DNA-Binding Proteins That Interact with the Long Terminal Repeat of Radiation Leukemia Virus

IZABELLA GORSKA-FLIPOT^{1†} AND PAUL JOLICOEUR^{1,2*}

Laboratory of Molecular Biology, Institut de recherches cliniques de Montréal, 110, avenue des Pins Ouest, Montreal, Quebec, Canada H2W 1R7,¹ and Département de microbiologie et d'immunologie, Faculté de médecine, Université de Montréal, Montreal, Quebec, Canada H3C 3J7²

Received 6 September 1989/Accepted 28 December 1989

We used the electrophoretic mobility shift assay to identify the interactions of nuclear proteins with the long terminal repeat of leukemogenic, thymotropic BL/VL3 radiation leukemia virus (RadLV). In the promoter region, we identified a CCAAT box-binding protein (CBP) that has the same binding characteristics as the CCAAT box-binding protein that binds to the Moloney murine sarcoma virus promoter and most likely represents the CP1 factor. In the upstream enhancer region unique to BL/VL3, we detected several sequence-specific complexes, one with T-lymphocyte extracts but not with fibroblast extracts. This U3 region, UEB, may be important for the T-cell specificity of BL/VL3 RadLV. In the enhancer, which has been uniquely rearranged in this virus, we identified three specific protein-binding sites. Two of them showed characteristics of the LVb and core binding sites previously described for other murine retroviruses. But one binding site, identified as Rad-1, is unique to BL/VL3 RadLV and was found downstream, only 1 nucleotide from the core sequence. Rad-1 has a corelike motif on the minus strand, and the factor that binds to it could be competed by a BL/VL3 core-containing fragment. Moreover, the protein-DNA contacts involve the typical three core Gs separated by one T. These results suggest that Rad-1 binds a factor identical or similar to the core-binding factor. Our data suggest that the LVb, core, and Rad-1 motifs may be sufficient for this enhancer, most likely in association with other U3 long terminal repeat sequences, to promote a high rate of transcription of BL/VL3 RadLV in its specific target cells (thymocytes).

BL/VL3 Kaplan radiation leukemia virus (RadLV) is a strain of ecotropic murine leukemia virus (MuLV) which shows a strong tropism for thymocytes *in vivo* and *in vitro* and is highly leukemogenic, inducing T-cell lymphomas (thymomas) in susceptible mice (9, 25, 34, 35, 45). BL/VL3 RadLV is one of the few ecotropic MuLVs which do not replicate in mouse fibroblasts *in vitro* (35, 45), and this phenotype has been shown to be determined by its long terminal repeat (LTR) (43). LTRs of MuLV carry the promoter and enhancer sequences (32). The BL/VL3 LTRs are significantly different from other MuLV LTRs, mainly in the U3 region (22, 43, 45). The U3 enhancer region has been rearranged and shows three copies of a 43-base-pair (bp) direct repeat separated by 11-bp direct repeats (43). Some U3 sequences remain homologous to sequences of nonleukemogenic endogenous ecotropic MuLVs, but others appear to be unique to BL/VL3 and are distinct enough that we were able to derive a BL/VL3-specific probe which does not detect endogenous viral sequences of the mouse genome (43).

The LTRs of various strains of MuLVs have been shown to harbor the determinants of tissue (3, 4, 12, 20, 21, 36, 46) and disease (5, 6, 10, 11, 13, 14, 21, 31, 33, 52) specificity, and these determinants have been shown to reside primarily in the enhancer direct repeats (3, 4, 11, 12, 20, 21, 33). Enhancers make up a particular combination of regulatory *cis* elements that are uniquely arranged as to number, type, and spatial array. These elements represent binding sites for sequence-specific transcription factors, which may activate transcription in a tissue- or cell-specific manner (for reviews,

see references 38, 41, and 42). Several such factors that interact with numerous upstream elements, including enhancer sequences of many eucaryotic viral or cellular genes, have already been identified and characterized. Recently, the factors that bind to the LTRs of some MuLVs have been studied (2, 17-20, 39, 48, 51).

In the present study, we investigated the protein-binding potential of the BL/VL3 RadLV LTR and compared it with the protein-binding potential of the LTR of the nonleukemogenic ecotropic BALB/c endogenous MuLV. The two LTRs are homologous, but their enhancer direct repeats are differently organized. We found that the BL/VL3 U3 region harbors two unique binding sites apparently not shared by other MuLV strains.

MATERIALS AND METHODS

Preparation of probes and competitor fragments. The *Pst*I-*Kpn*I fragments of the LTR of BL/VL3 (clone 13) RadLV (45) and of N-CI-35 MuLV (12) were subcloned in the polylinker of vector pEMBL18 to give VL3-pEMBL and N-CI-35-pEMBL, respectively. The fragment containing the BL/VL3 promoter sequences was excised from the VL3-pEMBL plasmid by cleavage with *Sca*I (at position 307 in the LTR) and the polylinker *Eco*RI site. This 232-bp fragment was isolated on 1% agarose by electroelution on an NA-45 DEAE membrane, followed by elution as recommended by the manufacturer (Schleicher & Schuell, Inc.). It was digested by *Hin*PI and 3' end labeled in digestion buffer by Klenow polymerase by standard techniques (37). The resulting three subfragments of 129, 55, and 48 bp were separated on 10% polyacrylamide gel, and the radioactive fragments were excised from the gel and purified as previously described (40). In typical labeling, the specific radioactivity was 3,000 to 7,000 cpm/pmol. The promoter frag-

* Corresponding author.

