

Dartmouth College

## Dartmouth Digital Commons

---

Dartmouth Scholarship

Faculty Work

---

4-1993

### Characterization of a Protein that Binds Multiple Sequences in Mammalian Type C Retrovirus Enhancers.

Wanwen Sun  
*Dartmouth College*

Mary M. O'Connell  
*Massachusetts Institute of Technology*

Nancy A. Speck  
*Dartmouth College*

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



Part of the [Medical Biochemistry Commons](#), and the [Virology Commons](#)

---

#### Dartmouth Digital Commons Citation

Sun, Wanwen; O'Connell, Mary M.; and Speck, Nancy A., "Characterization of a Protein that Binds Multiple Sequences in Mammalian Type C Retrovirus Enhancers." (1993). *Dartmouth Scholarship*. 1171.  
<https://digitalcommons.dartmouth.edu/facoa/1171>

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 Dartmouth Scholarship by an authorized administrator of Dartmouth Digital Commons. For more information, please contact [dartmouthdigitalcommons@groups.dartmouth.edu](mailto:dartmouthdigitalcommons@groups.dartmouth.edu).

## Characterization of a Protein That Binds Multiple Sequences in Mammalian Type C Retrovirus Enhancers

WANWEN SUN,<sup>1</sup> MARY O'CONNELL,<sup>2†</sup> AND NANCY A. SPECK<sup>1\*</sup>

*Department of Biochemistry, Dartmouth Medical School, Hanover, New Hampshire 03755,<sup>1</sup> and Center for Cancer Research, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139<sup>2</sup>*

Received 3 September 1992/Accepted 29 December 1992

**Mammalian type C retrovirus enhancer factor 1 (MCREF-1) is a nuclear protein that binds several directly repeated sequences (CNGGN<sub>6</sub>CNCG) in the Moloney and Friend murine leukemia virus (MLV) enhancers (N. R. Manley, M. O'Connell, W. Sun, N. A. Speck, and N. Hopkins, *J. Virol.* 67:1967–1975, 1993). In this paper, we describe the partial purification of MCREF-1 from calf thymus nuclei and further characterize the binding properties of MCREF-1. MCREF-1 binds four sites in the Moloney MLV enhancer and three sites in the Friend MLV enhancer. Ethylation interference analysis suggests that the MCREF-1 binding site spans two adjacent minor grooves of DNA.**

The murine type C retroviruses cause a variety of different diseases in their infected hosts. Multiple regions of the retroviral genome determine the pathogenic phenotype of a virus, and many investigators have identified viral genetic determinants that influence pathogenesis and continue to study the mechanisms by which these genetic regions exert their effects. One viral genetic determinant that has been shown in a number of studies to influence pathogenesis is the retroviral enhancer. In several replication-competent murine leukemia viruses (MLVs), the enhancer has been shown to influence such properties as the leukemogenicity of the virus, the latent period of disease onset, the organ tropism, and the disease specificity (5, 6, 8–11, 18, 26, 27).

An 18-bp sequence is particularly well conserved among mammalian type C retroviral enhancers [AAACAGGATA TCTG(T/C)GGT] (14). This sequence contains both the leukemia virus factor b (LVb) site (CAGGAT) and the core site [TG(T/C)GGT] (38). The LVb and core sites are important viral genetic determinants for pathogenesis of the SL3-3 MLV and Moloney MLV. Point mutations introduced into the core site in both the SL3-3 and Moloney MLV enhancers increased the latent period of disease onset by these viruses and, in the case of Moloney, altered disease specificity from thymic to predominantly erythroid leukemia (17, 39). Mutations in the LVb site in the Moloney MLV also caused an increase in the latent period and a small shift in disease specificity to erythroleukemia (39).

The LVb site was originally defined as the binding site for a protein found in crude B-cell nuclear extracts, LVb. It soon became apparent that more than one protein binds to the LVb site; thus, in this paper we use the name LVb to define a DNA sequence (CAGGAT) within the enhancer. The LVb site binds the Ets proteins Ets-1 and Ets-2 (15, 34) as well as a protein called LVt (28). The adjacent core site was first identified as a conserved sequence element in several viral enhancers (43); it also binds multiple proteins, including the CAAT/enhancer binding protein (C/EBP) (20), activating protein 3 (AP3) (32), SL3 core binding factor (S-CBF) (3), AKV core binding factor (A-CBF) (3), and a

protein known under several aliases, including the polyoma-virus enhancer binding factor 2 (PEBP2) (35), the SL3 enhancer factor 1 (SEF1) (40), the SL3 and AKV core binding factor (S/A-CBF) (3), and, simply, core binding factor (CBF) (42).

In the accompanying paper, we describe a protein in crude nuclear extracts that binds sequences in both the LVb and core sites, called mammalian type C retrovirus enhancer factor 1 (MCREF-1) (28). MCREF-1 also binds several sequences in the Friend MLV enhancer: the Friend virus factor a (FVa) and FVb1 sites, three sequences in the Moloney MLV enhancer, and one site in the GC-rich region immediately 3' to the Moloney MLV enhancer direct repeat. Here we describe the partial purification of MCREF-1 from calf thymus and further characterize the binding properties of MCREF-1.

### MATERIALS AND METHODS

**Biochemical assays.** (i) **Substrates for protein binding in electrophoretic mobility shift assays.** The origins of the oligonucleotides from the Moloney and Friend MLV enhancers that were used in biochemical assays are shown in Fig. 1, and the sequences are listed in Table 1. Complementary oligonucleotides were synthesized with the Biosearch Cyclone DNA Synthesizer at Dartmouth Medical School and an Autogen 6500 DNA synthesizer at the Center for Cancer Research, Massachusetts Institute of Technology. All oligonucleotides were purified by electrophoresis through 20% polyacrylamide–7 M urea gels.

Radioactive probes were made by labeling 100 pmol of one oligonucleotide, either the plus or minus strand of the binding site, with [ $\gamma$ -<sup>32</sup>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 (23). The double-stranded probes were then purified by electrophoresis through 20% native polyacrylamide gels. The specific activity of the probes was typically 4,000 to 10,000 cpm/fmol.

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

(ii) **Protein-DNA binding analysis.** MCREF-1 binding ac-

\* Corresponding author.

† Present address: Biozentrum der Universität Basel, CH-4056 Basel, Switzerland.

