Structure and Function of Endogenous Feline Leukemia Virus Long Terminal Repeats and Adjoining Regions

BRIAN T. BERRY,¹ ANANTA K. GHOSH,¹ D. VINOD KUMAR,¹ DAVID A. SPODICK,²
AND PRADIP ROY-BURMAN^{1,2*}

*Departments of Pathology¹ and Biochemistry,² University of Southern California School of Medicine,
Los Angeles, California 90033*

Received 9 May 1988/Accepted 28 June 1988

The nucleotide sequence of the 5' long terminal repeat (LTR) of three independent loci (CFE-6, CFE-16, and CF-14) of endogenous feline leukemia virus (FeLV) DNAs of the domestic cat genome was determined. The 3' LTR of the CFE-6 clone was also sequenced. The endogenous FeLV LTRs, which were very similar to each other in sequence and in organization of the functional domains, differed considerably from the exogenous FeLV LTR in the U3 region. The major differences in U3 included variations in sets of small (14 to 19 base pair) direct repeats, altered location of the simian virus 40 core enhancer-like sequence, and occurrence of three segments of largely nonhomologous sequences. There was extensive homology between endogenous and exogenous FeLV LTRs in sequences beginning from the TATA box through the R region down to the 3' end of the U5 region. The DNA sequence downstream of the 5' LTR encompassing the primer-binding site, leader, and almost to the end of the p15^{gag} coding region, a point up to which the sequencing was carried out, also revealed a high degree of conservation. However, the detection of frameshift and nonsense mutations in this region of a nearly full-length endogenous provirus sequence (CFE-6) predicted its defectiveness and correlated with the lack of infectivity of this DNA. The functional studies of the endogenous LTRs, based on linkage to the bacterial *cat* gene and transient expression in feline cell lines, indicated that although the basic characteristics for promotion and enhancement of transcription were retained in each LTR, there was a significant variation in the activity of the *cat* constructs. Reconstruction and deletion analyses with the CFE-6 5' LTR revealed the presence of strong transcription regulatory sequences in the 702-base-pair region immediately upstream of the 5' boundary of the endogenous LTR. These and related data suggest that in addition to the transcription-modulating elements occurring within the LTR, the *cis*-acting nucleotide sequences in the upstream cellular DNA may determine the overall efficiency of transcription of the defective endogenous FeLV provirus loci of the felid genome.

The cellular DNA of the domestic cat (*Felis catus*) contains multiple copies of endogenous proviruses related to the infectious feline leukemia virus (FeLV) (1, 3, 22). These inherited endogenous sequences are dispersed throughout the cat genome and are expressed as subgenomic transcripts in a tissue-specific manner (4, 17-19) but do not give rise to detectable infectious virus (3, 22). A number of the endogenous FeLV sequences were cloned and their structures were analyzed by restriction mapping (30, 31). These long terminal repeat (LTR)-flanked sequences demonstrated the existence of different size classes of endogenous loci in the cat genome, including those of nearly full length to those with major deletions in *gag* or *pol* or both *gag* and *pol* genes. While none of the endogenous clones were infectious in the DNA transfection studies, a differential expression was observed with the genes introduced into heterologous cells (31).

Since the 5' LTR of the integrated provirus regulates the expression of the viral genes through the interaction of various *cis*-acting elements and *trans*-acting factors, we asked if the variation in the transcription efficiency of the cloned endogenous FeLV DNAs might be the result of alterations in sequences in or around the 5' LTRs. In this paper, we examine the ability of several different 5' LTRs of the endogenous clones to promote or enhance transient expression of a linked gene in cells of the homologous species. We also present and compare the DNA sequences

of three LTRs, each representing a different class, for an analysis of the regions which are likely to influence their ability to drive transcription and to examine the evolution of these proviral sequences.

MATERIALS AND METHODS

Cell lines. H927 feline embryo fibroblasts (23), the feline T-lymphoid tumor cell line 3201B (29), and the NIH 3T3 mouse fibroblast cell line were maintained at 5% CO₂ in Dulbecco modified Eagle medium (DMEM) with high glucose, supplemented with 10% fetal bovine serum and antibiotics.

Transfections and CAT assays. H927 feline or NIH 3T3 mouse fibroblasts were transfected by a modification of the calcium phosphate method of Graham and Van der Eb (11). Briefly, 5×10^5 cells were seeded onto 100-mm dishes 1 day prior to transfection in DMEM plus 10% fetal bovine serum. Calcium phosphate DNA precipitates were prepared from 40 µg of plasmid DNA with 60 µg of sheared calf thymus carrier DNA in 1 ml of HEPES (*N*-2-hydroxyethylpiperazine-*N'*-2-ethanesulfonic acid)-buffered saline by constantly bubbling nitrogen gas through the solution while adding calcium chloride. After allowing at least 20 min for the precipitate to form, one-half (20 µg of plasmid DNA) of the precipitate was added to each dish and incubated at 37°C for 4 h. The cells were then glycerol shocked (15% glycerol in HEPES-buffered saline) for 4 min, washed with fresh medium, and incubated for 36 to 48 h at 37°C prior to harvesting (21, 32).

The feline T-lymphoid tumor cell line 3201B was transfect-

* Corresponding author.

ted with the plasmid DNAs by a modified DEAE-dextran protocol (2). Briefly, 2×10^7 cells were washed twice with serum-free DMEM and once with TBS (25 mM Tris hydrochloride [pH 7.4], 137 mM NaCl, 5 mM KCl, 0.6 mM Na_2HPO_4 , 0.7 mM CaCl_2 , 0.5 mM MgCl_2) by gentle centrifugation. The cells were then incubated at room temperature in 2 ml of TBS containing 500 μg of DEAE-dextran (Pharmacia) per ml and 20 μg of plasmid DNA for 90 min. The unadsorbed DNA was removed by washing twice with TBS and once with serum-free DMEM, and the cells were replated in 20 ml of serum-containing DMEM. After 48 h of incubation at 37°C, the cells were harvested for the chloramphenicol acetyltransferase (CAT) assay.