† Present address: Ontario Cancer Institute, Toronto, Ontario, Canada M4X 1K9.

ment of the N-CI-35 LTR was excised by digestion with *Sau3A* (at position 275 of the LTR) and *KpnI* (position 419). The resulting 138-bp fragment was isolated from agarose, cut with *HinPI*, and labeled as described above. The radioactive 58-, 49-, and 31-bp fragments were separated on polyacrylamide gel and purified.

Radioactive probes labeled at the 5' end were prepared by dephosphorylation and labeling with [γ - 32 P]ATP and 10 U of T4 polynucleotide kinase (Pharmacia) as described previously (37). The 5'-end-labeled fragments were generated by subsequent digestion with an appropriate restriction enzyme. The specific radioactivity obtained varied from 4,000 to 8,000 cpm/pmol. The competitor fragments were prepared as described above, but the labeling steps were omitted.

Preparation of synthetic oligonucleotides. Each oligonucleotide was synthesized as two complementary strands and was obtained commercially.

Single-stranded DNAs were 5' end labeled by T4 polynucleotide kinase (37), and equimolar amounts of complementary nonlabeled oligonucleotides were annealed. Annealed DNAs were purified through 12% polyacrylamide gel. These DNAs were used as templates in gel retardation and methylation interference assays. For cold competitor fragments, the labeling step was omitted.

Cell lines and nuclear extracts. The mouse Ti-6 lymphocyte cell line has already been described (44). The origin of NIH 3T3 and SIM.R fibroblasts has already been given (45). Nuclear extracts were usually prepared from 10^9 cells in the logarithmic growth phase by the procedure described originally by Dignam et al. (15) with the modifications proposed by Wildeman et al. (54). The concentration of proteins as measured by the standard Bradford technique varied from 3 to 7 mg/ml.

Gel retardation assay and competition. Analytical binding to the end-labeled fragments of the promoter region was performed with 2 to 3 pmol of the probe (usually 10,000 to 15,000 cpm) and 15 μ g of the nuclear extract in 20 μ l of 10 mM Tris hydrochloride (pH 7.5)–100 mM NaCl–1 mM EDTA–1 mM 2-mercaptoethanol–2 μ g of poly(dI · dC) (54). The samples were incubated at 28°C for 30 min and resolved on 4% native polyacrylamide gel as described by Strauss and Varshavski (49). The binding assay with the oligonucleotides was modified as follows. The concentration of the labeled templates was increased to 5 pmol, and the protein concentration was increased to 30 μ g per binding reaction; the NaCl concentration was decreased to 40 mM, and only 0.5 μ g of unspecific competitor poly(dI · dC) was added. In competition experiments, a 100-fold excess of the cold competitor fragment was mixed with the nuclear extract and the first incubation was for 10 min on ice. An appropriate radiolabeled fragment was then added and incubation was continued for the next 20 min at 28°C before loading on the acrylamide gel.

DNase I footprinting and methylation interference. DNase I footprinting and methylation interference procedures were performed essentially as described by Kovacs et al. (27) and Speck and Baltimore (48), respectively.

RESULTS

Characteristics of the LTR studies. N-tropic N-CI-35 MuLV is a fibrotropic, nonthymotropic, nonleukemogenic virus in most of the strains tested (12). BL/VL3 clone V13 RadLV is a nonfibrotropic, thymotropic, highly leukemogenic MuLV (43, 45). Its inability to replicate on mouse fibroblasts in vitro has been found to be caused by its LTR

(43). The enhancer region of its LTR has been rearranged significantly. It contains three nearly perfect reiterations of 43-bp repeats which are separated from each other by two identical 11-bp sequences. This 43-bp repeated sequence is homologous to the internal part of a 99-bp single repeat of the N-CI-35 LTR (Fig. 1). Besides their differences in the number of repeats, the LTRs of N-CI-35 and BL/VL3 are highly homologous (95%), except in the region immediately upstream of the first direct repeat of BL/VL3 (nucleotides 101 to 131 in Fig. 1A).

The promoter of BL/VL3 RadLV is a sequence typical of most murine retroviruses, except that its TATA box is preceded by a C-rich insertion forming two 6-bp direct repeats (45). These repeats are not present in N-CI-35 MuLV (12) but are present in some other endogenous proviruses (26). The R and U5 regions of the BL/VL3 RadLV LTR are distinct from those of N-CI-35 MuLV by some point mutations or point insertions.

