The Simian Foamy Virus Type 1 Transcriptional Transactivator (Tas) Binds and Activates an Enhancer Element in the *gag* Gene

MEL CAMPBELL, CELESTE ENG, AND PAUL A. LUCIW*

Department of Medical Pathology, University of California, Davis, California 95616

Received 26 April 1996/Accepted 26 June 1996

Simian and human foamy viruses (SFV and HFV) encode a transcriptional transactivator, Tas, which governs the levels of viral transcripts initiated by both the promoter in the long terminal repeat (LTR) and the internal promoter (IP) located within the *env* gene of these viruses. Tas-responsive target elements, TRE^{LTR} in the LTR and TRE^{IP} in the *env* gene, are located 5' of the TATA box in both viral promoters and function as orientation- and position-independent enhancers. We have identified a strong Tas-responsive element, designated TRE^{GP}, near the 3' end of the *gag* gene and preceding the *pol* gene of SFV-1. In transient-expression assays with plasmids containing reporter genes, a 59-bp DNA fragment containing TRE^{GP} (nucleotides 2224 to 2282) functioned as an enhancer element, dependent on Tas, in several cell types and in the context of a heterologous basal promoter. DNase footprinting revealed that the fusion protein glutathione S-transferase–Tas, purified from genetically engineered bacteria, interacts with about 40 bp (nucleotides 2237 to 2279) in the TRE^{GP}. A low degree of sequence homology was noted between TRE^{GP} and TRE^{IP}. In virus-infected cells, novel transcripts with 5' ends immediately upstream from the reverse transcriptase translation frame (nucleotides 2611 to 5778) were identified. Upstream of the start site for these transcripts is a TATA box (nucleotides 2575 to 2579), which was required for transcription in transient-expression assays. Although a spliced mRNA initiated in the viral LTR is implicated in the synthesis of the HFV Pol polyprotein which encodes protease, reverse transcriptase, and integrase, it is possible that SFV-1 contains a promoter within the *pol* gene for initiating a reverse transcriptase transcript. Taken together, these studies define a novel Tas-responsive enhancer element, which binds the viral transactivator, and a potential promoter within the *pol* gene.

Foamy viruses, members of the *Spumavirus* genus of retroviruses, have complex genomes encoding virion structural genes, a transcriptional activator, and one or two additional open reading frames (ORFs) of undefined function (reviewed in reference 29). Both human foamy virus (HFV) and the simian foamy viruses (SFVs) contain two promoters for initiation of viral transcripts (3, 19, 22). One promoter is located in the 5' long terminal repeat (LTR) and produces a full-length viral transcript, which is the viral genome and also the precursor to spliced subgenomic mRNAs. Additionally, an internal promoter (IP) is embedded in sequences in the 3' portion of the envelope gene and directs transcription of the viral transactivator and the ORF(s) (18). The transactivator has been designated *taf* for SFV and *bel-1* for HFV. However, at the First International Foamy Virus Meeting (November 1994), it was recommended that *taf* and *bel-1* be designated *tas* (transactivator of spumavirus), and it is henceforth referred to as *tas-1* for SFV-1.

Analysis of *cis*-acting regulatory sequences for transcription of both simple and complex retroviruses has defined sequence domains in the 5' LTR as key elements in the regulation of initiation of viral RNA synthesis (5). However, in recent years, it has been demonstrated that *cis*-acting transcriptional control elements of retroviruses are not limited solely to the LTR (reviewed in reference 32). Regulatory elements outside of the LTR have been identified in the genomes of avian retroviruses (1, 4), murine leukemia virus (20), human immunodeficiency virus type 1 (7, 30, 31), human T-cell lymphotropic virus type 1 (28), mouse mammary tumor virus (8, 26), and bovine leukemia virus (33). The significance of such regulatory elements for

virus replication is not fully understood. Foamy viruses are also an example of a group of retroviruses in which viral sequence elements outside of the LTR are important for transcriptional regulation; Tas-1 DNA target elements designated Tas-1-responsive elements (TREs) have been defined in the IP in the *env* gene (TRE^{IP}) as well as the LTR (TRE^{LTR}) (3, 35; reviewed in reference 29). Because HFV and SFV-1 contained the intragenic TRE^{IP} and because of the identification of intragenic enhancer elements in other retroviral genomes, we asked whether the SFV-1 genome encoded additional intragenic TREs. This report describes the characterization of a novel SFV-1 TRE, which maps to the extreme 3' end of the *gag* gene, immediately 5' of the *gagpol* overlap region. In transient-expression assays, this TRE, designated the TRE of *gagpol* (TRE^{GP}), exhibited a level of Tas-1 responsiveness that was equivalent to that observed with reporter gene constructs containing the TRE^{LTR} or TRE^{IP}. As previously shown for the TRE^{IP} (35), the TRE^{GP} bound the fusion protein glutathione S-transferase (GST)–Tas-1. Additionally, we obtained evidence that the TRE^{GP} could be an upstream element for an additional intragenic promoter which produces transcripts in SFV-1-infected cells that initiate within the *pol* gene.

MATERIALS AND METHODS

Cell culture and transfections. NIH 3T3 (murine fibroblast), Cf2Th (canine thymus fibroblast), and Cos-7 (simian virus 40 [SV40]-transformed African green monkey fibroblast) cell lines were obtained from the American Type Culture Collection, Rockville, Md. The cells were grown in Dulbecco's modified Eagle's medium supplemented with 10% fetal bovine serum. For transfections into all three cell lines, 1 µg of reporter plasmid expressing chloramphenicol acetyltransferase (CAT) and 2 µg of either pSP65 (control) or the effector plasmid pSV-tas-1 (formerly pSV-taf-1) were mixed with the cationic lipid DOTAP (Boehringer Mannheim, Indianapolis, Ind.) and applied to 3 × 10⁵ cells in a 35-mm-diameter plate. At 48 h after transfection, cell extracts were prepared and assayed for CAT activity by measuring the conversion of [³H]acetyl coenzyme A

* Corresponding author. Phone: (916) 752-3430. Fax: (916) 752-4548.

to [³H]acetylchloramphenicol (23, 27). All transfections were performed in duplicate, and the variation was generally less than 15%.