Transfected cells were washed three times with cold phosphate-buffered saline and lysed by freeze-thawing three times or by sonication in 100 μl of 250 mM Tris hydrochloride (pH 7.8). Cell debris was pelleted by centrifugation, and 40 μl of cellular supernatant was assayed for CAT activity as described by Gorman et al. (10). The percent acetylation of [^{14}C]chloramphenicol was determined by separating the acetylated and unacetylated forms by thin-layer chromatography, followed by liquid scintillation counting of spots scraped from the silica gel plates.

Plasmid constructions. Plasmids pSVO-CAT and pSV2-CAT have been described previously (10). The plasmid pSVO-CAT contains the bacterial *cat* gene linked to a simian virus 40 (SV40) polyadenylation signal but without any promoter or enhancer signal sequences. The plasmid pSV2-CAT is identical to pSVO-CAT except that it also contains the SV40 early promoter-enhancer region upstream of the *cat* gene. Restriction maps of six cloned endogenous FeLV DNA sequences were reported previously (31). Five endogenous FeLV proviral elements (CFE-6, CFE-8, CF-14, CFE-16, and CFE-54) were isolated from *EcoRI*-digested phage DNA. The 5' LTR-containing fragments from these five clones and from the exogenous Gardner-Arnstein strain of FeLV subgroup B (9) were isolated and placed upstream of the *cat* gene by blunt-end ligation into the *HindIII* site of the plasmid pSVO-CAT for analysis of promoter activity or the *BamHI* site of pSV2-CAT for analysis of enhancer activity. Recombinant plasmids which contained the inserted 5' LTR-containing fragment were isolated either in the correct 5' to 3' direction relative to the provirus, designated orientation A, or the incorrect (inverted) orientation, B. The mix-and-match plasmids involving CFE-6 and CFE-16 LTRs were created by replacing the 448-base-pair (bp) *EcoRV*-*BglII* fragment of CFE-6 with the 451-bp *EcoRV*-*BglII* fragment of CFE-16. Similarly, the 451-bp fragment of CFE-16 was replaced by the 448-bp fragment from CFE-6. Plasmids were purified by two CsCl centrifugations (16), and the orientations were verified by restriction analysis prior to use in transfections.

For generation of CFE-6 5' LTR deletion mutants, the 1.37-kilobase (kb) *BamHI*-*SstI* fragment encompassing the LTR was first cloned into the same restriction sites of pUC18. The recombinant plasmid was then cut with *PstI* and *PpuMI* to remove 524 nucleotides from the 5' end of the 702-nucleotide-long upstream flanking region of the LTR. Further unidirectional deletions from the *PpuMI* site were made by treating the linear plasmid with exonuclease III for different time periods, followed by removal of the single-stranded region by mung bean nuclease. After recircularization of the plasmids by T4 ligase, the 5' deletion endpoints were determined by the method of double-stranded sequencing with the 17-mer universal primer (Boehringer Mannheim guidelines for quick and simple plasmid sequencing, 1986).

Various deleted LTR fragments were then excised from the plasmids by digestion with *HindIII* and *SstI*, blunt-ended by T4 polymerase, and cloned upstream of the *cat* gene of pSVO-CAT.

DNA sequence analysis. Three endogenous FeLV 5' LTRs (CFE-6, CF-14, and CFE-16), each representing a different proviral size class, and the 3' LTR of the CFE-6 provirus were selected for DNA sequencing. For sequencing of the 5' LTRs, the *BamHI*-*SstI* fragment from CFE-6 and the *SstI* fragments of CF-14 and CFE-16 were blunt-end ligated into M13mp19 at the *SmaI* site. Similarly, the *HindIII*-*EcoRI* fragment encompassing the 3' LTR was ligated into the M13 vector. Both orientations were isolated, and overlapping clones were generated by the rapid deletion method (7) and sequenced by the dideoxy chain termination method (26). Analysis and alignment of the sequences were aided by the Bionet Resource.

RESULTS

Sequence of the endogenous FeLV LTRs and comparison with the corresponding exogenous sequence. To examine the structural relationship among various endogenous FeLV 5' LTRs, the sequences of three endogenous LTRs, each derived from a different size class of provirus loci, were determined. The clone CFE-6 represented the 5' LTR of an endogenous locus close to full length, while clones CFE-16 and CF-14 represented the truncated provirus loci deleted in *pol* or in both *gag* and *pol*, respectively (30, 31). The determination of the boundaries of each LTR and assignment of the U3, R, and U5 regions as well as functional domains for CAAT, TATA, and polyadenylation signal boxes were based on similarities to reported mammalian retroviral LTR consensus sequences (6) and comparison with the published sequences of exogenous FeLV LTRs (11, 20, 27). The R and U5 sequences of the endogenous LTRs were highly related among themselves as well as to different subgroups (27, 28) of exogenous FeLV (Fig. 1). Of 142 nucleotides encompassing the R and U5 regions, exogenous FeLVs showed differences at only 4 nucleotide positions due to a single nucleotide deletion or base substitution. Endogenous sequences also differed in 4 nucleotides, 2 being at different positions compared with the changes observed among exogenous LTRs. The U3 region of the endogenous LTRs, however, revealed significant divergence from the exogenous counterparts. The features contributing to the diversity were (i) the length of U3, the endogenous segment being approximately 30 to 60 bp larger than the exogenous unit; (ii) the presence of stretches of sequences, either homologous or largely nonhomologous, to exogenous U3 scattered on the endogenous element, beginning immediately after the 5' inverted repeat (IR) down to the start of the TATA box; (iii) the presence of three different sets of direct repeats either absent, partially present, or present only as single copies in the exogenous U3; and (iv) the location of the SV40 core enhancer-like sequence of endogenous U3 at a position about 90 bp downstream relative to its position in the exogenous U3. These major differences between endogenous and exogenous FeLV LTR U3s are schematically presented in Fig. 2. The 14-bp DR-1A of the endogenous U3 was not detected in the exogenous LTR. Its second copy (DR-1B), which contained a maximum of 2 base changes, however, was present in both exogenous and endogenous sequences partly overlapping with the 3' IR region (Fig. 1 and 2). The 19-bp DR-2A, located immediately upstream of the SV40 core enhancer-like sequence of the endogenous