Proteins that interact with promoters of BL/VL3 and N-CI-35 MuLVs. We performed gel retardation assays with DNAs derived from the promoter regions of the LTRs of BL/VL3 RadLV and BALB/c endogenous N-tropic (N-CI-35) MuLV. Nuclear extracts were prepared from the T-lymphoid Ti-6 cell line, which supports replication of both viruses, and from SIM.R fibroblasts, which are nonpermissive for BL/VL3 RadLV but support replication of BALB/c endogenous N-CI-35 MuLV (43, 45). The promoter region of each virus was divided into three subfragments containing the CCAAT box, the TATA box, or the R sequence, respectively (Fig. 1B). We were not able to show any significant interaction between nuclear factors and the TATA box, the C-rich short direct repeat localized in front of the BL/VL3 TATA box (55-bp subfragment), or the 48-bp R fragment (data not shown). However, both N-CI-35 and BL/VL3 promoter subfragments containing the CCAAT box sequence (58 and 129 bp, respectively) reacted with a protein(s) from both nuclear extracts, as shown by the formation of retarded bands (Fig. 2A, lane 1; data not shown). Both the BL/VL3 and N-CI-35 LTR subfragments formed complexes of the same relative mobility. The proteins that interacted with the BL/VL3 and N-CI-35 CCAAT box regions appeared to be similar, as revealed by competition experiments with homologous (BL/VL3) or heterologous (N-CI-35) DNA (data not shown).

To establish more precisely which nucleotides were involved in binding with the protein(s), we performed DNase I protection and methylation interference assays. The region protected from DNase I digestion on the BL/VL3 promoter covered the CCAAT sequence entirely and was the same for fibroblast and lymphocyte nuclear extracts (Fig. 2B). The pattern of protection of the N-CI-35 promoter was similar. The close contact between proteins and BL/VL3 DNA, as established by methylation interference assay, revealed that the Gs that interacted with the protein(s) are the two Gs on the minus strand (data not shown); these are complementary to the two cytosines of the CCAAT sequence (Fig. 2B). Together, these results indicate that (i) identical or very similar proteins interact with the N-CI-35 and BL/VL3 promoters and (ii) identical or very similar proteins are present in mouse lymphocytes and fibroblasts.

Characterization of the CCAAT-binding protein of BL/VL3 RadLV. The sequence recognized by the CCAAT-specific protein on the BL/VL3 promoter is identical, except for one base pair (at position 416), to the binding site of one CCAAT-binding protein, E/CBP (18, 19, 28), known to interact with the CCAAT sequence of Moloney murine sarcoma virus

A

B

FIG. 1. (A) Partial nucleotide sequence of the BL/VL3 (clone 13) RadLV LTR. For comparison, the corresponding region of BALB/c N-tropic endogenous (N-Cl-35) MuLV is shown. Both complete sequences have been published previously (12, 43). The nucleotides identical in both LTRs are indicated as dots, and spaces left for insertion in the corresponding region are indicated as dashes. The 43-bp (thin lines) and 11-bp (thick lines) direct repeats of BL/VL3 RadLV are indicated under the sequence by horizontal arrows. The synthetic oligonucleotides UEA, UEB, DRA, DRB, and DRC are indicated as open boxes above the sequences. (B) Schematic representation of the LTR structures of BL/VL3 RadLV (clone V13) and BALB/c N-tropic endogenous (N-Cl-35) MuLV. Lines represent the U3 and R regions; black boxes represent CCAAT and TATA boxes; open boxes represent direct repeats; shaded box represents an upstream enhancer region which does not share homology with the N-Cl-35 MuLV LTR. Arrows above the lines show the direction of the 99-, 43-, and 11-bp repeats. Arrows under the line represent the 6-bp repeats; the wavy line represents a polylinker. The positions of the N-Cl-35 138-, 58-, 49-, and 31-bp fragments and the BL/VL3 232-, 129-, 55-, and 48-bp fragments used in mobility shift assay, DNase I footprinting, and methylation interference are shown. The figure is not drawn to the scale. H, *Hin*PI; K, *Kpn*I; P, *Pst*I; Sc, *Sca*I; S, *Sau*3A. Not all sites are illustrated.

(MSV). To test whether the nuclear factor that bound to BL/VL3 was identical to E/CBP, we performed a competition binding assay. The binding of a nuclear factor(s) to the BL/VL3 promoter subfragment was effectively competed with an excess of the cold MSV DNA fragment containing the CCAAT box-binding protein (CBP) binding site (Fig. 2A, lane 2). In a separate experiment, the complexes that resulted from binding of a nuclear factor(s) to both the MSV and BL/VL3 CCAAT subfragments displayed identical electrophoretic mobilities. In addition, binding to MSV DNA could be competed by the cold CCAAT-containing BL/VL3 subfragment (data not shown). Together, these results strongly suggest that a protein(s) that interacts with the CCAAT regions of the BL/VL3 and N-Cl-35 MuLVs is identical or very similar to the CBP described by Graves et al. (19).

