Differential Protein Binding in Lymphocytes to a Sequence in the Enhancer of the Mouse Retrovirus SL3-3

ANDERS THORNELL, BENGT HALLBERG, AND THOMAS GRUNDSTRÖM*

Unit for Applied Cell and Molecular Biology, University of Umeå, S-901 87 Umeå, Sweden

Received 17 July 1987/Accepted 11 January 1988

An electrophoretic mobility shift assay was used to characterize interactions of nuclear proteins with a DNA segment in the enhancer element of the leukemogenic murine retrovirus SL3-3. Mutation of a DNA sequence of the 5'-TGTGG-3' type decreased transcription *in vivo* specifically in T-lymphocyte cell lines. Extracts of nuclei from different T-lymphocyte cell lines or cells from lymphoid organs resulted in much higher amounts of complexes *in vitro* with this DNA sequence than did extracts from other cell lines or organs tested. Differences were also found in the sets of complexes obtained with extracts from the different types of cells. The DNA sequence specificities of the different SL3-3 enhancer factor 1 (SEF1) protein complexes were found to be distinct from those of several other previously identified DNA motifs of the TGTGG type because of differences in several nucleotides critical for binding and because these other DNA motifs could not compete with the identified DNA sequence for binding of SEF1. Limited treatment with several different proteases cleaved the SEF1 proteins such that their DNA-binding domain(s) remained and created complexes with decreased and nondistinguishable electrophoretic mobility shifts and with new properties. These results indicate that the SEF1 proteins have a structure with a flexible and relatively vulnerable hinge region linking a DNA-binding domain(s) to a more variable domain(s) with other functions. We suggest that the binding of SEF1 is an essential factor for the T-cell tropism of SL3-3 and the ability of this virus to cause T-cell lymphomas.

Regulatory elements, denoted enhancers, that can potentiate transcription from a variety of promoters in a relatively distance- and orientation-independent manner were originally defined in viruses and subsequently also identified for chromosomal genes (for reviews, see references 8, 19, 31, and 61). Results of competition experiments *in vitro* (57, 62, 75) and *in vivo* (45, 59) have shown that specific *trans*-acting factors are involved in enhancer function. The enhancer activity of simian virus 40 (SV40) has been shown to result from the integrated function of multiple sequence elements (79) which bind such factors (76). Results of studies of several other enhancers both *in vivo* and *in vitro* have supported the notion of a modular organization of enhancers based upon shorter sequence elements (79). Homologies between such sequence elements have also been found for enhancers with very different cell type specificities. The first noted and most commonly found homology between enhancers is an element with 5'-TGTGG-3' as its most conserved sequence feature (23, 31, 40, 53, 72, 73).

Enhancers have been shown to display several distinct properties that are indicative of a high degree of flexibility. (i) They can function when separated from the promoter by a long distance of DNA. (ii) They can, without displaying symmetry in their DNA sequences, function irrespective of their orientations relative to the promoter. (iii) Their sequence motifs can be rearranged without loss of enhancer activity (8, 19, 31, 61, 79). However, a certain stereospecific alignment of an enhancer to a promoter is needed for the optimal level of enhancement (67). Because of these properties, Takahashi et al. (67) presented a hypothesis that proteins binding to the enhancer may be composed of two distinct functional domains, a DNA-binding domain and a protein-interacting domain, with the two domains linked by a flexible stem. Such a type of protein structure has been demonstrated for procaryotic DNA-binding proteins: the

lambda repressor CI (26, 50, 74), the lambda Cro protein (10, 68, and references therein), the catabolite activator protein CAP (44), and the LexA protein (39).

SL3-3 is a murine retrovirus which has been shown to induce T-cell lymphomas in several strains of mice, usually 2 to 4 months postinoculation (37, 51). The T-lymphocyte-specific leukemogenicity of this virus has been localized to its long terminal repeat (LTR) (36), which contains an enhancer which is preferentially active in T lymphocytes and is believed to be the leukemogenic determinant of this virus (7). This LTR contains a 72-base-pair (bp)-long duplicated DNA segment followed by a third copy of the first 34 bp of the repeat (36). We have localized the most important sequences for the enhancer activity in T lymphocytes to these repeated sequences (B. Hallberg and T. Grundström, manuscript in preparation). At least a part of this activity was shown to reside in the duplicated 38-bp DNA segment containing a sequence belonging to the TGTGG group of commonly found enhancer DNA motifs.

We report here on the interactions of nuclear proteins with the TGTGG sequence of SL3-3, as detected by DNase I footprinting and further characterized by an electrophoretic mobility shift assay (12, 16). This latter technique, which resolves nucleoprotein complexes from uncomplexed DNA by electrophoresis through polyacrylamide gels, has been used for the analysis of a number of procaryotic and eucaryotic DNA-binding proteins (6, 13, 14, 24, 25, 38, 52, 64, 66). In this report we show that when nuclear extracts from different cell lines are allowed to interact with the TGTGG sequence of the SL3-3 enhancer, a set of sequence-specific protein-DNA complexes with different electrophoretic mobilities are obtained. We found that the SL3-3 enhancer factor 1 (SEF1) proteins creating these complexes are much more abundant in cell lines of T lymphocytes and in lymphoid organs than in the other cell lines and organs tested. Mutation of the SEF1-binding site on the DNA decreased transcription *in vivo* in two T-lymphocyte cell lines, in

* Corresponding author.

contrast to results obtained with other cells. The results argue for the notion that the T-lymphocyte preference in the activity of the 38-bp repeat of the enhancer is due to the quantitative differences in the SEF1 proteins or differences in the sets of such proteins in the different cell types or both. We also show that limited treatments with proteases cleaved the different SEF1 proteins such that their DNA-binding domain(s) remained and created complexes with non-distinguishable electrophoretic mobilities and with new properties. These results indicate that the proteins we have studied have a flexible and vulnerable hinge region linking a DNA-binding domain(s) and a more variable domain(s) with other functions.

MATERIALS AND METHODS

Cell lines. The following cell lines were used: X-63 Ag8.653 (denoted in the text as X-63-1), a nonimmunoglobulin-secreting mouse myeloma (29, 34); K46-R, a subclone of the mouse B-cell lymphoma K46 (32); EL-4, a mouse T-helper-cell lymphoma (20, 46); CTL, a monoclonal cytotoxic mouse T-cell line (2); J558L, a mouse myeloma (49); BA-LENTL 13, a mouse T-helper-cell line (33); Jurkat, a human T-helper-cell line (18); and HeLa, a human carcinoma cell line. The lymphocyte cell lines were grown in culture flasks or Spinner flasks with RPMI 1640 medium supplemented with 5% fetal bovine serum, 5 μ g of insulin per ml, 5 μ g of transferrin per ml, 5 ng of selenium per ml, 37.5×10^6 U of streptomycin per ml, 5.0×10^7 U of benzylpenicillin per ml, 1 mM sodium pyruvate, and 20 mM β -mercaptoethanol. For CTL, which is dependent on interleukin-2, the medium was supplemented with 40 U of interleukin-2 per ml, which was prepared as described by Robb et al. (54) and assayed as described by Kappler et al. (28). HeLa cells were grown as described by Banerji et al. (4).

Preparation of nuclear extracts. For the preparation of nuclear extracts, 5×10^8 to 10×10^8 cells were harvested by centrifugation in a rotor (2,000 rpm, 10 min; GS-3; Sorvall), suspended in culture medium, pooled, centrifuged (2,000 rpm, 10 min), suspended in phosphate-buffered saline (140 mM NaCl, 2.7 mM KCl, 14 mM Na_2HPO_4 , 1.5 mM KH_2PO_4 , 0.5 mM MgCl_2 , 0.7 mM CaCl_2), and centrifuged again (as described above). Spleen and thymus cells were prepared from fresh organs of 20 CBA mice or 1 calf by slicing and grinding them through a metal net, followed by washing with phosphate-buffered saline. The cells were incubated on ice for 10 min, and then the supernatant was carefully decanted and centrifuged at 500 rpm for 10 min in a centrifuge (TJ-6; Beckman Instruments, Inc., Fullerton, Calif.). After a new decantation, the supernatant was centrifuged at 2,000 rpm for 10 min to pellet the spleen cells. Liver and kidney nuclei were prepared from fresh organs from four CBA mice as described by Graves et al. (21).

Nuclear extracts were prepared from the cell pellets as described by Dignam et al. (11), with the modifications of Wildeman et al. (75) and the additional modification that the protease inhibitor phenylmethylsulfonyl fluoride (PMSF; 1 mM) and, when nothing else is specified below, the protease inhibitors leupeptin (5 μ g/ml), pepstatin A (2.5 μ g/ml), aprotinin (12 μ g/ml), benzamidin (2 mg/ml), soybean trypsin inhibitor (12 μ g/ml), and chicken egg white ovomucoid containing ovoinhibitor (12 μ g/ml) were also added to the buffers. Protein concentrations of extracts were typically 5 to 15 μ g/ μ l.

