotides preceding the AUG initiator codon enhance translation in mammalian cells. J. Mol. Biol. **196**:947–950.
33. **Laemmli, U. K.** 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature (London) **227**:680–685.
34. **Lech, K., K. J. Reddy, and L. A. Sherman.** 1988. Preparing DNA from small-scale liquid lysates, p. 1.13.7–1.13.10. *In* F. M. Ausubel, R. Brent, R. E. Kingston, D. D. Moore, J. G. Seidman, J. A. Smith, and K. Struhl (ed.), Current protocols in molecular biology. Greene Publishing Associates and Wiley-Interscience, New York.
35. **Lee, C. C., and C. T. Caskey.** 1990. cDNA cloning using degenerate primers, p. 46–53. *In* M. A. Innis, D. H. Gelfand, J. J. Sninsky, and T. J. White (ed.), PCR protocols. Academic Press, Inc., San Diego, Calif.
36. **Masters, D. B., C. T. Griggs, and C. B. Berde.** 1992. High sensitivity quantification of RNA from gels and autoradiograms with affordable optical scanning. BioTechniques **12**:902–911.
37. **Melnikova, I., B. E. Crute, S. Wang, and N. A. Speck.** 1993. Sequence specificity of the core-binding factor. J. Virol. **67**: 2408–2411.
38. **Melnikova, I. N., and N. A. Speck.** Unpublished data.
39. **Miyoshi, H., K. Shimizu, T. Kozu, N. Maseki, Y. Kaneko, and M. Ohki.** 1991. t(8;21) breakpoints on chromosome 21 in acute myeloid leukemia are clustered within a limited region of a single gene, *AML1*. Proc. Natl. Acad. Sci. USA **88**:10431–10434.
40. **Mount, S. M.** 1982. A catalogue of splice junction sequences. Nucleic Acids Res. **10**:459–472.
41. **Olesen, J. T., and L. Guarente.** 1990. The HAP2 subunit of yeast CCAAT transcriptional activator contains adjacent domains for subunit association and DNA recognition: model for the HAP2/3/4 complex. Genes Dev. **4**:1714–1729.
42. **Prosser, H. M., D. Wotton, A. Gegonne, J. Ghysdael, S. Wang, N. A. Speck, and M. J. Owen.** 1992. A novel phorbol ester response element within the human T cell receptor β enhancer. Proc. Natl. Acad. Sci. USA **89**:9934–9938.
43. **Rauscher, F. J., III, P. J. Voulalas, B. R. Franza, Jr., and T. Curran.** 1988. Fos and Jun bind cooperatively to the AP-1 site: reconstitution in vitro. Genes Dev. **2**:1687–1699.
44. **Redondo, J. M., J. L. Pfohl, C. Hernandez-Munain, S. Wang, N. A. Speck, and M. Krangel.** 1992. Indistinguishable nuclear factor binding to functional core sites of the T cell receptor δ and murine leukemia virus enhancers. Mol. Cell. Biol. **12**:4817–4823.
45. **Redondo, J. M., J. L. Pfohl, and M. S. Krangel.** 1991. Identification of an essential site for transcriptional activation within the human T-cell receptor δ enhancer. Mol. Cell. Biol. **11**:5671–5680.
46. **Robbins, J., S. M. Dilworth, R. A. Laskey, and C. Dingwall.** 1991. Two interdependent basic domains in nucleoplasmin nuclear targeting sequence: identification of a class of bipartite nuclear targeting sequence. Cell **64**:615–623.
47. **Salinovich, O., and R. C. Montelaro.** 1986. Reversible staining and peptide mapping of proteins transferred to nitrocellulose after separation by sodium dodecylsulfate-polyacrylamide gel electrophoresis. Anal. Biochem. **156**:341–347.
48. **Sanger, F., S. Nicklen, and A. R. Coulson.** 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA **74**:5463–5467.
49. **Satake, M., M. Inuzuka, K. Shigesada, T. Oikawa, and Y. Ito.** 1992. Differential expression of subspecies of polyomavirus and murine leukemia virus enhancer core binding protein, PEBP2, in various hematopoietic cells. Jpn. J. Cancer Res. **83**:714–722.
50. **Siebenlist, U., and W. Gilbert.** 1980. Contacts between Escherichia coli RNA polymerase and an early promoter of phage T7. Proc. Natl. Acad. Sci. USA **77**:122–126.
51. **Smith, D. B., and L. M. Corcoran.** 1989. Expression and purification of glutathione-S-transferase fusion proteins, p. 16.7.1–16.7.8. *In* F. M. Ausubel, R. Brent, R. E. Kingston, D. D. Moore, J. G. Seidman, J. A. Smith, and K. Struhl (ed.), Current protocols in molecular biology. Greene Publishing Associates and Wiley-Interscience, New York.
52. **Speck, N. A., B. Renjifo, E. Golemis, T. N. Fredrickson, J. W. Hartley, and N. Hopkins.** 1990. Mutation of the core or adjacent LVb elements of the Moloney murine leukemia virus enhancer alters disease specificity. Genes Dev. **4**:233–242.
53. **Speck, N. A., B. Renjifo, and N. Hopkins.** 1990. Point mutations in the Moloney murine leukemia virus enhancer identify a lymphoid-specific viral core motif and 1,3-phorbol myristate acetate-inducible element. J. Virol. **64**:543–550.
54. **Spencer, D. M., Y.-H. Hsiang, J. P. Goldman, and D. H. Raulet.** 1991. Identification of a T-cell-specific transcriptional enhancer located 3′ of Cγ1 in the murine T-cell receptor γ locus. Proc. Natl. Acad. Sci. USA **88**:800–804.
55. **Strauss, W. M.** 1989. Preparation of genomic DNA from mammalian tissue, p. 2.2.1–2.2.3. *In* F. M. Ausubel, R. Brent, R. E. Kingston, D. D. Moore, J. G. Seidman, J. A. Smith, and K. Struhl (ed.), Current protocols in molecular biology. Greene Publishing Associates and Wiley-Interscience, New York.
56. **Thornell, A., B. Hallberg, and T. Grundstrom.** 1988. Differential protein binding in lymphocytes to a sequence in the enhancer of the mouse retrovirus SL3-3. Mol. Cell. Biol. **8**:1625–1637.
57. **Thornell, A., B. Hallberg, and T. Grundstrom.** 1991. Binding of SL3-3 enhancer factor 1 transcriptional activators to viral and chromosomal enhancer sequences. J. Virol. **65**:42–50.
58. **Tsai, S. Y., I. Sagami, H. Wang, M.-J. Tsai, and B. W. O'Malley.** 1987. Interactions between an DNA-binding transcription factor (COUP) and a non-DNA binding factor (S300-II). Cell **50**:701–709.
59. **Wang, S., and N. A. Speck.** Unpublished data.
60. **Wang, S., and N. A. Speck.** 1992. Purification of core-binding factor, a protein that binds the conserved core site in murine leukemia virus enhancers. Mol. Cell. Biol. **12**:89–102.
61. **Winoto, A., and D. Baltimore.** 1989. A novel, inducible and T cell-specific enhancer located at the 3′ end of the T cell receptor α locus. EMBO J. **8**:729–733.