Plasmid constructions. Plasmid p5'SFV-1 (24) served as the parent clone for all recombinant plasmids constructed in this study. p5'SFV-1 contains a 1.2-kb *Bam*HI-*Bam*HI (nucleotides [nt] 1025 to 2334) DNA fragment from the SFV-1 proviral clone (GenBank accession number X58484). All subsequent constructs, with the exception of pGP-15 and pGP-20 (Fig. 1), were PCR-amplified subregions contained within the *Bam*HI fragment of p5'SFV-1. The PCR-amplified fragments were synthesized with primer pairs that annealed to the SFV-1 template defined by the endpoints given in Fig. 1. Primers were synthetic oligonucleotides, 28 to 36 bases in length; a recognition site for *Bam*HI (5'CGGGATCCGG3') or *Asp* 718 (5'CGGGTACCCG3') was incorporated at the 5' end of each oligonucleotide for cloning purposes (underlined). Oligonucleotides were synthesized with a gene assembler (Pharmacia, Piscataway, N.J.), and the PCR conditions were as described previously (3). To test for enhancer activity, amplified viral DNA fragments were purified through agarose gels, digested with *Bam*HI, and inserted into the *Bgl*II site of the enhancerless reporter plasmid pCAT-promoter (Promega, Madison, Wis.). pCAT promoter contains the SV40 promoter upstream of the CAT reporter gene but lacks the SV40 72-bp repeats; these repeats are the enhancer element for this virus. The *Bgl*II site is 5' of the SV40 promoter and CAT. The PCR-amplified DNA fragments representing portions of SFV-1 DNA were digested with *Asp* 718 and *Bam*HI and cloned into the promoterless CAT construct pCAT-22A2s as previously described (3). The replacement constructs (see Table 1) were produced in two steps. For the minimal internal promoter construct pSIP-5(-E) (-E, minus enhancer), SFV-1 DNA encompassing nt 8374 to 8420 was inserted into pCAT-22A2s as an *Asp* 718-*Bam*HI fragment. The inserted fragment contained an engineered *Xho*I site 5' of SFV-1 nt 8374 for subsequent insertion, in either orientation, of GP-22 as an *Xho*I fragment, thus generating pR(S) or (AS)-SIP-5(-E). Similarly, to generate a minimal LTR-derived CAT reporter construct, SFV-1 DNA encompassing nt 11318 to 44 (-35 to +44) was inserted into pCAT-22A2s as an *Asp* 718-*Bam*HI fragment to produce plasmid pL(-E). GP-22 was then inserted into the *Asp* 718 site of pL(-E) as an *Asp* 718 fragment, in either orientation, to generate plasmid pR(S) or (AS)-L(-E). Clustered-site mutants of GP-22 (M1 to M5) were constructed by inserting mutated oligonucleotides into pCAT-promoter as described above. All mutations were G-to-T and A-to-C transversions. Mutants M1 to M3 contained point mutations covering the 5' end, middle, or 3' end of GP-22, respectively (see Fig. 2). Mutants M4 and M5 were constructed on the basis of the DNA alignment of TRE^{GP} and TRE^{IP} shown in Fig. 5B. Plasmid pGP-24 was constructed by inserting a PCR fragment (SFV-1 nt 2200 to 2611) into pCAT-22A2s as an *Asp* 718 fragment. pGP-24-M1 contains a 5-nt mutation designed to disrupt the putative TATA box contained at the 3' end of GP-24 (SFV-1 nt 2575 to 2579; 5' TATAT 3' to 5' CCCCC 3'). This mutation was introduced through the AS oligonucleotide in a PCR with the same sense primer as used in the construction of pGP-24. pGP-24-M2 was constructed by utilizing a sense primer which introduced the same mutations present in pGP-22 M2 into pGP-24. Finally, plasmid pGP-24-M2/M1, which contains both the pGP-22M2 mutations in the TRE^{GP} and a mutated TATA box, was constructed by using a PCR with both mutagenic primers described above. The PCR fragment was cloned into pCAT-22A2s as described for pGP-24. All inserts were verified by DNA sequencing. Plasmids pSV-tas-1, pSIP-6, pSIP-5, and pL-CAT-41 have been described previously (3, 25).