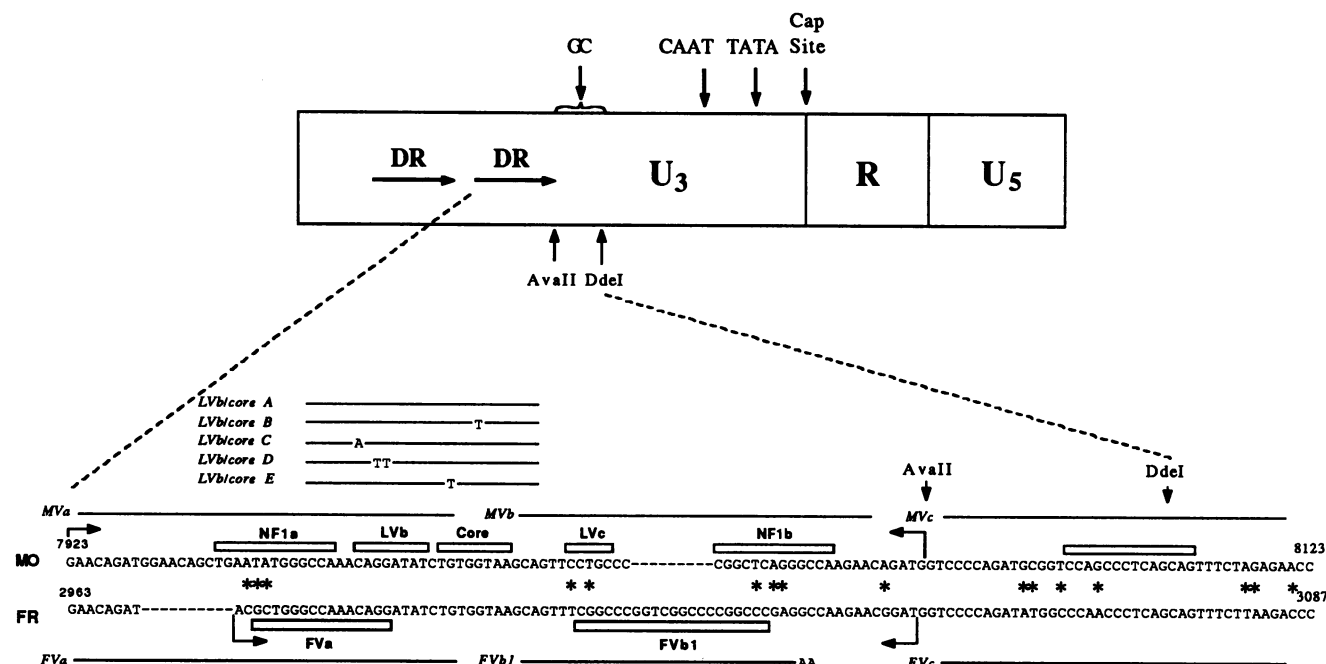


FIG. 1. Origins of oligonucleotides from the Moloney and Friend MLV enhancers used to detect MCREP-1. The upper portion of the figure is a schematic representation of the Moloney and Friend MLV long terminal repeat, showing the location of the enhancer direct repeat (DR) and adjacent GC-rich region. The sequences underneath are from the second copy of the direct repeat of the Moloney (MO) and Friend (FR) enhancers. Numbering of the Moloney MLV enhancer sequence is from the cap site at the 5' end of the viral genome, and the Friend MLV enhancer is numbered from the 5' end of the *env* gene, both by the numbering of Weiss et al. (44). Asterisks between the Moloney and Friend enhancer sequences indicate positions of sequence divergence. Dashes within the sequence indicate gaps in the alignment between the two enhancers. Binding sites for nuclear factors (NF1a and NF1b, designating the sites 5' and 3' to the LVb site, respectively; LVb; the core site; LVC; and sites for FVa and FVb1) are indicated by horizontal boxes above the sequence (29, 38). The 5' and 3' boundaries of the oligonucleotides used in the analysis derived from the Moloney virus enhancer are indicated by horizontal lines above the Moloney enhancer sequence. Oligonucleotides derived from the Friend MLV enhancer are shown below the Friend enhancer sequence. The LVb-core oligonucleotides are derived from sequences that are conserved between the Moloney and Friend MLV enhancers. Oligonucleotides containing point mutations in the LVb or core binding sites and an alteration in the 3' end of the FVb1 site are indicated by the substituted nucleotide within the horizontal line.

tivity was detected by the electrophoretic mobility shift assay (12, 13, 37). Binding reactions contained 10,000 cpm (2 to 5 fmol) of  $^{32}$ P-end-labeled probe, binding buffer (100 mM NaCl, 10 mM Tris [pH 7.4], 1 mM  $\beta$ -mercaptoethanol, 1 mM EDTA, 4% glycerol), 0.2 to 1.0  $\mu$ g of poly(dI-dC)-poly(dI-dC) (Pharmacia), and 1 to 10  $\mu$ l of protein sample, in a total volume of 15  $\mu$ l. To demonstrate the sequence specificity of

the protein-DNA complex, various amounts of unlabeled oligonucleotides were included in some of the binding reactions. After 15 min of incubation at room temperature, the reaction mixtures were fractionated by electrophoresis through a 5 to 6% native polyacrylamide gel containing 0.5 $\times$  TBE (22.5 mM Tris-HCl, 22.5 mM boric acid, 0.5 mM EDTA). Radioactivity was detected by autoradiography of

TABLE 1. Sequences of oligonucleotides used in the biochemical assays

Name	Sequence	Source	Coordinates
LVb/core A	CCAAACAGGATATCTGTGGTAAGCA	Moloney	7949-7973,8024-8048
LVb/core B	CCAAACAGGATATCTGTGTTAAGCA	Friend	2914-2938,2979-3003
LVb/core C	CCAAAAGGATATCTGTGGTAAGCA		
LVb/core D	CCAAACATTATATCTGTGGTAAGCA		
LVb/core E	CCAAACAGGATATCTTTGGTAAGCA		
MVa	AACAGATGGAACAGCTGAATATGGGCCAAACAGGATATCTG	Moloney	7923-7963,7999-8039
MVb	CAGTTCTGCCCCGGCTCAGGGCCAAGAA	Moloney	7972-8000,8047-8075
FVa	AACAGATACGCTGGGCCAAACAGGATATCTG	Friend	2964-2994
FVb1	CAGTTTCGGCCCGGTCGGCCCCGGCCGAGaa	Friend	3002-3031
MVc	CCCCAGATGCGGTCCAGCCCTCAGCAGTTTCTAGAGA	Moloney	8085-8120
FVc	CCCCAGATATGGCCCAACCCTCAGCAGTTTCTTAAGA	Friend	3048-3084
Ad-NF1	TTTTGGATTGAAGCCAATATGAG	Ad2 <sup>a</sup>	+22-+44 (left terminus)

<sup>a</sup> Adenovirus type 2 origin of replication, with point mutation at +28, as described by Chodosh et al. (7).

the dried gel. The amount of DNA in specific protein-DNA complexes was quantified by scintillation spectrometry of the protein-DNA complex bands excised from dried polyacrylamide gels, counted in the presence of Ecoscint A (National Diagnostics).

**Partial purification of MCREF-1. (i) Preparation of nuclear extracts.** Nuclear extracts were prepared from previously frozen calf thymus as described by Wang and Speck (42), with the following modification: after extracting proteins from the nuclear pellet in buffer E (250 mM sucrose, 400 mM NaCl, 50 mM *N*-2-hydroxyethylpiperazine-*N'*-2-ethanesulfonic acid [HEPES; pH 7.5], 10 mM  $\beta$ -mercaptoethanol, 1 mM phenylmethylsulfonyl fluoride) and collecting the nuclei by centrifugation, the proteins in the supernatant were precipitated from the nuclear extract by slow addition of ammonium sulfate to a final concentration of 35%, instead of 50% as reported previously for CBF (42). MCREF-1 was quantitatively precipitated with 0 to 35% ammonium sulfate (data not shown). The ammonium sulfate precipitate was collected by centrifugation, resuspended, and dialyzed against buffer B (50 mM NaCl, 20 mM HEPES [pH 7.5], 10 mM  $\beta$ -mercaptoethanol, 2 mM EDTA, 10% glycerol), as described previously (42). Typical yields were 0.5 g of nuclear extract per 100 g of calf thymus tissue.

**(ii) DE52 chromatography.** The resuspended and dialyzed 0 to 35% ammonium sulfate precipitate (2,200 mg of protein) was applied directly to a DE52 cellulose (Whatman) column (2.5 by 20 cm, 100 ml) that was equilibrated in buffer B. The column was washed with 3 column volumes of buffer B and then developed with a linear gradient of [NaCl] from 50 mM to 1.0 M in buffer B. Fractions (10 ml) were collected and assayed for MCREF-1 activity by electrophoretic mobility shift assays. Fractions were also analyzed for protein concentration and [NaCl]. Protein concentrations were determined by the method of Bradford (4) with reagents purchased from Bio-Rad. The concentration of NaCl in column fractions was determined by measuring conductivity in comparison to a standard curve.

Active fractions were pooled and dialyzed against buffer B (50 mM NaCl). The DE52 cellulose resin was regenerated with 5 column volumes of buffer B plus 2.5 M NaCl.

**(iii) Heparin-Sepharose chromatography.** Dialyzed pooled fractions from the DE52 column were applied onto a heparin-Sepharose column (2.5 by 13 cm, 70 ml) that was equilibrated with buffer B. The column was washed with 3 column volumes of buffer B and then developed with a 500-ml linear gradient of [NaCl] from 50 mM to 1.0 M in buffer B. Approximately 5.5-ml fractions were collected. Fractions were analyzed for MCREF-1 activity, protein concentration, and [NaCl]. Fractions with MCREF-1 activity were pooled and dialyzed against buffer B. Heparin-Sepharose was regenerated according to the manufacturer (Pharmacia).

**Recovery and renaturation of protein from SDS-polyacrylamide gels.** Sodium dodecyl sulfate (SDS)-polyacrylamide gel electrophoresis (PAGE) was performed as described by Laemmli (25). Protein samples (3 ml, 3.75 mg of protein) were precipitated with 9.1% trichloroacetic acid in the presence of 0.14 mg of deoxycholate per ml as a carrier. Denaturation-renaturation experiments were performed as described by Hager and Burgess (16), with several modifications. Following electrophoresis through a 9.3-cm SDS-8% polyacrylamide gel, the lanes containing the molecular weight markers were excised from the gel and stained with Coomassie brilliant blue. The remainder of the gel was then sliced from top to bottom into 22 equal 4.2-mm slices,

and each slice was then chopped into smaller pieces. Proteins were eluted from each gel slice by incubation in 200 to 300  $\mu$ l of a buffer containing 150 mM NaCl, 20 mM HEPES (pH 7.5), 5 mM dithiothreitol, 0.1 mM EDTA, 0.1% SDS, and 0.1 mg of bovine serum albumin per ml overnight at room temperature. The eluted protein was precipitated with 4 volumes of cold ( $-20^{\circ}\text{C}$ ) acetone and collected by centrifugation ( $16,000 \times g$  for 30 min). The pellet was washed with a solution containing 80% acetone-20% dilution buffer (150 mM NaCl, 20 mM HEPES [pH 7.5], 1 mM dithiothreitol, 0.1 mM EDTA, and 20% glycerol), dried, dissolved in 5  $\mu$ l of dilution buffer supplemented with 6 M guanidine-HCl, and incubated at room temperature for 30 min. The protein was then renatured by dilution with 250  $\mu$ l of dilution buffer without guanidine-HCl at  $4^{\circ}\text{C}$  overnight. Ten microliters of each protein sample was assayed for MCREF-1 binding activity by electrophoretic mobility shift assay, in the presence of 0.1  $\mu$ g of poly(dI-dC)-poly(dI-dC) (the protein eluted from the last [bottom] slice, slice 22, was not tested for binding activity).

