Characterization of Enhancer Elements and Their Mutations in the Long Terminal Repeat of Feline Endogenous RD-114 Proviruses

ANANTA K. GHOSH¹ AND PRADIP ROY-BURMAN^{1,2*}

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

Received 3 April 1989/Accepted 16 June 1989

To locate the enhancer regions of the feline endogenous RD-114 long terminal repeat (LTR), we examined expression of the chloramphenicol acetyltransferase gene driven by various segments of the U3 region from two different proviral loci (CRL3 and CR1). Transient expression assays demonstrated that the primary signal sequence for transcription enhancement was located within the 63-base-pair (bp) element of the CRL3 DNA occurring between positions -184 and -121 from the CAP site (+1), whereas the similar region of CR1 was almost inactive. This element from both CRL3 and CR1 contained a single 30-bp sequence (direct repeat [DR]-B2) found in duplicate tandem copies in the LTR of the infectious RD-114 provirus. Two 9-bp inverted repeats marked the DR-B unit of the active element, and a prominent base deletion in one of these repeats in CR1 DNA appeared to be related to loss of enhancer activity. Another segment of CRL3 (-296 to -184), also displaying enhancer function, contained tandem repeated sequences (DR-A1 and DR-A2). The DR-A2 unit, which lacked the 5' 20-bp sequence of the 47-bp DR-A1, could not function as an enhancer by itself, but it contributed to enhancer effects in cooperation with either the DR-A1 or DR-B2 region. The CR1 LTR contained a single DR-A1 sequence with extensive mutations, and the region (-313 to -181) containing this DR-A1 unit was nonfunctional, similar to the DR-B2 region of CR1. Site-directed mutagenesis analysis of another enhancer element, an octamer motif occurring between CAAT and TATA boxes of all RD-114 LTRs sequenced, revealed that this element was necessary for full enhancer function of the U3 region but with a variable effect, depending on the cell types in which chloramphenicol acetyltransferase expression was determined.

There are 15 to 20 copies of stably inherited RD-114 retroviral DNA per haploid genome of the domestic cat (1, 3, 27, 28, 32, 38). Apparently, most are defective in one or more viral genes, and it is probable that there is only a single locus with the capacity to code for a viable xenotropic RD-114 retrovirus (27, 38). However, several of the partially deleted RD-114 proviruses in the cat genome are strongly expressed in certain feline tissues and cell lines (5, 23-25, 30, 37). Previously, we showed that the 5' long terminal repeat (LTR) of an endogenous RD-114 locus (CRL3), partly deleted in the *env* gene, retained strong activity in promoting the transcription of a linked gene, although this activity was approximately threefold lower than that of the infectious RD-114 provirus (37). In contrast, the 5' LTR from another similarly deleted proviral locus (CR1) lacked this ability to support transcription. DNA sequencing of these LTRs revealed interesting differences in the U3 region, especially in the two sets of tandem direct repeats (DRs) found in the LTR of the infectious RD-114 provirus (37). None of the sequences within any of the RD-114 DRs or anywhere in U3 contained regions homologous to the core enhancers of similar virus 40 (SV40), mouse leukemia virus, or mouse mammary tumor virus (6, 7, 14, 18, 19, 41). In addition to these novel DRs or their mutated forms, all RD-114 proviral LTRs sequenced showed the presence of a conserved octameric sequence identical to the sequence functionally implicated in transcription regulation of several mammalian genes (9, 22, 26, 29, 35, 37).

In this report, we extend our endogenous RD-114 LTR studies to characterize the enhancer elements present in the

U3 region. A series of constructs linking various regions of the active (CRL3) or inactive (CR1) LTR to the bacterial chloramphenicol acetyltransferase (*cat*) gene has been generated. Study of these constructs derived from these two functionally opposite but structurally closely related variants has aided in the identification of important domains by correlating sequence changes with differences in activity. Here, we demonstrate the presence of transcriptional enhancer in domains of the U3 region of the CRL3 LTR and show that the similar domains of the CR1 LTR have been inactivated by naturally occurring mutations. We also provide evidence that the octamer motif, occurring between the CAAT and TATA boxes of this provirus, functions like its mammalian gene counterpart in enhancing transcription in a cell-type-dependent manner.

MATERIALS AND METHODS

Cell lines. The mouse fibroblast cell line NIH 3T3 and feline embryo fibroblast cell line H927 (30) were maintained in high-glucose Dulbecco modified Eagle medium supplemented with 10% fetal bovine serum. The human B-lymphoid tumor cell line Raji was cultured in RPMI 1640 medium supplemented with 10% fetal bovine serum.