DNase I footprinting experiments. Plasmid pESG106 (Hallberg and Grundström, in preparation), a derivative of plas-

mid pA51 (79), containing a synthesized 317-bp DNA segment from SL3-3 virus (positions 130 to 446; nucleotide numbers are those used in reference 36) including most of the U3 region of the LTR, was cleaved with the restriction endonuclease *KpnI* (a cleavage site constructed at position 130). Two complementary synthetic oligonucleotides, a 34-mer and an 18-mer, creating the sticky end of restriction endonuclease *KpnI* and an end with a 5' extension of 10 T residues, were prepared with the rapid microscale segmented synthesis technique (42) modified for phosphoramidite chemistry (H. W. D. Matthes, A. Staub, and P. Chambon, *Methods Enzymol.*, in press). The 34-mer was phosphorylated with T4 polynucleotide kinase (Boehringer GmbH, Mannheim, Federal Republic of Germany), the two oligonucleotides were annealed (37°C, 15 min), the *KpnI*-cleaved vector was added, the synthetic DNA segments were ligated with T4 DNA ligase, and then a fill-in reaction was performed with [α - 32 P]dATP (3,000 Ci/mmol; Amersham International) and the Klenow fragment of DNA polymerase I (Boehringer). The DNA was extracted with phenol-chloroform, precipitated with ethanol, cleaved with the restriction endonuclease *BamHI* (position 446), extracted with phenol-chloroform again, and separated on a native 5% polyacrylamide gel (acrylamide-bisacrylamide [19/1]). The gel was autoradiographed wet, and the proper DNA segment was excised, electroeluted (41), extracted with phenol-chloroform, and precipitated with ethanol. Footprinting reactions were based on the methods of Galas and Schmitz (15) and Ohlsson and Edlund (48), with the following slight modifications; 0.1 to 0.2 fmol of end-labeled DNA, 1 μ g of poly(dI-dC) (Pharmacia, Uppsala, Sweden), 3 to 10 μ l of nuclear extract (prepared in the presence of PMSF), and 12.5 to 400 ng of DNase I (Sigma Chemical Co., St. Louis, Mo.; adjusted empirically to produce an even pattern of cleavage products) were used in a total reaction volume of 20 μ l. The DNA fragments were separated on 6% polyacrylamide buffer gradient urea gels (5).

Construction of mutants and in vivo transcription assay. Plasmid pESG134 (Hallberg and Grundström, in preparation; Fig. 1A) is a derivative of plasmid pA51 (79), containing a version of the SL3-3 enhancer with two copies of the 34-bp repeat and one copy of the 38-bp segment in front of the SV40 early promoter region (positions 101 to 5171) (69) and the coding sequence of the rabbit β -globin gene (positions -9 to +1330) (70) in pBR322 between positions 4361 and 2296. The derivative of plasmid pESG134 containing mutation 2 was constructed by cassette mutagenesis using the shotgun ligation technique (22). In summary, the *PstI*-*ApaI* DNA segment containing the mutation (positions 37 to 203) (36) was assembled from the same set of 15 oligonucleotides as that employed for the synthesis of pESG134 (Hallberg and Grundström, in preparation), except that the proper nucleotides to create the substitution of three nucleotides were replaced. The assembly was confirmed by dideoxy sequencing, and a plasmid identical to pESG134, except for the mutation, was constructed by standard subcloning techniques. The cotransfected reference plasmid pESG003 (Hallberg and Grundström, in preparation) has a structure corresponding to that of pESG134, except that the enhancer is that of SV40 (positions 107 to 299) (69) and the promoter is that of rabbit β -globin (complete to position -113).

Transfections of lymphocytes were as described by Banerji et al. (3), with slight modifications; 4×10^6 to 6×10^6 cells were centrifuged (1,600 rpm, 5 min; TJ-6; Beckman), and the cell pellet was washed with 10 ml of Tris-buffered saline (TBS) (137 mM NaCl, 25 mM Tris hydrochloride, 5

mM KCl, 0.7 mM CaCl_2 , 0.5 mM MgCl_2 , 0.6 mM Na_2HPO_4 [pH 7.4]). After a second centrifugation (as described above), 2 to 4 μg of a mixture of recombinant plasmid DNA and reference plasmid DNA (ratio of 1/3) in TBS with 0.5 mg of DEAE-dextran (molecular weight, 2×10^6 ; Pharmacia) per ml was added to the cells. After 30 min of incubation, the transfection solution was removed by centrifugation. For X-63-1 cells, an incubation step was also performed with 1 ml of 10% dimethyl sulfoxide in TBS for 5 min. Lymphocytes were washed twice with TBS and incubated with fresh medium for 40 to 42 h. The cells were harvested by centrifugation (as described above), and preparations of RNA were made as described by Auffray and Rougeon (1), with some modifications. Frozen cells were suspended in 1 ml of 3 M LiCl–6 M urea–0.1% sarcosyl–100 mM β -mercaptoethanol–50 mM Tris hydrochloride–5 mM EDTA (pH 7.5) and sonicated at 0°C . The RNA was precipitated overnight at 4°C and centrifuged at $14,900 \times g$ (Heraeus microfuge) for 20 min. The pellet was dissolved in 10 mM Tris hydrochloride–1 mM EDTA–0.5% sodium dodecyl sulfate (pH 7.5), deproteinized by two successive phenol extractions and one chloroform extraction, and precipitated with ethanol. RNA (10 to 30 μg) was hybridized with an excess of single-stranded 5'-end-labeled DNA probe as described by Zenke et al. (79) and treated with S1 nuclease (80 U; Amersham) at 25°C for 3 to 5 h. The protected DNA sequences were resolved on 6% polyacrylamide–8 M urea gels.

Electrophoretic mobility shift assay. Complementary 19- to 31-base-long oligonucleotides, prepared as described above, were end labeled with [γ - ^{32}P]ATP ($>5,000$ Ci/mmol; Amersham)–T4 polynucleotide kinase (Boehringer), annealed (56°C for 15 min and 37°C for 10 min), extracted with phenol-chloroform, and electrophoresed on a native 10% polyacrylamide gel. The double-stranded DNA segments were excised, eluted from the gels by diffusion overnight, extracted with phenol-chloroform, and ethanol precipitated. Competitor DNA segments were prepared by the same protocol, except that a 25,000-fold excess of unlabeled ATP was added. Binding reactions were based on the method of Garner and Revzin (16), as modified by Schneider et al. (58), but with additional modifications; 0.5 to 1 fmol of end-labeled DNA segments was incubated for 15 min at room temperature together with 1 to 4 μg of nuclear extracts (specified in the figure legends) in 25 mM HEPES (N-2-hydroxyethylpiperazine-N'-2-ethanesulfonic acid)–1 mM EDTA–5 mM dithiothreitol–10% glycerol–150 mM NaCl–0.5 mM PMSF–0.1 pmol of heterologous single-stranded oligonucleotide (5'-GAAGTCTGACTGAATCAG-3')–50 ng of poly(dI-dC) (pH 7.5). The total reaction volumes were 6 μl . Nuclear extracts were always added last. The reaction mixtures were electrophoresed at 12 V/cm for 1 h on 6% native polyacrylamide gels (acrylamide-bisacrylamide [44/0.8]) in 65 mM Tris glycine (pH 8.5) with 40 mM Tris glycine (pH 8.5) in the buffer chambers. Gels were soaked in 5% glycerol (10 min), dried, and autoradiographed at -70°C with an intensifier screen.

Methylation interference experiments. DNA segments of 31 bp, which were labeled at one 5' end by annealing one labeled and one unlabeled oligonucleotide, were partially methylated at the guanine and adenine residues, as described by Siebenlist and Gilbert (63). The electrophoretic mobility shift assays were performed with extracts prepared in the presence of PMSF in a 20-fold larger scale than that described above. Bands were excised from the dried gels and eluted overnight in water containing 10 μg of salmon sperm carrier DNA per ml. The eluates were lyophilized to a

volume of 200 μl and extracted with phenol-chloroform. The DNA was applied to 100 μl of DEAE C50 (Pharmacia) columns, which were washed twice with 200 μl of 20 mM NaCl, twice with 200 μl of water, and then eluted three times with 200 μl of 30% triethylammonium acetate (pH 7.3). Eluates were lyophilized to dryness, suspended in 0.3 M sodium acetate, and ethanol precipitated. DNA pellets were incubated in 30 μl of 1 M piperidine at 90°C for 30 min, lyophilized, and analyzed on 15% polyacrylamide gels (acrylamide-bisacrylamide [19/1]) with 8 M urea, followed by autoradiography at -70°C with an intensifier screen.