Proteins that interact with the U3 sequences of BL/VL3

RadLV. To study the DNA-protein interactions in the U3 LTR of BL/VL3, we synthesized five oligonucleotides (Fig. 1A). These were used in an electrophoretic mobility shift assay with lymphocyte (Ti-6) and fibroblast (SIM.R and NIH 3T3) nuclear extracts.

UEA DNA formed complexes which migrated as a single major band with both extracts (Fig. 3, lanes 1 and 2). This band was also seen when UEA single-stranded radiolabeled DNA was used as the template for binding. This complex could be competed by minus-strand or plus-strand cold single-stranded UEA DNA (data not shown). UEB DNA formed complexes which migrated as three retarded bands (*a*, *b*, and *c*) with Ti-6 lymphocyte nuclear extracts (Fig. 3, lane 4) and two retarded bands (*a* and *b*) with SIM.R and NIH 3T3 fibroblast extracts (lanes 5 and 6). Band *a* could be competed by UEB single-stranded DNA, but bands *b* and *c*

FIG. 2. Interactions of nuclear proteins with the promoter region of BL/VL3 RadLV and N-Cl-35 MuLV. (A) Competition for binding of a nuclear factor(s) to the BL/VL3 promoter by DNA containing a CBP binding site. The 129-bp CCAAT-containing subfragment of the BL/VL3 LTR was incubated with a Ti-6 lymphocyte nuclear extract without (lane 1) or with a 100-fold molar excess of the cold 120-bp CCAAT-containing MSV promoter subfragment (lane 2) or an unrelated 220-bp *HpaII* pBR322 fragment (lane 3). (B) Comparison of DNA-protein interactions identified for the upstream promoter region of BL/VL3 RadLV and endogenous BALB/c N-Cl-35 MuLV. The areas protected by the nuclear factor(s) from the DNase I digestion are bracketed; closed circles show the sites of strong contacts, as determined by methylation interference. The nucleotide differences between the two sequences in this region are indicated by arrows. The CCAAT box is underlined.

could not (data not shown). The additional band, c, found with lymphocyte extracts was not detected in all of the experiments. We were unable to show any sequence-specific interaction by the methylation interference assay performed on both strands of all of the complexes formed with both oligonucleotides (data not shown). One possible explanation for this effect is a relatively low content of G+C in this region; indeed, the technique used is able to show close protein-DNA contact only for sequences containing Gs. Alternatively, the binding protein(s), in contrast to most known proteins, bound to methylated Gs.

DRA DNA, which contains the 11-bp spacer and the 5' part of one direct repeat, formed one major complex (Fig. 3, lanes 7 to 9). The major retarded complexes formed with DRB and DRC (covering the 3' part of a direct repeat) migrated slightly faster than the complexes formed with DRA (Fig. 3, lanes 10 to 15). Occasionally, fainter, faster migrating bands were also seen. To determine whether the same protein(s) interacted with DRB and DRC DNAs, competition experiments were performed. The cold DRA template could not disrupt the retarded complexes formed with DRB or DRC (Fig. 4, lanes 4 and 8), while cold DRC efficiently competed with DRB complex formation (Fig. 4, lane 7). Also, cold DRB DNA competed, although less efficiently, for binding to radiolabeled DRC DNA (Fig. 4, lane 3). These results strongly suggest that the same factor(s) binds to both DRC and DRB templates.

Methylation interference was used to define these DNA-

FIG. 3. Binding of nuclear proteins to synthetic oligonucleotides derived from the upstream enhancer element and the direct repeat of the enhancer U3 LTR of BL/VL3 RadLV. Electrophoretic mobility shift assay of nuclear extracts prepared from Ti-6 lymphocytes (lanes 1, 4, 7, 10, and 13); SIM.R fibroblasts (lanes 2, 5, 8, 11, and 14); NIH 3T3 fibroblasts (lanes 3, 6, 9, 12, and 15); and oligonucleotides UEA (lanes 1 to 3), UEB (lanes 4 to 6), DRA (lanes 7 to 9); DRC (lanes 10 to 12), and DRB (lanes 13 to 15). The bands indicated by arrows represent single-strand-specific binding. An asterisk at the left of a band represents DNA used for the methylation interference experiments shown in Fig. 5. An open circle represents a complex shown to be made with a proteolytic fragment of the protein responsible for major complex formation (see Fig. 5, lane 15, and the text). Bound (B) and free (F) fractions are indicated. a, b, and c are single-stranded, double-stranded, and lymphocyte-specific complexes, respectively.

protein interactions more precisely (Fig. 5). The protein that bound to DRA interacted with two Gs on the plus strand and one G on the minus strand (Fig. 5, lanes 1 to 6). This G was strongly diminished in the DNA extracted from the bound fraction (Fig. 5, lanes 2 and 5), compared with the DNA from the free fraction (lanes 1 and 4). These Gs define the sequence $\begin{smallmatrix} \text{CAGGA} \\ \text{GTCCT} \end{smallmatrix}$, which is homologous to the binding site of factor LVb previously described for the Moloney MuLV enhancer (48). The interference pattern from the major retarded complex formed with DRC DNA localized the DNA contacts to the three Gs on the plus strand and one of

FIG. 4. Competition for binding of BL/VL3 core nuclear factors. Ti-6 nuclear extract was incubated with radiolabeled DRC without (lane 1) or with a 100-fold molar excess of cold DRC (lane 2), DRB (lane 3), or DRA (lane 4) DNA. The same nuclear extract was incubated with radiolabeled DRB without (lane 5) or with a 100-fold molar excess of cold DRB (lane 6), DRC (lane 7), or DRA (lane 8) DNA.