Plasmid constructions. Plasmid pSVIX*cat* (12), provided by D. Celander, is identical to pSVI*cat* except that it contains an *Xho*I linker at the site of recircularization. This plasmid has the SV40 promoter, but not enhancer, sequences upstream of the *cat* gene (Fig. 1B). The endogenous RD-114 LTR putative enhancer regions were inserted upstream of the promoter sequences in pSVIX*cat*. Four 5' LTRs, clones CRL3, CR1, CR5, and CRL17 (38), were used. The 309-base-pair (bp) *Sau*3AI-*Hind*III fragment of the

* Corresponding author.

FIG. 1. Differential activity of the enhancer region from the 5' LTRs of two endogenous RD-114 proviruses, CRL3 and CR1. (A) Schematic of features, including U3, R, and U5 regions and certain restriction sites within or adjacent to each LTR. The bottom line for each LTR structure represents an expanded version of the 5' half of U3, with DR units shown as boxes. The LTR of the infectious RD-114 provirus contains two sets of tandem DRs (37). The first 47-bp DR (DR-A1) is followed by a second copy, DR-A2. Similarly, a 30-bp DR (DR-B1) is repeated as DR-B2. In CRL3, the 5' portion of DR-A2 is missing, but the 3' 27-bp region is retained. CR1 has only one DR-A1-related region, and both CRL3 and CR1 lack the first copy of the 30-bp DR. The DR boxes are shaded differently to emphasize sequence nonidentity between DR-A and DR-B domains.

CRL3 5' LTR, containing DR units of its U3 region as well as a portion of 5' flanking cellular DNA sequences, was inserted into the *XhoI* site of pSVIXcat (Fig. 1A and B). Similarly, the 193-bp *PstI*-*HindIII* fragment of each of the other three 5' LTRs (CR1, CR5, and CRL17) was cloned into that plasmid. The plasmid derivatives, named pCRL-3/SVIX, pCR-1/SVIX, pCR-5/SVIX, and pCRL-17/SVIX, were isolated in both orientations of the insert, the correct orientation being designated → and the inverted orientation being designated ← (Fig. 1B).

Plasmids pL3 and pCR1 (Fig. 2A) were constructed by inserting the 411-bp *HindIII*-*EcoRI* fragment of CRL3 or the corresponding 409 bp of the CR1 LTR into the *HindIII* site of pSVOCat (12). These constructs, lacking DR units but containing the promoter region (CAAT, octamer, and TATA boxes), were used as recipient plasmids for introduction of various enhancer sequence domains. The correctly oriented pL3 and pCR1 plasmids were cut at the *HindIII* site for insertion of either the 309-bp (-121 to -430 of CRL3) or 193-bp (-120 to -313 of CR1) fragment, using *HindIII* linkers. The resulting plasmids, containing homologous and mix-and-match sequences between CRL3 and CR1 LTRs, were termed pCRL-3/L3, pCR-1/L3, and pCRL-3/CR1 (Fig. 2A).

To generate DR deletion mutants, pCRL-3/L3 was first linearized by *NdeI* digestion, and the 5' upstream DNA was trimmed back by BAL 31 exonuclease. *HindIII* linkers were then added to the deletion ends and excised by *HindIII* digestion, and the isolated 5' deletion DNA fragments were inserted into the *HindIII* site of pL3. Two deletion mutants were isolated (Fig. 3A): one lacking most of the sequences upstream of the DR region (pL-3/L3) and the other with most of the DR-A1 sequence eliminated [pdL-3(-221)/L3]. For making other deletion mutants, pL-3/L3 was digested with both *AvaII* and *HindIII*, and two ensuing fragments (-296 to -184 and -184 to -121) were modified by *HindIII* linkers and ligated into the *HindIII* site of pL3. This process generated plasmids pdL-3(-184)/L3 and pdL-3(-296)/L3. Similarly, pdCR-1(-181)/CR1 and pdCR-1(-313)/CR1 were obtained from pCR-1/CR1 (Fig. 3A).

Two additional DR deletion mutants, pdCR-1(-181)/L3 and pL-3(-221)/L3, were also obtained. The *HindIII*

and the 27-bp repeats in CRL3 are indicated by dense stippling. A solid arrow represents the start of each repeat unit; the actual length of the unit (in base pairs) is indicated above the box. (B) Construction of cat plasmids, showing the pSVIXcat circle containing SV40 early promoters (■) and sequences encoding the *cat* gene (□). The plasmid sequences of the circle, although not shown, also contain SV40 intron and poly(A) addition sites immediately downstream of the *cat* gene. The diagrams below the circle depict LTR segments of different RD-114 proviruses (CRL3, CR1, CR5, and CR17) and the relevant restriction endonuclease cleavage sites. Abbreviations for restriction sites: B, *BamHI*; E, *EcoRI*; H, *HindIII*; Sa, *Sau3A1*; Sm, *SmaI*; P, *PstI*; X, *XhoI*. Results of CAT assays after transfection of these plasmids with correct (→) or inverted (←) orientation of the insert into NIH 3T3 cells are shown at the right. CAT activity is expressed as the level of conversion of chloramphenicol to its acetylated forms relative to that observed with pSVIXcat. Each value represents the average of three transfections; the variation between experiments was less than 30% of the average. (C) Representative autoradiogram of the CAT assay. Lanes contained transfected DNA from: a, pSVIXcat; b, pCRL-3/SVIX (→); c, pCRL-3/SVIX (←); d, pCR-1/SVIX (→); e, pCR-1/SVIX (←); f, pCRL-17/SVIX (→); g, pCRL-17/SVIX (←); h, pCR-5/SVIX (→); i, pCR-5/SVIX (←). Positions of migration of chloramphenicol (Cm) and its acetylated forms (1-Cm and 3-Cm) are indicated here and in Fig. 2B to 4B.