**Methylation and ethylation interference analyses.** Binding reaction mixtures for methylation interference analysis (75  $\mu$ l total volume) contained 100,000 cpm of end-labeled, methylated (31) probe, 1 to 5  $\mu$ g of poly(dI-dC)-poly(dI-dC), 15  $\mu$ l (19  $\mu$ g) of partially purified protein from pooled fractions from the heparin-Sepharose column, and 250 ng of unlabeled, nonspecific competitor DNA (LVb/core D [Fig. 1 and Table 1]) in the same binding buffer used for electrophoretic mobility shift assays. The binding reaction was electrophoresed through a 5% native polyacrylamide gel in  $0.5 \times$  TBE. 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) (2), subjected to 1 M piperidine cleavage, and analyzed by electrophoresis through a 15% polyacrylamide-7 M urea sequencing gel.

The probes for ethylation interference analysis were modified on phosphates with *N*-ethyl-*N*-nitrosourea (Sigma), as described by Siebenlist and Gilbert (36). The binding reactions, electrophoresis, and purification of the DNA from the native polyacrylamide gels were performed as described for methylation interference assays. Purified DNA from the protein-DNA complex and free DNA bands from the mobility shift assay were cleaved by alkali (36), and the cleavage products were analyzed by electrophoresis through a 15% polyacrylamide-7 M urea sequencing gel.

## RESULTS

**Assay for MCREF-1 activity.** We assayed for MCREF-1 by its ability to bind specifically the LVb-core region in the Moloney and Friend MLV enhancers by electrophoretic mobility shift assays (12, 13, 37). The sequences of the oligonucleotides that were used as probes and/or competitors are shown in Fig. 1 and Table 1. The probe is a 25-bp synthetic oligonucleotide (LVb/core E) that contains the wild-type LVb site and a mutated core site. Methylation interference analyses showed that this mutation in the core site (CCAAACAGGATATCTGTGGTAAGCA  $\rightarrow$  CCAAACAGGATATCTTTGGTAAGCA) disrupts binding of CBF (42) but does not affect MCREF-1 binding (28). The LVb/core E probe enables us to more easily detect MCREF-1 above the background of abundant CBF in calf thymus nuclear extracts. We performed competition assays on active fractions at each step of the purification with the

unlabeled LVb/core E oligonucleotide and with the double-stranded LVb/core D oligonucleotide that contains mutations in two guanines in the LVb site (CCAAACAT TATATCTGTGGTAAGCA). Together, competition with the LVb/core E and LVb/core D oligonucleotides identifies proteins that specifically bind sequences in the LVb site.

**Partial purification of MCREF-1.** An extract prepared from calf thymus nuclei was first fractionated on a DE52 cellulose column (Fig. 2A to C). Most of the MCREF-1 activity eluted from the DE52 cellulose column between 60 and 150 mM NaCl (fractions 57 to 65 [Fig. 2A and B]). There are two predominant protein-DNA complexes generated from the proteins in fractions 57 to 65 that appear to require sequences in the LVb site for binding (Fig. 2C). The MCREF-1 protein-DNA complex has a relatively low mobility (Fig. 2C, lanes 1, 4, and 7, open circle). The second protein-DNA complex has a much higher mobility (Fig. 2C, lane 4, solid circle). We have not characterized the protein that generates the higher-mobility protein-DNA complex and monitored only the low-mobility MCREF-1 protein-DNA complex throughout the purification.

The flowthrough fractions also contain proteins that bind to the LVb/core E probe (fractions 19 to 33 [Fig. 2A and B]), but MCREF-1 is a relatively minor component of this activity (Fig. 2C, lanes 1 to 3). Most of the activity in flowthrough fractions is not specific for the LVb site, since the formation of the intense, diffuse protein-DNA complex can be inhibited by the LVb/core D oligonucleotide (Fig. 2C, lane 3). The amount of MCREF-1 activity in the flowthrough fractions is difficult to quantify because of the high background of other proteins binding specifically or nonspecifically to the LVb/core E probe. Most of the binding activity in fractions 69 to 85 is also not specific for sequences in the LVb site (Fig. 2C, lanes 7 to 9), and only a small amount of MCREF-1 is present in these fractions.

We pooled fractions 57 to 65 from the DE52 cellulose column, which contain most of the MCREF-1 activity, and dialyzed these pooled fractions against buffer B (50 mM NaCl). Pooled fractions 57 to 65 from the DE52 cellulose column were subsequently chromatographed on a heparin-Sepharose column (Fig. 2D to F). Most of the MCREF-1 activity eluted from the heparin-Sepharose column between 380 and 470 mM NaCl (fractions 79 to 91). A small amount of MCREF-1 was present in flowthrough fractions 11 to 15 (Fig. 2E and F, lanes 1 to 3). Two proteins that gave rise to higher-mobility protein-DNA complexes specific for the LVb site also eluted from the heparin-Sepharose column in fractions 79 to 89 and 91 to 97 (Fig. 2E). The higher-mobility protein-DNA complex from fractions 91 to 97 (Fig. 2F, lanes 7 to 9, solid circle) has the same relative mobility as the protein-DNA complex generated from fractions 57 to 65 of the DE52 cellulose column (Fig. 2C, lanes 4 to 6, solid circle). The other protein that specifically binds the LVb site and elutes from the heparin-Sepharose column in fractions 79 to 89 (Fig. 2F, lanes 4 to 6, open triangle) cannot be clearly seen in the pooled fractions from the DE52 cellulose column.

We pooled fractions from the heparin-Sepharose column containing most of the MCREF-1 activity (fractions 78 to 93) and used these pooled fractions in all subsequent analyses. We refer to these pooled fractions as the heparin-Sepharose fraction. We achieved an approximately 18.7-fold purification of MCREF-1 following chromatography on both DE52 cellulose and heparin-Sepharose columns.

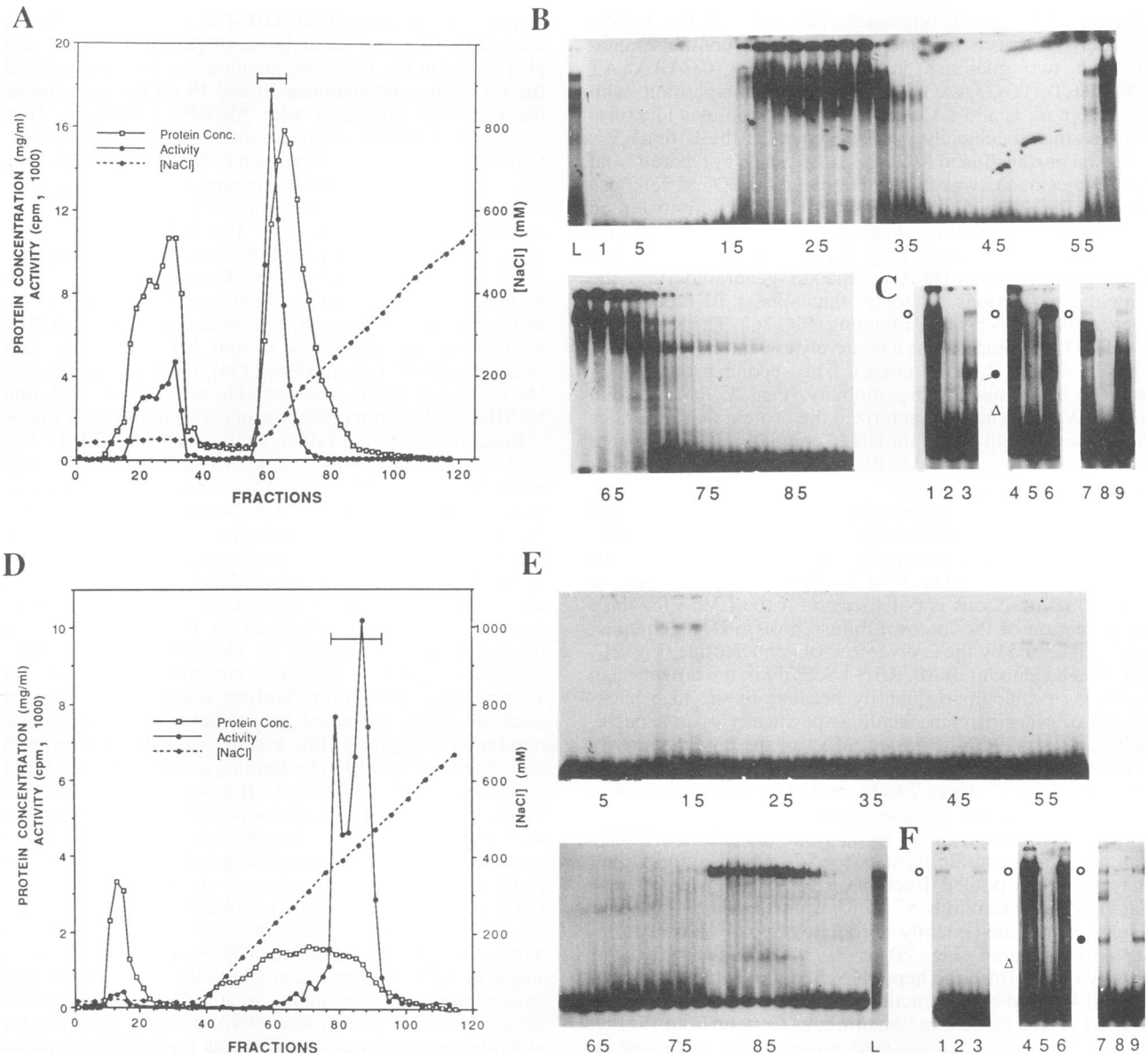
**MCREF-1 recognizes sequences in both the LVb and core sites.** We performed a methylation interference analysis to