FIG. 5. Methylation interference experiments to define the binding sites of nuclear factors that interact with the direct repeat of BL/VL3 RadLV. The experiment was performed with DRA (lanes 1 to 6), DRB (lanes 7 to 12), DRC (lanes 13 to 19), and Ti-6 lymphocyte nuclear extracts (shown by asterisks and an open circle in Fig. 3). Oligonucleotides were labeled on the plus (lanes 1 to 3, 7 to 9, and 13 to 16) or minus (lanes 4 to 6, 10 to 12, and 17 to 19) strand. Bands present in the DNA isolated from the free fraction (F) and absent or strongly diminished in the DNA isolated from the bound fraction (B) represent points of strong contact between the DNA and a protein(s) (arrowheads). Lane 15 represents the DNA isolated from the additional, faster migrating DRC complex, indicated by an open circle in the Fig. 3. The G or GA sequencing ladders of the oligonucleotides were included as markers (lanes 3, 6, 9, 12, 16, and 19).

the minus strand (Fig. 5, lanes 13 to 19). This site of protein-DNA contact is a core consensus sequence, $\begin{smallmatrix} \text{TGTGG} \\ \text{ACACC} \end{smallmatrix}$, shown also to bind a nuclear factor, core, on the Moloney MuLV direct repeat (48). Two Gs on the plus strand and three on the minus strand of DRB DNA were involved in the protein-DNA interaction, defining the binding sequence as $\begin{smallmatrix} \text{GCACCTG} \\ \text{CGTGGAC} \end{smallmatrix}$. This sequence is modified at three positions compared with the LVC factor-binding site of Moloney MuLV and is also different in the N-CI-35 direct repeat. Interestingly, this motif, designated Rad-1, on the second 43-bp repeat, also contains a corelike element located, however, on the minus strand. The site of close contact on the minus strand involves, as with the first core sequence, the same motif, 5'-GGTG-3', although in reverse orientation. This analogy of the binding site and the ability to compete for binding in the gel retardation assay indicates that an identical or very similar factor may bind to both sequences separated by only one nucleotide but located on opposite strands. On the first and third 43-bp repeats, this motif shows one point mutation. The binding capacity of these two slightly modified Rad-1 sequences remains to be tested.

The lower, fainter complexes found with DRC (Fig. 3, open circles) most likely represent products of proteolytic degradation of the factor responsible for major complex formation, since the pattern of G contacts is the same as for the major retarded complexes (Fig. 5, lane 15). A similar type of core factor proteolytic degradation has been described by Thornell et al. (51). Other, faster migrating complexes apparent sometimes in the band shift experiments

FIG. 6. Summary of DNA-protein interactions of the BL/VL3 direct repeat. The DNA-protein close contacts, as identified by G methylation interference assay, are indicated by dots. The boxed binding sites are comparable to the protein-binding sites of the Moloney MuLV enhancer direct repeat (48). Two sites (LVB and core) are conserved in both viruses. The binding site unique to BL/VL3 Rad-1 has a broken box and is defined by its external interacting G. The nucleotides identical in both sequences are capitalized. No binding site was found in the 11-bp spacer sequence separating the BL/VL3 direct repeats.

represent binding to single-stranded DNA, since they could be competed with single-stranded DNA (data not shown).

The 11-bp sequences separating the BL/VL3 direct repeats (5' part of oligonucleotide DRA) did not contain binding sites, as determined by the methylation interference assay (Fig. 5), suggesting that this stretch of DNA does not bind transcription factors but rather serves as a spacer necessary for the secondary structure of the enhancer.

The DNA-protein interactions found with the BL/VL3 direct repeat are summarized in Fig. 6 and are compared with those of a direct repeat of Moloney MuLV.

DISCUSSION

Our protein-binding studies revealed several binding sites on the BL/VL3 LTR. In the upstream region of its promoter, we documented a strong interaction with a CBP. The protein was present in two types of murine cells tested (fibroblasts and T lymphocytes). This CBP was found to bind to the BL/VL3 promoter as well as to the analogous sequence of the nonleukemogenic, fibrotropic N-CI-35 MuLV. This protein has the same sequence specificity as the CBP that binds to the MSV promoter (18, 19). We also showed that an MSV DNA fragment harboring the CCAAT sequence competed effectively for binding to the analogous BL/VL3 promoter subfragment. Since the recognition sequence of CBP is very well conserved among murine retroviruses, e.g., AKV, CasBR-E, Friend, SL3-3, and Moloney MuLVs and Moloney and Harvey MSVs (53), this CBP presumably has the ability to bind to the promoter of each of these MuLV strains. The characteristic protein-DNA contacts for CBP were found to be the same as those found for the CP1 factor of HeLa cells (7), confirming that the CBP described here belongs to the family of CBPs well conserved among species and present in different tissues (7, 8, 16, 19, 24).