RNA extraction, primer extension, and analysis by 5' RACE. Poly(A)⁺ RNA was prepared with Oligotex resin (Qiagen, Chatsworth, Calif.) under the conditions recommended by the manufacturer. Primer extensions were performed as previously described (3), except that poly(A)⁺ RNA was used in place of total RNA. The oligonucleotide used in the primer extension analysis (see Fig. 6C) was SFV 713AS (SFV-1 nt 2700 to 2677), 5' CTGCTCTTTGGCAAGGCTG TCTG 3'. 5' rapid amplification of cDNA ends (5' RACE) analysis of SFV-1 infected-cell RNA was performed as follows. cDNA was produced by reverse transcription of three 1-μg aliquots of poly(A)⁺ RNA from SFV-1-infected Cf2Th cells. Each of the three separate reactions, which were performed in parallel, used a primer designed to detect transcripts initiating in the LTR (SFV 473AS, SFV-1 nt 473 to 448), 5'TTGACATCTAAATCACCTTCTATTG3'; internal promoter (SFV 8922AS nt 8922 to 8900), 5'CGAAGGATTGGTAC CCAATCTC3'; or the *gagpol* region (SFV 3002 AS nt 3002 to 2975), 5' ATTT TATTGACCTCTCTATAGTCCAAC 3'. Following cDNA synthesis, oligomeric C tails were added to the purified cDNAs in a reaction with dCTP and terminal transferase. The tailed cDNAs were subjected to two (LTR and IP) or three (*gagpol*) rounds of 30 cycles of PCR, with each subsequent PCR amplification using nested primer pairs. The first round of PCR utilized the anchor primer (Gibco BRL, Gaithersburg, Md.) 5'CUACUACUACUAGGCCACGCGTCCGCTAGTACGGGIGGGIIGGGIIG3' and the AS primers (SFV425ASU nt 425 to 401) 5'CAUCAUCAUCAUGAATTATACCTTCCGCAGTGCACG3' (LTR), (SFV8806ASU nt 8806 to 8783) 5' CAUCAUCAUGATTTTCTTCA ATTTTTCATCTCCAG3' (IP), or (SFV2800ASU nt 2800 to 2774) 5'CAUCAUCAUTTGCAATGTTATGTGGTTTAATTCTCC3' (*gagpol*). Then 2 μl of the first-round PCR product was reamplified with the universal primer (Gibco BRL) 5'CUACUACUACUAGGCCACGCGTCCGCTAGTAC3' and the nested AS primers (SFV363ASU nt 363 to 339) 5' CAUCAUCAUATCC GCACTTATATTCGAGCCCCACG3' (LTR), (SFV 8781ASU nt 8781 to 8760) 5'CAUCAUCAUATTTTCTCCACAAAGAACCTTAG3' (IP), or (SFV

2773 ASU nt 2773 to 2448) 5'CAUCAUCAUTATGTCCTACTTGATT TCCCAATGC3' (*gagpol*). For the *gagpol* region, 2 μl of the second-round PCR was further reamplified with the universal primer and an additional nested AS primer (SFV2747ASU nt 2747 to 2725) 5'CAUCAUCAUTGCCAAAGT GCATCATACTTTA3'. Following amplification, the 5' RACE products were purified by electrophoresis on agarose gels, treated with uracil DNA glycosylase, and cloned into plasmid pAMP-1 under the conditions recommended by the manufacturer (Gibco BRL). The homopolymeric dCTP tail/SFV-1 cDNA junction was subsequently determined by DNA sequence analysis with the M13 upstream universal primer.

EMSA. Electrophoretic mobility shift assay (EMSA) was performed with double-stranded ³²P-end-labelled oligonucleotides, which had been labelled in a reaction with T4 polynucleotide kinase and [^γ-³²P]ATP. The binding-reaction conditions were as described previously (35), except that each reaction mixture contained 50,000 to 100,000 cpm of probe and 200 ng of either the fusion protein GST-Tas-1 or GST in a final reaction volume of 45 μl and double-stranded oligonucleotides representing the IP (nt 8325 to 8375) or wild-type or mutant *gagpol* region (nt 2224 to 2282). Binding was allowed to proceed for 60 min at room temperature, and then half of the reaction mixture was loaded on 4.5% nondenaturing polyacrylamide gels and electrophoresed to resolve free from bound probe. The gels were dried and exposed to X-ray film at -70°C. Competitor oligonucleotides were added at 35- and 350-fold molar excess prior to addition of the labelled probe. The nonspecific competitor shown in Fig. 4A is an oligonucleotide containing an SP-1-binding site (5'ATTCGATCGGGGCGGG GCGAGC3').

DNase I footprinting. Footprinting was performed as described previously (35) with the following ³²P-single-end-labelled probes. Plasmid GP-23 (SFV-1 nt 2175 to 2283 [not shown in Fig. 1]) was first digested with *Asp* 718 and *Sal*I. The *Asp* 718-*Sal*I fragment was gel purified, filled in with Klenow fragment in a reaction involving [^α-³²P]dGTP and cold dCTP, dTTP, and dATP, and subsequently digested a second time with either *Sma*I (to yield a probe with the sense strand labelled) or *Xba*I (to yield a probe with the antisense strand labelled). Following gel purification, 50,000 cpm of probe was incubated with either 1 or 4 μg of GST-Tas-1 at room temperature for 60 min. The reaction mixture was then treated with DNase I (Worthington, Freehold, N.J.), and the reaction products were phenol extracted, ethanol precipitated, washed twice with 70% ethanol, dried, and electrophoresed on 8% acrylamide gels under denaturing conditions. The gels were dried and exposed to X-ray film at -70°C.

RESULTS

TRE in the SFV-1 *gag* gene. Our laboratory has previously reported on the localization of TRE in the SFV-1 LTR (23) and IP (3, 22). Both the promoter-proximal TRE^{LTR} and TRE^{IP} behaved as conditional enhancer elements dependent on Tas-1. In this study, we screened portions of the SFV-1 *gag* gene for the presence of additional TRE by using a strategy designed to detect Tas-1-dependent enhancer activity. Portions of the SFV-1 *gag* gene were inserted into the enhancerless plasmid pCAT promoter, and the Tas-1 responsiveness of these constructs was examined in transient transfections in several tissue culture cell lines. Beginning with a 1.2-kb *Bam*HI fragment (SFV-1 nt 1125 to 2334; pGP-1 [Fig. 1]), an additional intragenic TRE was identified. Deletion analysis localized this element to a 59-bp region defined by pGP-22 (SFV-1 nt 2224 to 2282 [Fig. 1]). This region maps to the extreme 3' end of the SFV-1 *gag* gene and thus is located immediately 5' of the 86-nt *gagpol* overlap region (nt 2296 to 2381). This element was designated TRE^{GP}. The level of Tas-1 responsiveness of the TRE^{GP} was similar in magnitude to that of the TRE^{IP} (Fig. 1, pSIP-6), and the TRE^{LTR} (23). As with the other SFV-1 TRE, the TRE^{GP} functioned in either orientation and was absolutely dependent on the viral transactivator, Tas-1, for activity. Additionally, the TRE^{GP} exhibited enhancer properties when cloned in either orientation into the *Bam*HI site located 3' of the CAT reporter gene in the pCAT promoter (data not shown).