FIG. 2. Effect of enhancer region on homologous promoters. (A) Line drawings of *cat* plasmids containing only the promoter region from the CRL3 or CR1 LTR (pL3 or pCR1) or both promoter and upstream enhancer regions (pCRL3/L3). Of the two mix-and-match constructs, pCR-1/L3 contains promoters from CRL3 and upstream sequences from CR1, and pCRL-3/CR1 contains promoters from CR1 but the enhancer region from CRL3. The boxed regions mark DR units, and the relative positions of the CAAT, octamer, and TATA elements are indicated within the promoter region. CAT activity data are average values obtained from three transfection experiments in NIH 3T3 cells when the orientation of the enhancer (+E) region was in natural (→) or inverted orientation (←) with respect to the promoter region. -E, Values obtained without the enhancer region. All values are expressed relative to the activity of pL3 (normalized to 1.0). (B) Autoradiogram from a typical CAT assay. Lanes contained plasmid DNAs from: a, pSVOCat; b, pL3; c, pCRL-3/L3 (→); d, pCRL-3/L3 (←); e, pCR-1/L3 (→); f, pCR-1/L3 (←); g, pCR1; h, pCRL-3/CR1 (→); i, pCRL-3/CR1 (←).

fragment (-181 to -120), isolated from pdCR-1(-181)/CR1, when cloned into the *Hind*III site of pL3 provided pdCR-1(-181)/L3. To generate pL-3(-221)ID/L3, the *Nde*I-*Ava*II fragment of pdL-3(-221)/L3 was blunt ended into *Nde*I- and *Hind*III-cleaved pL3 (Fig. 3A). Deletion endpoints of all constructs and their orientations were determined by dideoxy sequencing after cloning of the adjoining region into an M13 vector (31).

Mutagenesis of the octamer sequence ATGCAAT was done by the method of Zoller and Smith (43) as modified by Geisselsoder et al. (11). Briefly, two oligonucleotides, octamer A (ATTCAACT; OA) and octamer B (ATTAAAAT; OB), were synthesized and purified by high-performance liquid chromatography. The *Hind*III-*Sma*I (-121 to +200) fragment of CRL3, containing the octamer motif, was cloned into the same restriction-digested M13mp18 vector and transformed into *Escherichia coli* CJ236 (*dut ung* strain) to yield uracil-containing single-stranded bacteriophage DNA (16). The phage DNA was then annealed with the kinase-treated mutant oligonucleotide, the complementary strand was extended by T4 DNA polymerase, and the ends were joined by T4 ligase. The heteroduplex DNA was used to transform MV1190, an *E. coli* strain containing active uracil *N*-glycosylase (16). The presence of this enzyme reduced the biological stability of the parental DNA strand and consequently led to accumulation of the mutant phage. After the mutations in the progeny phage were confirmed by DNA sequencing, the *Hind*III-*Sma*I fragment was excised from each mutant phage DNA and cloned into *Hind*III-*Sma*I-

cleaved pL3 or pL-3/L3 to generate pL3(OA) and pL3(OB) or pL-3/L3(OA) and pL-3/L3(OB), respectively (Fig. 4A).

All plasmids were purified by two CsCl gradient centrifugations before use in transfection analyses.

Transfections and CAT assays. Feline H927 or mouse NIH 3T3 fibroblasts were transfected by a modification of the calcium phosphate method of Graham and Van der Eb (13). Briefly, 5×10^5 cells were seeded onto 100-mm-diameter dishes 1 day before transfection in Dulbecco modified Eagle medium plus 10% fetal bovine serum. Calcium phosphate DNA precipitates were prepared from 20 μ g of plasmid DNA along with 30 μ g 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 20 min for the precipitates to form, the suspension 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 48 h at 37°C before harvesting (4).