FIG. 1. Nucleotide sequence comparison of the 5' LTRs from three endogenous FeLV provirus loci, clones CFE-6, CFE-16 and CF-14, with that of three exogenous FeLV isolates, GA-FeLV-B (13), FeLV-A/Glasgow (34), and FeLV-C/Sarma (25). Important structural features, namely, the SV40 core enhancer-like sequence, direct repeats (DR) of three different sets (DR-1, DR-2, and DR-3), CAAT, TATA, and the poly(A) signal sequences are boxed. Features marked by arrows are IRs delineating the LTRs, and the beginnings of the U3, R, and U5 regions. The first nucleotide of R is numbered +1 for all LTRs, and the numbers on the top indicate the nucleotide positions for the GA-FeLV-B LTR, while those underneath the nucleotide sequence mark the positions for the endogenous CF-14 LTR. Gaps are inserted to align homologous regions of the sequences. The region within the positions marked by asterisks exhibits partial homology to the 72-bp enhancer element of SV40 DNA, and the underlined sequences indicate the positions that are conserved in mammalian type C retroviruses (6).

U3, was duplicated (with a 1-bp change) in tandem only in the CF-14 clone. A very close version of this sequence occurred as a single copy in both CFE-6 and CFE-16, but only as a partially homologous single copy (8 of 19 nucleotides) in each of the three exogenous LTRs compiled for this

comparison. The DR-3A consisted of 14 bp, well conserved among the endogenous sequences, and repeated after a 7-bp spacer as DR-3B with 2 base changes. This second repeat, including the spacer sequences, was also retained in the exogenous LTR, while the DR-3A unit, like DR-2A, was

FIG. 2. Schematic representation of the divergence between exogenous and endogenous FeLV U3 nucleotide sequences. Regions of the endogenous LTR (EN-FeLV U3) that are homologous to the exogenous LTR (EX-FeLV U3) are indicated by white boxed areas, and the regions that are largely nonhomologous are shown as shaded areas. Solid regions in EN-FeLV U3 are portions of sequences not present in the EX-FeLV U3 within the homologous regions. Other features illustrated include thin-lined boxes to indicate the position of SV40 core enhancer-like sequences and bold-lined boxes to mark the locations of different direct repeats (DR), CAAT, and TATA boxes. The scale on the top indicates the approximate nucleotide positions of the sequence, and the asterisk marks the DR unit, which is present in the sequence of the endogenous LTR of the CF-14 clone but not in the other clones sequenced.

partially deleted in the exogenous unit (Fig. 2). In regard to the functional domains, which may be important for transcriptional promotion, the CAAT and TATA boxes were identical in sequence. They were, however, placed further apart relative to their positions on the 3' end of exogenous U3 because of insertions of three stretches of nucleotides (totaling 21 to 22 bp) into the endogenous U3 or, alternatively, by deletion of these regions from the exogenous U3 (Fig. 2).

The sequence of the 3' LTR of CFE-6 (data not shown) was nearly identical to its 5' LTR with only three point mutations. These mutations included an insertion of a C and a transversion mutation (C to A) in the DR-2A sequence, AGCCACCTGGCCTCAGAT, of the 5' LTR, altering it to AGCCACCTGGCCCTAAGAT in the 3' LTR. The other base change was a transition (T to C) in the core enhancer sequence. It was also noted that a putative primer sequence (AAGAAAAGGGGGAAA) was retained immediately upstream of the 3' LTR. This polypurine tract differed in only three positions from the corresponding exogenous FeLV sequence, AGAAAAGGGGGGAA, conserved in each of the A, B, and C virus subgroups.

It should be noted that the major characteristic sequence divergence between endogenous and exogenous FeLV LTRs, as seen from the comparison of endogenous sequences with those FeLV A, B, and C subgroups, remained unchanged when the comparison was extended to other isolates of pathogenic feline retroviruses, such as feline sarcoma virus (FeSV) or FeLV inducing fatal immunodeficiency disease in cats. The reason is that the LTRs of FeSV isolates, e.g., Gardner-Arnstein-FeSV, Synder-Theilen-FeSV, and Sarma-McDonough-FeSV, revealed extensive homology to each other as well as to FeLV (12, 13). These are, however, point mutations and small insertions which distinguish one FeSV isolate from the other, but they are considered minor in comparison to the differences observed with the endogenous FeLV LTRs. The largest insertion in ST-FeSV LTR relative to GA-FeSV is a stretch of four cytosines positioned immediately 3' to the first IR (13). Interestingly, these four C's were also found in the same position in each of the three endogenous LTRs sequenced (Fig. 1); similarly, a stretch of five C's was present in an isolate of FeLV-C (25). The feline AIDS (14)-associated viruses whose LTR sequences were determined also reveal a high degree of homology (94 to 97%) to prototype B and C subgroups of FeLV, with only a slightly higher (98%) homology to the FeLV-A subgroup (8, 20).

Comparison of sequences downstream of the 5' LTR. All three endogenous FeLV DNAs were sequenced to the +294 nucleotide position, and the sequencing continued to about +900 for the full-length clone CFE-6 (Fig. 3). There was extensive homology at the nucleotide level between the endogenous DNAs and between the endogenous and exogenous FeLV (ST-FeLV-B) in the internal region sequenced, except for scattered single base alterations, including substitutions, deletions, and insertions. In two locations, one at approximately +400 and the other at about +500, CFE-6 differed from the exogenous virus due to a 9-bp insertion and 2-bp deletion, respectively. Two of the endogenous clones (CFE-16 and CFE-6) were shown to match 18 of 18 nucleotides in the region of the putative primer-binding site that corresponded to proline tRNA. The third one varied in only one nucleotide due to a transition mutation. Splice donor and acceptor signals of the leader sequence were conserved, and the first six amino acids of the gPr80^{gag} were identical to those of the exogenous virus. A base deletion in CFE-6 after these six amino acids, however, caused a shift in the reading frame, which returned to homology with the exogenous FeLV after 42 amino acids. The homology continued and had encompassed more than half of the p15^{gag} polypeptide when in-frame base substitutions (beginning at the +767 nucleotide position), followed shortly by a base deletion, altered the amino acid sequence again. In addition to this large number of missense mutations, it appeared that synthesis of gPr80^{gag} would be blocked due to the generation of a termination codon at +422 position from frameshift mutations in the CFE-6 DNA.