We next examined a series of clustered-site mutations in the TRE^{GP} to define nucleotides within the 59-bp element that are required for transactivation by Tas-1 in transient-expression assays (Fig. 2). Mutants M1, M2, and M3 targeted either the 5' one-third, middle one-third, or 3' one-third of GP-22. Mutants M4 and M5 were designed on the basis of alignment of the

FIG. 1. Schematic presentation of the SFV-1 genome and viral plasmids evaluated in this study. Arrows indicate the transcriptional initiation site in the 5' LTR and IP. Solid lines beneath the genome represent SFV-1 sequences in reporter plasmids. Values in the -Tas-1 (cotransfection of reporter plasmid and pSP65) and +Tas-1 (cotransfection of reporter plasmid and pSV-Tas-1) columns represent the mean counts per minute of duplicate transfected cultures. Fold is the relative activity in the cotransfections with and without Tas-1. Sense and antisense refer to the orientation of the SFV-1 insert relative to the CAT gene.

TRE^{GP} with the TRE^{IP} shown in Fig. 5B. Mutant M4 disrupts the 6-nt region (SFV-1 nt 2264 to 2269 in *gag* and nt 8349 to 8354 in *env* [5'TGGAAA3']), which aligns in Fig. 5B and resides within the region previously reported to be protected from DNase I cleavage by binding to GST-Tas-1 (35). Mutant M5 targets the two nucleotides that align (SFV-1 nt 2261/2263 in *gag* and nt 8339/8341 in *env* [5'TxG3']). The two nucleotide changes represented in M5 are contained within a previously described mutation in the promoter-proximal TRE^{IP}; this mutation in TRE^{IP} eliminated GST-Tas-1 binding in vitro as measured by EMSA (35). All mutations were A-to-C or G-to-T transversions. As shown in Fig. 2, each clustered-site mutant in the TRE^{GP} displayed reduced Tas-1 responsiveness, as measured by CAT activity; mutants M2, M3, and M4 all exhibited baseline Tas-1 responsiveness, and mutants M1 and M5 retained a low level of residual activity. The CAT activity in transient transfections with these mutant constructs correlated well with the ability of GST-Tas-1 to bind to an oligonucleotide representing the mutant by EMSA analysis (see Fig. 3).

TRE^{GP} enhances transcription of the SFV-1 LTR and IP. The ability of the TRE^{GP} to substitute for the TRE^{LTR} and TRE^{IP} was examined by using reporter constructs that lacked any of the previously described TRE and contained only a TATA box as an upstream transcription element. These core promoters were derived from either the SFV-1 LTR [pL(-E), nt -35 to +44] or the IP [pSIP-5(-E), nt -35 to +13]. When cloned directly upstream of the corresponding TATA box, the TRE^{GP} was capable of restoring Tas-1 responsiveness to either of the SFV-1 core promoters (Table 1). The effect of the TRE^{GP} was independent of orientation but was, in general, unable to restore Tas-1-dependent CAT activity to the wild-type level observed in cells transfected with pSIP-5 or pLCAT-41. Thus, in the context of SFV-1 core LTR and IP, the TRE^{GP}

also displays the property of a conditional, Tas-1-dependent enhancer element.

GST-Tas-1 binds the TRE^{GP}. The ability of the fusion protein GST-Tas-1 to bind to the TRE^{GP} was examined by both EMSA and DNase I footprinting under binding conditions previously described to detect the binding of GST-Tas-1 to the proximal element in the internal promoter (TRE^{IP}) (35). As shown in Fig. 3, GST-Tas-1 (Fig. 3, lanes T) bound to oligonucleotides representing the wild-type GP-22 element whereas binding to the clustered-site GP-22 mutants was reduced (M1, M2, M4, and M5) or undetectable (M3). In competition experiments (Fig. 4A), unlabelled oligonucleotides representing the TRE^{GP} (lanes 4 and 5) or TRE^{IP} (lanes 6 and 7) competed for GST-Tas-1 binding to the TRE^{GP}. Likewise, unlabelled oligonucleotides representing either the TRE^{GP} (Fig. 4B, lanes 4 and 5) or TRE^{IP} (lanes 6 and 7) competed for binding of GST-Tas-1 to the TRE^{IP} (lane 3). An unlabelled nonspecific competitor containing a binding site for the mammalian transcription factor SP-1 displayed no capacity to compete for Tas-1 binding to these targets (Fig. 4A, lanes 8 and 9). In addition, purified Gst did not bind to any oligonucleotide evaluated in these EMSA experiments (Fig. 3, lanes G; Fig. 4, lanes 2).

To identify more precisely which nucleotides in the TRE^{GP} interact with GST-Tas-1, DNase footprinting analysis was performed. In previous studies, GST-Tas-1 was shown to protect approximately 26 bp of the IP from DNase I digestion (35). This protected region is located immediately 5' of the internal promoter TATA box (3) and was designated the promoter-proximal element (35). Under similar assay conditions, GST-Tas-1 also protected an extended region, about 40 bp, in the TRE^{GP} (Fig. 5A); the protected sequence extends from nt 2237 to 2279 in the TRE^{GP}. This extended footprint is in

FIG. 2. Mutational analysis of the TRE^{GP}. The nucleotide sequence of the TRE^{GP} is shown at the top (SFV-1 nt 2224 to 2282). Base pair changes in each mutant are given below the wild-type sequence. Periods indicate a wild-type base pair at that position. Values in the -Tas-1 (cotransfection of reporter plasmid and pSP65) and +Tas-1 (cotransfection of reporter plasmid and pSV-Tas-1) columns represent the mean counts per minute of duplicate transfected cultures. Fold is the relative activity in cotransfections with and without Tas-1.

agreement with the analysis of clustered-site mutations in TRE^{GP} in transient-expression assays (Fig. 2) and with the binding studies performed by EMSA (Fig. 3) in which clustered-site mutations over any third of the 59-bp GP-22 element resulted in low (M1) or basal (M2 and M3) CAT activity and reduced (M1 and M2) or undetectable (M3) Tas-1 binding. A comparison of both the homology between and the sequences protected in the TRE^{GP} and TRE^{IP} is shown in Fig. 5B.