The human Raji cells were transfected with the plasmid DNA by a modified DEAE-dextran protocol (2). Briefly, 2×10^7 cells were washed twice with serum-free medium 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 incubated at room temperature in 2 ml of TBS containing 500 μ g of DEAE-dextran (Pharmacia, Inc.) per ml and 20 μ g of plasmid DNA for 45 min. The unabsorbed DNA was removed by

FIG. 3. Structures and activities of mutants bearing 5' unidirectional deletions or deletions in the central region. (A) Line drawings numbered with deletion endpoints. The source of the enhancer region is indicated above each line by CRL-3 or CR-1. DR domains, promoter elements, and other symbols are as for Fig. 2. ID in construct designations indicates internal deletion. CAT activity data are the averages of at least three but in most cases up to eight independent transfection experiments. NT, No transfection done. (B) Representative CAT assay autoradiogram. Lanes contained plasmid DNAs from: a, pL3; b, pL-3/L3 ($\rightarrow$); c, pdL-3 (-221)/L3 ($\rightarrow$); d, pdL-3 (-184)/L3 ($\rightarrow$); e, pdL-3 (-296)ID/L3 ($\rightarrow$); f, pdL-3 (-221)ID/L3 ($\rightarrow$); g, pdCR-1 (-181)/L3 ($\rightarrow$); h, pCR-1/CR1 ($\rightarrow$); i, pdCR-1(-313)ID/CR1 ($\rightarrow$); j, pdCR-1 (-181)/CR1 ($\rightarrow$); k, pCR1. Lane l, Untransfected cell lysate.

two washes with TBS and one wash with serum-free medium, and the cells were replated in 20 ml of serum-containing medium. After 48 h of incubation at 37°C, the cells were harvested for the CAT assay.

Transfected cells were washed three times with cold phosphate-buffered saline and lysed in 100 μ l of 250 mM Tris hydrochloride (pH 7.8) by three cycles of freezing and thawing. The cellular debris was pelleted by centrifugation, and portions of the supernatant fluid equivalent to 150 μ g of protein were assayed for CAT activity as described by Gorman et al. (12). 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. The positions of the spots were located by autoradiography.

RESULTS

Activity of the enhancer regions from endogenous proviral DNAs. To characterize the enhancer elements, we introduced the LTR and adjoining cellular DNA region upstream

of the CAAT box and containing the DR units into the parental test plasmid pSVIX $_{cat}$ (Fig. 1A and B). Plasmid pCRL-3/SVIX comprised the 309-bp *HindIII-Sau3A1* (-121 to -430) fragment of the CRL3 clone placed 5' to SV40 promoter sequences. Similar constructions were made with the analogous 193-bp *HindIII-PstI* fragment occurring in each of the CR1, CR5, and CRL17 clones. These constructs, containing the insert in either correct or opposite orientation, were transfected into NIH 3T3 mouse fibroblasts and assayed for transient CAT expression. Average CAT activities relative to that of pSVIX $_{cat}$ are given in Fig. 1B, and a representative autoradiogram shown in Fig. 1C. The 309-bp region of the CRL3 DNA induced a severalfold increase in *cat* gene expression. In the correct orientation, the enhancement was 11-fold; the opposite orientation of the insert resulted in an approximately 5-fold increase. In contrast, the corresponding fragments of CR1, CR5, and CRL17 cloned DNAs were inactive in inducing *cat* gene expression over the background level of pSVIX $_{cat}$. Therefore, it appeared that the inactivity of the CR1 fragment was not an isolated case, as there were other loci that were similarly inactive.

FIG. 4. Structures and activities of mutants bearing altered octameric sequences. (A) The octamer sequence (ATGCAAAT) was replaced by either OA (ATTCAACT) or OB (ATTAAAAAT), and plasmids pL-3/L3 and pL3 containing these mutations were isolated. The wild-type and mutant plasmids were transfected into feline H927 fibroblasts, mouse NIH 3T3 fibroblasts, or human Raji B-lymphoma cells for *cat* gene expression analyses. CAT activity values are averages of the data obtained from three transfection experiments. Actual percent conversion of chloramphenicol to its acetylated forms is given in parentheses. (B) Autoradiogram from a typical CAT assay with Raji cell extracts. Lanes contained transfected DNAs from: a, pL3; b, pL3(OA); c, pL3(OB); d, pL-3/L3(OA); e, pL-3/L3(OB); f, pL-3/L3. Lane g, Untransfected cell lysate.