Protease treatment of nuclear extracts. For digestion with endoproteinase Glu-C *Staphylococcus aureus* V8 (V8 protease; Boehringer), proteinase K (Merck & Co., Darmstadt), papain (Sigma), or trypsin (Sigma), the nuclear extracts (containing EDTA and PMSF but without other protease inhibitors added) were diluted in electrophoretic mobility shift assay buffer to a protein concentration of 2 mg/ml. To 2 μl of nuclear extract (4 μg of protein), 1 μl of V8 protease (4 mg/ml in 20 mM Tris hydrochloride [pH 7.5]); 2 μl of papain (0.33 mg/ml in 20 mM Tris hydrochloride [pH 7.5]); 2 μl of trypsin (0.2 mg/ml in 20 mM Tris hydrochloride [pH 7.5]); or when nothing else is stated below, 2 μl of proteinase K (0.2 mg/ml in 20 mM CaCl_2) was added. The reaction mixtures were incubated for 10 to 15 min at 25°C before 2 μg of the digested nuclear extracts was added to the electrophoretic mobility shift assays.

RESULTS

DNase I footprints over the enhancer. To analyze whether a nuclear factor(s) bound to the LTR of SL3-3 virus, the enhancer was 3' end labeled with ^{32}P on its upstream side (at nucleotide position 130; numbering as in reference 36), incubated with nuclear extract from the murine cytotoxic T-lymphocyte cell line CTL, and subsequently treated with DNase I (see above). The nuclear extract very specifically modified the pattern of digestion characteristic of naked DNA (cf. lanes 3 and 4 of Fig. 1B). A 17-bp-long sequence in the duplicated 38-bp DNA segment (positions 178 to 193 and 251 to 265) was specifically protected against DNase I digestion. This DNA segment (5'-GATATCTGTGGT TAAGC-3') has an enhancer motif of the TGTGG type in its center. However, with extracts from other cell types, the footprinting technique did not demonstrate clear and reproducible interactions over this DNA segment (data not shown).

Effects of a mutation in the interaction sequence on transient expression. Deletion of the remaining 38-bp repeat from a derivative of the SL3-3 enhancer carrying one copy each of the 34- and 38-bp repeats decreases transcription about twofold in the T-lymphocyte cell lines CTL and EL-4, while it does not decrease transcription in the B-cell line X-63-1 or HeLa cells (Hallberg and Grundström, in preparation). To study whether the interaction with the identified DNA sequence was relevant for the selective effect of the 38-bp repeat sequence on transcription in vivo in T lymphocytes, the enhancement of transcription during transient expression of a mutant in the sequence protected against DNase I and the corresponding wild type were compared. Plasmid pESG134 (Fig. 1A; Hallberg and Grundström, in preparation), carrying the SL3-3 enhancer in front of the early promoters of SV40 and a rabbit β -globin gene, was used for these experiments. This plasmid has one 72-bp repeat of the LTR of SL3-3 deleted and therefore contains a single copy of the TGTGG sequence. This deletion has small effects on the

FIG. 1. (A) Structure of the plasmid pESG134, which contains most of a version of the U3 region of the LTR of SL3-3 (positions 1 to 446) (36), with two copies of the 34-bp repeat and one copy of the 38-bp segment in front of the SV40 early promoter region (positions 101 to 5171) (69), and the coding sequence of the rabbit β -globin gene (positions -9 to +1391) (70) in pBR322 between positions 4361 and 2296. Mutant 2 of plasmid pESG134 is identical, except for a 3-bp substitution of nucleotide positions 188 to 190 of SL3-3 (indicated with an asterisk; see also Fig. 5 and the text). (B) DNase I footprinting with nuclear extract from the cytotoxic T-cell line CTL over a 317-bp DNA segment of the LTR of SL3-3 that was end labeled at position 130 (numbering system used in reference 36). Lanes 1 and 2, Sequencing reactions of G+A and A+C, respectively; lane 3, cleavage products of the DNA segment digested with 25 ng of DNase I in the absence of nuclear extract; lane 4, cleavage products of the DNA segment digested with 150 ng of DNase I in the presence of 30 μ g of nuclear extract. The column to the left of panel B indicates nucleotide positions, with the filled and open boxes representing the 34- and 38-bp DNA segments repeated three and two times, respectively. The nucleotide numbers of the limits of the protected regions are indicated at the right, together with the positions of the sequence 5'-TGTGG-3'.

level of transcription in different cell lines (Hallberg and Grundström, in preparation) and greatly facilitated the construction, by shotgun gene synthesis, of a mutated derivative (see above). A 3-bp substitution in the central area of the footprint (mutation 2; see Fig. 5) was chosen for this analysis. Each of the three nucleotides was subsequently found to be important for specific protein-DNA interactions in vitro (see below and Fig. 4 to 6). The plasmid pESG134 or the corresponding plasmid carrying the mutation was co-transfected with a control plasmid followed by transient expression, RNA preparation, and quantification by S1 nuclease mapping of the amount of specific transcript relative to a reference transcript from the control plasmid (see above). The levels of transcription obtained in representative experiments in different lymphocyte cell lines is

shown in Fig. 2, and these results, together with additional measurements, are summarized in Table 1. As expected from the lack of effect of deleting the 38-bp repeat in the B-cell line X-63-1, we found no effect of the mutation on the level of transcription in this cell line. Interestingly, the mutation of the 38-bp repeat decreased transcription twofold in both T-lymphocyte cell lines tested, strongly arguing that the selective twofold effect of the 38-bp repeat on transcription in T cells is due to the interaction with the identified DNA sequence.

Binding of nuclear factors shift the electrophoretic mobility of the TGTGG-containing DNA sequence. The differential importance in T lymphocytes of the DNA sequence from which a DNase I footprint was obtained encouraged us to do a more detailed analysis, by the electrophoretic mobility shift assay, of factors from different cell types that bind to the DNA sequence. This assay involved incubation of an end-labeled DNA segment of a specific sequence plus unlabeled, nonspecific DNA with extracts of cell nuclei. DNA with nuclear factor(s) bound was then separated from free DNA by electrophoresis through a native polyacrylamide gel and visualized by autoradiography (see above). When a 32 P-labeled, 19-bp DNA segment (positions 177 to 195 or 249 to 267), including the DNase I footprints described above, was used in such an assay, complexes with slower migrations than that of free DNA were obtained with all the extracts tested. Different extracts from the same cell type reproducibly gave the same banding pattern. The results with typical extracts from different cells of the lymphocyte lineages and from the human carcinoma cell line HeLa,

FIG. 2. Quantitative S1 nuclease analysis of RNA prepared from transient expression in X-63-1, EL-4, and CTL cells by using plasmid pESG134 (wt, wild type) and plasmid mutant 2 (m2) of pESG134, which harbors a substitution of nucleotides from positions 188 to 190 (see Fig. 1A and 5). EES corresponds to RNA transcribed from the early early SV40 promoter start sites of the two plasmids, and Glob refers to RNA transcribed from the cotransfected reference recombinant (see text).

TABLE 1. Effects of mutation in the TGTGG sequence in different lymphocyte cell lines

Cell line	Relative transcription (%) of ^a :	
	pESG134	pESG134 mutant 2
X-63-1	100	101 (110, 123, 89, 84, 97)
EL-4	100	54 (49, 55, 64, 49)
CTL	100	44 (70, 46, 38, 38, 28)

^a Specific transcription was during transient expression in different lymphocyte cell lines of plasmids with the SL3-3 enhancer in front of the early SV-40 promoters and a rabbit β -globin gene (see text). Plasmid pESG134 and the construction of the derivative carrying mutation 2 (Fig. 5) are described in the text. The values correspond to the amounts of RNA (determined by quantitative S1 nuclease analysis), expressed relative to the value for pESG134, which was taken as 100%, in each cell line. Individual values are in parentheses.

which was used as a control, are shown in Fig. 3A. To study the DNA sequence specificities of the complexes with different mobilities that were obtained, we carried out binding reactions in the presence of several different unlabeled competitor DNA segments. When the homologous DNA segment was added as competitor (50 $\times$ molar excess), the bands of a set of complexes became drastically weakened, whereas for two other complexes with relatively small electrophoretic mobility shifts (one of which was found in higher amounts in the human cell lines than in the mouse cell lines), at most minor competition occurred (Fig. 3B and C). No complex was lost because of competition when two other DNA segments were used as competitors instead: the homologous sequence but with the mutation analyzed in vivo (see above; see mutation 2 in Fig. 5; GTT at positions 188 to 190 replaced by TGG) and the SL3-3 sequence of positions 202 to 226 (or positions 274 to 298) (36). Thus, two different types of complexes were obtained: complexes whose formation the homologous DNA sequence could efficiently and specifically compete against, probably representing binding of factors with high sequence specificities, and complexes whose formation could not be competed against with a limited molar excess of competitors, probably representing binding of factors with much lower specificities for the DNA sequence. Interestingly, all four T-lymphocyte cell lines tested (EL-4, a mouse T-helper lymphoma cell line; BALENTL 13, a mouse T-helper-cell line; CTL, a monoclonal cytotoxic mouse T-cell line, and Jurkat, a human T-helper-cell line) created much more of the complexes with high DNA sequence specificities than did the other cell lines (Fig. 3A; HeLa, a human carcinoma; X-63-1, a mouse myeloma [B-cell lineage]; J558L, another mouse myeloma cell line; and K46-R, a mouse B-cell lymphoma; note that four times more extract and a longer autoradiography were used for analysis of non-T cells [Fig. 3B] than for analysis of T cells [Fig. 3C]). In agreement with the results of the DNase I footprinting experiment, the highest amount of specific complex was obtained with extracts from the CTL cell line. All bands representing complexes disappeared upon boiling the extracts or treating them with high amounts of trypsin or papain, 0.5% sodium dodecyl sulfate, or 0.5% sarcosyl, while RNase A had no effect (data not shown). Based on these results and the fact that upon limited protease cleavages new complexes with decreased electrophoretic mobility shifts were formed (see below), the factors are believed to be proteins. We denote the proteins found to bind with high sequence specificities to the identified DNA sequence SL3-3 enhancer factor 1 (SEF1) proteins. Note that the distribution of bands due to different complexes with SEF1 proteins did

not appear to be identical for all cell lines (Fig. 3A to C). The cytotoxic T-cell line CTL gave complexes both with lower and slightly higher mobilities compared with those obtained in substantial amounts with the different T-helper-cell lines. Note also the limited stability during electrophoresis of the complexes with the SEF1 proteins, resulting in smearing between the positions of the complexes and the free DNA.