characterize binding of MCREF-1 to sequences in the LVb/core probe (Fig. 3). Methylation of guanines 8 and 9 on the plus strand in the LVb site, guanine 6 in the minus strand in the LVb site, and guanines 18 and 19 on the plus strand of the core site interferes with MCREF-1 binding. This is consistent with the consensus site defined by Manley et al. (28) on the basis of a comparison of MCREF-1 binding sites. The two pairs of guanines on the plus strand are separated by 8 bp, or one turn of the B-form DNA helix from guanines 8 and 9 to guanines 18 and 19. The MCREF-1 methylation interference pattern is distinct from that generated by Ets-1, Ets-2, or CBF (Fig. 3). Ets-1 and Ets-2 binding are disrupted by methylation of guanines 8 and 9 on the plus strand and 6 and 14 on the minus strand. The contacts made by MCREF-1 in the core site (guanines 18 and 19) are outside of the binding site for Ets-1 and Ets-2 (34). CBF contacts guanines 16, 18, and 19 on the plus strand in the core site and, unlike MCREF-1, does not contact sequences in the LVb site (42).

**Denaturation-renaturation analysis of MCREF-1.** Previously identified proteins that bind to the LVb and core region of the Moloney or SL3-3 MLV enhancers specifically recognize sequences in either the LVb site or the core site (21, 28, 34, 40-42). To determine whether the MCREF-1 protein-DNA complex comprises two distinct proteins, one binding to the LVb site and the other binding to the core site, we attempted to separate putative LVb and core binding proteins by SDS-PAGE. Proteins in the heparin-Sepharose fraction were fractionated by electrophoresis through an SDS-polyacrylamide gel. We cut the gel into 22 equal 4.2-mm slices from top to bottom, eluted the proteins from each gel slice, subjected the proteins to a denaturation-renaturation regimen (16), and assayed the renatured proteins from each gel slice for binding to the LVb/core A probe by electrophoretic mobility shift assay. We reasoned that if MCREF-1 consists of two cooperatively binding proteins, such as a member of the Ets protein family and CBF, then proteins isolated from individual gel slices either should not yield a protein-DNA complex or should generate protein-DNA complexes with higher mobilities than the relatively low-mobility MCREF-1 protein-DNA complex. On the other hand, if the MCREF-1 protein-DNA complex consists of a single protein, or a multimeric protein comprising homologous subunits, then proteins isolated from a single gel slice should give rise to a protein-DNA complex with the same mobility and sequence specificity as the MCREF-1 protein-DNA complex generated by proteins in the heparin-Sepharose fraction prior to separation by SDS-PAGE.

Three adjacent slices from the SDS-polyacrylamide gel contained proteins yielding a protein-DNA complex with a mobility similar to that obtained with the heparin-Sepharose fraction (Fig. 4A, gel slices 11 to 13). These proteins migrated in the SDS-polyacrylamide gel with an apparent molecular mass of 50 to 70 kDa. We do not know whether these are three distinct proteins or different proteolytic breakdown products of a common larger protein. The minimal molecular mass for MCREF-1 (50 kDa) is greater than that of the largest CBF polypeptide previously identified (35 kDa) (42).

We assayed for binding specificity of the renatured proteins by competition analysis with the LVb/core A oligonucleotide as a probe and a series of oligonucleotides containing mutations at one or more guanines in the LVb and core sites to compete for binding of MCREF-1 to the LVb/core A probe (LVb/core B to E [Fig. 1 and Table 1]). The competition pattern of the MCREF-1 protein-DNA complex from the SDS-PAGE-fractionated and renatured proteins is indis-



**FIG. 2.** Fractionation of MCREF-1. (A to C) Chromatography on DE52 cellulose. Nuclear extract was loaded onto a DE52 cellulose column. The column was developed with a linear gradient of [NaCl] from 50 mM to 1.0 M. (A) DE52 cellulose column profile. Alternate fractions were assayed for protein concentration, [NaCl], and MCREF-1 activity with the LVb/core E probe. The bracketed horizontal line indicates the fractions that were pooled and loaded onto the next column. (B) Electrophoretic mobility shift assay of column fractions. Fraction numbers are indicated on the bottom. Lanes: L, 10  $\mu$ l of nuclear extract; 1 to 91, 10  $\mu$ l of elution fractions from the DE52 cellulose column. (C) Competition analysis of binding activity in various fractions from the DE52 cellulose column. Lanes: 1 to 3, flowthrough fraction 29; 4 to 6, fraction 63; 7 to 9, fraction 69. The first lane for each fraction (1, 4, and 7) represents binding to the LVb/core E probe. The second lane (2, 5, and 8) represents binding to the LVb/core E probe in the presence of 50 ng of unlabeled double-stranded LVb/core E oligonucleotide. The third lane (3, 6, and 9) represents binding to the LVb/core E probe in the presence of 50 ng of unlabeled LVb/core D. Open circle, low-mobility MCREF-1 protein-DNA complex; closed circle, position of a distinct higher-mobility LVb-specific protein-DNA complex. (D to F) Chromatography on heparin-Sepharose. (D) Pooled fractions (57 to 65) from the DE52 cellulose column were dialyzed against buffer B and loaded onto a heparin-Sepharose column. Alternate fractions were assayed for specific DNA binding activity, for protein concentration, and for [NaCl]. The bracketed horizontal line indicates the pooled heparin-Sepharose fractions. (E) Electrophoretic mobility shift assay of column fractions. Lanes: L, 3  $\mu$ l of pooled DE52 fractions 57 to 65 that were loaded onto the column; 1 to 97, elution fractions from the heparin-Sepharose column. (F) Competition analysis of binding activity in selected fractions from the heparin-Sepharose column. Lanes: 1 to 3, flowthrough fraction 15; 4 to 6, fraction 83; 7 to 9, fraction 93. The first lane for each fraction (1, 4, and 7) represents binding to the LVb/core E probe. The second lane (2, 5, and 8) represents binding to the LVb/core E probe in the presence of 50 ng of unlabeled LVb/core E. The third lane (3, 6, and 9) represents binding to the LVb/core E probe in the presence of 50 ng of unlabeled LVb/core D. Open circle, low-mobility MCREF-1 protein-DNA complex; closed circle and triangle, positions of other higher-mobility LVb-specific protein-DNA complexes.

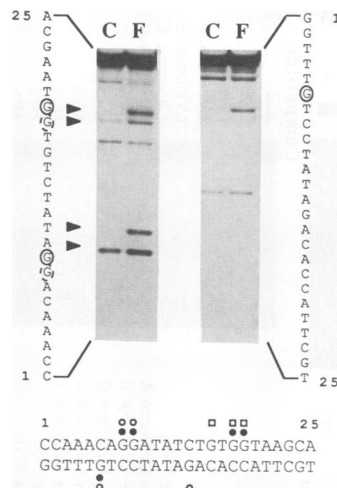


FIG. 3. Methylation interference analysis of MCREF-1 on the LVb/core A probe. (A) The autoradiogram of the methylation interference assay is in the upper portion of the figure. Methylation interference was performed with the heparin-Sepharose fraction on methylated probe LVb/core A. Left panel, plus strand of the LVb/core probe A; right panel, minus strand. Lanes C and F correspond to the protein-DNA complex and free DNA bands, respectively. Arrows show the methylated guanines specifically depleted from the DNA in the protein-DNA complex. Circles in the vertical sequence indicate the location of guanines whose methylation specifically inhibits MCREF-1 binding to the LVb/core A probe. Closed circles (solid line) indicate complete interference; broken circles (dashed lines) indicate partial interference. The bottom portion of the figure summarizes the methylation interference results. The sequence is numbered 1 to 25 according to the 5' end of the plus strand of the LVb/core A probe. Closed circles, guanine contacts for MCREF-1; open circles, contacts for Ets-1 (34); open squares, contacts for CBF (42).

tinguishable from that of the MCREF-1 protein-DNA complex obtained from the heparin-Sepharose fraction (Fig. 4B). Oligonucleotides with mutations in guanine 19 (plus strand, LVb/core B), guanine 6 (minus strand, LVb/core C), or guanines 8 and 9 (plus strand, LVb/core D) do not compete effectively for binding of MCREF-1 to the LVb/core A probe. An LVb-core oligonucleotide with a mutation in guanine 16 (plus strand, LVb/core E) does compete for binding of MCREF-1 to the LVb/core A probe. The competition pattern is consistent with the methylation interference data (Fig. 3).