These experiments with heterologous promoters were then extended to study the effect of the enhancer region on homologous promoters. Since the nucleotide sequences of the CRL3 and CR1 LTRs were known (37), all subsequent work was devoted to defining the enhancer elements within the 309-bp region of CRL3 and comparing them with those in the similar region of the functionally inactive CR1 LTR. The vectors containing the homologous promoters were created by inserting the 411-bp *HindIII-EcoRI* fragment (−121 to +290) of CRL3 or the corresponding 409 bp (−120 to +289) of CR1 5' of the *cat* gene in pSV Ocat (12), which did not contain any promoter or enhancer elements. The resulting plasmids, pL3 and pCR1, were used to construct mix-and-match sequences between individual promoter and enhancer regions (Fig. 2A). Transfection and CAT assays with these constructs revealed that CRL3 and CR1 DNAs differed in both promoter and enhancer activities. The promoter ability of pCR1 was approximately 50% of that observed with pL3 (Fig. 2). This reproducible drop in activity could not be attributed to the sequences of the functional domains such as CAAT, octamer, and TATA boxes, since CR1 sequences in these domains were identical to either the CRL3 LTR or the LTR of the infectious RD-114 virus (37). In fact, the entire 409-bp (−120 to +289) segment of CR1 was highly homologous to the corresponding portion of the other two DNAs, differing in only 22 scattered nucleotide changes unique to CR1. However, it should be noted that the 119-bp region immediately upstream of the CAP site (+1) of the CR1 LTR had five unique changes, including a deletion 40 bp upstream of the CAAT box (37).

The degree of inactivation of the enhancer function of CR1 seemed much higher than the loss of its promoter ability. The CR1 region in plasmid pCR-1/L3 did not significantly alter expression driven by the CRL3 promoters, whereas the CRL3 enhancer region stimulated expression from the CR1 promoters by approximately ninefold in the sense orientation and threefold in the antisense orientation (Fig. 2). Considering the relative strengths of the CR1 and CRL3 promoters, this level of enhancement was similar to that observed with plasmid pCRL-3/L3, the values obtained being sevenfold and threefold, respectively. Together, these results indicated that the inability of the CR1 LTR to promote transcription was primarily caused by its enhancer weakness. The results also revealed that the active CRL3 enhancers were partially dependent on orientation, the antisense orientation being threefold lower in activity than the native order.

Borders of the enhancer region. The DR domains frequently occurring in the U3 region of the retroviral LTR have been implicated as transcriptional enhancer elements (6, 8, 15, 17, 20, 21). Since the CRL3 LTR contained two copies of partly homologous DR sequences (DR-A1 and DR-A2) and a single copy of DR-B that is found in duplicate copies in the infectious RD-114 virus LTR (37), we wished to assess the relative strength of the DR units and examine their cooperativity. Several deletion mutants were constructed. The first product isolated by 5' unidirectional deletion was pL-3/L3, which lacked most of the flanking cellular DNA sequence. This plasmid expressed the *cat* gene with efficiency similar to that of the previous construct, pCRL-3/L3,

which retained about 150 bp of 5' flanking DNA (Fig. 3). Further 5' deletion of the insert to -221 removed DR-A1 sequences completely except for the last 3 nucleotides. This mutant, pDL-3 (-221)/L3, also expressed the *cat* gene at the same level seen with the pL-3/L3 control, indicating that the DR-A1-containing region added very little to the total enhancer activity. Additional deletion resulted in plasmid pDL-3 (-184)/L3, with an endpoint at -184. However, this mutant, lacking both DR-A1 and DR-A2 regions, showed a reduction in activity of 42% of the control level. Thus, it seemed that the majority of the enhancer effect was associated with the sequences (-184 to -121) that included the 30-bp DR-B2 domain. This point was strengthened by the observed significant drop in activity when the DR-B2 region (-184 to -121) was internally deleted. Plasmid pDL-3(-296)ID/L3 showed only 35% of the control level of activity. Additional 5' deletion of this plasmid, lacking the DR-B2 region, resulted in elimination of most of the DR-A1 sequences and consequent loss of all enhancer activity in plasmid pDL-3(-221)ID/L3. Taken together, the results implied that the major enhancer domain resided in the -184 to -121 DR-B2 region and that the DR-A1 and DR-A2 regions had a cooperative enhancer effect, although the DR-A2 region by itself could not function as an enhancer in the transient expression assay system.

To test whether the DR-B2-containing region of the CR1 LTR could stimulate transcription when separated from the rest of the upstream sequences, we linked the -181 to -120 region of CR1 to the CRL3 promoters, as indicated in the pDCR-1(-181)/L3 construct (Fig. 3A). This CR1 region showed only a slight increase (40%) over that of the promoter-alone construct, whereas the same region from CRL3 displayed an almost fivefold enhancement of activity. Deletion plasmids were also made with CR1 DR regions linked upstream of the homologous CR1 promoters. Three such constructs (bottom of Fig. 3A), containing either the DR-A1 or DR-B2 region or both, did not exhibit any positive enhancer effect. Rather, it appeared that the DR-A1 region of CR1 might have a negative action on transcription promotion. For example, whereas pDCR-1(-181)/CR1 (retaining only the DR-B2 region) had no detectable effect on *cat* gene expression by CR1 promoters, pCR-1/CR1 (containing both DR-A1 and DR-B2 regions) and pDCR-1(-313)ID/CR1 (containing only the DR-A1 region) showed reduced activity (40 and 20%, respectively) compared with the value for the CR1 promoter alone (Fig. 3A). This effect might be related to extensive mutations in the CR1 47-bp DR domain relative to the CRL3 or exogenous LTR element.