Promoter and enhancer functions of endogenous FeLV LTRs. In our previous transient expression studies in NIH 3T3 cells with the provirus DNA containing both 5' and 3' LTRs, we observed that some of the truncated elements were actively transcribed, while the nearly full-length clones were inactive (31). In order to define further the transcription promotion capacity of the various cloned endogenous 5' LTRs, we tested the ability of individual LTRs to direct transient expression of the *cat* gene when transfected into cells of the homologous species. All 5' LTR-containing fragments used for making CAT expression plasmids had similar structural features. As shown in Fig. 4, all of the five DNA fragments had a conserved *Sst*I site located about 152 bp downstream of the LTR, and three DNAs had a convenient *Sst*I or *Bam*HI site at approximately 0.3 kb upstream of the LTR. The CFE-16 clone contained the upstream *Sst*I site about 0.8 kb away from the LTR, while the CFE-6 LTR had

FIG. 3. Sequences 3' of the 5' LTR of endogenous FeLV DNAs. The numbering indicates the position of the nucleotides on the endogenous proviral DNA with reference to the viral RNA cap site marked +1 (as shown in Fig. 1). The amino acid sequence identities between endogenous and ST-FeLV-B exogenous (15) FeLVs have been underlined. The correspondences established are based on blocks of homologies. Features highlighted include the primer-binding site complementary to proline tRNA (PBS^{Pro}), the conserved splice donor and acceptor sites, the beginning of gPr80^{gag} polypeptide, a nonsense mutation, and the beginning of p15^{gag}. The sequence ends 17 amino acids prior to the end of the p15^{gag} protein of ST-FeLV-B.

FIG. 4. Construction of LTR-containing CAT plasmids. Restriction fragments containing 5' LTRs from exogenous GA-FeLV-B (30) and different endogenous FeLV proviral loci were isolated, end-repaired, and ligated into either the *Hind*III site upstream of the *cat* gene in pSV0-CAT or the *Bam*HI site downstream of the *cat* gene in pSV2-CAT. Both orientations of the insert were isolated, but only the 5' to 3' orientation of each hybrid DNA is presented here. The position of each LTR is indicated by a shaded box. Restriction sites: B, *Bam*HI; E, *Eco*RI; H, *Hind*III; Sm, *Sma*I; St, *Sst*I.

702-bp upstream flanking sequences. LTR fragments were inserted upstream of the *cat* gene of the plasmid pSV0-CAT and the correct orientation of the recombinant DNAs isolated and transfected into H927 feline fibroblast cells. The results of the transient *cat* expression assay are shown in

TABLE 1. Relative promoter activity of endogenous FeLV 5' LTRs as assayed by transient expression of CAT activity in H927 feline embryonic fibroblasts

Plasmid	Mean relative CAT activity ^a ± SD (no. of expt)
pSV0-CAT	0.08 ± 0.06 (7)
pSV2-CAT	1.00 (7)
pCAT-EX	3.55 ± 0.78 (4)
pCAT-6	0.07 ± 0.01 (3)
pCAT-8	0.08 ± 0.01 (3)
pCAT-14	0.07 ± 0.02 (3)
pCAT-16	4.87 ± 0.75 (3)
pCAT-54	0.07 ± 0.01 (3)

^a CAT activity is expressed as the level of conversion of chloramphenicol to acetylated forms relative to the conversion of pSV2-CAT. The values from all experiments are reported here as the mean plus or minus the standard deviation. The number of CAT assays is shown in parentheses.

FIG. 5. Comparison of the promoter and enhancer activities of feline LTR-*cat* gene constructs in feline embryonic fibroblasts (H927 cells). The cells were transfected with 20 µg of each recombinant *cat* plasmid and assayed for CAT activity. The autoradiogram from a typical assay is shown here. The positions of migration for chloramphenicol (Cm) and its acetylated forms (1-Cm, 3-Cm) are indicated. The plasmids used were: 1, pSV0-CAT; 2, pSV2-CAT; 3, pCAT-L3-A (a feline endogenous RD-114 5' LTR clone [32]); 4, pCAT-16-A; 5, pCAT-EX-A (exogenous GA-FeLV-B 5' LTR clone); 6, pCAT-EX-B; 7, pSV2-CAT-6-A; 8, pSV2-CAT-6-B; 9, pSV2-CAT-8-A; and 10, pSV2-CAT-8-B. Plasmids containing the inserted LTR fragment in the 5' to 3' direction are designated by A and the inverted orientation (3' to 5') by B.

Table 1. Only two of the various LTRs examined showed detectable transcriptional promoter activity. The 5' LTR of exogenous GA-FeLV-B (pCAT-EX) expressed approximately fourfold-higher activity than the control pSV2-CAT plasmid (Fig. 5, lanes 2 and 5, and Table 1). The endogenous LTR of CFE-16 locus (pCAT-16) showed activity even higher than that of pCAT-Ex (Fig. 5, lane 4, and Table 1), while the other endogenous clones tested failed to exhibit detectable activity (Table 1).

To assess the potential enhancer activity separately from the promoter function, the endogenous LTRs were linked downstream of the *cat* gene in the SV2-CAT plasmid (Fig. 4) for analysis of expression in the same feline fibroblast cell line. Some representative results of these experiments are illustrated in Fig. 5, and all data are summarized in Table 2. In contrast to the results of the promoter ability test, most of the endogenous LTRs examined indicated activity higher

TABLE 2. Relative enhancer activity of endogenous FeLV 5' LTRs as assayed by transient expression of CAT activity in H927 feline embryonic fibroblasts

Plasmid	Orientation ^a	Mean relative CAT activity ^b ± SD (no. of expt)
pSV0-CAT		0.08 ± 0.06 (7)
pSV2-CAT		1.0 (7)
pSV2-CAT-EX	A	5.4 ± 0.1 (3)
pSV2-CAT-EX	B	3.1 ± 1.4 (5)
pSV2-CAT-6	A	1.4 ± 0.8 (5)
pSV2-CAT-6	B	4.8 ± 2.8 (5)
pSV2-CAT-8	A	4.8 ± 0.5 (5)
pSV2-CAT-8	B	2.5 ± 1.2 (5)
pSV2-CAT-14	A	2.9 ± 1.0 (5)
pSV2-CAT-14	B	3.1 ± 1.2 (5)
pSV2-CAT-16	A	3.8 ± 1.3 (5)
pSV2-CAT-16	B	2.9 ± 1.6 (5)
pSV2-CAT-54	A	1.5 ± 0.6 (5)
pSV2-CAT-54	B	2.3 ± 1.2 (5)

^a Orientation indicates that the LTR is either in the correct 5'-3' position relative to *cat* (A) or in the opposite orientation (B).