We also tested the upstream enhancer region of BL/VL3 RadLV for the presence of protein-binding sites. In the gel retardation assay performed with the probes from this region, we demonstrated several sequence-specific complexes, but we were unable to determine the precise sites of DNA-protein contacts. The analogous A+T-rich sequence of the leukemogenic SL3-3 MuLV was shown to stimulate LTR-driven transcription (20). A+T-rich sequences are often present in front of eucaryotic genes. They are known to mediate a secondary structure of DNA distinct from the

typical beta helix and may act to modulate transcription by preventing nucleosome formation (50). It is tempting to speculate that the proteins that bind to the BL/VL3 upstream enhancer sequences stabilize an open structure of DNA, creating an entry site for other enhancer factors. This model is supported by the finding that two retarded complexes (Fig. 3, arrows) could be eliminated by the homologous single-stranded DNA, showing that these proteins have an affinity for DNA that remains in an open structure.

Three sequence-specific interactions in the BL/VL3 enhancer (direct repeats) were characterized. Two of them showed sequence specificity similar to that of the LVb and core factors described earlier for the Moloney, AKV, Friend, and SL-3 MuLV LTRs (2, 39, 48, 51). The BL/VL3 LVb binding site is identical to that of the AKV and Moloney MuLVs and presumably binds the same factor as Moloney MuLV. The BL/VL3 core sequence is identical to that of AKV MuLV but differs from that of SL-3 MuLV by one nucleotide and from that of Moloney MuLV by two nucleotides (Fig. 6). Several different factors that bind to the retroviral core sequence have been described (2, 23, 48, 51). By the criteria of cell distribution and homology of sequences surrounding the TGTGG core consensus, the BL/VL3 core-binding protein is similar to the ubiquitous S/A-CBF factor which binds to the core sequence of the leukemogenic SL3-3 MuLV and of the nonleukemogenic AKV MuLV (2), although the protein-DNA contact points for the S/A-CBF factor were not reported. By the same criteria, the BL/VL3 core-binding protein appears to be distinct from the SL-3 core-binding protein, SEF1 (51). Whether the BL/VL3 core-binding site binds the same factor as the Moloney core-binding site remains to be determined. Because the core motif of thymotropic BL/VL3 RadLV is identical to those of nonthymotropic AKV or N-CI-35 MuLV (12, 53), it appears that the core sequence itself is not sufficient, and its sequence does not have to be mutated, to generate a virus with very distinct tissue tropism, at least with these virus strains.

We have identified a third binding site unique to BL/VL3 direct repeats and located immediately downstream of the core consensus sequence. This sequence binds a factor that we have designated Rad-1. Interestingly, this sequence has a corelike motif on the minus strand, and the factor that binds to it can be competed by the BL/VL3 core-containing fragment. The protein-DNA contacts involve the typical three core Gs separated by one T, although in the opposite orientation. The factor that binds to this sequence might be the same as or similar to the core-binding factor. Since the core- and Rad-1-binding sites are separated by only one nucleotide, it remains to be determined whether both sites can be occupied simultaneously by two identical or similar factors. The proximity of these two binding sites could facilitate protein-protein interactions necessary for enhancer activity.

The direct repeats of the BL/VL3 LTR are short (43 bp) and remain the shortest repeats ever found in an LTR of an MuLV. Our binding studies with one such complete repeat revealed only three binding sites. Among the three binding sites identified, two have already been described and one appears to be unique to BL/VL3 RadLV. Therefore, the presence of these three binding sites (LVb, core, and Rad-1) within the enhancer seems to be sufficient to promote a high rate of transcription of BL/VL3 RadLV in its specific target cells (T lymphocytes). Consequently, the other binding sites, described for the Moloney MuLV enhancer (LVa, NF-1, LVc, or GRE) (48) and absent in the BL/VL3 enhancer, may

not play a critical role in T-cell-specific transcription. This conclusion is also supported by recent findings on Moloney and Friend MuLVs (39).

Only one of the LTR oligonucleotides that we studied (UEB) showed binding to a specific factor present in T lymphocytes and not in fibroblasts. Although its detection was not reproducible and we could not localize its binding sites precisely, this factor(s) might represent a tissue-specific transcription factor. All of the other binding factors that we studied were present in T-lymphocyte and fibroblast extracts. Nevertheless, some of them might still function as tissue-specific factors. Indeed, although many DNA-binding factors are ubiquitous and were initially considered general transcription factors, it appears that some of them are active only in specific cell types (1, 29, 30, 47, 51). Further experiments will be needed to determine the biological role, if any, of the factors that bind to the BL/VL3 core and Rad-1 motifs in promoting transcription of BL/VL3 RadLV in its natural target cells, thymocytes, and in inducing T-cell leukemia in inoculated mice. These factors may also represent factors that control transcription of T-cell-specific genes.