Identification of the 5' ends of transcripts initiated in the *pol* gene in SFV-1-infected cells. The localization of TRE^{GP} to the extreme 3' end of the *gag* gene and immediately upstream of the *pol* ORF is reminiscent of the TRE^{IP}, in which a TRE, including a Tas-1 binding site, is situated immediately 5' of the

ORF located at the 3' end of the SFV-1 genome (3, 22). However, inspection of the SFV-1 sequence between the TRE^{GP} and the first ATG in the *pol* ORF (encoded by nt 2332 to 2334), which has been shown to be essential for efficient HFV replication (11, 17), revealed no sequence resembling a canonical TATA box. We were nevertheless interested to determine if the TRE^{GP} could be an upstream element of a second intragenic viral promoter element, in addition to the previously described IP in the *env* gene. Accordingly, to identify a potential transcription initiation site in the *pol* gene, we analyzed poly(A)⁺ RNA isolated from SFV-1-infected Cf2Th cells by the 5' RACE method. Reverse transcriptase was used to produce cDNA from RNA obtained from infected cells, and

TABLE 1. TRE^{GP} Restoration of Tas-1 responsiveness to SFV-1 IP and LTR core promoters

Plasmid ^a	NIH 3T3 cells			Cf2Th cells			Cos-7 cells		
	Activity ^b		Fold ^c	Activity		Fold	Activity		Fold
	-Tas-1	+Tas-1		-Tas-1	+Tas-1		-Tas-1	+Tas-1	
pSIP-5(-E)	2,327	2,448	1	20,555	81,286	4	11,667	35,629	3
pSIP-5	2,191	37,936	17	10,527	2,114,336	201	13,041	1,217,108	93
pR(S)-SIP-5(-E)	2,097	26,938	13	35,020	1,491,264	43	27,817	540,884	19
pR(AS)-SIP-5(-E)	1,929	15,030	8	11,200	1,161,904	104	23,806	483,781	20
pL(-E) ^c	3,686	3,080	1	68,341	96,255	1	10,759	44,431	4
pLCAT-41	6,597	120,354	18	10,733	2,034,568	189	3,310	1,033,548	312
pR(S)-L(-E)	5,643	17,961	3	90,912	478,709	5	8,704	227,011	26
pR(AS)-L(-E)	3,125	21,397	7	49,929	720,313	14	20,972	274,028	13

^a pSIP-5(-E) and pL(-E) are CAT reporter constructs that contain the core promoters derived from the SFV-1 IP and LTR, respectively. Both contain only a TATA box as an upstream *cis*-acting element. pSIP-5 and pLCAT-41 contain the TRE^{IP} or TRE^{LTR} upstream of the homologous core promoter. pR(S/AS)-SIP-5(-E) and pR(S/AS)-L(-E) contain the TRE^{GP} in the sense or antisense orientation, respectively, relative to the core promoters.

^b Values in the -Tas-1 (cotransfection of reporter plasmid and pSP65) and +Tas-1 (cotransfection of reporter plasmid and pSV-Tas-1) columns represent the mean counts per minute from duplicate transfected cultures.

^c Fold represents the relative activity in cotransfection of reporter plasmid and Tas-1.

FIG. 3. Gel retardation analysis of wild-type and mutant TRE^{GP} oligonucleotide probes. Radiolabelled probes were incubated with purified recombinant GST (lanes G) or GST-Tas-1 (lanes T). These end-labelled probes correspond to the mutants shown in Fig. 2. Lanes labeled – contain probe without added protein.

this cDNA was amplified by PCR in parallel reactions with primers designed to detect transcripts initiated by the LTR, IP, or *gagpol* region. The 5' RACE products were purified by gel electrophoresis and cloned into a plasmid, and the junction between the homopolymeric C tail and viral cDNA was determined by DNA sequence analysis. Figure 6A shows a typical sequence analysis of the C-tail/cDNA junction for transcripts initiated in the 5' LTR, *pol* gene, and IP. A summary of the sequence analysis of a total of 57 clones produced by the 5' RACE method is shown in Fig. 6B. Transcripts from the LTR and IP exhibited a small amount of heterogeneity around the respective initiation site: 6 nt (–2 to +4) for the LTR, and 3 nt (8408 to 8410) for the IP. These data are in agreement with the +1 site previously reported for the SFV-1 IP (nt 8408) by

primer extension and in vitro transcription analysis (3) and for the SFV-1 LTR by primer extension analysis (unpublished results). Sequence analysis for transcripts initiating in the *pol* region revealed a clustering of initiation sites located at nt 2594 to 2600 (7 nt) and at nt 2607 to 2611 (5 nt). These two clusters of sites map 15 to 21 and 28 to 32 nt downstream of a putative TATA box (5'TATAT3') located at nt 2575 to 2579. The 5' end of transcripts initiated in *pol* was also independently mapped by primer extension analysis of poly(A)⁺ RNA isolated from SFV-1-infected Cf2Th cells (Fig. 6C). The major extension products were localized within (Fig. 6C; band c, nt 2611; band b, nt 2595) or adjacent to (band a, 2592) the clusters of 5' ends determined by the 5' RACE method (indicated by asterisks in Fig. 6B). In addition, primer extension experiments, involving the use of an antisense primer in the 5' end of the CAT coding region to analyze total RNA from Cf2Th cells cotransfected with pGP-24 and Tas-1 (see below), showed transcripts that initiated at SFV-1 nt 2606 and 2595 (data not shown). Taken together, these studies on RNA from infected cells and cells transfected with an expression plasmid provide evidence that viral transcripts are initiated in the *pol* region of SFV-1.