^b CAT activity is expressed as the level of conversion of chloramphenicol to its acetylated forms relative to the conversion obtained with pSV2-CAT. The number of CAT assays is shown in parentheses.

FIG. 6. Effect of flanking regions on *cat* expression. (A) Mix-and-match constructs were generated. (B) Levels of CAT activity monitored in H927 cells after transfection with (a) pSVO-CAT, (b) pSV2-CAT, (c) pCAT-6, (d) pCAT-16, (e) pCAT-6/16, and (f) pCAT-16/6. The orientation of each insert was in the 5' to 3' direction.

than the SV2-CAT control activity. The observed effect was generally bidirectional, with a variable degree of enhancement depending on the cloned DNA. Although the degree of enhancement was no greater than two- to five-fold over the pSV2-CAT control, the values obtained might be considered significant since the pSV2-CAT backbone of the constructs already contained a highly potent SV40 enhancer and the inserted FeLV sequences had to stimulate activity above this high background.

Sequences outside the endogenous LTRs may influence their transcription regulatory activity. Nucleotide sequence data revealed that except for the occurrence of DR-2B in CF-14, all three endogenous LTRs were very similar in their sequence and organization of the potential functional domains. This structural similarity, however, did not correlate with the observed promoter abilities of these LTRs. For example, pCAT-16 displayed marked promoter activity in the transient expression assays, while neither pCAT-6 nor pCAT-14 showed any detectable activity. CFE-16 and CFE-6 LTRs were, however, highly homologous, retaining over 97% sequence homology, with mostly single base alterations or deletions at 15 positions of the 518-bp length in the entire LTR. These alterations were scattered throughout the LTR and did not involve the conserved sequences of functional domains like the CAAT and TATA boxes. To examine the potential consequences of these nucleotide changes in the LTRs of the endogenous provirus loci, we made constructs in which the majority of the LTR sequence (U5, R, 252 to 255 bp of U3) from the *EcoRV* site (GATATC at -255 bp for CFE-16 or at -252 for CFE-6) to the *BglII* site at +196 bp of pCAT-16 and pCAT-6 were exchanged to create the hybrid plasmids pCAT-6/16 and pCAT-16/6 (Fig. 6A). The results of CAT expression assays in H927 cells transfected with these plasmid DNAs are illustrated in Fig. 6B. Replacement of the region encompassing -252 to +196 bp of the CFE-6 clone with the corresponding region of the CFE-16 element did not

restore its activity, while the inverse exchange (CFE-6 into CFE-16) did not eliminate the activity. Thus, this portion of the LTR leader sequence was interchangeable and apparently did not contain the sequences responsible for the drastic difference in the levels of activity between the two cloned DNAs. The differential transcription efficiency of the reconstructed DNAs was not restricted to H927 feline fibroblasts, since a similar effect was also observed with feline T-lymphoid tumor cells and NIH 3T3 mouse fibroblasts (Table 3).

It should be noted that only three single nucleotide changes between CFE-6 and CFE-16 in the 129-bp region of U3 upstream of the *EcoRI* site were not included in the mix-and-match constructs. It appeared unlikely that minor sequence variations in this shorter area could be sufficient to account for the dramatic differences in function. It was considered more likely that the sequences in the 5'-flanking region of CFE-6 might be dominant in the strong negative effect observed in the CAT assays. To examine this point further, the DNA of the CFE-6 clone was progressively deleted from the *BamHI* site at the 5' end of the insert. Deletion mutants were transfected into H927 cells and assayed for CAT activity. The 5' boundary of each deletion and the results of a typical CAT analysis with these constructs are shown in Fig. 7. The CAT activity of pCAT-6 increased significantly when the first 524 bp (-1074 to -550 bp) of the 5'-flanking DNA were deleted. Subsequent deletion of 124 bp (-550 to -426 bp) further restored the activity to a level comparable to or even higher than that of pSV2-CAT. Additional deletions extending to the interior sequences of the LTR led to progressive reductions of CAT activity. For example, sequential deletions of 134 bp (-426 to -292 bp) and 80 bp (-292 to -182 bp) resulted in relative activities of 59 and 9%, respectively. Since the 80-bp deletion eliminated the DR-1A and DR-2A repeats as well as the SV40 core enhancer-like sequence, the differential activity

TABLE 3. Influence of adjoining regions on the promoter activity of endogenous LTRs assayed by transient expression of CAT activity

Plasmid	Mean relative CAT activity ^a ± SD		
	H927 cells	3201B cells	NIH 3T3 cells
pSV2-CAT	1.00 (27 ± 18%)	1.00 (0.18 ± 0.02%)	1.00 (25 ± 16%)
pSVO-CAT	0.07 ± 0.03	0.08 ± 0.01	0.09 ± 0.05
pCAT-6	0.09 ± 0.04	0.07 ± 0.01	0.12 ± 0.09
pCAT-16	3.47 ± 2.09	3.75 ± 2.05	0.32 ± 0.14
pCAT-6/16	0.09 ± 0.01	0.09 ± 0.01	0.06 ± 0.06
pCAT-16/6	2.56 ± 1.38	2.77 ± 1.18	0.49 ± 0.27

^a Levels of CAT activity are expressed relative to the level obtained with pSV2-CAT in each cell type. Each value represents the mean from three sets of experiments for H927 or 3201B cells and from four sets of experiments for NIH 3T3 cells. The actual percent acetylation for extracts of pSV2-CAT-transfected cells is shown in parentheses.

between the 134-bp and 80-bp deletions also suggested that the region encompassing the DR-1A/DR-2A/core enhancer sequences contained at least a portion of the transcription enhancement ability that we observed with the LTRs linked downstream of pSV2-CAT. Furthermore, the reduction of relative activity from 127 to 59% with the change of the 5' boundary from -426 to -292 bp implied the presence of additional positive regulatory elements in this 134-bp region. This region contained 80 bp of the 5' terminus of the LTR and 54 bp of the immediately adjacent 5'-flanking cellular DNA.