ACKNOWLEDGMENTS

This work was supported by grants to P. J. from the Medical Research Council of Canada and from the National Cancer Institute of Canada. I.G.-F. was a recipient of a Fellowship of the Cancer Research Society Inc. (Montreal, Quebec, Canada).

LITERATURE CITED

1. Becker, P., S. Ruppert, and G. Schütz. 1987. Genomic footprinting reveals cell type specific DNA binding of ubiquitous factors. *Cell* 51:435-443.
2. Boral, A. L., S. A. Okenquist, and J. Lenz. 1989. Identification of SL3-3 virus enhancer core as a T-lymphoma cell-specific element. *J. Virol.* 63:76-84.
3. Böse, Z., H.-J. Thiesen, and F. Charnay. 1986. A transcriptional enhancer with specificity for erythroid cells is located in the long terminal repeat of Friend murine leukemia virus. *EMBO J.* 5:1615-1623.
4. 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.
5. Chatis, P. A., C. A. Holland, J. W. Hartley, W. P. Rowe, and N. Hopkins. 1983. Role of 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. Frederickson, N. Hopkins, and J. W. Hartley. 1984. A 3' end fragment encompassing the transcriptional enhancers of nondefective Friend virus confers erythroleukemogenicity on Moloney leukemia virus. *J. Virol.* 52:248-254.
7. Chodosh, L. A., A. S. Baldwin, R. W. Carther, and P. A. Sharp. 1988. Human CCAAT-binding proteins have heterologous subunits. *Cell* 53:11-24.
8. Cohen, R. B., M. Sheffery, and C. G. Kim. 1986. Partial purification of a nuclear protein that binds to the CCAAT box of the mouse α_1 -globin gene. *Mol. Cell. Biol.* 6:821-832.
9. Declève, A., C. Sato, M. Lieberman, and H. S. Kaplan. 1974. Selective thymic localization of murine leukemia virus-related antigens in C57BL/Ka mice after inoculation with radiation leukemia virus. *Proc. Natl. Acad. Sci. USA* 71:3124-3128.
10. DesGroseillers, L., and P. Jolicoeur. 1984. Mapping the viral sequences conferring the leukemogenicity and disease specificity in Moloney and amphotropic murine leukemia viruses. *J. Virol.* 52:448-456.
11. 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.
12. 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.
 13. DesGroseillers, L., E. Rassart, Y. Robitaille, and P. Jolicoeur. 1985. Retrovirus-induced spongiform encephalopathy: the 3' end long terminal repeat-containing viral sequences influence the incidence of the disease and the specificity of the neurological syndrome. *Proc. Natl. Acad. Sci. USA* **82**:8818-8822.
 14. DesGroseillers, L., R. Villemur, and P. Jolicoeur. 1983. 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.
 15. 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.
 16. Dorn, A., J. Bollekens, A. Staub, C. Benoist, and D. Mathis. 1987. A multiplicity of CCAAT box-binding proteins. *Cell* **50**:863-872.
 17. Flanagan, J. R., A. M. Krieg, E. E. Max, and A. S. Khan. 1989. Negative control region at the 5' end of murine leukemia virus long terminal repeats. *Mol. Cell. Biol.* **9**:739-746.
 18. Graves, B. J., R. N. Eisenman, and S. L. McKnight. 1985. Delineation of transcriptional control signals within the Moloney murine sarcoma virus long terminal repeat. *Mol. Cell. Biol.* **5**:1948-1958.
 19. Graves, B. J., P. F. Johnson, and S. L. McKnight. 1986. Homologous recognition of a promoter domain common to the MSV LTR and HSV tk gene. *Cell* **44**:565-576.
 20. Hallberg, B., and T. Grundström. 1988. Tissue specific motifs in the enhancer of leukemogenic mouse retrovirus SL3-3. *Nucleic Acids Res.* **16**:5927-5944.
 21. Holland, C. A., C. Y. Thomas, S. K. Chattopadhyay, C. Koehne, and P. V. O'Donnell. 1989. Influence of enhancer sequence on thymotropism and leukemogenicity of mink cell focus-forming viruses. *J. Virol.* **63**:1284-1292.
 22. Janowski, M., J. Merregaert, J. Boniver, and J. R. Maisin. 1985. Proviral genome of radiation leukemia virus: molecular cloning of biologically active proviral DNA and nucleotide sequence of its long terminal repeat. *J. Virol.* **55**:251-255.
 23. Johnson, P. F., W. M. Landschultz, 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.
 24. Jones, K. A., J. T. Kadonaga, P. J. Rosenfelt, T. J. Kelly, and R. Tjian. 1987. A cellular DNA-binding protein that activates eukaryotic transcription and DNA replication. *Cell* **48**:79-89.
 25. Kaplan, H. S. 1967. On the natural history of the murine leukemias: presidential address. *Cancer Res.* **27**:1325-1340.
 26. Khan, A. S., and M. A. Martin. 1983. Endogenous murine leukemia proviral long terminal repeats contain a unique 190-base-pair insert. *Proc. Natl. Acad. Sci. USA* **80**:2699-2703.