To determine whether the difference between T-cell and other cell lines in their contents of SEF1 proteins correlated with corresponding differences between cells of organs containing high amounts of T lymphocytes and cells from other organs, nuclear extracts were prepared from organs of CBA mice, and electrophoretic mobility shift experiments were performed. Primary thymocytes and splenocytes were found to be at least as rich in SEF1 proteins as T-cell lines, while

FIG. 3. Electrophoretic mobility shift assays of end-labeled 19-bp DNA segments (1 fmol; positions 177 to 195 of the sequence shown in Fig. 5) with nuclear extracts from different cell lines and different murine tissues (see text). (A) Complexes formed when 1 μ g of nuclear extracts from the human HeLa cell, different murine B cells (X-63-1, J558L, and K46-R), different murine T cells (EL-4, BALENTL 13, and CTL), and the human T-helper-cell line Jurkat were employed. (B) Competition analyses with 4 μ g of nuclear extracts from HeLa cells and different B-cell lines. A total of 50 fmol of 25-bp-long homologous (S; positions 174 to 198 or 246 to 270), mutation number 2 (1; see Fig. 5), or heterologous (2; positions 202 to 226 or 274 to 298 of SL3-3) DNA segments were used as competitors. (C) Competition analyses as described for panel B, but with 1 μ g of nuclear extracts from different T-cell lines. (D) Competition analyses as described for panel B, but with 1 μ g of nuclear extracts from different tissues of CBA mice. Minus signs indicate that there was no competitor, and -P indicates that no nuclear extract was added.

FIG. 4. Methylation interference experiment to define nucleotides on both strands (A and B) important for binding of SEF1 proteins from different cell lines to the 31-bp DNA segment (see text). Lanes 1 and 2, Unbound DNA; lanes 3 to 5, DNA complexes with SEF1 proteins (lane 3, X-63-1; lane 4, EL-4; lane 5, CTL). The nucleotide sequences are shown to the right of the panels, and solid triangles indicate the positions where methylation interfered with formation of the complex. The reaction of the methylating agent is much more efficient with guanines than with adenines (43). Nevertheless, the importance for binding could also be determined for many adenines (although for a few of these, they could be determined only on the original autoradiogram). The results are summarized in Fig. 5.

kidney and liver cells contained very little of these proteins (Fig. 3D). In a control experiment, the qualities of the nuclear extracts of the kidney and liver cells were confirmed by the fact that they contained a known transcription factor, NF1 (data not shown). The thymocytes were found to give a distribution of complexes corresponding to that of CTL, while the spleen-derived cells preferentially gave the faster-migrating complexes.

Methylation interference analysis of the complexes. To define contact residues on the DNA for the SEF1 proteins, a methylation interference analysis was performed. DNA segments labeled at one 5' end were partially methylated by dimethyl sulfate and incubated with nuclear extract, the complexes with SEF1 proteins as well as noncomplexed DNA segments were isolated from a native polyacrylamide gel, the DNA was cleaved with piperidine at the positions of methylation, and the DNA fragments were resolved on a sequencing type of polyacrylamide gel (see above). If any of the methyl groups introduced interfered with the binding of a specific protein, then the molecules of DNA methylated at that position would be selectively missing in the complex and, subsequently, in the corresponding sequencing gel

ladder. When the two strands of a DNA segment containing the TGTGG sequence were analyzed in such an experiment, the isolated free DNA generated a characteristic ladder of strong guanine bands (Fig. 4A and B, lanes 1 and 2). The methylating agent reacted with guanines and, to a much lower extent, with adenines (35). Nevertheless, in addition to guanine bands, many adenine bands were also identified in the ladder, although some of them were only identified on the original autoradiogram. The DNA in complexes within SEF1 was deficient in methylation at certain residues in the middle of the DNA segment (Fig. 4A and B, lanes 3 to 5; summarized in Fig. 5). Interestingly, the nucleotides found to be important for the formation of complexes with the SEF1 proteins were indistinguishable for the extracts of the three different lymphocyte cell lines. However, note that it cannot be excluded that methylation at an additional position(s) interfered with the formation of a certain SEF1 complex, because such an extra interference would not be detected in the mixture of other different complexes with the SEF1 proteins. The positions where methylation interferes with SEF1 binding include most, but not all, of the purines between positions 185 and 192, with methylation of the bordering adenine at position 192 causing only partial interference (Fig. 4 and 5).

Mutation analysis of the binding site. The binding of the SEF1 proteins to the DNA was analyzed further by testing the ability of the proteins to bind to DNA segments with the different mutations shown in Fig. 5. While the wild-type DNA segment efficiently competed with the labeled DNA segment, the mutated DNA segments either did not compete at all (mutation 2) or competed with decreased efficiencies (Fig. 6A and B, mutations 1 and 3 to 6). The use of labeled DNA segments with the same mutations in the electrophoretic mobility shift assay led to a decrease in the complexes with SEF1 in a fashion corresponding to the results of the competition experiments (data not shown). Note that none of the mutant DNA segments selectively eliminated by competition subset of the complexes formed with the SEF1 proteins. Thus, the DNA-binding sites of the SEF1 proteins that create the different complexes also appear to be the same by this assay.

No competition with recognition sequences for other enhancer-binding transcription factors. Many enhancer-binding

FIG. 5. Results from the methylation interference analysis (Fig. 4) of the 31-bp DNA segment corresponding to positions 171 to 201 (or 243 to 273). Six different mutations in this DNA segment are also indicated. Black boxes next to the nucleotides indicate residues whose methylation interfered with binding of SEF1, the open circle indicates a residue whose methylation partially interfered with protein-DNA interactions, and black dots indicate G and A residues whose methylation did not interfere with protein-DNA interactions. The A residues not marked with symbols were not methylated enough to enable analysis.

FIG. 6. (A and B) Competition analysis of binding of SEF1 proteins in electrophoretic mobility shift experiments (see text) in which 31-bp DNA segments were used with the different mutations (1 to 6; see Fig. 5) to show fine specificity of SEF1 protein binding to an end-labeled DNA segment (1 fmol) containing the central 19 bp (positions 177 to 195) of the sequence shown in Fig. 5. (A) Reactions with 4 μ g of nuclear extracts from cell lines HeLa and X-63-1 in which 50 fmol of the competing DNA segments indicated above the lanes was used. (B) Reactions with 1 μ g of nuclear extract from the cell lines EL-4 and CTL in which 25 fmol of the competing 25-bp DNA segments was used. Lanes: -P, no nuclear extract added; -, no competitor DNA segment added to the reaction; S, specific (homologous) sequence as competitor; 1 to 6, DNA segments with mutations 1 to 6, respectively (Fig. 5), as competitors. (C and D) Competition of binding of SEF1 proteins to 1 fmol of the end-labeled, 19-bp DNA segment with 50 fmol of 25-bp DNA segments harboring the identified binding sites for transcription factors. (C) Reactions with 4 μ g of nuclear extracts from HeLa and X-63-1 cells. (D) Reactions with 1 μ g of nuclear extracts from EL-4 and CTL cells. Lanes: S, specific (homologous) sequence as competitor; 1 and 2, GT-1 and GT-2 motifs of the SV40 enhancer (positions 256 to 280 and 237 to 261, respectively) (79); 3 and 4, μ 70 and μ 50 motifs of the immunoglobulin heavy-chain enhancer, respectively (60); 5, κ -3 motif of the immunoglobulin κ light-chain enhancer (60); 6, NF-1 motif (47); 7, *Sph*I and octamer motifs of the SV40 enhancer (positions 195 to 219) (79); -P, no protein added; -, no competitor added.

transcription factors that recognize different DNA sequences have been identified. Several of these recognition sequences contain a TGTGG sequence and, therefore, show a certain sequence homology to the SEF1-binding site. To determine whether SEF1 proteins are separate from several other identified enhancer-binding factors, we performed

competition experiments with a 50-fold molar excess of seven DNA segments containing recognition sequences for different enhancer factors (Fig. 6C and D; for the μ 70 motif [competitor 3], lack of competition was also confirmed with a 100-fold molar excess of the DNA segment). Only the homologous DNA segment was found to compete for binding, showing that the SEF1 proteins are distinct from the factors that bind to these seven other recognition sequences.