These results suggest that the MCREF-1 protein-DNA complex results from the binding of a single monomeric protein or a multimeric protein composed of homologous subunits to sequences in both the LVb and core sites. However, MCREF-1 could also be a heterodimeric protein consisting of subunits with similar molecular weights.

**MCREF-1 binds to multiple sites in both the Moloney and Friend MLV enhancers.** Binding of partially purified bovine MCREF-1 to the LVb-core sequence can be inhibited by two oligonucleotides from the Friend MLV enhancer containing the FVa and FVb1 sites (28). Here we extend this analysis to show the direct binding of bovine MCREF-1 to the FVa and FVb1 sites and also to the corresponding regions in the Moloney virus enhancer. Binding of MCREF-1 to the LVb/core A probe can be specifically inhibited with oligonucleotides containing the FVa and FVb1 sequences from the

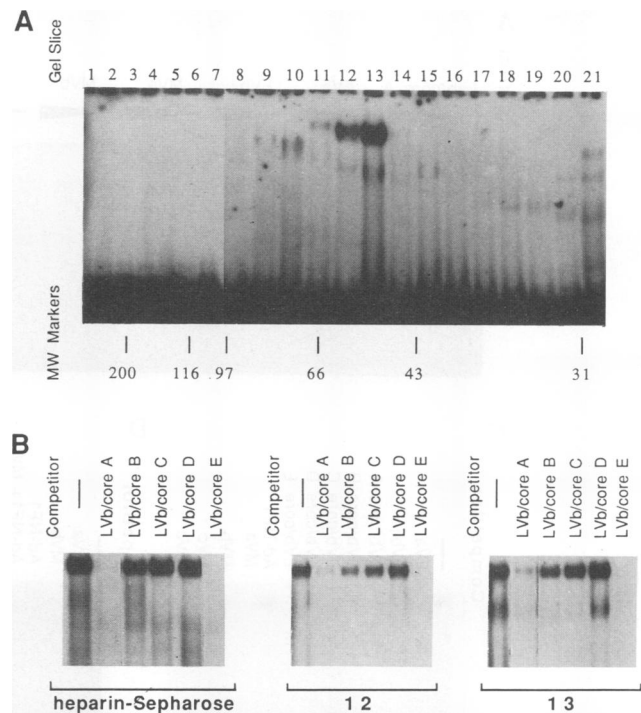


FIG. 4. Denaturation-renaturation analysis of MCREF-1. (A) Electrophoretic mobility shift assay of MCREF-1 following fractionation by SDS-PAGE. Proteins were eluted from SDS-polyacrylamide gel slices and subjected to a denaturation-renaturation regimen, as described in Materials and Methods. Renatured proteins were assayed for binding to the LVb/core A probe. The numbers on the top of each lane correspond to the gel slice (from top to bottom) from the SDS-polyacrylamide gel. Numbers on the bottom represent the location of molecular weight markers in the SDS-polyacrylamide gel (proteins eluted from gel slice 22 were not assayed). (B) Comparison of the sequence specificity of MCREF-1 from the heparin-Sepharose fraction to that of the renatured polypeptides. Binding to the LVb/core A probe was conducted in the absence (vertical line) or presence of 50 ng (500-fold molar excess) of LVb/core A, B, C, D, or E competitor oligonucleotides. The protein sample (either heparin-Sepharose or proteins from gel slices 12 and 13) is indicated on the bottom. The binding specificity of the protein from gel slice 11 was not analyzed.

sponding regions from the Moloney enhancer, MVa and MVb (Fig. 5A). MCREF-1 binding to the LVb/core A probe can also be inhibited by the Moloney virus GC-rich region located 3' to the enhancer direct repeat (MVc [Fig. 5A]) and less effectively by the corresponding region from the Friend virus (FVc [Fig. 5A]). The Friend virus FVa and FVb1 oligonucleotides appear to be the most effective competitors for MCREF-1 binding, the MVb and FVc oligonucleotides appear to be the least effective competitors, and a high-affinity nuclear factor 1 (NF-1) binding site from the adenovirus type 2 origin of replication (Ad-NF1) (7) does not compete for MCREF-1 binding to the LVb/core A probe.

We performed binding and competition analysis on the LVb/core A probe with a purified bacterially expressed 14-kDa N-terminal truncated form of Ets-1, kindly provided by Barbara Graves (34). The 14-kDa Ets-1 protein bound to the LVb/core A probe, and its binding could be inhibited by LVb-core probes with mutations in the core sites (LVb/core B and LVb/core E) and by the FVa and MVa oligonucle-



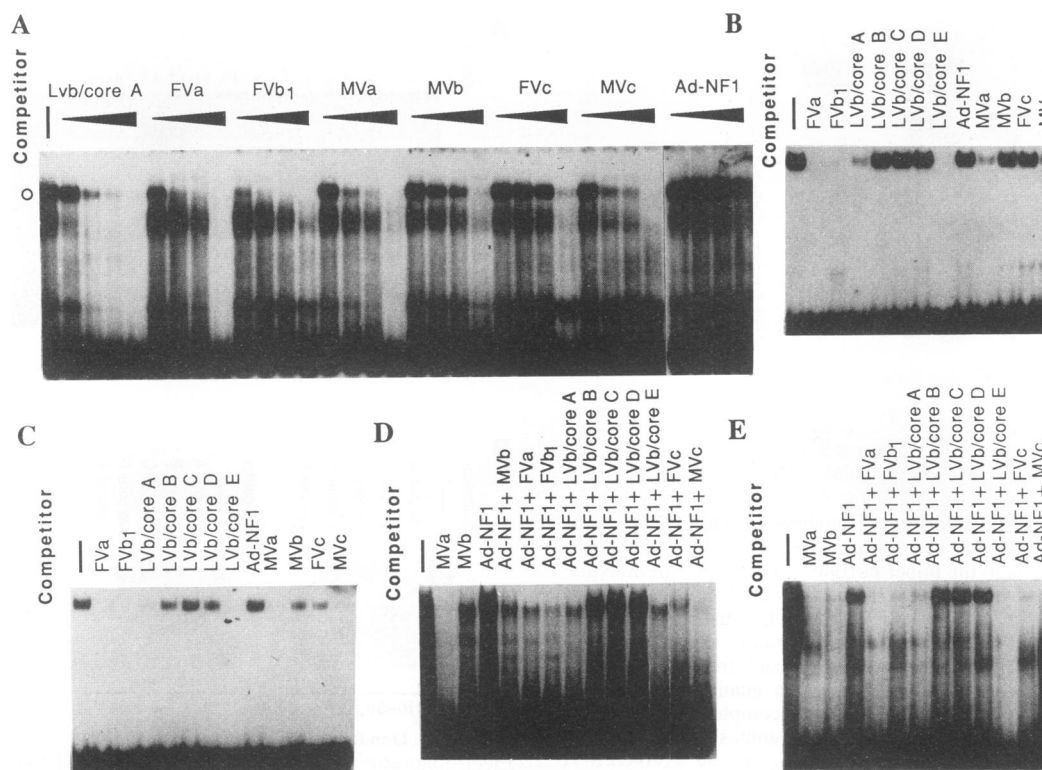


FIG. 5. Binding of MCREF-1 to sequences in the Friend and Moloney MLV enhancers. (A) Binding to the LVb-core A probe in the absence (vertical line) or presence of increasing amounts (1, 5, 10, or 50 ng) of LVb/core A, FVa, FVb1, MVa, MVb, FVc, MVc, and Ad-NF1 competitor oligonucleotides. The titration of increasing amounts of competitor DNA is indicated by the horizontal triangle. (B) Binding and competition on the FVa probe. Reaction conditions are as described for panel A, with 50 ng of competitor oligonucleotides. (C) Binding and competition on the FVb1 probe. (D) Binding to the MVa probe. (E) Binding to the MVb probe.

the LVb site (LVb/core C and LVb/core D) or with the FVb1 oligonucleotide (data not shown). This pattern of competition is distinct from that obtained for MCREF-1 and further supports the conclusion that MCREF-1 and Ets-1 are distinct proteins.