Based on the high degree of nucleotide homology between CFE-6 and CFE-16 within the 5' LTR sequence, it appeared very likely that transcriptional transactivation domains were the same in both LTRs. It is possible, however, that additional positive regulatory elements might be located in the 5'-flanking region of CFE-16.

The deletion mutation studies confirmed that the 5'-flanking sequences were largely responsible for the negative regulation of the pCAT-6 construct in the transient expression assays. The nucleotide sequence of the flanking DNA of

CFE-6 is shown in Fig. 8. This sequence and those of the 5'-flanking DNAs of the CFE-16 and CF-14 LTRs (data not shown) are not homologous to each other, consistent with the origin of the proviral clones from different loci. The 702-bp flanking DNA of CFE-6 could be divided into at least two domains on the basis of the CAT analysis. These domains are 524 bp (-1074 to -550 bp) and 124 bp (-550 to -426 bp) in length, each containing structurally different negative transcription regulatory elements. Further analysis would be required to define the precise boundaries of these *cis*-acting domains.

DISCUSSION

The interest of this work is threefold. First, the structure of the endogenous FeLV LTR has been defined at the nucleotide level. Second, the structural analysis provides a basis on which to explore the evolutionary relationship between the noninfectious proviral sequence of cat DNA and the genome of infectious FeLV. Third, the integrity of the transcription-regulating sequences and results of the

FIG. 7. CAT activity of 5' deletion mutants in transfected H927 cells. In the autoradiogram of a typical assay (top), the plasmids used were: 1, pSVO-CAT; 2, pSV2-CAT; 3, pCAT-16; 4, pCAT-6; 5, pCAT-6 deleted up to position -500; 6, pCAT-6 deleted to -426; 7, pCAT-6 deleted to -292; 8, pCAT-6 deleted to -182; and 9, untransfected cell lysate. A schematic diagram of pCAT-6 and its 5' deletion mutants with the indicated endpoints is shown below. The boxed region depicts the 5' LTR in which the positions of the direct repeats (DR-1A, DR-2A, DR-3A, and DR-3B), SV40 core enhancer-like sequence (CE), and CAAT and TATA boxes are indicated in the order of occurrence but not to scale. The relative activity value for each mutant, shown on the right, was the average of two experiments, with the variation between experiments less than 20% of the average. The data were calculated after adjusting for pSVO-CAT background and normalizing the pSV2-CAT value to 100%. The level of pCAT-6 activity was equal to the pSVO-CAT background activity (0.06% of pSV2-CAT activity).

FIG. 8. Nucleotide sequence of cellular DNA immediately upstream of the CFE-6 5' LTR. 5' deletion mutation analysis indicated the location of two sets of negative transcription regulatory elements, one set being in the region marked by *Bam*HI and *Ppu*MI restriction sites (-1074 to -550 bp), and the other within the *Ppu*MI site and the position marked by an asterisk (-550 to -426 bp).

functional studies with the endogenous LTRs suggest that although the basic characteristics for promotion and enhancement of transcription are retained, additional *cis*-acting elements may strongly influence the activity of the LTR. These three issues are discussed below.

Nucleotide sequence data show that all of the endogenous 5' LTRs examined are more closely related to each other than to the exogenous LTRs. Analogous to the diversity observed in the U3 region of the avian and murine endogenous and exogenous proviral LTRs (35), the variability between endogenous and exogenous FeLV LTRs is also extensive in this region. Three major segments in endogenous U3 are largely nonhomologous to the exogenous LTR, and these stretches of nucleotides probably confer the U3-specific hybridization used previously to distinguish between exogenous and endogenous FeLV sequences (5). Located within the endogenous U3, there are three sets of direct repeats (DR-1, DR-2, and DR-3), ranging in length from 14 to 19 bp, each except DR-2 occurring in duplicate in all three endogenous LTRs. A second copy of DR-2 is found in the CF-14 clone but not in the others. The exogenous LTR of FeLV subgroup A, B, and C has only one copy of DR-1, part of DR-2, and one and a half copies of DR-3. The DNA sequence of R-U5 and the region downstream of the 5' LTR encompassing the primer-binding site, leader, and almost to the end of p15^{gag} reveals a high degree of conservation. However, the presence of frameshift mutations and a termination codon predict altered gPr80^{gag} and p15^{gag} proteins, if any are produced at all.

Two possibilities may be considered to explain the observed divergence between the endogenous and exogenous

FeLV LTRs. The progenitor mammalian retrovirus, probably of rodent origin (3), which introduced the viral sequences into the feline germ line might have been similar to the present-day FeLV in the structure of its U3 region. The proviral loci in the host genome were subsequently subjected to modifications, including insertions and additional mutations in U3 and other regions, for selective advantages. Alternatively, the evolutionary progenitor of FeLV might have been more closely related to what is now retained in the endogenous elements, but the present-day FeLV, originating from the prototype virus, has undergone mutations over the evolutionary period. This second possibility may find support in the results of the present study. The major LTR variation between endogenous and exogenous FeLV sequences is located in the U3 region, implicated in replication and in species and tissue specificities (35). The divergence involves the direct repeats and several stretches of sequence that are conserved in replication-competent mammalian retroviruses. Deletion of direct repeats within LTRs is not uncommon in naturally occurring retroviruses (35), and it has been shown experimentally that large direct repeats carried within retroviral vectors are unstable and are frequently deleted during virus replication (24). Thus, the progenitor FeLV might have undergone deletions in certain duplicate sequences and acquired new sequences in the U3 region that conferred replication efficiency and other selective advantages to the virus. Both deletion and acquisition events might have been accomplished simultaneously by recombination with other mammalian retroviruses, although currently there is no direct evidence to substantiate this speculation on the origin of FeLV.

