Dartmouth College

Dartmouth Digital Commons

Open Dartmouth: Peer-reviewed articles by Dartmouth faculty

Faculty Work

1-1992

Purification of Core-Binding Factor, a Protein that Binds the Conserved Core Site in Murine Leukemia Virus Enhancers.

Shuwen W. Wang Dartmouth College

Nancy A. Speck Dartmouth College

Follow this and additional works at: https://digitalcommons.dartmouth.edu/facoa

Part of the Medical Cell Biology Commons, Medical Molecular Biology Commons, and the Virus

Diseases Commons

Dartmouth Digital Commons Citation

Wang, Shuwen W. and Speck, Nancy A., "Purification of Core-Binding Factor, a Protein that Binds the Conserved Core Site in Murine Leukemia Virus Enhancers." (1992). Open Dartmouth: Peer-reviewed articles by Dartmouth faculty. 1273.

https://digitalcommons.dartmouth.edu/facoa/1273

This Article is brought to you for free and open access by the Faculty Work at Dartmouth Digital Commons. It has been accepted for inclusion in Open Dartmouth: Peer-reviewed articles by Dartmouth faculty by an authorized administrator of Dartmouth Digital Commons. For more information, please contact dartmouthdigitalcommons@groups.dartmouth.edu.

Purification of Core-Binding Factor, a Protein That Binds the Conserved Core Site in Murine Leukemia Virus Enhancers

SHUWEN WANG AND NANCY A. SPECK*

Department of Biochemistry, Dartmouth Medical School, Hanover, New Hampshire 03756

Received 12 August 1991/Accepted 7 October 1991

The Moloney murine leukemia virus causes thymic leukemias when injected into newborn mice. A major genetic determinant of the thymic disease specificity of the Moloney virus genetically maps to two protein binding sites in the Moloney virus enhancer, the leukemia virus factor b site and the adjacent core site. Point mutations introduced into either of these sites significantly shifts the disease specificity of the Moloney virus from thymic leukemia to erythroleukemia (N. A. Speck, B. Renjifo, E. Golemis, T. Fredrickson, J. Hartley, and N. Hopkins, Genes Dev. 4:233-242, 1990). We have purified several polypeptides that bind to the core site in the Moloney virus enhancer. These proteins were purified from calf thymus nuclear extracts by selective pH denaturation, followed by chromatography on heparin-Sepharose, nonspecific double-stranded DNA-cellulose, and core oligonucleotide-coupled affinity columns. We have achieved >13.000-fold purification of the core-binding factors (CBFs), with an overall yield of approximately 19%. Analysis of purified protein fractions by sodium dodecyl sulfate (SDS)-polyacrylamide gel electrophoresis reveals more than 10 polypeptides. Each of the polypeptides was recovered from an SDS-polyacrylamide gel, and those in the molecular size range of 19 to 35 kDa were demonstrated to have core-binding activity. The purified CBFs were shown by DNase I footprint analyses to bind the core site in the Moloney virus enhancer specifically, and also to core motifs in the enhancers from a simian immunodeficiency virus, the immunoglobulin μ chain, and T-cell receptor γ-chain genes.

Many replication-competent murine leukemia viruses (MLVs) cause T-cell lymphomas and leukemias in newborn mice and rats (60). A number of genetic studies have identified the enhancer region of these viruses as an important determinant of both thymotropism and T-cell specificity of transformation (6, 8, 9, 12–15, 28, 35, 36, 57). The MLVs are thus an excellent model system for studying how viral transcription influences viral tropism and pathogenesis.

The enhancers from several MLVs have been studied in detail, with the goal of identifying host cell nuclear proteins that bind to the enhancers and thereby influence the leukemogenicity and disease specificity of these viruses (37, 39, 49, 55). The MLV enhancers are often arranged as directly repeated sequences, and, like all other viral enhancers studied to date, they consist of a complex array of binding sites for nuclear DNA-binding proteins (39, 49). Mutations introduced into several of these binding sites in the enhancer of the Moloney MLV reduced the incidence of thymic leukemias caused by the Moloney virus, with a concurrent appearance of erythroleukemias (Fig. 1A) (50). The binding sites that appear to contribute to the thymic disease specificity of the Moloney virus include the leukemia virus factor b (LVb) site and a conserved core motif that was originally identified by Weiher et al. (59) in the enhancers from the Moloney murine sarcoma virus, simian virus 40 (SV40), and the mouse polyomavirus. Both the LVb and core sites are highly conserved in the enhancers from mammalian C-type retroviruses (21). In addition, mutation of 2 bp near the leukemia virus factor c (LVc) site in the Moloney virus enhancer also caused a significant alteration in disease specificity (20). This latter sequence is not well conserved among mammalian C-type viruses, and a protein that binds to this site has not yet been identified.

Although several DNA-binding proteins contribute to thymic disease specificity of the Moloney MLV, we and others have been particularly interested in proteins that bind to the core site. Both in vivo transcription analyses and genetic studies indicate that the core site contributes more significantly than other protein binding sites to the T-cell transcriptional preference of the Moloney and SL3-3 MLV enhancers and to the T-cell disease specificity of the Moloney virus (37, 50, 51). For example, several studies demonstrated that point mutations in the core site in both the Moloney and SL3-3 MLV enhancers attenuate transcription specifically in hematopoietic cell lines (4, 38, 51, 55), whereas mutations in the LVb site in the Moloney MLV enhancer attenuate transcription in all cell types analyzed (51). Point mutations in the core site of the Moloney virus enhancer also caused the most dramatic shift in disease specificity of the virus, resulting in a 65% incidence of erythroleukemia, compared with the more modest 25% shift that occurred upon mutation of the adjacent LVb site (50).

Several protein-DNA complexes have been identified in T-cell nuclear extracts that bind to the core sequences in the Moloney, SL3-3, and AKV MLV enhancers (4, 49, 55). One of these protein-DNA complexes, variously called the SL3-3 and AKV core-binding factor (S/A-CBF) or the SL3-3 enhancer factor 1 (SEF1), binds to the slightly different core sequences from all three of these viruses (4, 55). Protein-DNA complexes specific for the SL3-3 core site (S-CBF) and AKV core site (A-CBF) have also been identified, and S-CBF has been purified to near homogeneity (4, 34).

We have purified several polypeptides that bind the core site in the Moloney MLV enhancer. We have named these proteins simply core-binding factors (CBFs), since the proteins bind not only the core sites in the SL3-3 and AKV MLV enhancers but also the core sites in the Moloney virus

^{*} Corresponding author.

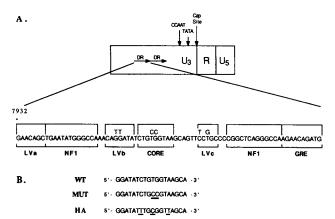


FIG. 1. Origins of oligonucleotides used to detect core-binding activity. (A) A schematic representation of the Moloney MLV long terminal repeat is shown at the top. The sequence at the bottom is from the promoter-distal copy of the direct repeat of the Moloney MLV enhancer (numbering is relative to the cap site at the 5' end of the Moloney MLV viral genome [61]). Binding sites for nuclear factors (leukemia virus factor a [LVa], nuclear factor I [NF1], LVb, core, LVc, and glucocorticoid regulatory element [GRE]) are shown in boxes (49). The point mutations in the LVb, core, and LVc binding sites that altered the disease specificity of the Moloney virus are shown above the sequence (20, 50). (B) Sequences of the top strands of the 18-bp oligonucleotides used as probes and competitors in the electrophoretic mobility shift assay for core-binding activity. The sequence differences between the MUT and HA binding site versus the WT binding site are underlined.

enhancer and in several cellular genes that are specifically transcribed in lymphocytes.

MATERIALS AND METHODS

Biochemical assays. (i) Wild-type, mutant, and high-affinity substrates for protein binding in electrophoretic mobility shift assays. Complementary 18-base oligonucleotides containing a high-affinity (HA) core site, wild-type (WT) core site, and mutant (MUT) core site (for sequences, see Fig. 1) were synthesized on a Biosearch Cyclone DNA synthesizer at Dartmouth Medical School and purified by electrophoresis through 20% polyacrylamide-7 M urea gels.

Radioactive probes were made by end-labeling 100 pmol of one oligonucleotide, either the top or bottom strand of the binding site, with $[\gamma^{-3^2}P]ATP$ (7,000 Ci/mmol; ICN) and T4 polynucleotide kinase (New England Biolabs) and then annealing the labeled oligonucleotide with 100 pmol of its complementary strand (31). The double-stranded probes were then purified by electrophoresis through a 20% native polyacrylamide gel. The specific activity of the probes was typically 4,000 to 10,000 cpm/fmol.

Competitor oligonucleotides were prepared by annealing equal amounts of unlabeled complementary oligonucleotides. The annealed oligonucleotides were used directly as competitors in binding reactions.

(ii) Substrates for methylation interference and DNase I footprinting analyses. The probes used in methylation interference and DNase I footprinting analyses were prepared from plasmid DNA. Two plasmids containing sequences from the wild-type Moloney MLV enhancer (PvuII₇₉₃₈-PvuII₈₀₁₃ and HaeIII₇₉₄₉-HaeIII₇₉₉₄, numbered [positions indicated by subscripts] according to Weiss et al. [61], subcloned into SP64) have been previously described (49). Both the PvuII₇₉₃₈-PvuII₈₀₁₃ probe (fragment I) and the

HaeIII₇₉₄₉-HaeIII₇₉₉₄ probe (fragment VII) were prepared by end labeling with $[\gamma^{-32}P]ATP$ in the presence of T4 polynucleotide kinase, either at the EcoRI or BamHI site in the SP64 polylinker, and recutting at either BamHI or EcoRI, respectively. A probe containing the Moloney MLV enhancer with a mutated core site was prepared by subcloning the PvuII₇₉₃₈-PvuII₈₀₁₃ fragment from plasmid pMo (core)CAT (51), which contains the Moloney MLV enhancer with a 2-bp point mutation in the core site (CTGTGGTA A→CTGCCGTAA), into the SmaI site of pUC13. A probe was prepared from this plasmid by cutting the pUC13 polylinker at either the EcoRI or BamHI site, end labeling, and recutting at BamHI or EcoRI as described above.

