Structural and Functional Analysis of Long Terminal Repeats of *Suncus murinus* Mammary Tumor Virus

SHIN-ICHI YANAGAWA,¹ AKIRA MURAKAMI,¹ MUNEMITSU HOSHINO,² AND HARUTAKA TANAKA^{1*}

Department of Viral Oncology, Institute for Virus Research, Kyoto University, Kyoto 606,¹ and Second Department of Pathology, Nagoya University School of Medicine, Nagoya 466,² Japan

Received 12 November 1987/Accepted 30 December 1987

A 2.7-kilobase (kb) cDNA sequence complementary to *Suncus murinus* mammary tumor virus (Sm-MTV) genomic DNA was prepared using purified virions produced by the Sm-MT cell line, which had been established from a spontaneous mammary tumor of *S. murinus*. It was found, by using this cDNA in Southern hybridization experiments, that Sm-MTV was endogenous to this animal and that some 50 copies of endogenous provirus were present per haploid cellular genome. In addition, a proviral Sm-MTV DNA sequence, 9.4 kb long (Sm-P-MTV10), was cloned from a Sm-MT cell genomic library, and its long terminal repeat was found to be 720 base pairs (bp) long, with the U3.R and U5 regions 574 and 146 bp long, respectively. The boundary between U3 and R was not determined with certainty, though in the cDNA, the U3 and R regions were 462 and 105 bp long, respectively. The overall homology between the U3.R regions in the cDNA and Sm-P-MTV was 75%. These two DNAs differed in such transcription regulatory signals as CCAAT and TATAA, the first being missing from the cDNA. Nevertheless, chloramphenicol acetyltransferase assays showed that the long terminal repeats of the cDNA and the Sm-P-MTV were transcriptionally active but not steroid hormone responsive. Like Mason-Pfizer monkey virus, Sm-MTV used tRNA^{Lys}_{1,2} as a primer for reverse transcription. In addition, the immunosuppressive peptide sequence common to many retroviruses was found in the *env* region of Sm-MTV. In these two points, Sm-MTV differed from mouse MTV.

Suncus murinus mammary tumor virus (Sm-MTV) is a novel retrovirus produced by an in vitro cultured cell line (Sm-MT), established from a spontaneously occurring mammary adenocarcinoma of a house musk shrew, *S. murinus* (30). Morphological, biochemical, and immunological studies have shown similarities and dissimilarities between this virus and other known retroviruses. Morphologically, mature Sm-MTV virions resemble to some extent those of type B (mouse mammary tumor virus [MMTV]) and type D (Mason-Pfizer monkey virus) retroviruses, and the occurrence of intracytoplasmic A particles is shared by all of these viruses. However, the Sm-MTV envelope is spikeless, and within it, the nucleoid is centrally located (30, 34). Sm-MTV reverse transcriptase requires Mg²⁺, as does this enzyme of the other two viruses mentioned above. This enzyme and some *gag* gene products are antigenically related to their type D retrovirus counterparts (30, 35; unpublished data). Experiments to confirm infectivity of Sm-MTV in vitro and in vivo have so far been unsuccessful, and thus a causal relationship with mammary tumorigenesis remains obscure.

S. murinus is a member of the family *Insectivora*, members of which are believed to be the most primitive of mammals, but which are much more closely related to the early primates (prosimians) than are rodents (21). For this reason, analysis of the phylogenetic relationship of Sm-MTV with other retroviruses is important, possibly providing data of more relevance to primate and human studies than can be obtained by using rodents.

This report is the first step in these studies and describes the cloning and nucleotide sequence analysis of Sm-MTV long terminal repeats (LTRs) and their transcriptional activities in comparison with other retroviruses.

MATERIALS AND METHODS

cDNA cloning of Sm-MTV. The Sm-MT cell line (30) was cultured in Dulbecco modified Eagle medium supplemented with 10% fetal calf serum, and Sm-MTV was purified from the culture medium by sucrose density gradient centrifugation, as previously described (30). cDNA was prepared by using RNA isolated from freshly harvested virions as the template. Recombinant plasmids containing cDNAs were constructed by the method of Okayama and Berg (16). Cells of competent *Escherichia coli* HB101 were transformed with these plasmids, and transformants were screened for Sm-MTV genomes by colony hybridization (9) with ³²P-labeled single-stranded Sm-MTV cDNA. Out of 100 transformants, the plasmid containing the largest DNA insert (2.7 kilobases [kb]) was chosen and designated psmMTV3.

Preparation of radioactive probes and molecular hybridization. The DNA fragments used as hybridization probes (Fig. 1, probes a, b, c, and d) were labeled with ³²P using the Amersham Multiprime DNA Labeling System (6) (Amersham, England), according to the manufacturer instructions. Specific activity of the probes was approximately 1.5 × 10⁹ dpm/μg of DNA.