Analysis of octamer deletion mutants. The octamer enhancer ATGCAAT, conserved as is or in its inverted form in several mammalian genes, also occurs between the CAAT and TATA boxes of RD-114 and baboon endogenous retrovirus but not in any other known retroviruses (37). To examine the function of this viral octamer, a set of constructs was made in which the octameric sequence was altered by oligonucleotide-directed mutagenesis. The mutants, OA and OB, each containing two nucleotide changes, had the sequences ATTCAACT and ATTAAAT, respectively (Fig. 4A). The mutated octamers in plasmid pL3 or pL-3/L3 were tested for *cat* gene expression in mouse NIH 3T3 and feline H927 fibroblasts and in human Raji B-lymphoma cells. *cat* gene expression was reduced by the octamer mutation to a variable degree, depending on the cell line used. Octamer mutants in the pL-3/L3 frame reduced the control activity to 55 to 75% in the fibroblastic cells (Fig. 4A). In the pL3 plasmids, the mutations lowered the level of

activity in NIH 3T3 cells by approximately 60%, but OB (not OA) caused significant reduction of *cat* gene expression in H927 cells. Of the three cell lines tested, Raji cells showed the most pronounced effect of the octamer mutations. In this cell line, pL3(OA) and pL3(OB) displayed only 20 and 11%, respectively, of the control level of activity (Fig. 4). Similarly, a reduced activity of 30 to 40% was seen with pL-3/L3(OA) and pL-3/L3(OB). These results implied that the viral octamer sequence, like the mammalian cellular gene counterpart, was necessary for efficient transcriptional activity and that the level of activity was dependent on the type of cell in which expression was measured.

DISCUSSION

We have analyzed the enhancer function of the U3 region of the LTR of endogenous RD-114 proviruses by using a *cat* gene expression system driven by either heterologous or homologous promoters. This analysis has yielded new information on the location of the enhancer comprising multiple functional elements. Although the boundary of each element has not been rigorously determined, the approximate location of each is deduced from analysis of the various deletion mutants. The element with the most pronounced activity that can function independently is located within a 63-bp sequence occurring between nucleotide positions -184 and -121. This region alone of the proviral clone CRL3 can stimulate the level of *cat* gene expression driven by its homologous promoters by fivefold. The similar region of another clone, CR1, located between -181 and -120, could not induce expression above the level obtained only with its promoters. The extent of inactivation of this CR1 enhancer element is apparently quite striking, especially in the presence of extensive homology between CRL3 and CR1 in the respective 63- and 61-bp stretches of sequence (37). There are only six base substitutions and two base deletions scattered in this region. The element, however, contains a copy of the 30-bp DR unit (DR-B2) (37). Of the eight base mutations, two distinguish the CR1 DR-B2 domain from that of CRL3: one deletion 5 bp 3' of its 5' border and one transversion mutation (A → T) 5 bp 5' of its 3' border. The base deletion is positioned near the middle of a 9-bp sequence, which occurs again in inverted form, with 3 nucleotides between copies (Fig. 5). In the LTR of infectious RD-114 (Ex-LTR in Fig. 5), the inverted repeats within the DR-B2 domain are perfectly matched except for a single base mutation. These repeats have three base mutations in CRL3, but the deletion in CR1 disrupts the sequence of the first domain, whereas its inverted repeat retains the same sequence as in CRL3. Since crucial sequences for retroviral transcriptional enhancers reside largely in the DRs of the U3 region (42), it is probable that DR-B domain has a significant role in the observed function. The palindromic sequence present within DR-B may be essential for the binding of a transactivating dimeric protein factor, and we speculate that partial disruption of this dyad symmetry in this sequence of CR1 may be a major determinant of its considerable loss of enhancer activity.

Two other segments of the CRL3 LTR that display enhancer activity also contain DR sequences. The DR-A1-containing region, located between positions -296 and -221, has all of the 47-bp DR-A1 sequence except for the last 3 nucleotides. Whereas the DR-A1 motif is tandemly repeated as DR-A2 in the LTR of the infectious RD-114 DNA, the 5' portion of DR-A2 is deleted in CRL3, resulting in an imperfect DR-A2 of 27 nucleotides. The DNA segment

FIG. 5. Comparison of the DR-B2 nucleotide sequence present in all three RD-114 LTRs. The sequence of the infectious or exogenous proviral LTR (Ex-LTR) as well as the sequences of CRL3 and CR1 LTRs were reported previously (37). The boxed areas represent the 9-bp inverted repeats; arrows indicate orientation. Symbols: ●, sequence identity between the DNAs; –, deletion found in CR1.