Electrophoretic mobility shifts with nuclear extracts partially digested with protease. To further characterize the different SEF1 proteins, a limited treatment with protease was performed on different nuclear extracts, which were prepared with EDTA and PMSF as the only protease inhibitors (and not the six additional inhibitors otherwise added), prior to the electrophoretic mobility shift analysis (see above). Note that for many of the cell types, the extracts prepared in the presence of a decreased number of protease inhibitors gave complexes with decreased mobility shifts even before protease treatments (Fig. 7), indicating that EDTA plus PMSF is not sufficient in all cell types to protect the SEF1 proteins from protease attack during preparation. Treatments of the extracts with proteinase K or protease V8 generated new nucleoprotein complexes with decreased electrophoretic mobility shifts. Proteinase K treatment yielded one major complex for each of the cell lines. A 50-fold molar excess of the homologous DNA segment could compete with the labeled DNA segment in the formation of this complex but a heterologous DNA segment could not, indicating that the sequence specificities of the SEF1 proteins are retained after cleavage. Also, cleavage of the extracts with protease V8 yielded one complex for each cell type against which an excess of the homologous DNA segment could compete but a heterologous DNA segment could not (Fig. 7B). Interestingly, while complexes with several different mobilities were obtained with uncleaved extracts, the decreased electrophoretic mobility shifts obtained with both proteases were to one position on the gel with the extracts from the different cell lines (cf. Fig. 7 with Fig. 3 and 6). The same effects of protease treatments were also seen with nuclear extracts of both murine and bovine spleen cells (Fig. 7).

Treatments of preformed complexes with proteinase K or protease V8 yielded corresponding results, as when protease treatments were before complex formation, although higher amounts of the proteases were needed (data not shown).

Site preferences of the protease cleavage of the SEF1 proteins. Preferential protease cleavage of the SEF1 proteins at a certain position(s) was further analyzed by treatments with different amounts of the proteases. A 100-fold variation in the quantities of proteinase K used only led to minor differences in the complexes that were obtained (Fig. 8). While small amounts of a larger complex remained with the lowest quantity of proteinase K, increasing amounts of the protease only slightly decreased the abundance of the specific complex. Bands that could be competed against and that had the same migration as those obtained with proteinase K could also be obtained both with 0.1 and 1 μ g of protease V8 (data not shown). Thus, there is a very high preference for cleavage with these two proteases at a certain position(s) in these proteins.

The protease quantities which created the preferred cleavage products did not have to correspond to the amounts needed to digest a protease inhibitor-free extract to the same extent because the protease inhibitor PMSF was used as a protection against protease attack during preparation of these nuclear extracts. However, for this study, the relative

FIG. 7. Electrophoretic mobility shift assays with protease-digested nuclear extracts from different cell lines or spleen cells. Proteinase K (A) and protease V8 (B) were employed in the digestions of the nuclear extracts. The lanes contained 0.5 fmol of end-labeled, 31-bp DNA segments (positions 171 to 201 or 243 to 273 of SL3-3) and 25 fmol of unlabeled, 31-bp homologous (s) or heterologous (ns; positions 103 to 133 of the SL3-3 LTR) competitor DNA segments. A total of 2 μ g of either nondigested or protease-digested nuclear extract (prepared in the presence of EDTA and PMSF; see text) was added to each reaction as indicated by a plus or a minus sign. Lanes: H, HeLa; X, X-63-1; E, EL-4; C, CTL; M, murine spleen cells (CBA mice); B, bovine spleen cells; and -P, no nuclear extract added. The sequence-specific complexes after protease treatment are indicated with an arrow.

protease sensitivities, rather than the absolute values, within the SEF1 proteins were of importance.

To address further the question of whether limited treatment of the different SEF1 proteins with different proteases created DNA-binding products with much less variable mobilities, we performed prolonged (3 h) electrophoresis of complexes obtained by treatment of extracts with proteinase K, protease V8, and two other proteases, papain and trypsin (Fig. 9). The electrophoretic mobilities were not separated for the specific complexes formed with the different SEF1 proteins after cleavage of the extracts of the different cell types with proteinase K, protease V8, or papain. Also, trypsin-digested nuclear extracts from the different cell types generated SEF1 protein complexes with equal mobilities, although with slightly larger electrophoretic mobility shifts compared with those obtained with the other three proteases.

The SEF1 protein complexes obtained after the nuclear extracts were treated with proteases were not only found to

create more distinct bands than with extracts not treated with protease but also to create complexes which were more stable during electrophoresis (cf. Fig. 3, 6, and 9). Prolonged electrophoresis led untreated extracts to have a much higher degree of loss of band strength and smearing of bands between the positions of the complexes and the free DNA compared with those for protease-treated extracts (data not shown).

Analysis of possible protein-protein interactions. The most direct interpretation of the effects of the limited protease treatments on the electrophoretic mobilities of the complexes is that parts that did not bind to DNA were cleaved off from the SEF1 proteins. However, an alternative possibility would be that one protein binds to the DNA sequence and then other, more protease-sensitive proteins can interact specifically with the DNA-binding protein and create complexes with different mobilities. However, the intermediary bands seen with extracts prepared with insufficient protease inhibition or after treatment with very low amounts of

FIG. 8. Electrophoretic mobility shift assay with 0.5 to 1 fmol of an end-labeled, 31-bp DNA segment (positions 171 to 201 or 243 to 273 of SL3-3) and 2 μ g of nuclear extracts from different cell lines digested with different amounts of proteinase K, as indicated above the lanes. Lanes: H, HeLa; X, X-63-1; E, EL-4; C, CTL; -, no nuclear extract added. The sequence-specific complexes are indicated with an arrow.

FIG. 9. Electrophoretic mobility shift assay in which prolonged migration in the gel (3 h) was used to compare the relative distribution of nucleoprotein complexes between nuclear extracts from different cell lines digested with protease V8, proteinase K, papain, and trypsin and the end-labeled, 31-bp DNA segment from the SL3-3 enhancer. Each lane contained 2 μ g of protease-digested nuclear extract and 0.5 fmol of end-labeled, 31-bp DNA segments. Lanes: H, HeLa; X, X-63-1; E, EL-4; C, CTL; M, murine (CBA) spleen cells; B, bovine spleen cells. The sequence-specific complexes are indicated with an arrow.

protease (Fig. 7; data not shown) make this alternative less likely. In addition, no band indicative of complexes with only such a DNA-binding protein could be seen with extracts that were not treated with protease (Fig. 3, 6, 7, and 8). Furthermore, no such band could be provoked when the binding reactions were performed in the presence of 1.5% of the detergents Nonidet P-40 or Triton X-100, which destabilize hydrophobic interactions (data not shown). To test further for the possible existence of proteins that were very protease sensitive, which would then have to be in molar excess and bind very efficiently to the DNA-binding protein, the proteinase K in treated nuclear extracts was inactivated by prolonged incubation with the protease inhibitor PMSF, and then more extract was added. This procedure did not result in a redistribution of material from the more rapidly migrating complex to the more slowly migrating complexes (data not shown), also arguing against the possibility that there were such very protease-sensitive proteins.

DISCUSSION

We detected interactions of the nuclear protein SEF1 with a specific DNA sequence that included but was not limited to the region 5'-CTGTGGTTAA-3' of the 72-bp DNA repeat in the enhancer of the T-lymphocyte-specific mouse retrovirus SL3-3. We found that such SEF1 proteins are much more abundant in cell lines of T lymphocytes and in lymphoid organs than in the other cell lines or organs tested, in agreement with the T-lymphocyte preference found in the activity of this segment of the enhancer. We also found that limited protease treatments cleave the SEF1 proteins such that their DNA-binding domain(s) remains and creates complexes with indistinguishable electrophoretic mobilities and new properties.

Results of the DNase I footprinting assay that was initially used demonstrated that a dominant interaction with the 72-bp DNA repeat was located at the sequence that was studied, at least in the T-lymphocyte cell line CTL. Interactions with this DNA sequence were more difficult to demonstrate for other cell lines tested by the DNase I footprinting technique. This could be due to the lower concentration

of interacting SEF1 proteins in the extracts of the other cell lines (Fig. 3A). The concentrations of SEF1 proteins could be too low in those extracts to achieve the saturation of the DNA molecules with protein, which is necessary to avoid DNase I cleavage of free DNA obscuring footprints. Another possibility, which cannot be excluded at present, is that the SEF1 proteins from CTL, which create complexes both with lower and slightly higher mobilities than those found in substantial amounts in the other cell lines, have such DNA-binding properties that they create DNase I footprints more efficiently than do the other SEF1 proteins. Several protein-DNA interactions that were possible to analyze by electrophoretic mobility shift assay but not by DNase I footprinting, because of an equilibrium between free and bound DNA in solution that is too rapid (less than 1 min), have previously been reported in the immunoglobulin heavy-chain enhancer system (60). During the course of this study, a protein from HeLa cells which binds to SV40 DNA in a gel retardation assay but not in a DNase I footprinting assay was also reported (55).