We assayed for direct binding of MCREF-1 to the Friend virus FVa and FVb1 probes and the Moloney virus MVa and MVb oligonucleotide probes. The concentration of oligonucleotide competitors used in these experiments (Fig. 5B to E) is equivalent to the greatest amount (50 ng) used for the LVb/core A probe (Fig. 5A). MCREF-1 generates a protein-DNA complex with a similar mobility on the FVa (Fig. 5B) and FVb1 (Fig. 5C) probes that can be specifically inhibited with the FVa, FVb1, LVb/core A, LVb/core E, MVa, and MVc oligonucleotides but not with the LVb/core B, C, or D oligonucleotide, the Ad-NF1 site, or the MVb or FVc oligonucleotide. Although 50 ng of either the MVb or the FVc oligonucleotide inhibits MCREF-1 binding to the LVb/core A probe (Fig. 5A), 50 ng of the MVb and FVc oligonucleotides does not inhibit MCREF-1 binding to the apparently higher-affinity FVa and FVb1 probes.

The MCREF-1 protein-DNA complex on the MVa and MVb probes is obscured by NF1 activity in the heparin-Sepharose fraction, which binds to the NF1 sites present in these two probes (Fig. 5D and E). Inclusion of the unlabeled Ad2-NF1 oligonucleotide in the binding reaction mixture effectively inhibits NF-1 binding, enabling us to visualize the MCREF-1 protein-DNA complex. The competition pattern of the MCREF-1 protein-DNA complex on the MVa probe is

difficult to discern because of the presence of a residual protein-DNA complex specific for the MVa probe that migrates at a mobility similar to that of MCREF-1. However, it is clear that the MCREF-1 protein-DNA complex on the MVa probe cannot be inhibited by the Ad-NF1, LVb/core B, LVb/core C, or LVb/core D oligonucleotide. On the MVb probe, which contains a relatively low-affinity binding site for MCREF-1, the competition pattern is similar to that obtained with the LVb/core A probe. Similar results were also obtained with the MVc probe (42a).

**Characterization of MCREF-1 binding sites.** We identified the guanine contacts for MCREF-1 on its various binding sites by methylation interference analysis (Fig. 6). We were not able to obtain a satisfactory methylation interference on the MVb probe because of its relatively low affinity for MCREF-1. Comparison of the contacts made by MCREF-1 on the LVb/core A, FVa, MVa, FVb1, and MVc probes (Fig. 6B) reveals several common features, some of which were noted previously (28, 29). (i) Binding of MCREF-1 to each CNGG repeat is generally characterized by a partial interference on the first guanine and a complete interference on the second guanine. (ii) The first CNGG repeat is separated by 6 bp from the second CNGG repeat (10 bp from center to center, or one turn of the B-form DNA helix). The 6-bp spacing between the CNGG repeats appears to be critical for binding. Note that the MVc probe contains, in addition to the bona fide MCREF-1 binding site, two consensus CNGG sequences separated by 7 instead of 6 bp (CTGGACCGCATCTGG, positions 3 to 17 [Fig. 6B]).



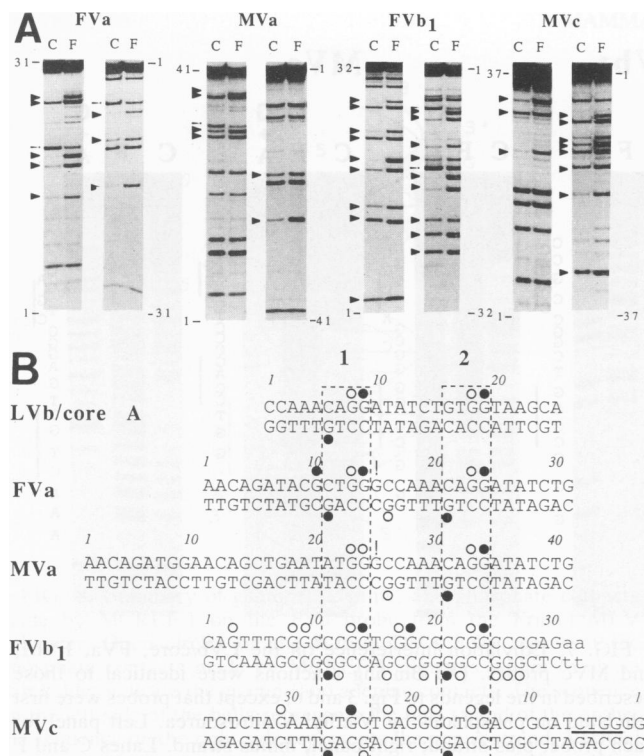


FIG. 6. Methylation interference assay on the FVa, MVa, FVb1, and MVc oligonucleotides. (A) Binding reactions were performed as described in the legend to Fig. 3. Left panel for each probe, plus strand; right panel, minus strand. Lanes C and F correspond to the protein-DNA complex and free DNA bands, respectively. Arrowheads indicate the location of guanines whose methylation specifically inhibits the binding of MCREF-1. The sideways exclamation points indicate enhanced cleavage seen in the protein-DNA complex. The numbers on the side of each lane correspond to the 5' end of the plus strand of the probe. (B) Summary of methylation interference results. Rectangles indicate the two CNGG direct repeats in the MCREF-1 binding site (1 and 2). Closed circles, guanine contacts for MCREF-1 at which complete interference was seen; open circles, partial interference; exclamation points, positions of enhanced cleavage. The MVc probe is shown in reverse orientation, with the plus strand on the bottom. The numbers indicate the position of the 5' end of each probe relative to its orientation in the viral genome. The underlined sequence in the MVc probe indicates the third CNGG repeat.

MCREF-1 does not bind to this site but instead binds to a site consisting of a 3' CNGG sequence separated by 6 bp from a sequence that deviates from the CNGG consensus (CTGCTGAGGGCTGG, positions 14 to 27). (iii) In those sites that contain a cytosine at the first position of the CNGG repeat (CNGGN<sub>6</sub>CNGG), the guanine on the other strand at that position is also a contact for MCREF-1. (iv) In CNGG sequences that are directly followed by a guanine (FVa, position 15; MVa, position 35; FVb1, positions 10 and 20 [minus strand]), modification of this guanine results in enhanced binding by MCREF-1 or alternatively enhanced cleavage by an endonuclease in the partially purified MCREF-1 preparation. (v) Deviation from the consensus CNGG sequence in one repeat of an MCREF-1 binding site is compatible with MCREF-1 binding if the other CNGG repeat conforms to the consensus. For example, the MCREF-1 site in the MVa oligonucleotide deviates from the

consensus CNGG sequence at one position in the first repeat (ΔTGGN<sub>6</sub>CAGG). In contrast, when the same C → A transversion is introduced into the LVb-core site to generate the LVb/core C oligonucleotide (ΔAGGN<sub>6</sub>GTGG) it disrupts MCREF-1 binding. Unlike the MVa oligonucleotide, the LVb/core C oligonucleotide does not contain a consensus CNGG sequence in the second repeat.

Dimethylsulfate methylates DNA on N-7 of guanine, which projects into the major groove, and N-3 of adenine in the minor groove. Methylation can affect protein binding by disrupting the formation of a hydrogen bond between the protein and the N-7 or N-3 atom of guanine or adenine or by sterically interfering with protein binding because of the introduction of a bulky methyl group. Methylation can also disrupt protein binding by indirect effects, for example, by altering the conformation of DNA or increasing the free energy required for the DNA to assume a particular conformation in the protein-DNA complex. The extensive contacts made by MCREF-1, particularly within the N<sub>6</sub> spacer between the conserved CNGG sequences, probably results from a combination of these effects of methylation.

We only note the methylated guanines that interfere with MCREF-1 binding in Fig. 6. Several methylated adenines also interfere with MCREF-1 binding. Adenine contacts made by MCREF-1 in the FVa site are summarized in Fig. 8.

We identified the phosphate contacts made by MCREF-1 on the DNA backbone in the LVb/core, FVa, and FVb1, and MVc probes by ethylation interference (36) (Fig. 7). Ethylation of phosphates can sterically interfere with binding of the protein or can disrupt electrostatic interactions between the protein and the negatively charged phosphates on the DNA backbone. MCREF-1 contacts between 2 and 5 phosphates on both strands in each CNGG sequence. The 5' and 3' boundaries of the phosphate interference pattern are difficult to determine precisely, since the backbone cleavage can occur on either side of the phosphate (36). The multiple cleavage products resolve from each other when these short oligonucleotide probes are electrophoresed through denaturing polyacrylamide gels and thus appear as several bands. Despite the difficulty in determining the precise boundaries of the phosphate interference pattern, we can conclude that the pattern is centered over the CNGG sequences and that it is staggered in the 3' direction, characteristic of proteins that bind in the minor groove of DNA.