In plaque, Southern, and dot hybridization analyses, hybridization and subsequent washing were performed under the following conditions. Filters were hybridized with ³²P-labeled probes (2.5 ng/ml) in a solution containing 50% formamide, 5× Denhardt solution, 5× SSPE (1× SSPE is 180 mM NaCl, 10 mM sodium phosphate, and 1 mM EDTA, pH 7.4), 0.1% sodium dodecyl sulfate (SDS), and 100 μg of salmon sperm DNA per ml for 48 h at 42°C. After hybridization, filters were washed three times in 1× SSC (0.15 M NaCl plus 0.015 M sodium citrate)–0.1% SDS at 65°C and once in 0.1× SSC–0.1% SDS at 65°C.

Molecular cloning of Sm-MTV-related provirus. High-mo-

* Corresponding author.

FIG. 1. Restriction maps of cloned cDNA (psmMTV3) and proviruses of Sm-MTV (λ sm-P-MTV9 and λ sm-P-MTV10). The relative positions of the restriction enzyme sites were determined by digestion of subcloned DNAs with the following enzymes: *Pst*I (P), *Eco*RI (E), *Xba*I (X), *Pvu*II (Pu), *Sma*I (S), *Hind*III (H), *Sal*I (SL), and *Bam*HI (B). These clones were aligned according to the common restriction enzyme sites and nucleotide sequences. The 5' ends of the genes are to the left of the figure. The LTRs and the U3.R region of psmMTV3 are boxed, and the EMBL3 arms are indicated by an open-ended bar. Shaded bars indicate Okayama-Berg vector sequences used for cDNA cloning. Probes used in this experiment are indicated: probe a (401-bp *Sma*I-*Alu*I fragment of psmMTV3 used as an LTR probe), probe b (1.5-kb *Hind*III fragment of psmMTV3), probe c (2.3-kb *Pst*I fragment of psmMTV3), and probe d (2.5-kb *Hind*III fragment of λ sm-P-MTV10). The horizontal lines with arrows (1 to 4) represent sequenced regions. (A)_n, Poly(A) tail.

molecular-weight DNA from Sm-MT cells were partially digested with *Sau*3AI and size fractionated by density gradient centrifugation, as previously described (13). DNA fragments (0.5 μ g), 15 to 22 kb long, were ligated to 1 μ g of EMBL3 arms, predigested with *Bam*HI and *Eco*RI (STRATAGENE, La Jolla, Calif.), according to the manufacturer directions. The DNA was packaged into bacteriophages by using a commercially prepared extract, according to the supplier instructions (Gigapack; STRATAGENE), and added to dish cultures of *E. coli* Q359. About 500,000 plaques were screened by the plaque hybridization technique of Benton and Davis (1), using probe c (Fig. 1).

Positive clones were plaque purified twice, and λ DNAs containing Sm-MTV related sequences were digested with *Hind*III, *Sma*I, *Sal*I, and *Bam*HI. The resulting fragments were subcloned in plasmids pUC118 or pUC119 and used for restriction mapping, Southern hybridization, and DNA sequencing. Proviruses thus cloned from the Sm-MT cell genome were designated Sm-P-MTV.

DNA sequencing. For analysis of psmMTV3, appropriate fragments were subcloned into M13mp18 or M13mp19 vectors and sequenced by the M13 dideoxy method described by Sanger (14, 22).

For analysis of Sm-P-MTV, a 2.1-kb *Sma*I fragment (λ sm-P-MTV10 5' LTR [Fig. 1]), a 2.4-kb *Sal*I-*Sma*I fragment (λ sm-P-MTV9 5' LTR), and a 2.6-kb *Sma*I-*Bam*HI fragment (λ sm-P-MTV10 3' LTR) were subcloned into either pUC118 or pUC119, plasmids which carry the M13 intergenic region (14). Deletion clones were then prepared using exonuclease III, mung bean nuclease, and Klenow enzyme, as described by Henikoff (10).

Cells of *E. coli* MV1304 transformed with these deletion plasmids were infected with M13K07, a bacteriophage designed to enable production of single-stranded plasmid DNA from pUC vectors carrying the intergenic region. Single-stranded plasmid DNAs packaged in this phage were isolated and used for dideoxy sequencing.

Construction of plasmids expressing the CAT gene. The chloramphenicol acetyltransferase (CAT) plasmids used in this experiment are shown in Fig. 2. The plasmid pSV2CAT contains the replication origin of pBR322 and the simian virus 40 early promoter, linked to the coding sequence of the bacterial CAT gene (8). In another plasmid, pSVOCAT, the entire simian virus 40 promoter region has been deleted and replaced with a unique *Hind*III restriction site (8). psm-LTR-CAT was constructed by gap filling and blunt-end ligation insertion of probe a, shown in Fig. 1, into the *Hind*III site of pSVOCAT.