The probe from the immunoglobulin μ-chain enhancer was obtained from plasmid pμXDB (kindly provided by Ranjan Sen), which contains an XbaI-to-DdeI fragment of the enhancer (nucleotides 1 to 521, according to Gillies et al. [19]) that was linkered with BamHI linkers and subcloned into the BamHI site of SP64. The plasmid was cut and end labeled as described above, either within the immunoglobulin μ-chain enhancer insert at PvuII₃₈₅ or in the XbaI site of the polylinker of SP64, and recut at XbaI or PvuII to generate the top- and bottom-strand probes for DNase I footprinting.

The T-cell receptor γ -chain (TCR γ) enhancer probe was prepared from plasmid J21-.20RR (52), kindly provided by David Raulet. J21-.20RR contains a 196-bp RsaI-RsaI fragment from the TCR γ enhancer subcloned into the polylinker of plasmid J21 (62). The plasmid was cut and end labeled in the polylinker of J21 at either the XhoI or ClaI site and then recut at ClaI or XhoI.

The probe from the simian immunodeficiency virus enhancer, clone 142 (SIVmac142) (7), was prepared from plasmid pSIV(-258)CAT, described by Renjifo et al. (45). Plasmid pSIV(-258)CAT was cut and end labeled at the *HindIII* site in the polylinker or at the *RsaI*₋₇₁ site and then recut at *RsaI* and *HindIII*, respectively.

All end-labeled probes were purified and eluted from 8 to 12% native polyacrylamide gels. Typical specific activities were approximately 1,000 to 4,000 cpm/fmol of 5' ends.

(iii) Protein-DNA binding analysis. Core-binding activity was detected by the formation of protein-DNA complexes in the electrophoretic mobility shift assay (16, 18, 48). Typically, binding reaction mixtures contained 10,000 to 20,000 cpm (2 to 5 fmol) ³²P-end-labeled probe, binding buffer (100 mM NaCl, 10 mM Tris (pH 7.4), 1 mM β-mercaptoethanol, mM EDTA, 4% glycerol), 0 to 1.0 µg of poly(dIdC) poly(dI-dC) (Pharmacia), 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, 50 ng (500-fold molar excess) of unlabeled oligonucleotides was included in some of the binding reaction mixtures. After 15 min of incubation at room temperature, the reaction mixtures were fractionated by electrophoresis through a 6% native polyacrylamide gel in $0.5 \times TBE$ (45 mM Tris-HCl, 45 mM boric acid, 1 mM EDTA) at room temperature. Radioactivity was detected by autoradiography of the dried gel. The amount of DNA present in specific protein-DNA complexes was quantified by scintillation spectrometry of the protein-DNA complex bands excised from dried polyacrylamide gels and counted in the presence of Ecoscint A (National Diagnostics).

(iv) Protein assay. Protein concentrations were determined by the method of Bradford (5), using reagents purchased from Bio-Rad. Bovine serum albumin (BSA) was used as a standard for determining the protein concentration.

(v) Measurement of NaCl concentration. The concentration

of NaCl in column fractions was determined by measuring conductivity in comparison with a standard curve.

(vi) SDS-PAGE. Sodium dodecyl sulfate (SDS)-polyacrylamide gel electrophoresis (PAGE) was performed essentially as described by Laemmli (33). For very dilute protein samples (<100 μ g/ml), 9.1% trichloroacetic acid (TCA) was used to precipitate the proteins in the presence of deoxycholate (0.14 mg/ml) as carrier.

A modified procedure of silver staining (63) was used to visualize the protein in the gel. Briefly, following electrophoresis, the gel was fixed with 50% methanol for 2 h, washed with distilled water for 30 min, stained with a silver nitrate solution (0.075% NaOH, 0.46% NH₄OH, 0.4% AgNO₃) for 15 min, washed with distilled water for 15 min, and developed in a solution consisting of 0.005% citric acid and 0.035% formaldehyde. The staining was stopped by 50% methanol. For Coomassie brilliant blue staining, the gel was fixed for 15 min in 10% acetic acid-50% methanol and stained for 1 h with 10% acetic acid-50% methanol-0.05% Coomassie brilliant blue R-250 (Bio-Rad). The gel was destained with 10% acetic acid.

(vii) Recovery and renaturation of protein from SDS-polyacrylamide gels. Denaturation/renaturation experiments were performed essentially as described by Hager and Burgess (23), with several modifications. Briefly, protein samples were prepared and electrophoresed as described above, and gels were stained with Coomassie brilliant blue. Each of the visible polypeptide bands was excised from the gel and chopped into small pieces. Proteins were eluted by incubating the gel slice overnight at room temperature in 200 to 300 µl of a buffer containing 150 mM NaCl, 20 mM N-2-hydroxyethylpiperazine-N'-2-ethanesulfonic (HEPES; pH 7.5), 5 mM dithiothreitol (DTT), 0.1 mM EDTA, 0.1% SDS, and 0.1 mg of BSA per ml. The eluted protein was precipitated with 4 volumes of cold (-20°C) acetone and collected by centrifugation $(16,000 \times g \text{ for } 30)$ min). The pellet was washed with a solution containing 80% acetone and 20% dilution buffer (150 mM NaCl, 20 mM HEPES [pH 7.5], 1 mM DTT, 0.1 mM EDTA, 20% glycerol), dried, dissolved in 50 µl of dilution buffer supplemented with 6 M guanidine-HCl, and incubated at room temperature for 30 min. The protein was then renatured by dialysis on a 0.025-\u00e4m-pore-size filter disk (Millipore) (10) against 4 ml of dilution buffer in one well of a 12-well microtiter plate for 1 h at room temperature. The protein samples were then transferred to Eppendorf tubes and assayed for core-binding activity by electrophoretic mobility shift assay.

(viii) Methylation interference analysis. The 45-bp HaeIII₇₉₄₉-HaeIII₇₉₉₄ fragment from the Moloney MLV enhancer was used as probe in methylation interference analyses (25). The binding reaction (50 μ l, total volume) contained 100,000 cpm of end-labeled methylated (40) probe and 25 µl of affinity-purified protein in the same buffer used for electrophoretic mobility shift assays. The binding reaction was electrophoresed through a 5% native polyacrylamide gel. Following overnight exposure of the gel, the bands corresponding to the protein-DNA complex and free DNA were excised from the gel, and the DNA was purified by electroelution onto NA45 membranes (Schleicher & Schuell) as described by Baldwin (2), subjected to 1 M piperidine cleavage, and analyzed by electrophoresis through a 12% polyacrylamide-7 M urea sequencing gel.

(ix) DNase I footprint analyses. DNase I footprinting assays (17) were performed by using probes derived from the enhancers from Moloney MLV, the immunoglobulin μ-chain gene, the TCRγ gene, and SIVmac142. Binding reactions

contained 2% polyvinyl alcohol, 0.8 mg of BSA per ml, $0.5 \times$ buffer Z (see below), 20,000 cpm of end-labeled probe, and various amounts of affinity-purified CBF, in a total volume of 50 μ l. The binding reaction mixtures were incubated on ice for 1 h. The DNA was then digested with 25 mU of DNase I (DPFF; Worthington Biochemicals) in the presence of 5 mM CaCl₂ for 1 min, at which point the reactions were terminated by the addition of 90 μ l of DNase I stop buffer (20 mM EDTA [pH 8.0], 1.0% SDS, 0.2 M NaCl, 250 μ g of tRNA per ml). The reaction mixture was extracted with phenol-chloroform, and the DNA was recovered by ethanol precipitation. The samples were electrophoresed through 8 to 10% polyacrylamide–7 M urea sequencing gels along with chemical sequencing tracts (40).

Purification of CBF. (i) Buffers. Homogenization buffer (buffer H) contains 250 mM sucrose, 50 mM HEPES (pH 7.5), 25 mM KCl, 10 mM β -mercaptoethanol, 5 mM MgCl₂, and 1 mM phenylmethylsulfonyl fluoride (PMSF). Extract buffer (buffer E) contains 250 mM sucrose, 400 mM NaCl, 50 mM HEPES (pH 7.5), 10 mM β -mercaptoethanol, and 1 mM PMSF. Buffer B contains 50 mM NaCl, 20 mM HEPES (pH 7.5), 10 mM β -mercaptoethanol, 2 mM EDTA, and 10% glycerol. Buffer Z contains 100 mM NaCl, 25 mM HEPES (pH 7.5), 1 mM DTT, 20% glycerol, 1 mM EDTA, and 0.1% Nonidet P-40.

(ii) Preparation of the HA and WT core site affinity columns. Affinity resins were prepared by Kadonaga and Tjian's modification of their original method (31). Complementary oligonucleotides with the following sequences were used to prepare the affinity columns:

WT: 5'-ATATCTGTGGTAAGCGG-3' 3'-CGCCTATAGACACCATT-5'

HA: 5'-GGATATTTGCGGTTAGCA-3'
3'-TAAACGCCAATCGTCCTA-5'

Oligonucleotides were synthesized and purified as described above. Each of the two complementary strands of either the WT core site or HA core site (440 μ g of each; 880 μ g total) was annealed, phosphorylated, concatenated, and coupled to 5 ml of Sepharose CL-2B (Pharmacia), immediately following activation of the resin with CNBr (Aldrich), as described by Kadonaga and Tjian (31). The efficiency of coupling was >90%.