## DISCUSSION

We have partially purified and further characterized the binding properties of MCREF-1, a protein that binds multiple sites on the Friend and Moloney MLV enhancers. Comparison of the sequence of the binding sites for MCREF-1 and the methylation and ethylation interference data suggest that the MCREF-1 binding site consists of a direct repeat of the sequence CNGG, with the two CNGG sequences separated by one complete turn of the B-form DNA helix. Independent mutations in each of the CNGG repeats disrupt binding of MCREF-1; therefore, the protein must contact both of the CNGG repeats simultaneously. The MCREF-1 protein-DNA complex can be generated from proteins isolated from single slices from an SDS-polyacrylamide gel. This indicates that MCREF-1 proteins are either single subunit proteins whose binding site spans one turn of the DNA helix or that MCREF-1 proteins are multimeric proteins consisting of subunits with similar sizes.

The LVb site forms the 3' CNGG repeat in the FVa site and the 5' CNGG repeat in the LVb-core site. MCREF-1

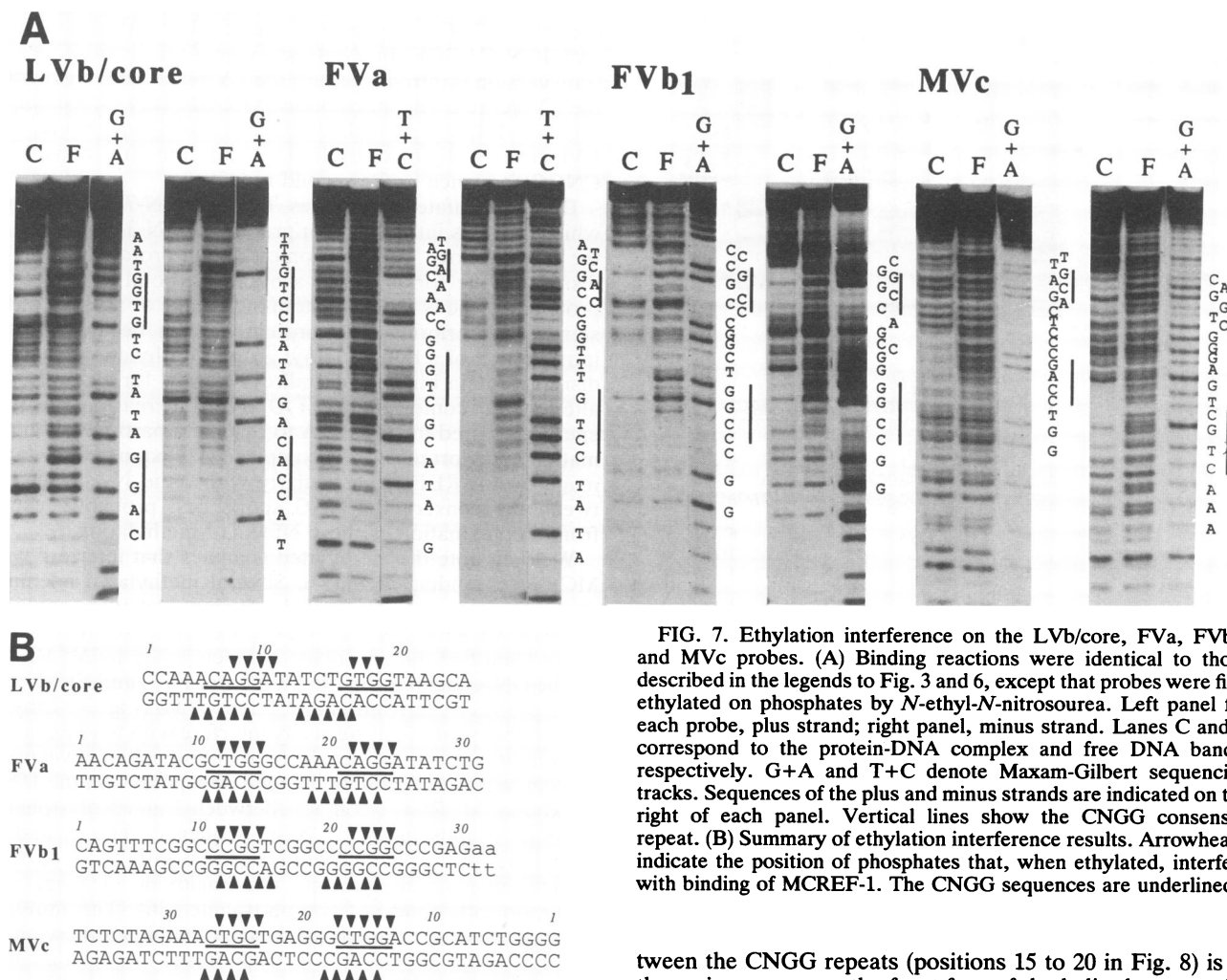


FIG. 7. Ethylation interference on the LVb/core, FVa, FVb1, and MvC probes. (A) Binding reactions were identical to those described in the legends to Fig. 3 and 6, except that probes were first ethylated on phosphates by *N*-ethyl-*N*-nitrosourea. Left panel for each probe, plus strand; right panel, minus strand. Lanes C and F correspond to the protein-DNA complex and free DNA bands, respectively. G+A and T+C denote Maxam-Gilbert sequencing tracks. Sequences of the plus and minus strands are indicated on the right of each panel. Vertical lines show the CNGG consensus repeat. (B) Summary of ethylation interference results. Arrowheads indicate the position of phosphates that, when ethylated, interfere with binding of MCREF-1. The CNGG sequences are underlined.

could theoretically form multimers and bind all three CNGG repeats in the FVa and LVb-core sites simultaneously or could preferentially bind either the FVa or LVb-core sites. The latter situation seems to be the case. When an oligonucleotide containing all three CNGG repeats in the FVa and LVb-core sites (CTGGGCCAAACAGGATATCTGTGG) is used in methylation interference assays, purine contacts are found only in the higher-affinity FVa site (CTGGGCCAAACAGGATATCTGTGG [data not shown]). Mutation of the 5' repeat in the FVa site (CTTTGCCAAACAGGATATCTGTGG) shifts the MCREF-1 contacts to the LVb-core site (data not shown).

Ethylation interference data suggest that MCREF-1 binds DNA primarily across the minor groove. This can be seen most clearly on a display of the chemical interference data on a B-form DNA helix (Fig. 8). Depicted are the MCREF-1 contacts on the FVa oligonucleotide derived from the Friend MLV enhancer. In this model, we designate the side of the helix facing the reader as the front face. MCREF-1 contacts phosphates in each CNGG sequence across two adjacent minor grooves primarily on the front face of the DNA helix. Both CNGG sequences are within the regions of phosphates contacted by the protein, which supports the derived consensus sequence of a direct repeat separated by a 6-bp spacer (CNGG<sub>6</sub>CNGG). The degenerate 6-bp spacer be-

tween the CNGG repeats (positions 15 to 20 in Fig. 8) is in the major groove on the front face of the helix, between the areas of minor groove contacts. MCREF-1 contacts one guanine in the major groove in this 6-bp spacer in the FVa probe (position 16) and one guanine 5' to the first CNGG sequence (position 10). MCREF-1 contacts guanines in the major groove in the 6-bp spacer between the CNGG repeats in the MvA, FVb1, and MvC probes. MCREF-1 also contacts guanines flanking the 5' or 3' side of the CNGG repeats in the FVa, FVb1, and MvC probes. The degeneracy of the DNA sequence between and flanking the CNGG repeats suggests that MCREF-1 does not directly read the base pair sequences in the major groove in these regions.

Methylation of adenines in the minor groove in the putative half-sites also interferes with MCREF-1 binding, which is consistent with the hypothesis that MCREF-1 binds in the minor groove. However, adenine interference data are inconclusive, since the structure of DNA in the more narrow minor groove is more easily perturbed by the introduction of a bulky methyl group, which could indirectly disrupt MCREF-1 binding.

Although the ethylation interference pattern suggests that MCREF-1 docks primarily across two adjacent minor grooves on the front face of the B-form DNA helix, this model does not completely account for all the chemical interference data. The methylation interference pattern indicates that MCREF-1 also contacts three guanines in the major groove on the opposite side (the back face) of the helix within the CNGG repeats. To accommodate both the gua-

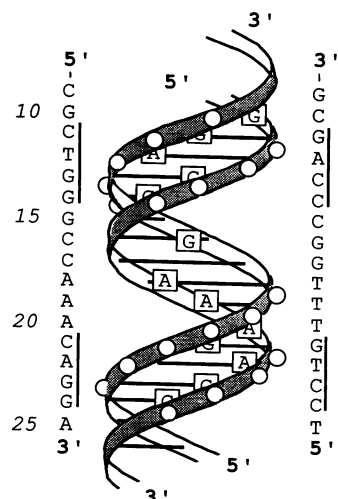


FIG. 8. Summary of guanine, adenine, and phosphate contacts made by MCREF-1 on the FVa probe from the Friend MLV enhancer. Open circles on the B-form DNA helix represent the phosphate contacts made by MCREF-1. Squares within the base pair ladder indicate the guanine and adenine contacts. The sequence on the left side of the helix represents the plus strand of the probe; the sequence on the right side represents the minus strand. Vertical bars next to the sequence indicate the position of the CNGG sequences.

nine and phosphate contacts, we propose that MCREF-1 docks primarily on the front face of the helix, but some part of the MCREF-1 protein reaches around the back of the helix to contact guanines in the major groove.