In λ sm-P-MTV10, *Sma*I sites, suitable for CAT plasmid construction, existed 53 base pairs (bp) downstream from the 5' LTR and 1.4 kb upstream from it (Fig. 1). To construct psmP-LTR-CAT, the 2.1-kb *Sma*I fragment thus prepared was inserted into pSVOCAT as described above. LTRs were inserted in the sense orientation with respect to the CAT genes in psm-LTR-CAT-1 and psmP-LTR-CAT-1.

pMMTVCAT is a steroid hormone-inducible CAT plasmid which was constructed by insertion of the 1.4-kb *Hind*III fragment of pMDSG (11), containing almost the entire MMTV LTR, into the *Hind*III site of pSVOCAT in the correct orientation.

Transfections and CAT assays. T47D cells, a human mammary tumor cell line whose cells carry high constitutive levels of progesterone receptors in addition to glucocorticoid, androgen, and estradiol receptors (3, 18), were grown in RPMI 1640 medium supplemented with 10% fetal calf serum. Cells (10^6) were seeded in a 60-mm dish 1 day before transfection.

Plasmid DNA (10 μ g) was transfected into T47D cells by the DEAE-dextran method, as previously described (12). Transfected cells were incubated for 12 h in standard medium in the presence or absence of 10 nM β -estradiol. At the end of culture, the estradiol-free medium was replaced with fresh medium, with or without 10 μ M dexamethasone, while the β -estradiol-containing medium was replaced with me-

FIG. 2. Construction of plasmids expressing the CAT gene. Details of plasmid construction are given in Materials and Methods.

dium with or without 100 nM synthetic progesterone R5020 (NEN Research Products, Boston, Mass.). Incubation was continued for a further 36 h.

XC cells (rat fibroblasts) and HeLa cells were cultured in Dulbecco modified Eagle medium supplemented with 10% fetal calf serum. Cells were plated the day before transfection at 3×10^5 cells per 60-mm dish, and plasmid DNA (10 μ g) was transfected into the cells by the calcium phosphate precipitation method, as previously described (7). At 12 h after transfection, the medium was replaced with fresh medium with or without 10 μ M dexamethasone and the cells were incubated for a further 36 h.

All cells were harvested 48 h after transfection and were lysed by three cycles of freeze-thawing and ultrasonication for 15 s in 200 μ l of 0.25 M Tris hydrochloride buffer, pH 7.8. After low-speed centrifugation, 10 μ l of supernatant was assayed for CAT activity, as described by Gorman et al. (8). The assay incubation time was 1 h, and under these conditions the rate of product formation was linear.

RESULTS AND DISCUSSION

cDNA cloning of Sm-MTV. Figure 1 shows the restriction map for psmMTV3. Several DNA fragments were subcloned in M13 vectors, and the nucleotide sequence of the region extending from the *Hind*III site to the poly(A) tail (horizontal line 1 in Fig. 1) was determined. The results showed that psmMTV3 contained such essential components of the LTR as the TATAA box, the polyadenylation signal, and the polypurine tract (Fig. 3).

To identify the viral RNA species present in virus-producing Sm-MT cells, total and poly(A)⁺ RNAs were analyzed by Northern hybridization, using probe c from psmMTV3. As shown in Fig. 4A, Sm-MT cells synthesized 8.8- and 4.5-kb poly(A)⁺ RNAs hybridizable with probe c, corresponding to genomic and subgenomic viral RNAs, respectively.

Detection of endogenous virus sequences related to Sm-MTV in the *S. murinus* genome. Cellular DNAs were prepared from livers of captive wild *S. murinus* specimens and from Sm-MT cells and were subjected to Southern hybridization analysis with probe c. Figure 4B shows that 30 to 40 Sm-MTV-related bands were detected in all DNA samples, clearly indicating that Sm-MTV was endogenous. Although the captive wild *S. murinus* specimens were not inbred, many bands were common to all the animals, implying that the incorporation of Sm-MTV sequences into the germ line of the *S. murinus* genome was a relatively ancient and repeated event. On the other hand, bands unique to each sample showed that the animals are even now continuously exposed to Sm-MTV infection. Unfortunately, whether or not the spontaneous mammary adenocarcinoma-derived Sm-MT cells had acquired extra copies of the Sm-MTV genome was not clarified, since no liver DNA sample was available from the animal from which these cells had been derived. However, the number of bands in the Sm-MT cells was comparable with that detected in normal liver DNA samples of other animals, indicating that few, if any, extra copies had been acquired by the Sm-MT cells. Dot hybridization showed that in the Sm-MT cells about 50 copies of

FIG. 3. Nucleotide sequences of Sm-MTV and Sm-P-MTV LTRs. Homologous nucleotides of psmMTV3 and the 3' LTR of Sm-P-MTV10 are indicated with asterisks. In Sm-P-MTV10-5' LTR and Sm-P-MTV9-5' LTR, nucleotides identical to those of Sm-P-MTV10-3' LTR are marked with dots. The empty spaces introduced for alignment are shown by dashes. The putative TATAA box, CCAAT box, and polyadenylation signal are boxed. Bars on the psmMTV3 sequence show the core sequence of the hormone-responsive element and enhancer (see text). In psmMTV3, the *Sma*I and *Alu*I sites, used for the preparation of probe a and the construction of psm-LTR-CAT plasmid, are indicated. Abbreviations: DR, direct repeats; IR, inverted repeats; PBS, primer-binding site, PPT, polypurine tract.