When the remaining 38-bp repeat is deleted from a derivative of the SL3-3 enhancer carrying one copy each of the 34- and 38-bp repeats, the stimulation of transcription decreases about twofold in the T-lymphocyte cell lines CTL and EL-4, while no decrease in transcription has been seen in a cell line of the B-cell lineage, X-63-1, or in HeLa cells (Hallberg and Grundström, in preparation). Here we have reported that a mutation of the recognition sequence for the SEF1 proteins decreases transcription *in vivo* twofold in the T-lymphocyte cell lines EL-4 and CTL, in contrast to the X-63-1 cell line, in which no decrease was observed (Fig. 2 and Table 1). This strongly argues that the SEF1-binding site mediates the enhancer activity in T cells of the duplicated 38-bp sequence.

The observed twofold effect on transcription is significant for the assay method used (79), but it is far from drastic. However, enhancer activity has been shown to result from the integrated function of multiple sequence elements (79; see also references therein) which bind *trans*-acting factors (9, 76). Most of the individual protein-DNA interactions of the SV40 enhancer affected transcription *in vivo* in HeLa

cells less than fourfold, and all of them affected transcription less than sevenfold, although the effect of the complete enhancer was several thousand-fold (76, 79). In light of these results, the level of the effect of mutating the SEF1-binding sequence of the SL3-3 enhancer studied here may turn out to be rather common for individual sequence elements of intact enhancers.

Homologies between several of the sequence elements have been pointed out in different enhancers (79). The most commonly found sequence elements share the sequence 5'-TGTGG-3'. Such elements have even been called the core of the enhancers (31, 73). The range of enhancers in which TGTGG sequences have been found includes also those with very high cell-type specificities (53, 71). We have found that the TGTGG sequence of the SL3-3 enhancer is important for transcription *in vivo* in T-lymphocyte cell lines but not in other tested cells (Fig. 2 and Table 1) (Hallberg and Grundström, in preparation). How could such a common type of DNA sequence mediate different effects in different cell types? An explanation supported by the results of the present study is that the TGTGG sequence is only a common part of a family of longer DNA sequences that interact with proteins of different sequence specificities. Hence, we have found several different nucleotide positions outside of the TGTGG sequence of SL3-3 in which methylation of a purine or a nucleotide substitution inhibits the binding of the SEF1 proteins (Fig. 4, 5, and 6A and B). Similar sequences in many other enhancers differed from those of SL3-3 in several of these positions (79). We also found that in the formation of complexes between the SEF1 proteins and the SL3-3 sequence studied here, no competition was observed with any one of a set of binding sites for different DNA-binding proteins that affect transcription (Fig. 6C and D), including the two TGTGG-type sequences of the SV40 enhancer. These results indicate that the SEF1 proteins are distinct from several other proteins that bind to related sequences. During the course of this study, the two TGTGG-type sequences of the SV40 enhancer were shown to bind nuclear proteins that were distinct from each other (77), despite the fact that these two nucleotide sequences were more related to each other than to the SEF1-binding sequence. Thus, the SEF1 proteins appear to belong to a TGTGG motif family of proteins that bind to related but distinct DNA sequences.

A trivial explanation for the qualitative or quantitative differences between complexes formed with extracts of different cell lines could be a difference in the protease contents, leading to various attacks *in vitro* on the DNA-binding proteins. To minimize such problems, we used a mixture of eight protease inhibitors in the preparations of the nuclear extracts (see above). If, instead, the extracts were prepared with only EDTA and PMSF as protease inhibitors, the complexes obtained with the CTL and myeloma cell lines demonstrated clear signs of protease attack on the SEF1 proteins (decreased amounts of complex and complexes with decreased electrophoretic mobility shifts; Fig. 7 and data not shown). Also, with the first five of the eight added inhibitors listed earlier in this report, problems with protease attack on SEF1 remained for the CTL cell line (data not shown). Thus, addition of sufficient protease inhibitors is essential when cell type differences in the contents of transcription factors in nuclei are studied. With the eight inhibitors added, the CTL cell line yielded the same major complexes as the cell lines that showed the smallest protease problems, arguing that sufficient protease inhibition was obtained. The CTL cell line, in fact, yielded a complex with a higher electrophoretic

mobility shift than those found in substantial amounts in other cell lines. This indicates that the set of complexes obtained with the different cell lines represent different SEF1 proteins rather than proteins cleaved during preparation. However, the possibility of remaining artifact bands due to protease cleavage cannot be excluded. This is especially the case for the tissues, in which, in addition to the complexes obtained with the cell lines, complexes with lower electrophoretic mobility shifts were also obtained.

In the present study we have demonstrated that, on limited protease treatment, the SEF1 proteins yield DNA-binding domains with less variable electrophoretic mobility shift effects on the DNA than the complexes with the nontreated proteins. Cleavage of the extracts with proteinase K, protease V8, or papain resulted in DNA sequence-specific complexes with nonseparated mobilities during electrophoresis for all the cell lines tested. Trypsin treatment of the extracts yielded complexes with slightly higher mobilities than those of the other complexes. This difference could indicate that the site(s) most efficiently cleaved with trypsin is at a position(s) that is slightly separated from the predominant cleavage sites of the other three proteases. Despite the stabilizing conditions of the electrophoretic mobility shift assay (12–14, 16, 66), the complexes with the SEF1 proteins were not found to be completely stable; this was observed as smearing and a lower amount of complexed DNA, especially during prolonged electrophoresis. In contrast, the complexes with protease-cleaved SEF1 were much more stable (cf. the smearing after 1 h of electrophoresis of SEF1 in Fig. 3 and after 3 h of electrophoresis of protease-cleaved SEF1 in Fig. 9). The higher stabilities and the much less variable mobilities during electrophoresis of the protease-cleaved SEF1 proteins are, together with the differential protease sensitivities of the different parts of the proteins, intriguing in the context of current ideas concerning the structure of enhancer and promoter-binding proteins. A favored hypothesis for the ability of enhancer and promoter structures to tolerate changes of distances between protein-binding sites on the DNA and to show functional bidirectionality without apparent symmetry of these binding sites is a structure of enhancer-binding proteins composed of two functional domains linked by a flexible stem: a DNA-binding domain and a domain that interacts with the transcription machinery (67). A structure with separate domains linked by a flexible stem is known to exist in several procaryotic DNA-binding proteins: the lambda repressor cI (26, 50, 74), the lambda Cro protein (10, 68 and references therein), the catabolite activator protein CAP (44), and the LexA protein (39). Support for the existence of such a type of structure among DNA-binding proteins also comes from results of studies of an upstream-activating sequence in yeasts (30) and receptors of steroid hormones (17, 78). During the course of this study, a CCAAT box-binding protein was also shown to have a structurally distinct DNA-binding domain (27). The demonstrated effects of proteinase K, protease V8, papain, and trypsin experimentally support the hypothesis that the enhancer-binding SEF1 proteins also have such a structure. The flexible stem would then correspond to the protein structure found to be most vulnerable to protease; and the domain(s) of protein-protein interactions with other components of the transcription machinery, with its ability to move relative to the rest of the complex, would then destabilize it in the gel.

The primary criteria for dividing important control elements for the initiation of transcription of eucaryotic protein-coding genes into enhancer and promoter elements have

been the ability of enhancer elements to function also when they are located far from the point of initiation of transcription and to support different types of steric restrictions. The results presented here argue that the SEF1 proteins have a flexible hinge region between their DNA-binding domain(s) and a domain(s) that mediates interactions that are important for other properties of these proteins. It is tempting to speculate that the degree of the flexibility of such a hinge region in enhancer-binding proteins could be a parameter of major importance for creating the steric flexibility used to distinguish enhancer elements from promoter elements.

The differences observed in the electrophoretic mobilities of the SEF1 proteins were found mainly before a domain(s) that did not bind to the DNA sequence was cleaved off. Cell type-specific proteins with highly homologous DNA-binding domains but with differences in a domain(s) with other functions could exert quite different effects on transcription in a certain cell type or together with a certain promoter, depending on their ability to cooperate with other, cell-specific or more general proteins that interact with other components of the enhancer, with promoter elements, or both. With such different *trans*-activating binding proteins, a single DNA sequence could mediate different effects on transcription, depending on the set of such proteins expressed and made available in different cells. During the course of this study, different proteins, some of which have only been found in a narrow group of cell types, have been described (35, 55, 65) that bind with very related DNA specificities to a 5'-ATTGTCAT-3' sequence found in the immunoglobulin κ gene promoter and several enhancers. Thus, different enhancer-binding proteins with cell type differences in their expression from another system are known to share very related DNA sequence specificities.