(iii) Regeneration of chromatography resins. Heparin-Sepharose was regenerated according to the manufacturer (Pharmacia). Nonspecific calf thymus DNA-cellulose (Sigma) was regenerated by washing with buffer B containing 2.5 M NaCl. Affinity resins were regenerated as described by Kadonaga and Tjian (31). The regenerated resins were all equilibrated with 10 volumes of the appropriate running buffers before samples were applied to the columns.

(iv) Preparation of nuclear extracts. Calf thymus was obtained from newly slaughtered 16-week-old calves at a slaughterhouse in Hyde Park, Vt. The thymuses either were used directly after slaughter or were diced into approximately 1-cm cubes, snap frozen in liquid N_2 , and stored at -80° C until needed. Frozen thymus was thawed in ice-cold buffer H before homogenization.

All steps were performed at 4°C. Nuclei were prepared from calf thymus essentially by the method of Halligan and Desiderio (24). Four hundred grams (wet weight) of diced calf thymus tissue was homogenized in a Waring blender (two pulses of 30 s each at highest speed) in 1 liter of buffer H. The crude lysate was passed through two, four, and eight successive layers of cheesecloth to remove unblended ma-

terial and connective tissue. The nuclei were collected by centrifugation at 1,800 \times g for 7 min in a GSA rotor. The nuclear pellet was resuspended by hard shaking in 800 ml of buffer E and incubated on ice for 15 min. Nuclei were then pelleted by centrifugation at $12,000 \times g$ in a GSA rotor for 10 min. Proteins in the supernatant were further precipitated by slow addition of ammonium sulfate over 20 min with constant stirring to a final concentration of 50%, and the reaction mixture was allowed to equilibrate with constant stirring for an additional 30 min. The ammonium sulfate precipitate was then collected by centrifugation at $16,000 \times g$ for 45 min in a GSA rotor. The pellet was resuspended in 100 ml of buffer B and dialyzed twice against 2 liters of buffer B overnight. A precipitate that formed during dialysis was removed by centrifugation at $10,000 \times g$ for 10 min in an SS34 rotor. The supernatant from this centrifugation was designated the nuclear extract. The protein concentration in the nuclear extract was typically 25 to 30 mg/ml. Typical yields were 1 g of nuclear extract per 100 g of calf thymus tissue.

- (v) Selective pH denaturation of the nuclear extract. A pH titration for the selective pH denaturation step was performed for each independent preparation of calf thymus nuclear extract. Phosphoric acid (5 M) was used to lower the pH of 15 ml of the nuclear extract. The pH of the nuclear extract was titrated from pH 7.0 to 2.5, and 1-ml aliquots were removed in increments of 0.5 pH units. The proteins were allowed to precipitate overnight at 4°C. Precipitated proteins were removed by centrifugation at $16,000 \times g$ for 10 min, and the pH of the supernatant was then readjusted to 7.0 by the addition of 14.8 M NH₄OH. Electrophoretic mobility shift assays were used to determine the yield of core-binding activity in each sample, and Bradford assays were used to determine the protein concentration. Generally, pH 4.5 to 5.0 was found to be the optimal pH for this step. For the purification summarized in Table 1, 130 ml of nuclear extract (3,120 mg of protein) was selectively pH denatured at pH 4.6.
- (vi) Heparin-Sepharose chromatography. The supernatant following pH denaturation was applied directly onto a heparin-Sepharose column (2.5 by 8.1 cm, 40 ml) that was equilibrated with buffer B. The column was washed with 3 column volumes of buffer B, and the core-binding activity was eluted with a 400-ml linear gradient of NaCl from 50 mM to 1.0 M in buffer B, and approximately 5-ml fractions were collected. Fractions were analyzed for core-binding activity, protein concentration, and NaCl concentration. Fractions with core-binding activity were pooled and dialyzed against buffer B.
- (vii) Nonspecific DNA-cellulose chromatography. Dialyzed pooled fractions from the heparin-Sepharose column were applied to a nonspecific double-stranded DNA-cellulose column (2.5 by 4.0 cm, 20 ml) that was equilibrated with buffer B supplemented with 5 mM EDTA. The column was washed with 3 column volumes of buffer B (plus 5 mM EDTA) and developed with a 200-ml linear gradient of NaCl from 50 mM to 1.0 M in buffer B plus 5 mM EDTA, and approximately 4-ml fractions were collected. Fractions containing specific core-binding activity were pooled and dialyzed against buffer Z.
- (viii) Recognition site affinity chromatography. The HA core site and WT core site DNA affinity columns (1 ml of resin) were equilibrated with buffer Z. Pooled fractions from the nonspecific DNA-cellulose column were adsorbed to either the HA core site affinity column or the WT core site affinity column. The column was washed with 10 column volumes of buffer Z, and the core-binding activity was step

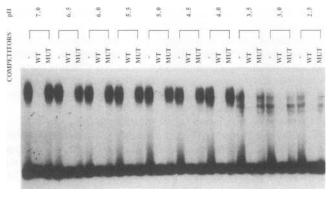


FIG. 2. Effect of pH denaturation on core-binding activity. Electrophoretic mobility shift assay to detect core-binding activity following selective pH denaturation. Phosphoric acid was used to lower the pH of the nuclear extract from 7.0 to 2.5 in increments of 0.5 pH units as described in Materials and Methods. Each binding reaction mixture contained 20,000 cpm of $^{32}\text{P-end-labeled WT}$ probe, 1 µg of poly(dI-dC) \cdot poly(dI-dC), and 10 µl of a 1:10 dilution of nuclear extract following pH denaturation, in the absence (–) or presence of 50 ng (500-fold molar excess) of the WT or MUT competitor oligonucleotide.

eluted either with 0.3 and 1.0 M NaCl in buffer Z for the HA core site affinity column or with 0.3 and 0.6 M NaCl for the WT core site affinity column. Fractions with core-binding activity were pooled and dialyzed against buffer Z, and the column was regenerated. The dialyzed sample was then loaded again onto the same regenerated column and eluted as described above. This procedure was repeated two additional times, for a total of four successive passages on either the HA or WT core site affinity column.

RESULTS

Assay for core-binding activity. Core-binding activity was detected by electrophoretic mobility shift assays (16, 18, 48). The sequences of the three oligonucleotides that were used as probes or competitors are shown in Fig. 1. Two different probes were used in the electrophoretic mobility shift assay. The first probe is a synthetic 18-bp oligonucleotide that contains the WT core site from the Moloney virus enhancer. The second probe is a deduced HA core sequence (56) with greater binding affinity for CBF (data not shown). The protein-DNA complex obtained with the Moloney virus WT probe can be competed for by the unlabeled HA binding site (data not shown). Likewise, the protein-DNA complex obtained with the HA probe can be competed for with the WT binding site (data not shown). Therefore, the same proteins appear to bind both the WT and HA probes. This is consistent with the original results reported by Thornell et al. (56) for the binding specificity of the analogous protein that they call SEF1 to the core site from the SL3-3 virus enhancer and the deduced core site.

The WT and HA probes have been used interchangeably in the electrophoretic mobility shift assay throughout the purification. We performed competition assays on active fractions in each step of the purification with unlabeled WT and HA oligonucleotides, and also with the MUT binding site which contains the 2-bp mutation in the Moloney virus core site that alters the disease specificity of the Moloney virus from thymic to erythroid leukemia (50).

Purification of CBF. We purified the CBFs from calf thymus nuclei by a combination of conventional chromatog-

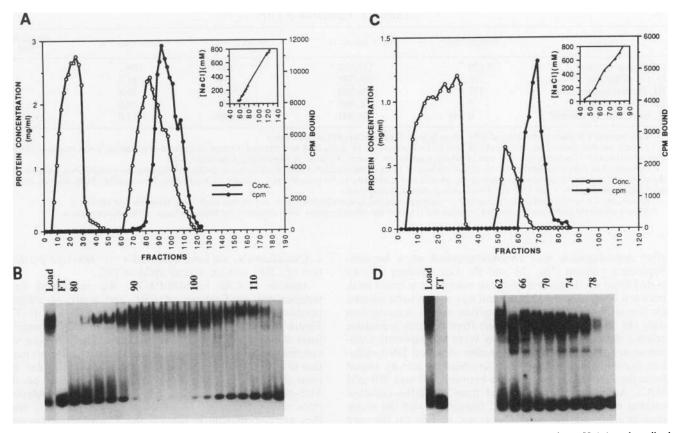


FIG. 3. Fractionation of CBF. (A and B) Chromatography on heparin-Sepharose. Nuclear extract was denatured to pH 4.6 as described in Materials and Methods. The supernatant from the selective pH denaturation step was then loaded onto a 40-ml heparin-Sepharose column equilibrated in buffer B (50 mM NaCl). The column was developed with a linear gradient of NaCl from 50 mM to 1.0 M. (A) An aliquot (5 μl) of each 5-ml fraction was assayed for core-binding activity by mobility shift assay, using the WT probe (10,000 cpm per reaction), in the presence of 200 ng of poly(dI-dC) · poly(dI-dC). The amount of radioactivity (•) in the active fractions was quantified by scintillation fluorography. The protein concentration (○) was determined by Bradford assay. (Inset) Conductivity measurement of the concentration of NaCl in column fractions across the gradient. (B) Autoradiograph showing the DNA-binding activity in the active fractions. Lanes: Load, 1 μl of the loaded fraction; FT, 5 μl of flowthrough fractions; 80 to 110, 5 μl of eluate fractions from the heparin-Sepharose column. (C and D) Chromatography on the nonspecific DNA-cellulose column. Pooled fractions (87 to 107) from the heparin-Sepharose column were dialyzed against buffer B and loaded on a 20-ml double-stranded calf thymus DNA-cellulose column in buffer B (supplemented with 5 mM EDTA). The column was washed and developed with a linear gradient of NaCl from 50 mM to 1.0 M in buffer B (5 mM EDTA). (C) An aliquot (5 μl) of each 2-ml fraction was assayed for specific DNA-binding activity (•) in the presence of 50 ng of poly(dI-dC) · poly(dI-dC) and for protein concentration (○) as described above. (Inset) Conductivity measurement of the concentration of NaCl in each column fraction. (D) Autoradiograph showing the core-binding activity in the active fractions. Lanes: Load, 5 μl of loaded pooled heparin-Sepharose column fractions; FT, 5 μl of the flowthrough fractions; 62 to 78, elution fractions from the nonspecific DNA-cellulose column.

raphy and recognition site affinity chromatography (31, 46). Calf thymus is an excellent source of CBF because it is inexpensive, available in unlimited supply, and an enriched source of T cells. Furthermore, thymus is the target organ for leukemia induced by the Moloney virus.