Sm-MT provirus were present per haploid genome (Fig. 5), about the same as the detected number of endogenous proviruses.

Under moderately stringent conditions, in which hybridization was performed under the usual highly stringent conditions but in which subsequent washing was only in $1 \times \text{SSC}-0.1\%$ SDS at 50°C , mouse BALB/c 3T3 cell and human HeLa cell DNAs showed no reactions with probes a or d, indicating that Sm-MTV is phylogenetically far from both MMTV and human endogenous retrovirus, two viruses which do show some homology with each other (17).

Cloning of provirus related to Sm-MTV. A genomic library established from Sm-MT cells was screened for Sm-MTV-

related sequences by plaque hybridization by using probe c. Of 5×10^5 plaques screened, approximately 200 positive clones were obtained. Of the positive clones, 20 were chosen and further analyzed by using probe a. One clone, designated λ sm-P-MTV10, was found by digestion with either *Sma*I or *Hind*III to contain two LTR-hybridizing restriction fragments: 2.1- and 7.7-kb fragments after *Sma*I digestion and 4.8- and 8.0-kb fragments after *Hind*III digestion (Fig. 1). Thus, we assumed that λ sm-P-MTV10 carried a full-length Sm-MTV provirus genome. Another clone, designated λ sm-P-MTV9, was found to have a single LTR. However, this clone shared several major restriction fragments with λ sm-P-MTV10: 2.8- and 3.9-kb *Sma*I fragments

and 4.8-, 2.5-, and 1.2-kb *Hind*III fragments (Fig. 1). This implied that the provirus genome in λ sm-P-MTV10 did not have any major deletions. These two clones were thus chosen for further analyses by restriction mapping.

Figure 1 shows restriction maps for these clones compared with that of psmMTV3. Nucleotide sequence analysis revealed that Sm-P-MTV10 was a 9.4-kb provirus with 720-bp LTRs at both ends. In λ sm-P-MTV9, the 3' LTR had been deleted, since the *Sau*3AI site present in the 3' LTR was used for ligation with an EMBL3 arm.

Sequences of the LTRs and their flanking regions. The nucleotide sequences of the LTRs are presented in Fig. 5. Immediately upstream from the 3' LTRs of psmMTV3 and Sm-P-MTV10, a 22-bp-long polypurine tract was found, serving as the initiation site for synthesis of plus-strand viral DNA, and at both ends of the LTRs, an inverted repeat of 5'-TG...CA-3', commonly present in retrovirus LTRs, was also detected. Cellular direct 6-bp repeats (5'-TCAGAC-3') were identified immediately outside the Sm-P-MTV10 LTRs; from those positions it was possible to define the boundaries of the LTRs. Thus, the LTRs of Sm-P-MTV10 and of Sm-P-MTV9 were 720 bp in length, while the U3.R region of psmMTV3 was 567 bp long. The sequences in the 5' and 3'

FIG. 4. Analyses of viral RNA and provirus DNA in Sm-MT cells. (A) Northern hybridization analysis of Sm-MTV RNA. Total RNA was prepared from Sm-MT cells by the method described previously (13). Poly(A)⁺ RNA was selected by oligo(dT)-cellulose column chromatography. RNA was denatured with glyoxal and dimethyl sulfoxide and separated on a 0.8% agarose gel. RNA transfer and hybridization were done according to the method of Thomas (29). The ³²P-labeled probe c was used for hybridization. Lanes: 1, total RNA from Sm-MT cells (10 μ g); 2, poly(A)⁺ RNA (1 μ g). (B) Detection of Sm-MTV-related sequences in *S. murinus* liver DNAs. High-molecular-weight DNAs were prepared from Sm-MT cells (lane 1) and livers of captive wild *S. murinus* specimens (lanes 2 to 5) trapped on Okinawa (lane 2), Izenajima (lane 3), and Tokunoshima islands (lane 4) and at Nagasaki (lane 5) in southwest Japan. *Eco*RI-digested DNAs (10 μ g) were separated on a 0.8% agarose gel, and Southern blot analysis was done by using probe c.