The results of functional activity studies with constructs having similar structural features except for cellular flanking sequences show that most of the endogenous FeLV 5' LTRs carry the potential to enhance the transcription of a gene linked 5' to the LTR. In contrast to the observed enhancing effect, the promoter ability of most of the endogenous LTR constructs in the *in vitro* test system is poor, except for one cloned DNA (CFE-16) that displays an activity as robust as that of the exogenous LTR construct. The CFE-16 locus is a truncated provirus, and the CAT activity of its 5' LTR in feline cells is consistent with the transcription efficiency of this proviral DNA in mouse cells (31). Transcriptional inefficiency of the cloned DNAs of the nearly full-length proviruses (31) is also consistent with the lack of promoter activity of their 5' LTRs in the transient expression assays in feline cells. In view of the similarity of the sequences within the 5' LTRs of these endogenous FeLV DNAs, and since all LTR constructs contained variable lengths of flanking DNA sequences, it is likely that the overall efficiency of the *cat* constructs is influenced by adjoining proviral DNA sequences. In fact, reconstruction and deletion analyses indicate the presence of negative regulatory elements outside the 5' LTR of the CFE-6 full-length proviral clone. Since the results of CAT assays with the mix-and-match constructs are independent of the cell types used, the negative regulatory elements might then be interacting with non-tissue-specific factors. The interaction must be very strong to override the effect of other transcription-initiating interactions involving the common functional domains within the endogenous FeLV LTRs.

Although there is a good correlation between promoter ability of the 5' LTR region and transcription efficiency of the proviral DNA in the CFE-16 and CFE-6 clones, there is an apparent lack of correlation in the data obtained with the CF-14 DNA. CF-14 is a highly truncated locus, its genome

size being only 4.8 kb (31). The 5' LTR region of CF-14 did not support *cat* expression in either feline cells (described here) or mouse cells (data not shown). Also not shown is the RNA dot blot analysis with cells transfected with CFE-16 and CF-14 *cat* constructs (pCAT-16 and pCAT-14). This RNA analysis confirmed that pCAT-14 did not support detectable transcription of the *cat* gene, while *cat* transcripts could be readily seen in cells transfected with pCAT-16. However, the corresponding CF-14 proviral DNA, containing both the 5' and 3' LTRs, was transcriptionally active in mouse cells (31). Since the CF-14 LTR sequence is very similar to those of CFE-16 and CFE-6, it is likely that its transcription-promoting inefficiency may be related to potential negative regulatory elements present in the 5'-flanking DNA, in analogy to the CFE-6 clone. This inhibition is then apparently overcome by the regulatory elements in the downstream viral and cellular sequences of this truncated provirus. It seems very likely that the 3' LTR present only at a distance of 4.0 kb from the 5' LTR in the CF-14 provirus may play a significant positive role in its transcription. This proposal on the interaction between positive and negative regulatory elements in the vicinity of the CF-14 provirus, however, awaits further analysis.

In conclusion, the nucleotide homology of the three 5' LTRs derived from three independent endogenous FeLV proviral loci strongly supports the concept that most members of this provirus family retain the ability to promote and enhance transcription. Differences in their *in vitro* promoter function are probably not indicative of minor sequence divergence but reflect an effect of *cis*-acting regulatory elements present in the adjacent cellular DNA. This implies that expression of endogenous FeLV proviruses is tightly regulated by proximal host DNA sequences. Likewise, demonstration of enhancer functions residing in these LTRs supports the idea that endogenous LTRs may influence the expression of cellular genes which may be located in the vicinity of the proviral loci.

ACKNOWLEDGMENTS

We thank B. Howard for plasmids pSV0-CAT and pSV2-CAT, A. Roy-Burman for computer-assisted artwork, E. Mader for critical reading, and R. Potasi and L. Doumak for careful typing of the manuscript.

This work was supported by Public Health Service grant CA40590 from the National Institutes of Health.