Promoter activity of a DNA fragment encompassing TRE^{GP} in transient-expression assays. We next examined the promoter activity of the region encompassing TRE^{GP} in transient-transfection assays to address the hypothesis that this enhancer element serves as an upstream *cis*-acting sequence to augment transcription from initiation complexes that form at the downstream TATA box (nt 2575 to 2579), which was implicated as functional by the 5' RACE method and primer extension analysis. The entire region beginning upstream of the TRE^{GP} (at nt 2200) to immediately 3' of the putative TATA box and the apparent transcription initiation sites deduced in Fig. 6 (to nt 2611), was cloned into the promoterless CAT construct pCAT-22A2s to generate pGP-24. In transient-expression assays, pGP-24 produced high levels of CAT activity only in the pres-

FIG. 4. (A) Gel retardation and competition analysis for GST-Tas-1 binding to the TRE^{GP} . End-labelled TRE^{GP} (GP-22) oligonucleotide was incubated alone (–; lane 1), with purified GST (G; lane 2), or with purified GST-Tas-1 (T; lanes 3 to 9). Unlabelled competitor oligonucleotides were added at 35-fold (lanes 4, 6, and 8) and 350-fold (lanes 5, 7, and 9) molar excess over the labelled probe. (B) Gel retardation and competition analysis for GST-Tas-1 binding to the TRE^{IP} . End-labelled TRE^{IP} oligonucleotide was incubated alone (–; lane 1), with purified GST (G; lane 2) or with purified GST-Tas-1 (T; lanes 3 to 7). Unlabelled competitor oligonucleotides were added at 35-fold (lanes 4 and 6) and 350-fold (lanes 5 and 7) molar excess over the labelled probe.

FIG. 5. (A) DNase I protection analysis of a TRE^{GP} probe. A restriction fragment encompassing the TRE^{GP} was uniquely end labelled on the coding or noncoding strand with ³²P and incubated alone (F [free; lanes 2, 5, 7, and 10]) or with purified GST-Tas-1 at 1 μ g (B [bound; lanes 3 and 8]) or 4 μ g (B [lanes 4 and 9]). Lanes 1 and 6 contain G+A chemical sequencing ladders (21). Protected sequences for the coding and noncoding strands are shown. (B) Sequence relationship of TRE^{GP} and TRE^{IP}. All alignments were performed with the GeneWorks Program Version 2.4, provided by IntelliGenetics, Campbell, Calif.). Nucleotide sequences conserved between TRE^{GP} and TRE^{IP} are shaded. Vertical hatch marks above and below the alignment indicate the GST-Tas-1 protected area seen in each element. The 6-nt mutation in mutant M4 and the 2-nt mutation in M5 are indicated by the single and double underline, respectively.

ence of Tas-1 (pGP-24, Table 2). Additionally, mutation of either the TATA box (pGP-24-M1, Table 2) or the TRE^{GP} (pGP-24-M2) (Table 2) or both the TATA box and the TRE^{GP} (pGP-24-M2/M1) (Table 2) resulted in a stepwise reduction of CAT activity in transient-expression assays. Also tested was a plasmid containing nt 2200 to 2611 cloned into a promoterless vector with the secreted alkaline phosphatase reporter gene. Transient-expression assays with Cf2Th cells revealed that the TRE^{GP} and TATA box (nt 2575 to 2579) were both required for high levels of secreted alkaline phosphatase expression (data not shown). Taken together, these observations indicate the importance of enhancer activity of TRE^{GP} and suggest that the TATA box at nt 2575 to 2579 controls initiation of transcription in the *pol* gene.

DISCUSSION

In this report, the SFV-1 genome was shown to contain a TRE at the end of the *gag* gene. This element, designated TRE^{GP}, was mapped through deletion analysis to a 59-bp region (nt 2224 to 2282) at the 3' end of the *gag* gene and immediately preceding the *gagpol* overlap region which begins at nt 2296 (Fig. 7). TRE^{GP} is similar to the other two viral TRE, TRE^{LTR} and TRE^{IP}, in that it functions by an enhancer-

like mechanism and exhibits a large response to Tas-1 in transient-expression assays. TRE^{GP} partially substituted for the TRE in the LTR and IP; thus, it is possible that the enhancer function of the TRE^{GP} in the context of the proviral genome is important for augmentation of viral transcription in the LTR and/or IP. Interaction in *cis* between the HFV LTR and IP has been reported (16).

In vitro binding of the fusion protein, GST-Tas-1 to TRE^{GP} was demonstrated by both EMSA (Fig. 3) and DNase footprinting (Fig. 5A). Similar studies by these methods have also revealed that GST-Tas-1 interacts with TRE^{IP} (35) and target elements in the SFV-1 LTR (unpublished results). A HFV GST-Bel-1 fusion protein has also been recently reported to interact with DNA targets in the HFV LTR and internal promoter (12). The DNase footprints on both the TRE^{GP} (43 bp [Fig. 5A]) and the TRE^{IP} (26 bp) (35) extend over approximately three to four turns of alpha-helical DNA. Although there is no direct evidence for multimerization of the Tas-1 protein, these extended footprints are consistent with the hypothesis that Tas-1 binding to the two targets involves multimer formation. Extended regions of in vitro protection by GST-Bel-1 on the HFV LTR and IP elements have also been reported (12). Oligonucleotides representing TRE^{GP} and TRE^{IP} competed for Tas-1 binding to either target element