FIG. 5. Estimation of number of Sm-P-MTV genes in Sm-MT cells by dot hybridization. *Eco*RI-digested DNA was spotted onto a nitrocellulose filter by using a Bio-Dot Microfiltration Apparatus (Bio-Rad Laboratories, Richmond, Calif.). DNA was denatured by heating to 100°C for 5 min, followed by immediate cooling in an ice bath, and was immobilized on a nitrocellulose filter by baking. Hybridization and washing were carried out under the stringent conditions described in Materials and Methods. Probe d was used for hybridization. Dots: 1, salmon sperm DNA (5 μ g); 2, 13 pg of probe d (equivalent to 5 copies per haploid genome) plus salmon sperm DNA (5 μ g); 3, 26 pg of probe d plus salmon sperm DNA; 4, 130 pg of probe d plus salmon sperm DNA; 5, 650 pg of probe d plus salmon sperm DNA; 6, 2 μ g of Sm-MT DNA plus 3 μ g of salmon sperm DNA; 7, 5 μ g of Sm-MT DNA. The radioactivity on dot 7 (2,800 cpm) was 2.5-fold higher than that on dot 6 (1,170 cpm), indicating that the radioactivity hybridized was proportional to the amount of Sm-MT DNA spotted. The radioactivity on dot 7 equaled that on dot 4 (2,680 cpm).

LTRs of Sm-P-MTV10 were identical and differed from that of the Sm-P-MTV9 LTR by only two bases. When the U3.R region of psmMTV3 was compared with the corresponding region of Sm-P-MTV10 LTR, the latter differed from the former by 15-bp insertions, 8-bp deletions, and 120-bp substitutions. Thus the overall homology of this region was 75%.

Immediately downstream from the 5' LTRs of Sm-P-MTV10 and -9, a primer-binding site, involved in the synthesis of minus-strand viral DNA, was identified. Of the 18 bases at this site, 15 (5'-TGGCGCCCAACGTGCGTG-3') were found to be complementary to the last 18 nucleotides of tRNA_{Lys}³ (20). In this sense, Sm-MTV is similar to Mason-Pfizer monkey virus (28), the human endogenous retrovirus HERV-K (17), and visna virus (27), viruses which all use the same tRNA as a primer, but different from MMTV (19) and human immunodeficiency virus (31), both of which use tRNA_{Lys}³.

The end of the R region is always defined by the dinucleotide CA, 10 to 20 bp downstream from the polyadenylation signal AATAAA. In all clones, this latter sequence was present at positions 560 to 565 (numbered positions refer to those in Fig. 5), and CA was found 15 bp downstream. The U3.R region of the Sm-P-MTV LTR was therefore 574 bp long, while the remaining 146-bp sequence constituted the U5 region.

The R region always starts with a G nucleotide positioned about 30 bp downstream from the TATAA box. In psmMTV3, a TATAATT sequence, a typical TATAA box, was found at positions 448 to 454; so if a G at position 477 is the correct cap site, then the U3 and R regions of psmMTV3 were 462 and 105 bp long, respectively. No CCAAT box was found in this LTR. Unexpectedly, the positions of the essential promoter elements were found to be markedly different in Sm-P-MTV. Four putative TATAA boxes, or their homologs, were identified at positions 286 to 291, 325 to 329, 419 to 424, and 528 and 532. A CCAAT box was found

42 bp upstream from the second TATAA box, a distance which is common in other eucaryotic promoters. Another CCAAT was identified at positions 449 to 454, which surprisingly corresponded to the position of the TATAATT box in psmMTV3. Thus, the structure of the promoter region in Sm-P-MTV is complicated, leaving unambiguous identification of the cap site undetermined.

The LTRs of Sm-MTV and Sm-P-MTV had neither direct repeat enhancer sequences nor the consensus sequence 5'-GGG_{TA}CG_{TAAT}GGG-3' for Sp1 recognition (2). However, two types of interesting sequences were found in psmMTV3. First, 8 bp upstream from the TATAATT box was a sequence, 5'-TGTTCT-3', known to be the core sequence for a steroid hormone-responsive element found in the MMTV LTR and in other glucocorticoid-inducible promoters. Second, 62 bp upstream from the TATAATT box, a 14-bp sequence, 5'-TGTTCTGTGGTCAG-3', was found, of which the first 6 bp are again the hormone-responsive element core, while the remaining GTGGTCAG is highly homologous with the enhancer core sequence 5'-GTGG_{TTT}AAAG-3' (32). This 14-bp sequence is also homologous with the sequence 5'-TGTTATGTGCTAAG-3' present in the LTR of the human endogenous retrovirus HERV-K (17). Recently, Ono et al. (18) reported that in human breast cancer cell line T47D, the expression of the HERV-K genome could be stimulated by progesterone after estradiol treatment. This led us to assess by CAT assay the transcriptional activities and the hormone responsiveness of Sm-MTV LTRs.

Transcriptional activity of LTR. Figure 6 shows that, in the absence of hormones, psm-LTR-CAT-1 and psmP-LTR-CAT-1, in which the LTRs had been inserted in the same orientation as the CAT gene, were transcriptionally active (lanes 8, 10, 14, and 16). From the percent conversion of chloramphenicol, they were calculated to be half as active as the simian virus 40 promoter (lanes 4, 8, and 10). These data are consistent with the speculation that the U3 of Sm-MTV is responsible for viral RNA transcription. It is interesting that psmP-LTR-CAT-1 showed transcriptional activity similar to or even higher than that achieved by psm-LTR-CAT-

1, even though the structure of the promoter region in Sm-P-MTV profoundly differed from that in Sm-MTV. In contrast, psm-LTR-CAT-2 and psmP-LTR-CAT-2, in which the LTRs had been inserted in the opposite direction, did not show any CAT activity (lanes 1 and 2); pSVOCAT was also inactive (lane 3).