(-296 to -184) containing both DR-A1 and DR-A2 sequences of CRL3, when separated from the DR-B2 region (-184 to -121) and linked upstream of the *cat* gene with CRL3 promoters, increases *cat* gene expression by threefold. Thus, the -296 to -184 region has significant enhancer activity. Further deletion of this construct to remove the -296 to -221 segment, however, results in almost complete loss of enhancer function, indicating that the DR-A2-containing region (-221 to -184) has little or no enhancer activity by itself. However, when it is placed upstream of the DR-B2-containing element, as in construct pdL-3(-221)/L3, it increases the DR-B2 enhancer effect by almost another twofold. It is also noteworthy that deletion of the DR-A1 region (-296 to -221) from the entire enhancer does not reduce enhancer function.

Taken together, these results imply that the DR-A2 region contributes to enhancer effect when it pairs with either the DR-A1 or DR-B2 region but that as an individual element it lacks enhancer activity. It also appears that the presence of both DR-A1 and DR-A2 regions may not be required for full function of the entire enhancer. However, this observation is limited to the assay system used, and it remains possible that DR-A1 and DR-A2 together are critical for expression of the endogenous proviral sequences in specific tissues of the domestic cat. Since the DR-A2 domain lacks the first 20 bp of the DR-A1 sequence, the 3' half of this repeated sequence may be functionally more important than the 5' half. However, the issue is also complicated by the demonstrated cooperativity between the DR-A1 and DR-A2 regions of CRL3, and it is unclear whether the 5' region present in DR-A1 but not in DR-A2 plays a role in the enhancement of activity by DR-A1 and DR-A2 together. In contrast to the enhancer ability of the DR-A region of the CRL3 clone, the DR-A1 region of CR1 appears to be completely inactivated by the natural mutations. The mutations in this DR involve a 6-bp insertion near the 5' end and a 5-bp deletion of the 3'-terminal sequence (37).

The octamer sequence conserved in all RD-114 LTRs sequenced constitutes another enhancer element. Distinct elements containing this octamer motif have been implicated in both ubiquitous and tissue-specific enhancer function of mammalian genes (10, 33, 34, 36, 39, 40). The octamer sequence is bound by the ubiquitous octamer-binding factor, as well as by a B-cell-specific factor found only in cells of the B-lymphoid series. Sen and Baltimore (33) have proposed that *in vivo*, a given segment of DNA is perhaps initially activated to make the chromatin region accessible to binding factors. The tissue-specific binding proteins may play the role of activation and thereby open the DNA to interaction with ubiquitous transcription-enhancing proteins. Our site-directed mutagenesis experiments reveal a differential effect of mutations in the proviral octamer sequence. Similar to

findings for immunoglobulin gene octamer, mutations maximally affect the expression of the linked gene in B cells. Thus, it seems likely that this octamer motif plays an important role in tissue-specific expression of the RD-114 endogenous loci (5, 23–25, 30). This observation also indicates a potential utility of RD-114 vectors for gene transfer and expression in target cells. The newly defined enhancer elements, combined with the built-in octamer motif and the unusual glycine-tRNA primer-binding site (37) in the RD-114 vector, may add to the choice of vectors for gene transfer into desired cell types to augment or replace a defective gene.

The mechanisms by which the RD-114 enhancer elements function to stimulate transcription remain to be elucidated. Additional studies will be required to characterize the cognate sites in the enhancer elements that interact with viral or cellular DNA-binding proteins and to understand how protein-protein interactions may be responsible for the enhancer sequence elements that function in combination.

ACKNOWLEDGMENTS

We thank B. Bachman for expert technical assistance as well as for all computer-generated artwork. We also thank D. V. Kumar for helpful comments, F. Miyagawa for typing, and E. Mader for critical reading of the manuscript.

This study was supported by Public Health Service grant CA 40590 from the National Institutes of Health.