 27. Kovesdi, J., R. Reichel, and Y. R. Neving. 1986. Identification of cellular transcription factor involved in E1A trans-activation. *Cell* **45**:218-228.
 28. Landschultz, W. H., P. F. Johnson, E. Y. Adashi, B. Y. Graves, and S. L. McKnight. 1988. Isolation of a recombinant copy of the gene coding C/EBP. *Genes Dev.* **2**:786-800.
 29. Lee, W., P. Mitchell, and R. Tjian. 1987. Purified transcription factor AP-1 interacts with TPA-inducible enhancer element. *Cell* **49**:741-752.
 30. Lenardo, M. J., and D. Baltimore. 1989. NF- κ B: a pleiotropic mediator of inducible and tissue-specific gene control. *Cell* **58**:227-229.
 31. Lenz, J., D. Celander, R. L. Crowther, R. Patar, D. W. Perkins, and W. A. Haseltine. 1984. Determination of the leukemogenicity of a murine retrovirus by sequences within the long terminal repeat. *Nature (London)* **308**:467-470.
 32. Levinson, B., G. Khoury, G. Vande Woude, and P. Gruss. 1982. Activation of SV40 genome by 72-base-pair tandem repeats of Moloney sarcoma virus. *Nature (London)* **295**:568-572.
 33. 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**:693-700.
 34. Lieberman, M., A. Declève, and H. S. Kaplan. 1978. Rapid in vitro assay for thymotropic, leukemogenic murine C-type RNA viruses. *Virology* **90**:274-278.
 35. Lieberman, M., O. Niwa, A. Declève, and H. S. Kaplan. 1973. Continuous propagation of radiation leukemia virus on a C57BL mouse embryo fibroblast line with attenuation of leukemogenic activity. *Proc. Natl. Acad. Sci. USA* **70**:1250-1253.
 36. LoSardo, J. E., L. A. Cupelli, M. K. Short, J. W. Berman, and J. Lenz. 1989. Differences in activities of murine retroviral long terminal repeats in cytotoxic T lymphocytes and T-lymphoma cells. *J. Virol.* **63**:1087-1094.
 37. Maniatis, T., E. F. Fritsch, and J. Sambrook. 1982. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
 38. Maniatis, T., S. Goodbourn, and J. A. Fischer. 1987. Regulation of inducible and tissue-specific gene expression. *Science* **236**:1237-1245.
 39. 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.
 40. Maxam, A. M., and W. Gilbert. 1980. Sequencing end-labeled DNA with base-specific chemical cleavages. *Methods Enzymol.* **65**:499-560.
 41. McKnight, S. L., and R. Tjian. 1986. Transcription selectivity of viral genes in mammalian cells. *Cell* **46**:795-805.
 42. Mitchell, P., and R. Tjian. 1989. Transcriptional regulation in mammalian cells by sequence-specific DNA binding proteins. *Science* **245**:371-378.
 43. Rassart, E., Y. Paquette, and P. Jolicoeur. 1988. Inability of Kaplan leukemia virus to replicate on mouse fibroblasts is conferred by its long terminal repeat. *J. Virol.* **62**:3840-3848.
 44. Rassart, E., P. Sankar-Mistry, G. Lemay, L. DesGroseillers, and P. Jolicoeur. 1983. New class of leukemogenic ecotropic recombinant murine leukemia virus isolated from radiation-induced thymomas of C57BL/6 mice. *J. Virol.* **45**:565-575.
 45. Rassart, E., M. Shang, Y. Boie, and P. Jolicoeur. 1986. Studies on emerging radiation leukemia virus variants in C57BL/Ka mice. *J. Virol.* **58**:96-106.
 46. 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.
 47. Schreider, C., A. Heguy, and R. G. Roeder. 1987. Identification and purification of a human lymphoid-specific octamer binding protein (OTF-2) that activates transcription of an immunoglobulin promoter *in vitro*. *Cell* **51**:783-793.
 48. 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.
 49. Strauss, F., and A. Varshavski. 1984. A protein binds to a satellite DNA at three specific sites that would be brought into mutual proximity by DNA folding in the nucleosome. *Cell* **37**:889-901.
 50. Struhl, K. 1985. Naturally occurring poly(dA-dT) sequences are upstream promoter elements for constitutive transcription in yeast. *Proc. Natl. Acad. Sci. USA* **82**:8419-8423.
 51. 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.
 52. Vogt, M., C. Hagblom, S. Swift, and M. Haas. 1985. Envelope gene and long terminal repeat determine the different biological properties of Rauscher, Friend, and Moloney mink cell focus-inducing viruses. *J. Virol.* **55**:184-192.
 53. Weiss, R., N. Teich, H. Varmus, and J. Coffin. 1985. Molecular biology of tumor viruses: RNA tumor viruses. Part II. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
 54. 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 a specific trans-acting factor. *EMBO J.* **3**:3129-3133.