Cato et al. (3) showed that CAT activity directed by MMTV LTR was inducible by dexamethasone or R5020 in T47D cells, a finding which was confirmed by the present study (lanes 6, 7, 12, and 13). On the other hand, neither hormone had any stimulatory effects on the promoter activities of Sm-MTV and Sm-P-MTV, in spite of the presence in the Sm-MTV LTR of the hormone-responsive element core sequence and of the 14-bp sequence described above. The same results were obtained in XC cells and HeLa cells (data not shown). These data are consistent with our previous observation that 1 μ M dexamethasone had no effect on the production of Sm-MTV by Sm-MT cells. The presence of CAT mRNA, directed by Sm-P-MTV LTR, will allow us to determine the cap site in the Sm-P-MTV genome by primer extension experiments, using oligodeoxynucleotides complementary to the 5' end of the CAT gene.

Nucleotide sequence analysis of the putative *env* gene. In addition to the LTRs, we sequenced the 3' portion of the putative *env* gene, which in psmMTV3 was the region extending to the *Hind*III site present 1,123 bp upstream from the 3' LTR (horizontal line 1 in Fig. 1) and in λ sm-P-MTV10 was the sequence extending to the *Sma*I site present 973 bp upstream from its 3' LTR (horizontal line 2 in Fig. 1). Comparison of the two resulting sequences showed that nucleotide sequence homology over this region was 93%, significantly higher than the homology over the U3.R regions of the LTR (75%). The differences which were found consisted in Sm-P-MTV10 of a 4-bp insertion, a 3-bp deletion, and a 64-bp substitution. An open reading frame extending from 108 bp upstream of the LTR to the end of the sequenced region (*Hind*III site) was identified in psmMTV3. In Sm-P-MTV10, on the other hand, nucleotide changes resulted in frame shifts and possibly premature termination.

FIG. 6. CAT activity directed by LTRs in T47D cells. The plasmids shown in Fig. 2 were used for transfection. Cm and AcCm represent chloramphenicol and acetylated chloramphenicol, respectively. T47D cells were transfected with psm-LTR-CAT-2 (lane 1), psmP-LTR-CAT-2 (lane 2), pSVOCAT (lane 3), pSV2CAT (lanes 4 and 5), pMMTVCAT (lanes 6, 7, 12, and 13), psm-LTR-CAT-1 (lanes 8, 9, 14, and 15), and psmP-LTR-CAT-1 (lanes 10, 11, 16, and 17). Transfected cells were either treated with 10 μ M dexamethasone (Dex) for 36 h (lanes 5, 7, 9, and 11), or pretreated with 10 nM β -estradiol for 12 h (lanes 12 to 17) followed by 100 nM R5020 treatment for 36 h (lanes 13, 15, and 17). At 48 h after transfection, cells were harvested and the CAT assay was performed as described in Materials and Methods. CAT activity was expressed as the percent conversion of [14 C]chloramphenicol to acetylated derivatives per hour.

Sm-MTV VASLAENVLQNRRLNLAPMKEGGICAAALKECCIFKDETGLVRDSIKRVNSLEERRRE-ID
 * * * * *
 MPMV VDSLAEVVLQNRRLDLLTAEQGGICLALQEKCCFYANKSGIVRDKIKNIQDDLERRRRLID
 * * * * *
 Mo-MLV L TSLSEVVLQNRRLDLLFLKEGGLCAALKECCFYADHTGLVRDSMAKLRERLNQRQLPES
 * * * * *
 REV-A IDSLAEVVLQNRRLDLLTAEQGGICLALQEKCCFYANKSGIVRDKIKNIQDDLARRKALYD
 * * * * *
 FeLV L TSLSEVVLQNRRLDILFLQEGGLCAALKECCFYADHTGLVRDNMAKLRERLQKQQLFDS
 * * * * *
 RSV VTSIRHAVLQNRRAIDFLLLAHGHCEDVAGCCPNLSDHSESIQKFKQLMKKHVNIIGVDS
 * * * * *
 HTLV-1 LLKIAQYAAQNRRLDLLFVEQGGICLALQEKCCFYANKSGIVRDKIKNIQDDLARRKALYD
 * * * * *
 HTLV-2 ILRVAQYAAQNRRLDLLFVEQGGICLALQEKCCFYANKSGIVRDKIKNIQDDLARRKALYD

FIG. 7. Alignment of the immunosuppressive peptides of a number of retroviral transmembrane protein sequences. The sequence in Sm-MTV described here was found 158 to 95 residues upstream from the carboxy terminus of the putative *env* gene product. Amino acids in the immunosuppressive peptides of other viruses which are identical with those in the Sm-MTV peptide are marked with an asterisk. The other retroviruses cited are as follows: MPMV, Mason-Pfizer monkey virus (28); Mo-MLV, Moloney murine leukemia virus (26); REV-A, reticuloendotheliosis-associated virus (33); FeLV, feline leukemia virus (15); RSV, Rous sarcoma virus (23); HTLV-1 and HTLV-2, human T-cell leukemia viruses types 1 and 2 (24, 25).