FIG. 6. (A) DNA sequence analysis of representative 5' RACE clones. The junctions between the homopolymeric C tail and viral cDNA for the LTR, *pol* region, and IP are shown. There is some inherent ambiguity in the 5' RACE method, because the homopolymeric tail may contain the same base as the putative start site as indicated by the brackets for the LTR and *pol* region; further refinement in this method will require tailing with other homopolymers. Arrows designate the unambiguous determination of the 5' end of transcripts initiated in the IP. (B) Summary of DNA sequence analysis of 5' RACE clones. The SFV-1 sequence is shown for 40 to 50 bp encompassing the TATA box in the *pol* region, IP, and LTR. Arrows and brackets indicate the 5' end of viral transcripts as deduced from DNA sequence analysis of 5' RACE clones. Numbers above the arrows and brackets indicate the number of these clones with sequences with the corresponding +1 site. The asterisks above the *pol* sequence correspond to the 5' ends independently verified by primer extension analysis (C). (C) Primer extension analysis of SFV-1-infected Cf2Th cell RNA. Poly(A)⁺ RNA (5 μ g) isolated from uninfected Cf2Th (lane 1) or SFV-1-infected Cf2Th cells (lane 2) was primer extended with ³²P-end-labeled oligonucleotide SFV-1 713AS. The primer extension products were analyzed on 6% acrylamide gels under denaturing conditions. Lanes G, A, T, and C are a dideoxy sequencing ladder of p5'SFV-1 primed with oligonucleotide 713AS. The positions of the major cDNAs produced are indicated by arrows a (nt 2592), b (nt 2595), and c (nt 2611). The 5' ends of these cDNAs are indicated by asterisks in panel B.

TABLE 2. Promoter activity of a DNA fragment encompassing TRE^{GP} in transient-expression assays

Plasmid	TRE ^{GP} ^a	TATA ^a	Cf2TH cells			NIH 3T3 cells		
			Activity ^b		Fold ^c	Activity		Fold
			-Tas-1	+Tas-1		-Tas-1	+Tas-1	
pGP-24	+	+	5,083	2,201,195	433	789	91,202	116
pGP-24-M1	+	mu	5,915	274,195	46	823	18,573	23
pGP-24-M2	mu	+	6,177	101,406	16	1,083	14,470	13
pGP-24-M2/M1	mu	mu	1,994	24,543	12	1,056	2,451	2

^a A wild-type (+) or mutated (mu) TRE^{GP} and TATA box in each reporter plasmid is indicated.

^b Values in the -Tas-1 (cotransfection of reporter plasmid and pSP65) and +Tas-1 (cotransfection of reporter plasmid and pSV-Tas-1) columns represent the mean counts per minute from duplicate transfected cultures.

^c Fold represents the relative activity in cotransfection of reporter plasmid and Tas-1.

(Fig. 4); however, the precise identification of nucleotides contacted by Tas-1 requires additional analysis. We have constructed two mutants in TRE^{GP} (M4 and M5) (Fig. 2 and 5B) which contain relatively small nucleotide changes; these mutants displayed both reduced Tas-1 binding *in vitro* and reduced Tas-1 responsiveness in the context of reporter gene constructs in transient-expression assays. Thus, these mutations may define critical contact points for interaction with Tas-1. Because of the low degree of homology between Tas-responsive targets in the SFV-1 genome (3) (Fig. 5B), it appears that the transactivator of this virus is a DNA-binding protein that is capable of recognizing several different target sequences.

The location of the TRE^{GP} is highly reminiscent of those of both the TRE^{IP} and the promoter-proximal TRE^{LTR}; for the LTR and IP, a Tas-1 target element is located immediately 5' of an ORF(s). Therefore, the effect of TRE^{GP} on a potential promoter in *pol* was evaluated. Analysis of viral RNA in infected cells by the 5' RACE method revealed transcripts initiated within the *pol* gene (Fig. 6A and B) with 5' ends mapping to two clusters of sites approximately 15 to 30 nt downstream of a putative TATA box located at nt 2575 to 2579. The validity of this method was established by additional analysis of the same RNA; we correctly identified the 5' ends of viral transcripts initiated in the LTR and internal promoter. Additionally, the primer extension method was used to analyze viral transcripts in infected cells; this method verified the 5'

ends of viral transcripts initiated in the *pol* region (Fig. 6C). Transcripts derived from *pol* initiate downstream of the first two potential start codons (ATG) in the *pol* ORF but are upstream of both the third ATG in the *pol* frame and the conserved proline residue (encoded by nt 2836 to 2638), which has been shown to be essential for *pol* activity in HFV (15). Although the exact 5' end of the Pol polypeptide of foamy viruses remains undefined, the novel transcripts that we have detected in SFV-1-infected cells would bypass the first ATG in the *pol* frame (Fig. 7). This ATG, conserved among HFV, SFV-1, SFV-3, and SFV_{cpz}, is essential for production of high levels of infectious HFV and is considered a likely candidate for translational initiation of the HFV Pro-Pol polypeptide (11, 17). In addition, it has recently been reported that the HFV Pol polypeptide is derived from a spliced mRNA which initiates in the 5' LTR and is generated through a splicing event which joins sequences at the major splice donor in R-U5 to a splice acceptor in the capsid domain of the HFV *gag* gene (34). Although we have not examined SFV-1-infected-cell RNA for the presence of this novel spliced transcript, these splice sites are conserved among HFV and SFV-1, and thus a splicing event may be the major pathway in the synthesis of a *pol* mRNA in SFV-1-infected cells.