In the course of our studies, we found that the corebinding activity is relatively stable at low pH. We therefore exploited this property of CBF and used selective pH denaturation as the initial purification step. We first determined the optimal pH with which the greatest purification of CBF was obtained with an acceptable yield of activity. The pH of the calf thymus nuclear extract was titrated from pH 7.0 to 2.5 in increments of 0.5 pH units, the extract was centrifuged to remove the denatured precipitated proteins, the supernatant was adjusted to pH 7.0, and the extract was assayed for protein concentration and core-binding activity (Fig. 2).

The DNA-binding activity that remains following selective pH denaturation is specific for the core sequence, as shown

by the ability of a 500-fold molar excess of the WT but not the MUT core site to compete for binding to the WT probe (Fig. 2). The mobility of the protein-DNA complex increases slightly as the pH is lowered, possibly as a result of partial proteolysis of the proteins by activated acid proteases. The protein-DNA complex is relatively wide; inclusion of less core-binding activity in the binding reaction mixtures (or shorter exposures of the mobility shift gels) reveals that this large complex is composed of multiple protein-DNA complexes with different mobilities (data not shown).

The yield and purification of CBF following selective pH denaturation varied somewhat between different preparations of nuclear extracts; therefore, an independent pH titration was performed for each nuclear extract that was used in the purification. In each of the independent CBF purifications, pH denaturation was performed at either pH 4.6 or 5.0, depending on the starting material, and we obtained a three- to sevenfold purification in this first step.

Following the selective pH denaturation, the supernatant

TABLE	1.	Purification	of	CBF^{u}
-------	----	--------------	----	-----------

Fraction	Protein (mg) ^b	Total activity (U) ^c	Sp act (U/mg)	Total yield (%)	Purification (fold)
I. Nuclear extract ^d	3,120	312,000	100	100	
II. pH denaturation	556.5	191,300	344	61.3	3.4
III. Heparin-Sepharose	131.6	124,000	942	39.8	9.4
IV. DNA-cellulose	5.2	62,000	12,020	20.0	120
V. Core affinity column ^e	0.044	58,900	1,338,640	19.0	13,386

^a Data represent a single preparation of CBF, starting with 300 g (wet weight) of calf thymus.

d Represents a 0 to 50% (NH₄)₂SO₄ precipitate of proteins extracted in 400 mM NaCl from calf thymus nuclei (see Materials and Methods).

after centrifugation was chromatographed on a heparin-Sepharose column (Fig. 3A and B). Core-binding activity eluted from the heparin-Sepharose column in a broad peak between 400 and 600 mM NaCl and was shown to be specific for the core sequence by competition assays (competition data not shown). Pooled fractions from heparin-Sepharose column with core-binding activity were subsequently chromatographed on a nonspecific double-stranded DNA-cellulose column (Fig. 3C and D). Core-binding activity eluted from the DNA-cellulose column between 390 and 610 mM NaCl. Additional activity eluted from the DNA-cellulose column in 150 to 390 mM NaCl, coinciding with the major protein peak, but this activity was not specific for the core sequence (competition data not shown). We achieved an approximately 3-fold purification of CBF on heparin-Sepharose and a 13-fold purification on the DNA-cellulose column, resulting in an overall purification of 120-fold at this step (Table 1).

Active fractions from the nonspecific DNA-cellulose column were further purified by recognition site affinity chromatography (31) on columns that were prepared with either catenated WT or HA core binding sites (Fig. 4). Pooled fractions containing core-binding activity from the first adsorption on the affinity column were dialyzed and readsorbed three additional times on the regenerated column. The HA core site affinity column was developed with two successive step elutions of 300 mM and 1 M NaCl. Corebinding activity did not elute from the HA site matrix in 300 mM NaCl, whereas >90% of core-binding activity was recovered in the 1.0 M NaCl eluate (Fig. 4A). Inclusion of an intermediate NaCl wash (600 mM NaCl) split the corebinding activity into two fractions: one that eluted from the HA core site column in 600 mM NaCl, and a second fraction that eluted in 1.0 M NaCl (data not shown).

Adsorption on the WT core site affinity column resulted in a different elution profile, in that a significant proportion of the core-binding activity eluted from the WT core site column in 300 mM NaCl and the remaining core-binding activity eluted from the column in 600 mM NaCl (Fig. 4B). We also note that the proteins that preferentially elute from the WT core site column in 300 mM NaCl give rise to protein-DNA complexes that migrate faster on native polyacrylamide gels than do the protein-DNA complexes obtained with the 600 mM NaCl eluate.

We estimate that a >100-fold purification of CBF was obtained from four successive passages on either the WT or HA core site column, with 95% yield of core-binding activity in this step. The purification of CBF is summarized in Table

1. Cumulatively, we have achieved a >13,000-fold purification of CBF, with an overall yield of 19%.

Analysis of CBF by SDS-PAGE. We performed three independent purifications of CBF and analyzed affinitypurified fractions from each purification by SDS-PAGE (33). Figure 5A shows the SDS-PAGE analysis of protein samples from each step of one CBF purification. The HA core site column was used in the final step of this particular preparation of CBF. Multiple polypeptides ranging in molecular size from approximately 19 to 40 kDa were detected on the SDS-polyacrylamide gel (Fig. 5A, lane 5). The two polypeptides of approximately 60 kDa are keratin proteins, since they are also present in lanes 6 and 7, which contain only precipitated buffer and sample buffer, respectively. We had anticipated that there would be multiple polypeptides in the affinity-purified fractions, since this was consistent with the observation that there are multiple protein-DNA complexes on native polyacrylamide gels, and also that the core-binding activity could be split into two fractions by elution in different NaCl concentrations from affinity columns.

The second preparation of CBF was also purified on the HA core site column. Multiple polypeptides in the affinity purified fractions were again detected by SDS-PAGE, but the pattern of polypeptides was not identical to that obtained in the first purification (Fig. 5B, lanes 3 and 4). Although there are polypeptides common to affinity-purified CBF from both preparations, there are also a number of polypeptides in the first preparation of CBF (lane 4) that are either absent or insufficiently abundant to be detected in the second CBF preparation (lane 3).

In the third purification of CBF, the pooled fractions with core-binding activity from the nonspecific DNA-cellulose column were divided into two aliquots, and each of these aliquots was independently purified by recognition site affinity chromatography. One aliquot of the pooled DNAcellulose eluate was chromatographed on the HA core site affinity column, and the second aliquot was chromatographed on the WT core site column. A sample of the 1.0 M NaCl eluate from the HA core site column was electrophoresed through lane 5 (Fig. 5B) of the SDS-polyacrylamide gel, and the 300 and 600 mM NaCl eluates from the WT core site column were electrophoresed through lanes 6 and 7, respectively. There is clearly variability in the number and abundance of polypeptides between independent preparations of CBF. Note, however, that several polypeptides consistently appear in every preparation of CBF. For example, a doublet that migrates at 42 kDa is in every CBF preparation, as is a doublet that migrates just below the

^b Protein was measured by the method of Bradford (5) for fractions I to IV, using BSA as a standard. Protein concentration in fraction V was roughly estimated from silver-stained SDS-polyacrylamide gels by adding together the amount of protein in individual polypeptide bands.

c One unit of activity is defined as the amount of binding by 10 μg of the nuclear extract. Binding was quantified by excising the protein-DNA complex from the dried polyacrylamide gel and determining the amount of radioactivity in the protein-DNA complex by scintillation fluorography. Each reaction mixture contained approximately 2 to 5 fmol of ³²P-end-labeled oligonucleotide probe.

Affinity chromatography was performed by using the HA core site affinity column. Data represent the fourth passage on the affinity column.