The Sm-MTV open reading frame corresponded to 338 amino acids and seemed to encode the 3' portion of the *env* gene transmembrane protein.

It is interesting that the immunosuppressive peptide sequence common to many retroviral transmembrane proteins, but not to MMTV (4, 5), was found in this region (Fig. 7). In Sm-MTV, its amino acid sequence (Fig. 7) started 158 residues upstream of the carboxy terminus of the open reading frame. The same sequences found in Mason-Pfizer monkey virus and Moloney murine leukemia virus are highly homologous with that of Sm-MTV (36 amino acids identified out of a total of 63 amino acids; 57% homology). However, antigens common to Sm-MTV and Mason-Pfizer monkey virus have not been found in the *env* gene products but in those of the *gag* gene (unpublished data).

At residue positions 227, 287, 299, and 312 upstream from the carboxy terminus of the open reading frame, potential N-glycosylation sites have been detected, further confirming that this open reading frame constitutes the 3' portion of the gene encoding the transmembrane protein.

A better understanding of the structure of Sm-MTV requires complete sequencing of the provirus.

ACKNOWLEDGMENT

This work was partly supported by a grant-in-aid for Cancer Research, 62015038, from the Ministry of Education, Science, and Culture of Japan.

LITERATURE CITED

- Benton, W. D., and R. W. Davis. 1977. Screening λ gt recombinant clones by hybridization to single plaques *in situ*. *Science* 196:180-182.
- Briggs, M. R., J. T. Kadonaga, S. P. Bell, and R. Tjian. 1986. Purification and biochemical characterization of the promoter-specific transcription factor, Sp1. *Science* 234:47-52.
- Cato, A. C. B., R. Miksicek, G. Schutz, J. Arnemann, and M. Beato. 1986. The hormone regulatory element of mouse mammary tumour virus mediates progesterone induction. *EMBO J.* 5:2237-2240.
- Cianciolo, G. J., T. D. Copeland, S. Oroszlan, and R. Snyderman. 1985. Inhibition of lymphocyte proliferation by a synthetic peptide homologous to retroviral envelope proteins. *Science* 230:453-455.
- Copeland, T. D., W. P. Tsai, and S. Oroszlan. 1985. Antibody to a synthetic peptide detects a conserved region of retrovirus transmembrane proteins. *Biochem. Biophys. Res. Commun.* 126:672-677.
- Feinberg, A. P., and B. Vogelstein. 1983. A technique for radiolabeling DNA restriction endonuclease fragments to high specific activity. *Anal. Biochem.* 132:6-13.
- Gorman, C. M., G. T. Merlino, M. C. Willingham, I. Pastan, and B. H. Howard. 1982. The Rous sarcoma virus long terminal repeat is a strong promoter when introduced into a variety of eucaryotic cells by DNA-mediated transfection. *Proc. Natl. Acad. Sci. USA* 79:6777-6781.
- 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.
- Hanahan, D., and M. Meselson. 1980. Plasmid screening at high colony density. *Gene* 10:63-67.
- Henikoff, S. 1984. Unidirectional digestion with exonuclease III creates targeted breakpoints for DNA sequencing. *Gene* 28:351-359.
- Lee, F., R. Mulligan, P. Berg, and G. Ringold. 1981. Glucocorticoids regulate expression of dihydrofolate reductase cDNA in mouse mammary tumour virus chimaeric plasmids. *Nature (London)* 294:228-232.
- Lopata, M. A., D. W. Cleveland, and B. Sollner-Webb. 1984. High level transient expression of a chloramphenicol acetyltransferase gene by DEAE-dextran mediated DNA transfection coupled with a dimethyl sulfoxide or glycerol shock treatment. *Nucleic Acids Res.* 12:5707-5717.
- Maniatis, T., E. F. Fritsch, and J. Sambrook. 1982. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
- Messing, J., and J. Vieira. 1982. A new pair of M13 vectors for selecting either DNA strand of double-digested restriction fragments. *Gene* 19:269-276.
- Nunberg, J. H., M. E. Williams, and M. A. Innis. 1984. Nucleotide sequences of the envelope genes of two isolates of feline leukemia virus subgroup. *B. J. Virol.* 49:629-632.
- Okayama, H., and P. Berg. 1982. High-efficiency cloning of full-length cDNA. *Mol. Cell. Biol.* 2:161-170.
- Ono, M. 1986. Molecular cloning and long terminal repeat sequences of human endogenous retrovirus genes related to types A and B retrovirus genes. *J. Virol.* 58:937-944.
- Ono, M., M. Kawakami, and H. Ushikubo. 1987. Stimulation of expression of the human endogenous retrovirus genome by female steroid hormones in human breast cancer cell line T47D. *J. Virol.* 61:2059-2062.
- Peters, G., and C. Glover. 1980. tRNA's and priming of RNA-directed DNA synthesis in mouse mammary tumor virus. *J. Virol.* 35:31-40.
- Raba, M., K. L. Limberg, M. Burghagen, J. R. Katze, M. Simsek, J. E. Heckman, U. L. Rajbhandary, and H. J. Gross. 1979. Nucleotide sequence of three isoaccepting lysine tRNAs from rabbit liver and SV40 transformed mouse fibroblast. *Eur. J. Biochem.* 97:305-318.
- Romer, A. S., and T. S. Parsons. 1977. The vertebrate body, 5th ed., p. 76-91. The W. B. Saunders Co., Philadelphia.
- Sanger, F., S. Nicklen, and A. R. Coulson. 1977. DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. USA* 74:5463-5467.
- Schwartz, D. E., R. Tizard, and W. Gilbert. 1983. Nucleotide sequence of Rous sarcoma virus. *Cell* 32:853-869.
- Seiki, M., S. Hattori, Y. Hirayama, and M. Yoshida. 1983. Human adult T-cell leukemia virus: complete nucleotide sequence of the provirus genome integrated in leukemia cell DNA. *Proc. Natl. Acad. Sci. USA* 80:3618-3622.
- Shimotohno, K., Y. Takahashi, N. Shimizu, T. Gojobori, D. W.