Which molecular mechanisms could mediate the existence of SEF1 proteins with similar or identical DNA-binding domains but with more variety in other domains? This could occur at many different levels. It could be based on the existence of families of very closely related genes, differential gene rearrangements, differential splicing of the mRNA, or different types of posttranslational modifications, such as differential protease cleavages, glycosylations, phosphorylations, or some other type of modification. Purification and characterization of the SEF1 proteins and cloning of the gene(s) will be of crucial importance for an understanding of the molecular basis behind the differences between them.

AKV, a virus very closely related to SL3-3, does not preferentially infect thymocytes (56), does not efficiently cause T-cell leukemia (36), and has an enhancer element that is not preferentially active in T cells (7) despite the fact that it only has a few sequence differences compared with the enhancer element of SL3-3 (36). Interestingly, the enhancer of AKV is different from that of SL3-3 in the SEF1-binding site; there is a C instead of a T at position 190 (Fig. 5). We have shown here that a DNA segment with the SL3-3 sequence binds SEF1 much better than a segment with the AKV sequence, which is represented by mutation 4 in Fig. 6A and B. Thus, the differences in T-cell preference, both in the viruses and in their enhancers, correlates with differences in the efficiencies of the DNA sequences of SL3-3 and AKV studied here to bind to a transcription factor, SEF1, that is preferentially made in T cells. It is therefore tempting to speculate that the binding of SEF1 is an essential factor for the T-cell tropism of SL3-3 and the ability of this virus to cause leukemia. This hypothesis should be tested in additional experiments, both in vitro and in different in vivo systems.

ACKNOWLEDGMENTS

We thank Martin Gullberg, Helena Ohlsson, and Thomas Edlund for helpful discussions; Christine Grundström and Pia Nilsson for skillful technical assistance; Håkan Persson, Thomas Leandersson, and Fredrik Ivars for gifts of cell lines; and Peter Brodin for help with illustrations.

This study was supported by grant Dnr 4892-104 from the Swedish Natural Science Research Council, grant Dnr 2235-B87-01XA from the Swedish Cancer Society, and grant Dnr 84-5413B from the National Swedish Board for Technical Development. T. Grundström was the recipient of special researcher position grant Dnr 4892-107 from the Swedish Natural Science Research Council.

LITERATURE CITED

1. Auffray, C., and F. Rougeon. 1980. Purification of mouse immunoglobulin heavy-chain messenger RNAs from total myeloma tumor RNA. *Eur. J. Biochem.* **107**:303-314.
2. Baker, P. E., S. Gillis, and K. A. Smith. 1979. Monoclonal cytolytic T-cell lines. *J. Exp. Med.* **149**:273-278.
3. Banerji, J., L. Olson, and W. Schaffner. 1983. A lymphocyte-specific enhancer is located downstream of the joining region in immunoglobulin heavy chain genes. *Cell* **33**:729-740.
4. Banerji, J., S. Rusconi, and W. Schaffner. 1981. Expression of a β -globin gene is enhanced by remote SV40 DNA sequences. *Cell* **27**:299-308.
5. Biggins, M. D., T. J. Gibson, and G. F. Hong. 1983. Buffer gradient gels and ^{35}S label as an aid to rapid DNA sequence determination. *Proc. Natl. Acad. Sci. USA* **80**:3963-3965.
6. Carthew, R. W., L. A. Chodosh, and P. A. Sharp. 1985. An RNA polymerase II transcription factor binds to an upstream element in the adenovirus major late promoter. *Cell* **43**:439-448.
7. Celander, D., and W. A. Haseltine. 1984. Tissue-specific transcription preference as a determinant of cell tropism and leukemogenic potential of murine retroviruses. *Nature (London)* **312**:159-162.
8. Chambon, P., A. Dierich, M. P. Gaub, S. Jakowlev, J. Jongstra, A. Krust, J. P. LePennec, P. Oudet, and T. Reudelhuber. 1984. Promoter elements of genes coding for proteins and modulation of transcription by oestrogens and progesterone, p. 1-42. *In* R. O. Greep (ed.), *Proceedings of the Laurentian Hormone Conference*, vol. 40. Academic Press, Inc., New York.
9. Davidson, I., C. Fromenthal, P. Augereau, A. Wildeman, M. Zenke, and P. Chambon. 1986. Cell-type specific protein binding to the enhancer of simian virus 40 in nuclear extracts. *Nature (London)* **323**:544-548.
10. de Crombrughe, B., S. Busby, and H. Buc. 1984. Review. Cyclic AMP receptor protein: role in transcription activation. *Science* **224**:831-838.
11. Dignam, J. D., R. M. Lebovitz, and R. G. Roeder. 1983. Accurate transcription initiation by RNA polymerase II in a soluble extract from isolated mammalian nuclei. *Nucleic Acids Res.* **11**:1475-1489.
12. 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.
13. Fried, M. G., and D. M. Crothers. 1984. Kinetics and mechanism in the reaction of gene regulatory proteins with DNA. *J. Mol. Biol.* **172**:263-282.
14. Fried, M. G., and D. M. Crothers. 1984. Equilibrium studies of the cyclic AMP receptor protein-DNA interaction. *J. Mol. Biol.* **172**:241-262.
15. 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.
16. Garner, M. M., and A. Revzin. 1981. A gel electrophoresis method for quantifying the binding of proteins to specific DNA regions: applications to components of the *Escherichia coli* lactose operon regulatory system. *Nucleic Acids Res.* **9**:3047-3060.
17. Giguère, V., S. M. Hollenberg, M. G. Rosenfeld, and R. M. Evans. 1986. Functional domains of the human glucocorticoid receptor. *Cell* **46**:645-652.
18. Gillis, S., and J. Watson. 1980. Biochemical and biological