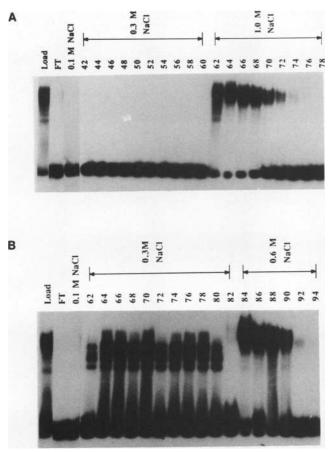


FIG. 4. Electrophoretic mobility shift assays of the core-binding activity following affinity chromatography. (A) Affinity chromatography on the HA core site column. The autoradiograph shows core-binding activity in eluate fractions from the fourth passage on the HA core site affinity column. Pooled active DNA-cellulose fractions (62 to 77) were dialyzed against buffer Z (0.1 M NaCl) and adsorbed onto an affinity column containing 1 ml of the HA core site affinity resin. The column was then washed with buffer Z (0.1 M NaCl) and step eluted with buffer Z containing 0.3 and 1.0 M NaCl. Binding reaction mixtures contained 5 µl of eluate fractions and 10,000 cpm of HA probe in a total volume of 15 µl. Lanes: Load, loaded material; FT, flowthrough fractions; 0.1 M NaCl, 0.1 M NaCl eluate; 42 to 60, 0.3 M NaCl eluate fractions; 62 to 78, fractions eluted with 1.0 M NaCl. (B) Affinity chromatography on the WT core site column. The active DNA-cellulose fractions were adsorbed onto and eluted from the WT column four times successively as described for panel A except that 0.3 and 0.6 M NaCl were used to step elute the core-binding activity. Shown is the assay of fractions from the fourth adsorption on the WT affinity column. Lanes: Load, loaded material; FT, flowthrough fractions; 0.1 M NaCl, 0.1 M NaCl eluate; 62 to 82, fractions eluted with 0.3 M NaCl; 84 to 94, fractions eluted with 0.6 M NaCl.

31-kDa size marker, another doublet at 25.5 kDa, and a single band at 22.5 kDa (Fig. 5B, arrows).

Denaturation/renaturation analysis of CBF. To determine which of the multiple polypeptides that were seen by SDS-PAGE possessed core-binding activity, we excised each protein band that was visible from a Coomassie brilliant blue-stained SDS-polyacrylamide gel, eluted the proteins from each gel slice, and subjected the proteins to a denaturation/renaturation regimen described by Hager and Burgess (23). Each of the renatured proteins was then assayed for

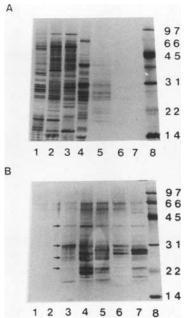


FIG. 5. Characterization of CBF by SDS-PAGE. (A) SDS-PAGE analysis of fractions from one purification of CBF. Seven micrograms each of crude extract, pH-denatured extract, and heparin-Sepharose activity pool were mixed with SDS sample buffer. Protein samples from active pooled fractions from the nonspecific DNA-cellulose and affinity columns were first concentrated by TCA precipitation. Protein samples were electrophoresed through a 15% SDS-polyacrylamide gel and visualized by silver staining. Lanes: 1, 7 μg of nuclear extract; 2, 7 μg of pH-denatured nuclear extract (pH 4.6); 3, 7 µg of pooled heparin-Sepharose fractions; 4, 100 µl of pooled DNA-cellulose fraction; 5, 250 µl of pooled HA core site affinity column fractions; 6 and 7, buffer Z and SDS sample buffer, respectively; 8, low-molecular-weight markers (positions indicated in kilodaltons). The two bands that migrate with an apparent molecular mass of 60 to 66 kDa in the affinity-purified CBF fraction (lane 5) are keratin proteins and are also present in control buffers (lanes 6 and 7). (B) Affinity-purified fractions from several independent purifications of CBF, starting with different calf thymus nuclear extracts. Eluate fractions from the HA and WT core site affinity columns (500 µl) were TCA precipitated and electrophoresed through a 15% SDS-polyacrylamide gel, and the gel was stained with silver. Lanes: 1, buffer Z; 2, SDS sample buffer; 3 to 5, three independent purifications of CBF eluted from the HA core site affinity column in 1.0 M NaCl (lane 4 is from the first, lane 3 is from the second, and lane 5 is from the third CBF purification); 6 and 7, 0.3 and 0.6 M NaCl eluates of CBF from the WT core site affinity column, respectively; 8, low-molecular-weight markers (positions indicated in kilodaltons). Arrows designate polypeptides that are common to all CBF preparations. The samples in lanes 5 to 7 are derived from one purification starting with the same calf thymus nuclear extract and divided only at the affinity chromatography step.

core-binding activity by electrophoretic mobility shift assay (Fig. 6).

Some of the polypeptide bands in the Coomassie brilliant blue-stained SDS-polyacrylamide gel were very faint; they cannot be easily seen in the photograph but could be seen by the naked eye on the stained gel (Fig. 6A). The arrows to the right of each lane in Fig. 6A mark the position of each polypeptide that was excised from the gel. The keratin proteins (the two uppermost proteins in each lane) were also excised and used as negative controls. The polypeptides were labeled alphabetically, from top to bottom of each lane (largest to smallest). The letters assigned to each polypeptide

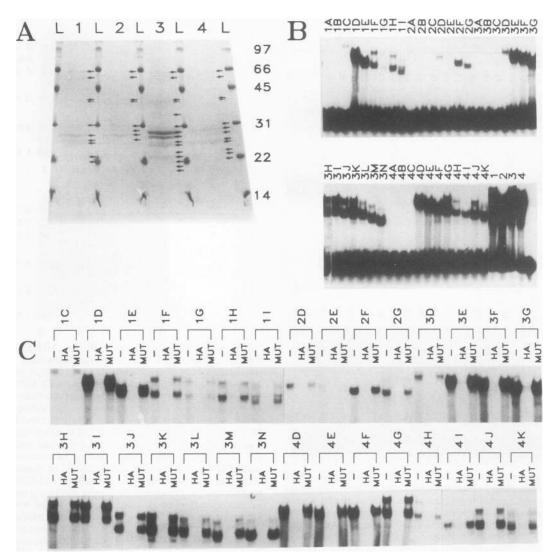


FIG. 6. Renaturation of core-binding activity from SDS-polyacrylamide gels. (A) Fractionation of polypeptides by SDS-PAGE. Shown is the Coomassie brilliant blue-stained SDS-polyacrylamide gel from which the different polypeptides were excised. CBF preparations are the same as those shown in lanes 4 to 7 in Fig. 5B. One milliliter of each affinity-purified fraction was TCA precipitated and electrophoresed through a 15% SDS-polyacrylamide gel, and the gel was stained with Coomassie brilliant blue. Lanes: 1 and 4, proteins eluted from the HA core site column from two independent purifications (corresponding to lanes 4 and 5 in Fig. 5B); 2 and 3, 0.3 and 0.6 M NaCl eluates from the WT core site affinity column (corresponding to lanes 6 and 7 in Fig. 5B); L, low-molecular-weight markers (positions indicated in kilodaltons). Arrows indicate the bands that were excised from the Coomassie brilliant blue-stained SDS-polyacrylamide gel. These bands were labeled sequentially by letters from top to bottom (1A to 1H, 2A to 2G, 3A to 3N, and 4A to 4K). The two keratin proteins (top two bands, A and B in each lane) were also excised from each lane and used as negative controls for core-binding activity. The doublet migrating at approximately 42 kDa (large and small arrows) was excised in a single gel slice (polypeptide C). (B) Electrophoretic mobility shift assay of core-binding activity from renatured polypeptides. The excised polypeptides were eluted from the SDS-polyacrylamide gel slices and subjected to a denaturation/renaturation regimen as described in Materials and Methods. Ten microliters of each renatured protein was assayed for core-binding activity, using 10,000 cpm of the HA probe, in the absence of poly(dI-dC). The number in each lane corresponds to the lanes 1 to 4 from the SDS-polyacrylamide gel shown in panel A. The letter corresponds to the individual polypeptides (A to N) marked by arrows from top to bottom in each lane in the SDS-polyacrylamide gel (i.e., 1A corresponds to the top polypeptide A excised from lane 1 in the Coomassie brilliant blue-stained SDS-polyacrylamide gel in panel A). Samples numbered 1 to 4 represent 5 µl of each affinity-purified CBF sample prior to fractionation by SDS-PAGE. (C) Sequence specificity of the renatured polypeptides. Ten microliters of each of the renatured polypeptides was assayed, using the HA probe as described for panel B, in the absence (-) or presence of the HA or MUT competitor.

do not necessarily correspond to the same polypeptide in each lane (e.g., polypeptide 1E in lane 1, which is the fifth-largest polypeptide in this sample, is not equivalent to polypeptide 2E, the fifth-largest polypeptide in lane 2). In each lane, however, the two keratin proteins were the two largest proteins and thus are labeled A and B in each sample.

All of the polypeptides that were recovered from the SDS-polyacrylamide gel, in the size range of 19 to 35 kDa, had core-binding activity (Fig. 6B). These correspond to polypeptides 1D to 1H, 2D to 2G, 3D to 3N, and 4D to 4K in the mobility shift assay. Neither of the keratin proteins (polypeptides A and B in each lane) had core-binding activ-

ity. Polypeptide C, which is the third-largest polypeptide in each sample and migrates as a doublet at 42 kDa (large and small arrows), also had no core-binding activity in all but one sample (lane 1 from Fig. 6A; polypeptide 1C in Fig. 6C). We suspect that the 42-kDa doublet comigrates with a polypeptide with core-binding activity that was present in this particular CBF preparation. The remainder of the polypeptides (1D to 1H, 2D to 2G, 3D to 3N, and 4D to 4K) bind the core site. The mobility of the protein-DNA complexes increased with decreasing molecular weight of the renatured proteins, and the distance between complexes with the lowest and highest mobilities is approximately equivalent to the width of the protein-DNA complex before fractionation by SDS-PAGE (compare, for example, lanes 4D to 4K with lane 4). We also note that many polypeptides gave rise to two protein-DNA complexes with different mobilities. We suspect that the larger, lower-mobility protein-DNA complex results from binding of two molecules of CBF to the oligonucleotide probe. Competition assays established that each of the polypeptides with DNA-binding activity binds specifically to the core site (Fig. 6C).