LITERATURE CITED

- Baluda, M. A., and P. Roy-Burman. 1973. Partial characterization of RD-114 virus by DNA-RNA hybridization studies. *Nature (London) New Biol.* **244**:59-62.
- Banerji, J., L. Olson, and W. Schaffner. 1983. A lymphocyte-specific cellular enhancer is located downstream of the joining region in immunoglobulin heavy chain genes. *Cell* **33**:729-740.
- Benveniste, R. E., C. J. Sherr, and G. J. Todaro. 1975. Evolution of type C viral genes: origin of feline leukemia virus. *Science* **190**:886-888.
- Busch, M. P., B. G. Devi, L. H. Soe, B. Perbal, M. A. Baluda, and P. Roy-Burman. 1983. Characterization of the expression of cellular retrovirus genes and oncogenes in feline cells. *Hematol. Oncol.* **1**:61-75.
- Casey, J. W., A. Roach, J. I. Mullins, K. B. Burck, M. O. Nicolson, M. B. Gardner, and N. Davidson. 1981. The U3 portion of feline leukemia virus DNA identifies horizontally acquired proviruses in leukemic cats. *Proc. Natl. Acad. Sci. USA* **78**:7778-7782.
- Chen, H. R., and W. C. Barker. 1984. Nucleotide sequences of the retroviral long terminal repeats and their adjacent regions. *Nucleic Acids Res.* **12**:1767-1778.
- Dale, R. M. K., B. A. McClure, and J. P. Houchins. 1985. Rapid single-stranded cloning strategy for producing a sequential series of overlapping clones for use in DNA sequencing: application to sequencing the corn mitochondrial 18S rDNA. *Plasmid* **13**:31-40.
- Donahue, P. R., E. A. Hoover, G. D. Beltz, N. Riedel, V. M. Hirsch, J. Overbaugh, and J. I. Mullins. 1988. Strong sequence conservation among horizontally transmissible, minimally pathogenic feline leukemia viruses. *J. Virol.* **62**:722-731.
- Elder, J. H., and J. I. Mullins. 1983. Nucleotide sequence of the envelope gene of Gardner-Arnstein feline leukemia virus B reveals unique sequence homologies with a murine mink cell focus-forming virus. *J. Virol.* **46**:871-880.
- Gorman, C. M., L. F. Moffat, and B. H. Howard. 1982. Recombinant genomes which express chloramphenicol acetyltransferase in mammalian cells. *Mol. Cell. Biol.* **2**:1044-1051.
- Graham, F. L., and A. J. Van der Eb. 1973. A new technique for the assay of human adenovirus 5 DNA. *Virology* **52**:456-467.
- Guilhot, S., A. Hampe, L. D'Auriol, and F. Galibert. 1987. Nucleotide sequence analysis of the LTRs and *env* genes of SM-FeSV and GA-FeSV. *Virology* **161**:252-258.
- Hampe, A., M. Gobet, J. Even, C. J. Sherr, and F. Galibert. 1983. Nucleotide sequence of feline sarcoma virus long terminal repeats and 5' leaders show extensive homology to those of other mammalian retroviruses. *J. Virol.* **45**:466-472.
- Hoover, E. A., J. I. Mullins, S. L. Quackenbush, and P. W. Gasper. 1987. Experimental transmission and pathogenesis of feline AIDS. *Blood* **70**:1880-1892.
- Laprevotte, I., A. Hampe, C. J. Sherr, and F. Galibert. 1984. Nucleotide sequence of the *gag* and *gag-pol* junction of feline leukemia virus. *J. Virol.* **50**:884-894.
- Maniatis, T., E. F. Fritsch, and J. Sambrook. 1982. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
- Niman, H. L., M. Akhavi, M. B. Gardner, J. R. Stephenson, and P. Roy-Burman. 1980. Differential expression of two distinct endogenous retrovirus genomes in developing tissues of the domestic cat. *J. Natl. Cancer Inst.* **64**:587-594.
- Niman, H. L., M. B. Gardner, J. R. Stephenson, and P. Roy-Burman. 1977. Endogenous RD-114 virus genome expression in malignant tissues of domestic cats. *J. Virol.* **23**:578-596.
- Niman, H. L., J. R. Stephenson, M. B. Gardner, and P. Roy-Burman. 1977. RD-114 and feline leukemia virus genome expression in natural lymphomas of domestic cats. *Nature (London)* **266**:357-360.
- Overbaugh, J., P. R. Donahue, S. L. Quackenbush, E. A. Hoover, and J. I. Mullins. 1988. Molecular cloning of a feline leukemia virus that induces fatal immunodeficiency disease in cats. *Science* **239**:906-910.
- Parker, B., and G. R. Stark. 1979. Regulation of simian virus 40 transcription: sensitive analysis of the RNA species present early in infections by virus or viral DNA. *J. Virol.* **31**:360-369.
- Quintrell, N., H. E. Varmus, J. M. Bishop, M. O. Nicolson, and R. M. McAllister. 1974. Homologies among the nucleotide sequences of the genomes of C-type viruses. *Virology* **58**:569-575.
- Rasheed, S., and M. B. Gardner. 1980. Characterization of cat cell cultures for expression of retrovirus, FOCMA and endogenous sarc genes, p. 393-400. In W. D. Hardy, Jr., M. Essex, and A. J. McClelland (ed.), *Proceedings of the Third International Feline Leukemia Virus Meeting*. Elsevier/North-Holland Publishing Co., New York.
- Rhode, B. W., M. Emerman, and H. M. Temin. 1987. Instability of large direct repeats in retrovirus vectors. *J. Virol.* **61**:925-927.
- Riedel, N., E. A. Hoover, P. W. Gasper, M. O. Nicolson, and J. I. Mullins. 1986. Molecular analysis and pathogenesis of the feline aplastic anemia retrovirus, feline leukemia virus C-SARMA. *J. Virol.* **60**:242-250.
- Sanger, F., A. R. Coulson, B. G. Barrell, A. J. H. Smith, and B. A. Roe. 1980. Cloning in single-stranded bacteriophage as an aid to rapid DNA sequencing. *J. Mol. Biol.* **143**:161-178.
- Sarma, P. S., and T. Log. 1971. Viral interference in feline

- leukemia-sarcoma complex. *Virology* **44**:352-358.
28. **Sarma, P. S., T. Log, S. Skuntz, S. Krishnan, and K. Burkley.** 1978. Experimental horizontal transmission of feline leukemia viruses of subgroups A, B and C. *J. Natl. Cancer Inst.* **60**:871-874.
29. **Snyder, H. W., W. D. Hardy, E. E. Zuckerman, and E. Fleissner.** 1978. Characterization of a tumour-specific antigen on the surface of feline lymphosarcoma cells. *Nature (London)* **275**:656-658.
30. **Soe, L. H., B. G. Devi, J. I. Mullins, and P. Roy-Burman.** 1983. Molecular cloning and characterization of endogenous feline leukemia virus sequences from a cat genomic library. *J. Virol.* **46**:829-840.
31. **Soe, L. H., R. W. Shimizu, J. R. Landolph, and P. Roy-Burman.** 1985. Molecular analysis of several classes of endogenous feline leukemia virus elements. *J. Virol.* **56**:701-710.
32. **Spodick, D. A., A. K. Ghosh, S. Parimoo, and P. Roy-Burman.** 1988. The long terminal repeat of feline endogenous RD-114 retroviral DNAs: analysis of transcription regulatory activity and nucleotide sequence. *Virus Res.* **9**:263-283.
33. **Spodick, D. A., L. H. Soe, and P. Roy-Burman.** 1984. Genetic analysis of the feline RD-114 retrovirus-related endogenous elements. *Virus Res.* **1**:543-555.
34. **Stewart, M. A., M. Warnock, A. Wheeler, N. Wilkie, J. I. Mullins, D. E. Onions, and J. C. Neil.** 1986. Nucleotide sequence of a feline leukemia virus subgroup A envelope gene and long terminal repeat and evidence for the recombinational origin of subgroup B viruses. *J. Virol.* **58**:825-834.
35. **Weiss, R., N. Teich, H. Varmus, and J. Coffin (ed.).** 1985. RNA tumor viruses: molecular biology of tumor viruses, 2nd ed., supplement and appendixes. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.