- characterization of lymphocyte regulatory molecules. *J. Exp. Med.* **152**:1709-1719.
19. Glutzman, Y., and T. Shenk. 1983. Current communications in molecular biology: eukaryotic gene expression. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
 20. Gorer, P. A. 1950. Studies in antibody response of mice to tumour inoculation. *Br. J. Cancer* **4**:372-379.
 21. Graves, B. J., P. F. Johnson, and S. L. McKnight. 1986. Homologous recognition of a promoter domain common to the MSV LTR and the HSV tk gene. *Cell* **44**:565-576.
 22. Grundström, T., M. W. Zenke, M. Wintzerith, H. W. D. Matthes, A. Staub, and P. Chambon. 1985. Oligonucleotide-directed mutagenesis by microscale "shot-gun" gene synthesis. *Nucleic Acids Res.* **13**:3305-3316.
 23. Hen, R., E. Borrelli, P. Sassone-Corsi, and P. Chambon. 1983. An enhancer element is located 340 base pairs upstream of the adenovirus-2 E1A capsite. *Nucleic Acids Res.* **11**:8747-8760.
 24. Hendrickson, W., and R. Schlieff. 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.
 25. Hendrickson, W., and R. F. Schlieff. 1984. Regulation of the *Escherichia coli* L-arabinose operon studied by gel electrophoresis DNA binding assay. *J. Mol. Biol.* **178**:611-628.
 26. Hochschild, A., N. Irwin, and M. Ptashne. 1983. Repressor structure and the mechanism of positive control. *Cell* **32**:319-325.
 27. Hooft van Huijsduijnen, R. A. M., J. Bollekens, A. Dorn, C. Benoist, and D. Mathis. 1987. Properties of a CCAAT box-binding protein. *Nucleic Acids Res.* **15**:7265-7282.
 28. Kappler, J. W., B. Skidmore, J. White, and P. Marrack. 1981. Antigen-inducible, H-2-restricted, interleukin-2-producing T cell hybridomas. *J. Exp. Med.* **153**:1198-1214.
 29. Kearney, J. F., A. Radbruch, B. Liesegang, and K. Rajewsky. 1979. A new mouse myeloma cell line that has lost immunoglobulin expression but permits the construction of antibody-secreting hybrid cell lines. *J. Immunol.* **123**:1548-1550.
 30. Keegan, L., G. Gill, and M. Ptashne. 1986. Separation of DNA binding from the transcription-activating function of a eucaryotic regulatory protein. *Science* **231**:699-704.
 31. Khory, G., and P. Gruss. 1983. Enhancer elements. *Cell* **33**:313-314.
 32. Kim, K. J., C. Kanellopoulos-Langevin, R. M. Merwin, D. H. Sachs, and R. Asofsky. 1979. Establishment and characterization of BALB/C lymphoma lines with B cell properties. *J. Immunol.* **122**:549-554.
 33. Kim, K. J., F. I. Weinbaum, B. J. Mathieson, P. E. McKeever, and R. Asofsky. 1978. Characteristics of BALB/c T cell lymphomas grown as continuous *in vitro* lines. *J. Immunol.* **121**:339-344.
 34. Köhler, G., and C. Milstein. 1975. Continuous cultures of fused cells secreting antibody of predefined specificity. *Nature (London)* **256**:495-497.
 35. Landolfi, N. F., J. D. Capra, and P. W. Tucker. 1986. Interaction of cell-type-specific nuclear proteins with immunoglobulin V_H promoter region sequences. *Nature (London)* **323**:548-551.
 36. Lenz, J., D. Celander, R. L. Crowther, R. Patarca, D. W. 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.
 37. Lenz, J., R. Crowther, S. Klimenko, and W. A. Haseltine. 1982. Molecular cloning of a highly leukemogenic, ecotropic retrovirus from an AKR mouse. *J. Virol.* **43**:943-951.
 38. Levinger, L. F. 1985. D1 protein of *Drosophila melanogaster*. *J. Biol. Chem.* **260**:14311-14318.
 39. Little, J. W., and S. A. Hill. 1985. Deletions within a hinge region of a specific DNA-binding protein. *Proc. Natl. Acad. Sci. USA* **82**:2301-2305.
 40. Lusky, M., L. Berg, H. Weiher, and M. Botchan. 1983. Bovine papilloma virus contains an activator of gene expression at the distal end of the early transcription unit. *Mol. Cell. Biol.* **3**:1108-1122.
 41. Maniatis, T., E. F. Fritsch, and J. Sambrook. 1982. Molecular cloning: a laboratory manual, p. 164-165. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
 42. Matthes, H. W. D., W. M. Zenke, T. Grundström, A. Staub, M. Wintzerith, and P. Chambon. 1984. Simultaneous rapid chemical synthesis of over one hundred oligonucleotides on a microscale. *EMBO J.* **3**:801-805.
 43. Maxam, A. M., and W. Gilbert. 1977. A new method for sequencing DNA. *Proc. Natl. Acad. Sci. USA* **74**:560-564.
 44. McKay, D. B., and T. A. Steitz. 1981. Structure of catabolite gene activator protein at 2.9 Å resolution suggests binding to left-handed B-DNA. *Nature (London)* **290**:744-749.
 45. Mercola, M., J. Goverman, C. Mirell, and K. Calame. 1985. Immunoglobulin heavy-chain enhancer requires one or more tissue specific factors. *Science* **227**:266-270.
 46. Mohit, B., and K. Fan. 1971. Hybrid cell line from a cloned immunoglobulin-producing mouse myeloma and a nonproducing mouse lymphoma. *Science* **171**:75-77.
 47. Nagata, K., R. A. Guggenheimer, and J. Hurwitz. 1983. Specific binding of a cellular DNA replication protein to the origin of replication of adenovirus DNA. *Proc. Natl. Acad. Sci. USA* **80**:6177-6181.
 48. Ohlsson, H., and T. Edlund. 1986. Sequence-specific interactions of nuclear factors with the insulin gene enhancer. *Cell* **45**:35-44.
 49. Oi, V. T., S. L. Morrison, L. A. Herzenberg, and P. Berg. 1983. Immunoglobulin gene expression in transformed lymphoid cells. *Proc. Natl. Acad. Sci. USA* **80**:825-829.
 50. Pabo, C. O., R. T. Sauer, J. M. Sturtevant, and M. Ptashne. 1979. The λ repressor contains two domains. *Proc. Natl. Acad. Sci. USA* **76**:1608-1612.
 51. Pederson, F. S., R. L. Crowther, D. Y. Tenney, A. M. Reimhold, and W. A. Haseltine. 1981. Novel leukaemogenic retroviruses isolated from cell line derived from spontaneous AKR tumor. *Nature (London)* **292**:167-170.
 52. Piette, J., M. H. Kryszke, and M. Yaniv. 1985. Specific interaction of cellular factors with the B enhancer of polyoma virus. *EMBO J.* **4**:2675-2685.
 53. Queen, C., and J. Stafford. 1984. Fine mapping of an immunoglobulin gene activator. *Mol. Cell. Biol.* **4**:1042-1049.
 54. Robb, R. J., A. Munck, and K. A. Smith. 1981. T cell growth factor receptors. Quantitations, specificity and biological relevance. *J. Exp. Med.* **154**:1455-1474.
 55. Rosales, R., M. Vigneron, M. Macchi, I. Davidson, J. H. Xiao, and P. Chambon. 1987. *In vitro* binding of cell-specific and ubiquitous nuclear proteins to the octamer motif of the SV40 enhancer and related motifs present in other promoters and enhancers. *EMBO J.* **6**:3015-3025.
 56. Rosen, C. A., W. A. Haseltine, J. Lenz, R. Ruprecht, and M. W. Cloyd. 1985. Tissue selectivity of murine leukemia virus infection is determined by long terminal repeat sequences. *J. Virol.* **55**:862-866.
 57. Sassone-Corsi, P., A. Wildeman, and P. Chambon. 1985. A trans-acting factor is responsible for the simian virus 40 enhancer activity *in vitro*. *Nature (London)* **313**:458-463.
 58. Schneider, R., I. Gander, U. Müller, R. Mertz, and E. L. Winnacker. 1986. A sensitive and rapid gel retention assay for nuclear factor I and other DNA-binding proteins in crude nuclear extracts. *Nucleic Acids Res.* **14**:1303-1324.
 59. Schöler, H. R., and P. Gruss. 1984. Specific interaction between enhancer-containing molecules and cellular components. *Cell* **36**:403-411.
 60. Sen, R., and D. Baltimore. 1986. Multiple nuclear factors interact with the immunoglobulin enhancer sequences. *Cell* **46**:705-716.
 61. Serfling, E., M. Jasin, and W. Schaffner. 1985. Enhancers and eucaryotic gene transcription. *Trends Genet.* **1**:224-230.
 62. Sergeant, A., D. Bohmann, H. Zentgraf, H. Weiher, and W. Keller. 1984. A transcription enhancer acts *in vitro* over distances of hundreds of base-pairs on both circular and linear templates but not on chromatin-reconstituted DNA. *J. Mol. Biol.* **180**:577-600.
 63. 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.

64. Singh, H., R. Sen, D. Baltimore, and P. A. Sharp. 1986. A nuclear factor that binds a conserved sequence motif in transcriptional control elements of immunoglobulin genes. *Nature (London)* 319:154-158.
65. Staudt, L. M., H. Singh, R. Sen, T. Wirth, P. A. Sharp, and D. Baltimore. 1986. A lymphoid-specific protein binding to the octamer motif of immunoglobulin genes. *Nature (London)* 323:640-643.
66. Strauss, F., and A. Varshavsky. 1984. A protein binds to a satellite DNA repeat at three specific sites that would be brought into mutual proximity by DNA folding in the nucleosome. *Cell* 37:889-901.
67. Takahashi, K., M. Vigneron, H. Matthes, A. Wildeman, M. Zenke, and P. Chambon. 1986. Requirement of stereospecific alignments for initiation from the simian virus 40 early promoter. *Nature (London)* 319:121-126.
68. Takeda, Y., D. H. Ohlendorf, W. F. Anderson, and B. W. Matthews. 1983. DNA-binding proteins. *Science* 221:1020-1026.
69. Tooze, J. 1982. DNA tumor viruses, 9th ed. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
70. van Ooyen, A., J. van der Berg, N. Mantel, and C. Weissmann. 1979. Comparison of total sequence of a cloned rabbit β -globin gene and its flanking regions with a homologous mouse sequence. *Science* 206:337-344.
71. Walker, M. D., T. Edlund, A. M. Boulet, and W. J. Rutter. 1983. Cell-specific expression controlled by the 5'-flanking region of insulin and chymotrypsin genes. *Nature (London)* 306:557-561.
72. Weiher, H., and M. R. Botchan. 1984. An enhancer sequence from bovine papilloma virus DNA consists of two essential regions. *Nucleic Acids Res.* 12:2901-2916.
73. Weiher, H., M. König, and P. Gruss. 1983. Multiple point mutations affecting the simian virus 40 enhancer. *Science* 219:626-631.
74. Wharton, R. P., E. L. Brown, and M. Ptashne. 1984. Substituting an α -helix switches the sequence-specific DNA interactions of a repressor. *Cell* 38:361-369.
75. Wildeman, A. G., P. Sassone-Corsi, T. Grundström, M. Zenke, and P. Chambon. 1984. Stimulation of *in vitro* transcription from the SV 40 early promoter by the enhancer involves specific trans-acting factors. *EMBO J.* 3:3129-3133.
76. Wildeman, A. G., M. Zenke, C. Schatz, M. Wintzerith, T. Grundström, H. Matthes, K. Takahashi, and P. Chambon. 1986. Specific protein binding to the simian virus 40 enhancer *in vitro*. *Mol. Cell. Biol.* 6:2098-2105.
77. Xiao, J. H., I. Davidson, D. Ferrandon, R. Rosales, M. Vigneron, M. Macchi, F. Ruffenach, and P. Chambon. 1987. One cell-specific and three ubiquitous nuclear proteins binds *in vitro* to overlapping motifs in the domain B1 of the SV40 enhancer. *EMBO J.* 6:3005-3013.
78. Yamamoto, K. R. (ed.). 1985. Steroid receptor regulated transcription of specific genes and gene networks. *Annu. Rev. Genet.* 19:209-252.
79. Zenke, M., T. Grundström, H. Matthes, M. Wintzerith, C. Schatz, A. G. Wildeman, and P. Chambon. 1986. Multiple sequence motifs are involved in SV 40 enhancer function. *EMBO J.* 5:387-397.