Characterization of CBF binding sites. A methylation interference assay (25) was conducted to identify the purines in the core site that are contacted by CBF (Fig. 7). Dimethylsulfate methylates guanine at the N-7 position, which projects into the major groove. Methylation of three guanines on the top strand within the conserved core motif (CTGTGGTAA) interferes with protein-DNA complex formation (asterisks in lane C [protein-DNA complex]). In addition, we note that methylation at three adenines in the minor groove (arrows) also interferes with CBF binding, either by modifying the binding site for CBF or by distorting the DNA helix in the minor groove.

We also performed DNase I footprinting analyses (17) to confirm that CBF binds to the core site on the intact Moloney virus enhancer. The DNase I footprint generated by CBF is located directly over the core site in the 75-bp repeat of the Moloney MLV enhancer (Fig. 8A and B). Mutations in the core site that altered the thymic disease specificity of the Moloney virus to erythroleukemia (CTGTGGTAA \rightarrow CTGCCGTAA) significantly disrupt the binding of CBF to the Moloney MLV enhancer (Fig. 8C and D). There appears to be only one binding site for CBF in one copy of the Moloney virus direct repeat and thus presumably two sites in the intact enhancer. We have not assayed for CBF binding sites in the long terminal repeat outside of the direct repeat region.

Core consensus sequences have been identified in a large number of viral and cellular enhancers (19, 21, 52, 56, 59). We analyzed several enhancers by DNase I footprinting to determine which of these core motifs are bona fide binding sites for CBF (Fig. 9). CBF binds three core sites in the enhancer from the immunoglobulin μ -chain gene and two core sites in the enhancer from the TCRy gene (Fig. 9A and B). The core sites in the immunoglobulin μ-chain gene were originally identified by Gillies et al. (19) and were later shown not to contribute to the transcriptional activity of the immunoglobulin µ-chain enhancer in B cells (30). The contribution of these core sites to transcription from the immunoglobulin µ-chain enhancer in T cells has not been determined. The core sites in the TCRy gene are within a 200-bp minimal enhancer fragment shown by Spencer et al. (52) to be necessary for transcription in T cells.

CBF also binds one core site located between nucleotides -147 to -165 in the enhancer from SIVmac142 (Fig. 9C). This core site is located in a region of the SIVmac142

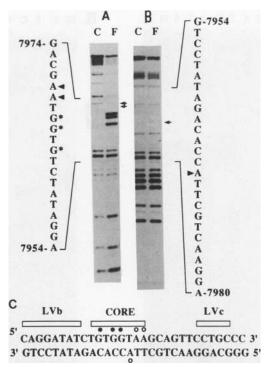


FIG. 7. Methylation interference analysis of CBF on the Moloney MLV enhancer. (A and B) CBF was purified from an HA affinity column as described in the legend to Fig. 4. SDS-PAGE analysis of the CBF preparation used in the methylation interference analysis is shown in Fig. 5B, lane 5. The probe used in the reaction contains a 45-bp HaellI-HaellI fragment from the Moloney MLV enhancer (nucleotides 7949 to 7994). A typical preparative reaction mixture contained 100,000 cpm of probe in a final volume of 50 µl, with 25 µl of affinity-purified CBF. (A) Top strand of the Moloney enhancer; (B) bottom strand. Lanes C and F correspond to the protein-DNA complex and free DNA bands, respectively. Asterisks indicate the locations of guanines whose methylation specifically inhibits the binding of CBF to the core site. Arrows indicate methylated adenines that interfered with CBF binding. (C) Summary of methylation interference results. ●, guanine contacts; O, adenines that when methylated interfere with CBF binding.

enhancer that was demonstrated by 5' deletion analysis to contribute to transcription of the enhancer in Rat-1 fibroblasts and to transcriptional stimulation by activators of protein kinase C in Jurkat T cells (45). A binding site for CBF is also present in the U3 region of the human immunodeficiency virus, at positions -391 to -408 relative to the transcriptional start site (42) (data not shown).

We also performed DNase I footprinting assays on the enhancer from the SV40 virus, which contains two core motifs, GT-I and GT-IIC. CBF did not protect either the GT-I or GT-IIC motif from DNase I digestion (data not shown). Thornell et al. (55) also found that the binding of the protein that they call SEF1 to the SL3-3 MLV core site could not be competed for by a 50-fold molar excess of either the GT-I or GT-IIC motif from the SV40 enhancer.

Figure 10 is a summary of the methylation interference and DNase I footprinting analyses presented in Fig. 7 to 9.

DISCUSSION

We have purified several polypeptides from calf thymus nuclei that bind the core site in the Moloney MLV enhancer. Analysis of affinity-purified CBF preparations by SDS-

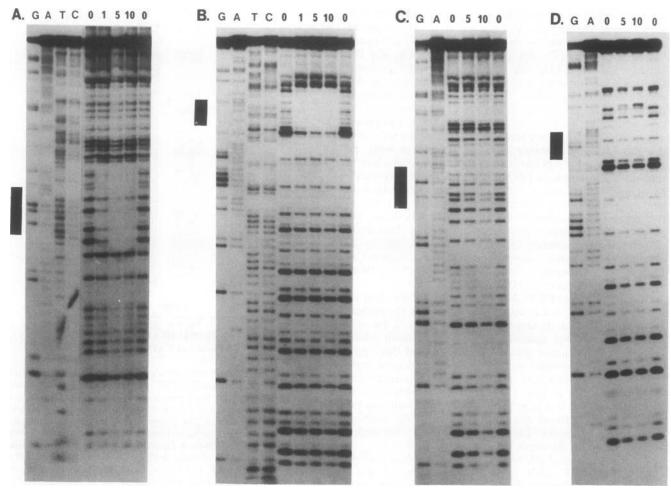


FIG. 8. Binding of CBF to the Moloney MLV enhancer. (A) Binding to the top strand of the Moloney virus enhancer. The Moloney MLV enhancer probe contains one copy of the 75-bp repeat of the enhancer, from PvuII₇₉₃₈ to PvuII₈₀₁₃. Each reaction mixture contained 10,000 cpm of end-labeled probe, 20 μg of BSA, and various amounts of CBF purified on the HA core site column. The SDS-PAGE analysis of the CBF preparation used in all DNase I footprinting analyses is shown in lane 5 of Fig. 5B. DNase I was added to the binding reaction mixture, and the partially cleaved Moloney MLV enhancer fragment was electrophoresed through a 10% polyacrylamide-7 M urea sequencing gel. Numbers above the lanes (0, 1, 5, and 10) represent the amounts (in microliters) of affinity-purified CBF added to the binding reaction mixture. G, A, T, and C represent chemical sequencing tracks. Boxes at the left of the sequence indicate the position of the core motif (CTGTGGTAA) in the Moloney MLV enhancer. (B) Binding of CBF to the bottom strand of the Moloney MLV enhancer. (C) Binding of CBF to the top strand of the PvuII₇₉₃₈-PvuII₈₀₁₃ fragment from the Moloney MLV enhancer, containing the 2-bp mutation in the core site (CTGCCGTAA) that altered the disease specificity of the Moloney virus (50). (D) Binding of CBF to the bottom strand of Moloney MLV enhancer with mutations in the core site.

PAGE and denaturation/renaturation experiments reveals more than 10 polypeptides that have core-binding activity. There are several potential sources of the multiple polypeptides in affinity-purified preparations of CBF. One possible explanation is that CBF is a family of transcription factors. These proteins could be encoded by different genes, as in the case of the octamer-binding proteins (27) or the AP-1 family of transcription factors (11). Alternatively, the different CBFs could result from multiple splicing of RNA transcribed from a single gene, as in the case of the family of CCAAT -box-binding proteins (CTF/NF1) (47) or the immunoglobulin transcription factors E2-5, E12, and E47 (26, 43).

Another explanation is that the multiple polypeptides represent proteolytic cleavage products of a single CBF precursor. We believe that proteolysis is a contributing factor to the multiplicity of polypeptides for several reasons. First, the heterogeneity in the composition of CBF preparations with respect to the number and abundance of polypep-

tides is suggestive of proteolytic cleavage that is occurring to variable extents from one purification to the next. Second, only the irreversible serine protease inhibitor PMSF was used in preparing nuclear extracts, and extracts were often made from previously frozen thymus tissue. When instead we used calf thymuses obtained within 10 min of slaughter and added a cocktail of protease inhibitors to the cell lysis buffer (PMSF [10^{-7} M], pepstatin A [$2.5 \mu g/ml$], leupeptin [5μg/ml], aprotinin [12 μg/ml], SBTI [soybean trypsin inhibitor] [12 μ g/ml], benzamidine [0.5 mg/ml], and E64 [5 mM]), we could detect several larger polypeptides in the affinitypurified CBF preparations by SDS-PAGE, in addition to the polypeptides seen in previous CBF preparations (data not shown). These larger polypeptides ranged in molecular mass from 32 to 41 kDa and also had core-binding activity, as determined by denaturation/renaturation analysis (data not shown). These two observations, the heterogeneity of polypeptides in different CBF preparations and the appearance of