The chemical interference pattern of MCREF-1 is unique from that of other DNA binding proteins that have been analyzed, including the Ets family of proteins, the yeast protein GAL4, the homeodomain proteins, bZIP proteins as represented by C/EBP, and the helix-turn-helix proteins. The Ets proteins bind one major groove on one helix face (34). GAL4 binds two successive major grooves on one helix face (30). The homeodomain proteins bind one major groove and one adjacent minor groove on one face of the helix (1, 24, 45). Lambda repressor, which is a helix-turn-helix protein, binds two successive major grooves on one helix face (19, 22). C/EBP contacts a major groove for one full helical turn (33). Ethylation interference indicates that MCREF-1 binds two successive minor grooves on one helix face. Since the various members of a family of DNA binding proteins interact with the DNA helix in a characteristic manner, MCREF-1 does not appear to belong to the Ets, homeodomain, bZIP, or helix-turn-helix family of proteins. Therefore, MCREF-1 may represent a new protein structure that binds DNA.

MCREF-1 binding sites are found in a large number of mammalian type C retrovirus enhancers. The MCREF-1 site in the LVb-core region is conserved in at least 32 of 35 independent retrovirus isolates. The MCREF-1 site immediately 5' to and including the LVb site (corresponding to the FVa site in the Friend enhancer) is also conserved in a significant number of viruses, although there are several variations in the sequence of the site: CTGGGCCAAACAGG (Friend MLV, second copy of the direct repeat), CCGGGCCAAACAGG (gibbon ape leukemia virus, San Francisco isolate), CAGGGCCAAACAGG (simian sarcoma virus), ATGGGCCAAACAGG (Harvey, Moloney, and my-

eloproliferative sarcoma viruses; Moloney and Abelson MLV), and TTGGGCCAAACAGG (Rauscher mink cell focus-forming virus, Rauscher spleen focus-forming virus, NS6MCF virus, Friend spleen focus-forming virus [polycythemia inducing], first copy of the Friend MLV enhancer, Lake Casitas brain E neurotropic virus, amphotropic murine retrovirus clone 4070a, Ho wild mouse leukemia virus, Rauscher spleen focus-forming virus) (see reference 14 for sequence alignment). All of the sequences in the half-sites (CTGG, CCGG, CAGG, ATGG, and TTGG) have been found in bona fide MCREF-1 binding sites (28, 29; this paper). Eight additional viruses contain the sequence TAGGGCCAAACAGG. Since a T is permissible in either of the first two positions in other MCREF-1 binding sites, we predict that MCREF-1 will also bind to this site. MCREF-1 may therefore contribute to the transcription of a large number of mammalian type C retroviruses in vivo. Mutations in the LVb site in the Moloney virus which should disrupt MCREF-1 binding increased the latent period of disease caused by the Moloney virus (39). However, this mutation would disrupt not only MCREF-1 binding but also binding of LVt and the proteins in the Ets family. No mutations in this MCREF-1 site that would selectively disrupt MCREF-1 binding have yet been analyzed.

What might be the role of MCREF-1 in specifying the pathogenic phenotypes conferred by mammalian type C retrovirus enhancers? Simply the presence or absence of binding sites for MCREF-1 cannot be a determining factor, since binding sites for MCREF-1 are distributed in multiple locations on many retrovirus enhancers. If MCREF-1 does influence pathogenesis by these enhancers, then the mechanism must involve the number or distribution of MCREF-1 binding sites on these enhancers, the relative affinity of MCREF-1 for these sites, and/or the potential for interaction with other proteins on the enhancer. Interaction with other proteins could take the form of cooperative binding or steric interference between MCREF-1 and proteins binding to overlapping or adjacent sites.

#### ACKNOWLEDGMENTS

We thank Shuwen Wang for valuable advice during the course of these experiments. We also thank Nancy Hopkins and Barbara Graves for critically reading the manuscript.

This work was supported by Public Health Service grant CA51065-01A1 from the National Cancer Institute, in part by grant DMB-9009098 from the National Science Foundation, 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.

#### REFERENCES

1. Affolter, M., A. Percival-Smith, M. Muller, W. Leupin, and W. J. Gehring. 1990. DNA binding properties of the purified Antennapedia homeodomain. *Proc. Natl. Acad. Sci. USA* **87**: 4093-4097.
2. 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.
3. 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.
4. 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.

5. Chatis, P. A., C. A. Holland, J. W. Hartley, 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.
6. 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.
7. Chodosh, L. A., A. S. Baldwin, R. W. Carthew, and P. A. Sharp. 1988. Human CCAAT-binding proteins have heterologous subunits. *Cell* **53**:11-24.
8. 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.
9. 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.
10. 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.
11. 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.
12. Fried, M., and D. M. Crothers. 1981. Equilibrium and kinetics of Lac repressor operator interactions by polyacrylamide gel electrophoresis. *Nucleic Acids Res.* **9**:6505-6525.
13. 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.
14. 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.
15. Gunther, C. V., and B. Graves. Unpublished data.
16. 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.
17. 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.
18. 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.
19. Johnson, A. 1980. Mechanism of action of the Cro protein. Ph.D. thesis. Harvard University, Cambridge, Mass.
20. 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.
21. Jones, K. A., J. T. Kadonaga, P. J. Rosenfeld, T. J. Kelly, and R. Tjian. 1987. A cellular DNA-binding protein that activates eukaryotic transcription and DNA replication. *Cell* **48**:79-89.
22. Jordon, S. R., and C. O. Pabo. 1988. Structure of the lambda complex at 2.5 Å resolution: details of the repressor-operator interactions. *Science* **242**:893-899.
23. Kadonaga, J. T., and R. Tjian. 1986. Affinity purification of sequence-specific DNA binding proteins. *Proc. Natl. Acad. Sci. USA* **83**:5889-5893.
24. Kissinger, C. R., B. Liu, E. Martin-Blanco, T. B. Kornberg, and C. O. Pabo. 1990. Crystal structure of an engrailed homeodomain-DNA complex at 2.8 Å resolution: a framework for understanding homeodomain-DNA interactions. *Cell* **63**:579-590.
25. Laemmli, U. K. 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature (London)* **227**:680-685.
26. 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.
27. 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.
28. Manley, N. R., M. O'Connell, W. Sun, N. A. Speck, and N. Hopkins. 1993. Two factors that bind to highly conserved sequences in mammalian type C retroviral enhancers. *J. Virol.* **67**:1967-1975.
29. 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.
30. Marmorstein, R., M. Carey, M. Ptashne, and S. Harrison. 1992. DNA recognition by GAL4: structure of a protein-DNA complex. *Nature (London)* **356**:408-414.
31. Maxam, A. M., and W. Gilbert. 1980. Sequencing end-labeled DNA with base-specific chemical cleavages. *Methods Enzymol.* **65**:499-560.
32. 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.
33. Nye, J. A., and B. J. Graves. 1990. Alkylation interference identifies essential DNA contacts for sequence-specific binding of the eukaryotic transcription factor C/EBP. *Proc. Natl. Acad. Sci. USA* **87**:3992-3996.
34. Nye, J. A., J. M. Petersen, C. V. Gunther, M. D. Jonsen, and B. J. Graves. 1992. Interaction of murine Ets-1 with GGA-binding sites establishes the ETS domain as a new DNA-binding motif. *Genes Dev.* **6**:975-990.
35. 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.
36. 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.
37. 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.
38. 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.
39. 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.
40. Thornell, A., B. Hallberg, and T. Grundström. 1988. Differential protein binding in lymphocytes to a sequence in the enhancer of the mouse retrovirus SL3-3. *Mol. Cell. Biol.* **8**:1625-1637.
41. Thornell, A., B. Hallberg, and T. Grundström. 1991. Binding of SL3-3 enhancer factor 1 transcriptional activators to viral and chromosomal enhancer sequences. *J. Virol.* **65**:42-50.
42. 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.
- 42a. Wang, Q., and N. A. Speck. Unpublished data.
43. Weiher, H., M. Zonig, and P. Gruss. 1983. Multiple point mutations affecting the simian virus 40 enhancer. *Science* **219**:626-631.
44. 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.
45. Wolberger, C., A. K. Vershon, B. Liu, A. D. Johnson, and C. O. Pabo. 1991. Crystal structure of a MATα2 homeodomain-operator complex suggests a general model for homeodomain-DNA interactions. *Cell* **67**:517-528.