- Golde, I. S. Y. Chen, M. Miwa, and T. Sugimura. 1985. Complete nucleotide sequence of an infectious clone of human T-cell leukemia virus type II: an open reading frame for the protease gene. *Proc. Natl. Acad. Sci. USA* **82**:3101-3105.
26. Shinnick, T. M., R. A. Lerner, and J. G. Sutcliffe. 1981. Nucleotide sequence of Moloney murine leukemia virus. *Nature (London)* **293**:543-548.
 27. Sonigo, P., M. Alizon, K. Staskus, D. Klatzmman, S. Cole, O. Danos, E. Retzel, P. Tiollais, A. Haase, and S. Wain-Hobson. 1985. Nucleotide sequence of the visna lentivirus: relationship to AIDS virus. *Cell* **42**:369-382.
 28. Sonigo, P., C. Barker, E. Hunter, and S. Wain-Hobson. 1986. Nucleotide sequence of Mason-Pfizer monkey virus: an immunosuppressive D-type retrovirus. *Cell* **45**:375-385.
 29. Thomas, P. S. 1980. Hybridization of denaturated RNA and small DNA fragments transferred to nitrocellulose. *Proc. Natl. Acad. Sci. USA* **77**:5201-5205.
 30. Tsutsui, Y., S. Nagayoshi, S. Saga, M. Takahashi, A. Aoyama, J. Yamashiro-Malavasi, S. Hayakawa, T. Yokoi, S. Oda, and M. Hoshino. 1985. A new retrovirus produced by tissue culture cell line from mammary tumor of a house musk shrew, *Suncus murinus*. *Virology* **144**:273-278.
 31. Wain-Hobson, S., P. Sonigo, O. Danos, S. Cole, and M. Alizon. 1985. Nucleotide sequence of the AIDS virus, LAV. *Cell* **40**:9-17.
 32. Weiher, H., M. Konig, and P. Gruss. 1983. Multiple point mutations affecting the simian virus 40 enhancer. *Science* **219**:626-631.
 33. Wilhelmssen, K. C., K. Eggleton, and H. Temin. 1984. Nucleic acid sequences of the oncogene *v-rel* in reticuloendotheliosis virus strain T and its cellular homolog, the proto-oncogene *c-rel*. *J. Virol.* **52**:172-182.
 34. Yamashiro-Malavasi, J., Y. Tsutsui, S. Nagayoshi, S. Saga, M. Takahashi, A. Aoyama, S. Hayakawa, T. Yokoi, S. Oda, A. Yokoyama, and M. Hoshino. 1986. Ultrastructural characterization of the retrovirus particles (Sm-MTV) liberated from the mammary tumor cell line (Sm-MT) of a house musk shrew, *Suncus murinus* (*Insectivora*). *Cancer Res.* **46**:3128-3137.
 35. Yokoi, T., H. Nakane, S. Hayakawa, S. Saga, S. Nagayoshi, Y. Tsutsui, A. Aoyama, M. Hoshino, and K. Ono. 1986. Biochemical and immunological characterization of *Suncus murinus* mammary tumor virus DNA polymerase. *Virology* **148**:268-274.