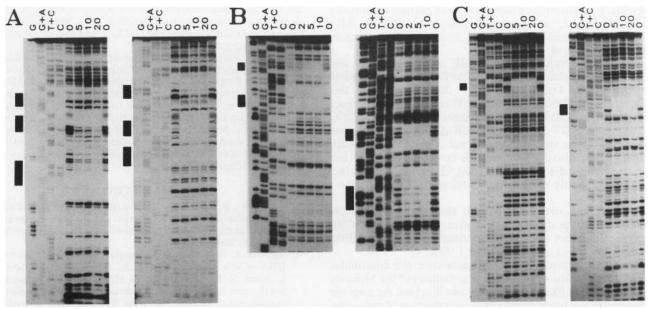


FIG. 9. Binding of CBF to the enhancers from the immunoglobulin μ-chain and TCRγ genes and from SIVmac142. (A) Binding to the core sites on the top and bottom strands of the immunoglobulin μ-chain enhancer. The immunoglobulin μ-chain enhancer probe contains a 136-bp sequence from PvuII₃₈₅ to DdeI₅₂₁ (numbering according to Gillies et al. [19]), encompassing the three core sites identified by Gillies et al. (19). Samples were electrophoresed through an 8% polyacrylamide-7 M urea sequencing gel. Assay conditions and symbols are the same as in Fig. 8. (B) Binding of CBF to the top and bottom strands of the TCRγ enhancer. The TCRγ enhancer probe contains a 196-bp restriction fragment (nucleotides 284 to 480, numbered according to Spencer et al. [52]). Samples were electrophoresed through an 8% polyacrylamide-7 M urea sequencing gel. (C) Binding of CBF to the top and bottom strands of the SIVmac142 enhancer. The SIVmac142 enhancer probe contains sequences from Sau3AI₋₂₅₈ to RsaI₋₇₁ (numbering relative to the transcriptional start site [7, 45]).

larger polypeptides upon addition of more protease inhibitors, suggest that proteolysis contributes to the multiplicity of polypeptides in CBF preparations. Proteolysis may not be the only contributing factor, however, and several distinct core-binding proteins may be present in affinity-purified CBF. Since inclusion of additional protease inhibitors dis-

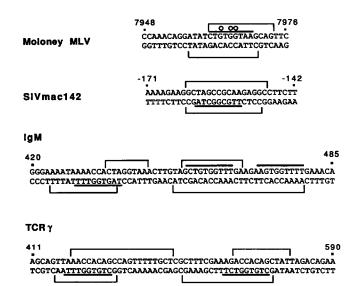


FIG. 10. Summary of CBF binding site analyses. Brackets indicate boundaries of protection by CBF from DNase I cleavage on the Moloney MLV, immunoglobulin μ -chain (IgM), TCR γ , and SIV-mac142 enhancers. The core sequences are indicated by heavy horizontal bars. Guanine contacts in the Moloney MLV enhancer core site are shown by open circles.

tributed CBF activity among more polypeptides, and our ultimate goal is to purify sufficient quantities of one or more polypeptides for microsequence analysis, we routinely include only PMSF during purification for CBF.

We have since obtained evidence that two of the CBF polypeptides, a 21- and a 24-kDa protein, are very closely related in primary structure. High-pressure liquid chromatography analysis of peptides obtained from tryptic digests of the 21- and 24-kDa proteins yielded superimposable peptide peaks, with the exception of one extra peptide peak in the larger (24-kDa) protein (58).

Two other proteins that bind the core consensus sequence in the Moloney virus enhancer have been purified to homogeneity: the CAAT/enhancer-binding protein (C/EBP) (29) and activating protein 3 (AP-3) (41). For several reasons, we believe that CBF is distinct from both C/EBP and AP-3. First, the tissue distribution of CBF activity does not coincide with the tissue-specific expression of C/EBP. We detect CBF activity in nuclear extracts prepared from calf thymus and spleen but not from liver, kidney, lung, or heart, using the WT Moloney enhancer core site as a probe (data not shown). The tissue distribution of core-binding activity in calf tissues is identical to that reported by Thornell et al. in mouse tissues (55). Thus, the core-binding activity that we and Thornell et al. detect appears to be confined to, or at least enriched in, hematopoietic tissues. C/EBP, on the other hand, was purified from rat liver and appears to be expressed in cells that metabolize lipids and cholesterol-related compounds at rapid rates, including liver, fat, intestine, lung, adrenal gland, and placenta, with no detectable expression in the spleen (3). Thus, in three tissues, liver, lung, and spleen, there is no overlap of C/EBP expression and CBF activity. The other protein that was demonstrated to bind the Molo-

ney MLV enhancer core site, AP-3, was purified from HeLa cells. However, an extensive characterization of AP-3 expression has not been reported; thus, we cannot compare its expression with that of CBF.

Another difference between CBF and both C/EBP and AP-3 is that AP-3 and C/EBP bind the GT-I motif in the SV40 enhancer (29, 41), whereas CBF does not protect this sequence from DNase I digestion. There is a 2-bp difference between the Moloney MLV enhancer core site (CTGTGG TAA) and the GT-1 motif (GTGTGGAAA); apparently one or both of these substitutions are not compatible with CBF binding. CBF also does not bind the homologous GT-IIC site on the SV40 enhancer (CTGTGGAAT); thus, it is distinct from the recently cloned transcriptional enhancer factor 1 (TEF-1) (64).

Finally, and most importantly, the sequence of a 26-amino-acid tryptic peptide fragment of the 24-kDa CBF polypeptide bears no significant homology to either C/EBP or transcriptional enhancer factor 1 (58).

Since the Moloney MLV enhancer core site contributes significantly to the thymic disease specificity of the Moloney virus, it is quite likely that the proteins that bind the core site also contribute to the transcriptional regulation of genes that are expressed in T cells. As a first step in testing this hypothesis, we have examined the enhancers from two cellular genes specifically transcribed in lymphocytes for CBF binding sites: the TCRy gene and the immunoglobulin μ-chain gene. We have identified two bona fide binding sites for CBF in the TCRy enhancer. These two sites are within a 196-bp fragment of the TCRy enhancer that comprise the minimal enhancer fragment (52). Both of these sites were protected from DNase I digestion by proteins in a 0.5 M KCl eluate of a T-cell nuclear extract from a heparin-Sepharose column (52). A third potential binding site for CBF in the TCRy enhancer is located 145 bp upstream of the two confirmed CBF binding sites (52); however, binding to this third site was not analyzed. A putative CBF binding site is also present in the enhancer from the TCRô gene, within a 250-bp region of the enhancer that contributes to transcription in T cells (44). Other investigators have similarly noted core consensus sequences in the enhancers from other genes specifically transcribed in T cells, including the $TCR\alpha$, TCRβ, CD3ε, and CD3δ genes (52, 56).

CBF also binds three closely spaced core sites in the immunoglobulin μ -chain enhancer. Although expression of the completely rearranged immunoglobulin heavy-chain gene (V_H - D_H - J_H) is confined to B lymphocytes, transcription from the unrearranged or partially rearranged (D_H - J_H) immunoglobulin heavy-chain locus also occurs in T cells (1, 32), and there is significant expression from introduced immunoglobulin heavy-chain genes in the thymus of transgenic mice (22, 53). CBF may thus contribute to the transcription of both the TCR γ and immunoglobulin μ -chain gene in T cells, but proof of this awaits functional analyses.

We believe that CBF is analogous to the factors called SEF1 by Thornell et al. (55) and S/A-CBF by Boral et al. (4). SEF1 has recently been purified from calf thymus; affinity-purified preparations of SEF1 also contain multiple polypeptides with core-binding activity (54).

ACKNOWLEDGMENTS

We thank Stan Froehner, Gus Lienhard, and Jack Bodwell for invaluable advice during the protein purification, Barbara Graves for many helpful discussions, and Jack Lenz for communicating unpublished results. We thank Winship Herr, Ranjan Sen, and David Raulet for kind gifts of plasmids, Steve Desiderio for the protocol for

calf thymus nuclear extracts, and the people at Green Mountain Meats for cooperation in providing calf thymus. We also thank Ed Levin for help with graphics, and Gus Lienhard and Ranjan Sen for critically reading the manuscript.

This work was supported by the Leukemia Research Foundation and also in part by BRSG S07RR05392 from the Biomedical Research Grant Program of Research Resources, National Institutes of Health, by Public Health Service grant CA51065-01A1 from the National Cancer Institute, and by a Cancer Center CORE grant (CA23018) from the National Cancer Institute to the Norris Cotton Cancer Center. N.A.S. is a recipient of a Junior Faculty Research Award from the American Cancer Society, and S.W. is an awardee of a predoctoral fellowship from the Norris Cotton Cancer Center at the Mary Hitchcock Hospital.

REFERENCES

- Alt, F. W., N. Rosenberg, V. Enea, E. Siden, and D. Baltimore. 1982. Multiple immunoglobulin heavy-chain gene transcripts in Abelson murine leukemia virus-transformed lymphoid cell lines. Mol. Cell. Biol. 2:386-400.
- 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, vol. 2. Greene Publishing Associates and Wiley-Interscience, New York.
- Birkenmeyer, E. H., B. Gwynn, S. Howard, J. Garry, 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.
- 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.
- Bradford, M. M. 1976. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. Anal. Biochem. 72:248-254.
- Celander, D., and W. A. Haseltine. 1984. Tissue-specific transcription preference as a determinant of cell tropism and leuke-mogenic potential of murine retroviruses. Nature (London) 312:159-162.
- Chakrabarti, L., M. Guyader, M. Alizon, M. D. Daniel, R. C. Desrosiers, P. Tiollais, and P. Sonigo. 1987. Sequence of simian immunodeficiency virus from macaque and its relationship to other human and simian retroviruses. Nature (London) 328:543

 547.
- Chatis, P. A., C. A. Holland, J. W. Hartiey, W. P. Rowe, and N. Hopkins. 1983. Role for the 3' end of the genome in determining disease specificity of Friend and Moloney murine leukemia viruses. Proc. Natl. Acad. Sci. USA 80:4408-4411.
- Chatis, P. A., C. A. Holland, J. E. Silver, T. N. Fredrickson, N. Hopkins, and J. W. Hartley. 1984. A 3' end fragment encompassing the transcriptional enhancer of nondefective Friend virus confers erythroleukemogenicity on Moloney murine leukemia virus. J. Virol. 52:248-259.
- Chodosh, L. 1989. Purification of DNA-binding proteins using biotin/streptavidin affinity systems, p. 12.6.5. 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, vol. 2. Greene Publishing Associates and Wiley-Interscience, New York.
- Curran, T., and B. R. Franza, Jr. 1988. Fos and jun: the AP-1 connection. Cell 55:395–397.
- DesGroseillers, L., and P. Jolicoeur. 1984. Mapping the viral sequences conferring leukemogenicity and disease specificity in Moloney and amphotropic murine leukemia viruses. J. Virol. 52:448-456.
- DesGroseillers, L., and P. Jolicoeur. 1984. The tandem direct repeats within the long terminal repeat of murine leukemia viruses are the primary determinant of their leukemogenic potential. J. Virol. 52:945-952.
- DesGroseillers, L., E. Rassart, and P. Jolicoeur. 1983. Thymotropism of murine leukemia virus is conferred by its long