LITERATURE CITED

1. **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.
2. **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.
3. **Benveniste, R. E., and G. J. Todaro.** 1974. Evolution of C-type viral genes: inheritance of exogenously acquired viral genes. *Nature (London)* **252**:456-459.
4. **Berry, B. T., A. K. Ghosh, D. V. Kumar, D. A. Spodick, and P. Roy-Burman.** 1988. Structure and function of endogenous feline leukemia virus long terminal repeats and adjoining regions. *J. Virol.* **62**:3631-3641.
5. **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.
6. **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.
7. **DeFranco, D., and K. Yamamoto.** 1986. Two different factors act separately or together to specify functionally distinct activities at a single transcriptional enhancer. *Mol. Cell. Biol.* **6**:993-1001.
8. **DesGroselliers, 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.
9. **Falkner, F. G., and H. B. Zachau.** 1984. Correct transcription of an immunoglobulin kappa gene requires an upstream fragment containing conserved sequence elements. *Nature (London)* **310**:71-74.
10. **Fletcher, C., N. Heintz, and R. G. Roeder.** 1987. Purification and characterization of OTF-1, a transcription factor regulating cell cycle expression of a human histone H2B gene. *Cell* **51**:773-781.
11. **Geisselsoder, J., F. Witney, and P. Yuckenberg.** 1987. Efficient site-directed *in vitro* mutagenesis. *BioTechniques* **5**:786-791.
12. **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.
13. **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.
14. Gruss, P., R. Dhar, and G. Khoury. 1981. Simian virus 40 tandem repeated sequences as an element of the early promoter. *Proc. Natl. Acad. Sci. USA* 78:943-947.
 15. Ishimoto, A., M. Takimoto, A. Adachi, M. Kakuyama, S. Kato, K. Kakimi, K. Fukuoka, T. Ogiu, and M. Matsuyama. 1987. Sequences responsible for erythroid and lymphoid leukemia in the long terminal repeats of Friend-mink cell focus-forming and Moloney murine leukemia viruses. *J. Virol.* 61:1861-1866.
 16. Kunkel, T. A. 1985. Rapid and efficient site-specific mutagenesis without phenotypic selection. *Proc. Natl. Acad. Sci. USA* 82:488-492.
 17. Laimins, L. A., P. Gruss, R. Pozzatti, and G. Khoury. 1984. Characterization of enhancer elements in the long terminal repeat of Moloney murine sarcoma virus. *J. Virol.* 49:183-189.
 18. Lenz, J., D. Calender, G. 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.
 19. 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.
 20. 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.
 21. 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.
 22. Mason, J. O., G. T. Williams, and M. S. Neuberger. 1985. Transcription cell type specificity is conferred by an immunoglobulin V_H gene promoter that includes a functional consensus sequence. *Cell* 41:479-487.
 23. 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.
 24. 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.
 25. 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.
 26. Parslow, T. G., D. L. Blair, W. J. Murphy, and D. K. Granner. 1984. Structure of the 5' ends of immunoglobulin genes: a novel conserved sequence. *Proc. Natl. Acad. Sci. USA* 81:2650-2654.
 27. Reeves, R. H., W. G. Nash, and S. J. O'Brien. 1985. Genetic mapping of endogenous RD-114 retroviral sequences of domestic cats. *J. Virol.* 56:303-306.
 28. Reeves, R. H., and S. J. O'Brien. 1984. Molecular genetic characterization of the RD-114 gene family of endogenous feline retroviral sequences. *J. Virol.* 52:164-171.
 29. Roebuck, K. A., R. J. Walker, and W. E. Stumph. 1987. Multiple functional motifs in the chicken U1 RNA gene enhancer. *Mol. Cell. Biol.* 7:4185-4193.
 30. Roy-Burman, P., M. P. Busch, S. Rasheed, M. B. Gardner, and M. M. C. Lai. 1980. Oncodevelopmental gene expression in feline leukemia, p. 361-372. *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.
 31. 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.
 32. Sarma, P. S., J. Tseng, Y. K. Lee, and R. V. Gilden. 1973. Virus similar to RD-114 virus in cat cells. *Nature (London) New Biol.* 244:56-59.
 33. Sen, R., and D. Baltimore. 1986. Multiple nuclear factors interact with the immunoglobulin enhancer sequences. *Cell* 46:705-716.
 34. Singh, H., R. Sen, D. Baltimore, and P. A. Sharp. 1986. A nuclear factor that binds to a conserved sequence motif in transcriptional control elements of immunoglobulin genes. *Nature (London)* 319:154-158.
 35. Sive, H. L., N. Heintz, and R. G. Roeder. 1986. Multiple sequence elements are required for maximal in vitro transcription of a human histone H2B gene. *Mol. Cell. Biol.* 6:3329-3340.
 36. Sive, H. L., and R. G. Roeder. 1986. Interaction of a common factor with conserved promoter and enhancer sequences in histone H2B, immunoglobulin and U2 small nuclear RNA (SnRNA) genes. *Proc. Natl. Acad. Sci. USA* 83:6382-6386.
 37. 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.
 38. 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.
 39. Staudt, L. M., R. G. Clerc, H. Singh, J. H. LeBowitz, P. A. Sharp, and D. Baltimore. 1988. Cloning of a lymphoid-specific cDNA encoding a protein binding the regulatory octamer DNA motif. *Science* 241:577-580.
 40. 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.
 41. Weiher, H., M. Konig, and P. Gruss. 1983. Multiple point mutations affecting the simian virus 40 enhancers. *Science* 219:626-631.
 42. 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.
 43. Zoller, M., and M. Smith. 1983. Oligonucleotide-directed mutagenesis of DNA fragments cloned into M13 vectors. *Methods Enzymol.* 100:468-500.