- terminal repeat. Proc. Natl. Acad. Sci. USA 80:4203-4207.
- 15. DesGroseillers, L., R. Villemur, and P. Jolicoeur. 1984. The high leukemogenic potential of Gross passage A murine leukemia virus maps in the region of the genome corresponding to the long terminal repeat and to the 3' end of env. J. Virol. 47:24-32.
- Fried, M., and D. M. Crothers. 1981. Equilibria and kinetics of Lac repressor operator interactions by polyacrylamide gel electrophoresis. Nucleic Acids Res. 9:6505-6525.
- Galas, D., and A. Schmitz. 1978. DNase footprinting: a simple method for the detection of protein-DNA binding specificity. Nucleic Acids Res. 5:3157-3170.
- Garner, M. M., and A. Revzin. 1981. A gel electrophoresis method for quantifying the binding of proteins to specific DNA regions: application to components of the E. coli lactose operon regulatory system. Nucleic Acids Res. 9:3047-3060.
- Gillies, S. D., S. L. Morrison, V. T. Oi, and S. Tonegawa. 1983.
 A tissue specific transcription enhancer element is located in the major intron of a rearranged immunoglobulin heavy chain gene. Cell 33:717-728.
- Golemis, E., Y. Li, T. N. Fredrickson, J. W. Hartley, and N. Hopkins. 1989. Distinct segments in the enhancer region collaborate to specify the type of leukemia induced by nondefective Friend and Moloney viruses. J. Virol. 63:328-337.
- Golemis, E., N. A. Speck, and N. Hopkins. 1990. Alignment of U3 region sequences of mammalian type C viruses: identification of highly conserved motifs and implications for enhancer design. J. Virol. 64:534-542.
- Grosschedl, R., D. Weaver, D. Baltimore, and F. Constantini. 1984. Introduction of a μ immunoglobulin gene into the mouse germ line: specific expression in lymphoid cells and synthesis of functional antibody. Cell 38:647-658.
- 23. 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.
- 24. Halligan, B. D., and S. V. Desiderio. 1987. Identification of a DNA binding protein that recognizes the nonamer recombinational signal sequence of immunoglobulin genes. Proc. Natl. Acad. Sci. USA 84:7019-7023.
- Hendrickson, W., and R. F. Schlief. 1985. A dimer of AraC protein contacts three adjacent major groove regions of the Aral DNA site. Proc. Natl. Acad. Sci. USA 82:3129-3133.
- Henthorn, P., M. Kiledjian, and T. Kadesch. 1990. Two distinct transcription factors that bind the immunoglobulin enhancer μΕ5/κΕ2 motif. Science 247:467-470.
- Herr, W., R. A. Sturm, R. G. Clerc, L. Corcoran, D. Baltimore, P. A. Sharp, H. H. Ingraham, M. G. Rosenfeld, M. Finney, G. Ruvkin, and H. R. Horvitz. 1988. The POU domain: a large conserved region in the mammalian pit-1, oct-1, oct-2, and Caenorhabditis elegans unc-86 gene products. Genes Dev. 2:1513-1516.
- Ishimoto, A., A. Adachi, K. Sakai, and M. Matsuyama. 1985.
 Long terminal repeat of Friend-MCF virus contains the sequence responsible for erythroid leukemia. Virology 141:30-42.
- 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.
- Kadesch, T., P. Zervos, and D. Ruezinsky. 1986. Functional analysis of the murine IgH enhancer: evidence for negative control of cell-type specificity. Nucleic Acids Res. 14:8209-8221.
- Kadonaga, J. T., and R. Tjian. 1986. Affinity purification of sequence-specific DNA binding proteins. Proc. Natl. Acad. Sci. USA 83:5889-5893.
- Kemp, D. J., A. W. Harris, S. Cory, and J. M. Adams. 1980.
 Expression of the immunoglobulin Cμ gene in mouse T and B lymphoid and myeloid cell lines. Proc. Natl. Acad. Sci. USA 77:2876-2880.
- Laemmli, U. K. 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature (London) 227:680-685.

- Lenz, J. (Albert Einstein College of Medicine). Personal communication.
- Lenz, J., D. Celander, R. L. Crowther, R. Patarca, D. E. Perkins, and W. A. Haseltine. 1984. Determination of the leukaemogenicity of a murine retrovirus by sequences within the long terminal repeat. Nature (London) 308:467-470.
- Li, Y., E. Golemis, J. W. Hartley, and N. Hopkins. 1987. Disease specificity of nondefective Friend and Moloney murine leukemia viruses is controlled by a small number of nucleotides. J. Virol. 61:696-700.
- LoSardo, J. E., A. L. Boral, and J. Lenz. 1990. Relative importance of elements within the SL3-3 virus enhancer for T-cell specificity. J. Virol. 64:1756-1763.
- LoSardo, J. E., L. A. Cupelli, M. K. Short, J. W. Berman, and J. Lenz. 1989. Differences in activities of murine retroviral long terminal repeats in cytotoxic T lymphocytes and T-lymphoma cells. J. Virol. 63:1087-1094.
- Manley, N. R., M. A. O'Connell, P. A. Sharp, and N. Hopkins. 1989. Nuclear factors that bind to the enhancer region of nondefective Friend murine leukemia virus. J. Virol. 63:4210– 4223
- Maxam, A. M., and W. Gilbert. 1980. Sequencing end-labeled DNA with base-specific chemical cleavages. Methods Enzymol. 65:499-560.
- Mercurio, F., and M. Karin. 1989. Transcription factors AP-3 and AP-2 interact with the SV40 enhancer in a mutually exclusive manner. EMBO J. 8:1455-1460.
- Meusing, M. A., D. H. Smith, C. D. Cabradilla, C. V. Benton, L. A. Lashy, and D. J. Capon. 1985. Nucleic acid structure and expression of the human AIDS/lymphadenopathy retrovirus. Nature (London) 313:450-458.
- Murre, C., M. Schonleber, and D. Baltimore. 1989. A new DNA binding and dimerization motif in immunoglobulin enhancer binding, daughterless, MyoD, and myc proteins. Cell 56:777– 783
- Redondo, J. M., S. Hata, C. Brocklehurst, and M. S. Krangel. 1990. A T cell-specific transcriptional enhancer within the human T cell receptor δ locus. Science 247:1225-1229.
- Renjifo, B., N. A. Speck, S. Winandy, N. Hopkins, and Y. Li. 1990. cis-acting elements in the U3 region of a simian immunodeficiency virus. J. Virol. 64:3130-3134.
- Rosenfeld, P. J., and T. J. Kelly. 1986. Purification of nuclear factor 1 by DNA recognition site affinity chromatography. J. Biol. Chem. 261:1398-1408.
- 47. Santoro, C., N. Mermod, P. C. Andrews, and R. Tjian. 1988. A family of human CCAAT-box-binding proteins active in transcription and DNA replication: cloning and expression of multiple cDNAs. Nature (London) 334:218-224.
- Singh, H., R. Sen, D. Baltimore, and P. Sharp. 1986. A nuclear factor that binds to a conserved sequence motif in transcriptional control elements of immunoglobulin genes. Nature (London) 319:154-158.
- Speck, N. A., and D. Baltimore. 1987. Six distinct nuclear factors interact with the 75-base-pair repeat of the Moloney murine leukemia virus enhancer. Mol. Cell. Biol. 7:1101-1110.
- 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.
- 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.
- 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.
- 53. Storb, U., C. Pinkert, B. Arp, P. Engler, K. Gollahon, J. Manoz, W. Brady, and R. L. Brinster. 1986. Transgenic mice with μ and κ genes encoding anti-phophorylcholine antibodies. J. Exp. Med. 164:627-641.
- Thornell, A., and T. Grundstrom. 1991. Characterization of SEF1 transcriptional activators critical for T-lymphomagenicity

of murine leukemia viruses, abstr. 259, p. 259. Abstr. RNA Tumor Viruses.

- 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.
- 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.
- Vogt, M. 1982. Virus cloned from Rauscher virus complex induces erythroblastosis and thymic lymphoma. Virology 118: 226-236.
- 58. Wang, S., N. Speck, and W. Lane (Harvard Microchemistry Facility). Unpublished data.
- Weiher, H., M. Zonig, and P. Gruss. 1983. Multiple point mutations affecting the simian virus 40 enhancer. Science 219: 626-631.

- Weiss, R., N. Teich, H. Varmus, and J. Coffin (ed.). 1982. RNA tumor viruses, vol. 1. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
- Weiss, R., N. Teich, H. Varmus, and J. Coffin (ed.). 1985. RNA tumor viruses, vol. 2. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
- 62. 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.
- Wray, W., T. Bouklikas, V. P. Wray, and R. Hancock. 1981.
 Silver staining of proteins in polyacrylamide gels. Anal. Biochem. 118:197-203.
- 64. Xiao, J.-H., I. M. Davidson, J.-M. Garnier, and P. Chambon. 1991. Cloning, expression, and transcriptional properties of the human enhancer factor TEF-1. Cell 65:551-568.