Additional observations suggested the importance of TRE^{GP} in the synthesis of viral transcripts. The Tas-1 responsiveness of reporter gene plasmids with hybrid promoters containing TRE^{GP} was strong and of the same magnitude as that

FIG. 7. Schematic presentation of TRE in the SFV-1 genome. Shown are the locations of the 5' ends of transcripts initiated in the *pol* region relative to the first three translational start codons in the *pol* frame. The splice acceptor in the *gag* gene essential for HFV *pol* production (34) is shown. The TRE are shown by hatched boxes.

of the TRE^{IP} (Fig. 1) and the promoter-proximal TRE^{LTR} (23) in transient-expression assays. When the region from nt 2200 to 2611, containing both the TRE^{GP} and a putative TATA box at 2575 to 2579, was cloned into promoterless CAT or secreted alkaline phosphatase reporter constructs, high levels of Tas-1-dependent promoter activity were observed (Table 2). The 5' end of these transcripts mapped to sites contained within the cluster of +1 sites defined by the 5' RACE method and primer extension analysis of poly(A)⁺ RNA from SFV-1-infected cells. The promoter activity of the region containing the TRE^{GP} and the TATA box in *pol* appeared to be weak in the context of the viral genome in infected cells but was strong when tested in transient-expression assays with reporter gene plasmids. Various mechanisms of promoter suppression or interference have been proposed for lack of promoter activity of the 3' LTR of retroviruses (2, 6, 13, 14) or internal promoters cloned in retroviral vectors (9, 10). It is unknown if a similar mechanism might cause the suppression of transcripts initiating in the *pol* region of SFV-1. It is possible that the promoter activity of this region in *pol* is also under temporal regulation and is less active at late times in infection, when the RNA used in this study was isolated for analysis. Also, the promoter in *pol* may function efficiently only in certain cell types in vivo. Additional studies are needed to understand the role of the TRE^{GP} in the SFV-1 replication cycle and to evaluate the significance of initiation of viral transcripts in the *pol* gene.

ACKNOWLEDGMENTS

We thank June Zou for providing the GST-Tas-1 expression plasmid. The technical assistance of Karen Shaw and Chihong Zhou in preparation of synthetic oligonucleotides is gratefully acknowledged.

This research was supported by grants from the National Institutes of Health: RO1-AI30525 to P.A.L. and base grant RR00169 to the California Regional Primate Research Center.

REFERENCES

- Arrigo, S., M. Yun, and K. Beemon. 1987. *cis*-acting regulatory elements within the *gag* genes of avian retroviruses. *Mol. Cell. Biol.* 7:388-397.
- Boerkoel, C. F., and H. Kung. 1992. Transcriptional interaction between retroviral long terminal repeats (LTRs): mechanism of 5' LTR suppression and 3' promoter activation of *c-myc* in avian B-cell lymphomas. *J. Virol.* 66:4814-4823.
- Campbell, M., L. Renshaw-Gegg, R. Renne, and P. A. Luciw. 1994. Characterization of the internal promoter of simian foamy viruses. *J. Virol.* 68:4811-4820.
- Carlberg, K., T. A. Ryden, and K. Beemon. 1988. Localization and footprinting of an enhancer within the avian sarcoma virus *gag* gene. *J. Virol.* 62:1617-1624.
- Coffin, J. M. 1995. Retroviridae: the viruses and their replication, *In* p. 763-843. B. N. Fields (ed.), *Fundamental virology*. Raven Press, New York.
- Cullen, B. R., T. Lomedico, and G. Ju. 1984. Transcriptional interference in avian retroviruses—implications for the promoter insertion model of leukemogenesis. *Nature (London)* 307:241-245.
- El Kharroubi, A., and E. Verdin. 1994. Protein-DNA interactions within DNase I-hypersensitive sites located downstream of the HIV-1 promoter. *J. Biol. Chem.* 269:19916-19924.
- Elliott, J. F., B. Pohajdak, D. J. Talbot, J. Shaw, and V. Paetkau. 1988. Phorbol diester-inducible, cyclosporine-suppressible transcription from a novel promoter within the mouse mammary tumor virus *env* gene. *J. Virol.* 62:1373-1380.
- Emerman, M., and H. M. Temin. 1984. Genes with promoters in retrovirus vectors can be independently suppressed by an epigenetic mechanism. *Cell* 39:459-467.
- Emerman, M., and H. M. Temin. 1986. Comparison of promoter suppression in avian and murine retrovirus vectors. *Nucleic Acids Res.* 14:9381-9395.
- Enssle, J., I. Jordan, B. Maurer, and A. Rethwilm. 1996. Foamy virus reverse transcriptase is expressed independently from the gag protein. *Proc. Natl. Acad. Sci. USA* 93:4137-4141.
- He, F., W. S. Blair, J. Fukushima, and B. R. Cullen. 1996. The human foamy virus Bel-1 transcription factor is a sequence-specific DNA binding protein. *J. Virol.* 70:3902-3908.
- Herman, S. A., and J. M. Coffin. 1986. Differential transcription from the long terminal repeats of integrated avian leukosis virus DNA. *J. Virol.* 60:497-505.
- Klaver, B., and B. Berkhout. 1994. Comparison of 5' and 3' long terminal repeat promoter function in human immunodeficiency virus. *J. Virol.* 68:3830-3840.
- Kogel, D., M. Aboud, and R. M. Flugel. 1995. Mutational analysis of the reverse transcriptase and ribonuclease H domains of the human foamy virus. *Nucleic Acids Res.* 23:2621-2625.
- Lochelt, M., M. Aboud, and R. M. Flugel. 1993. Increase in the basal transcriptional activity of the human foamy virus internal promoter by the homologous long terminal repeat in *cis*. *Nucleic Acids Res.* 21:4226-4230.
- Lochelt, M., and R. M. Flugel. 1996. The human foamy virus *pol* gene is expressed as a pro-pol polyprotein and not as a gag-pol fusion protein. *J. Virol.* 70:1033-1040.
- Lochelt, M., R. M. Flugel, and M. Aboud. 1994. The human foamy virus internal promoter directs the expression of the functional bel-1 transactivator and bet protein early after infection. *J. Virol.* 68:638-645.
- Lochelt, M., W. Muranyi, and R. M. Flugel. 1993. Human foamy virus genome possesses an internal, bel-1-dependent and functional promoter. *Proc. Natl. Acad. Sci. USA* 90:7317-7321.
- Loh, T. P., L. L. Sievert, and R. W. Scott. 1988. Negative regulation of retrovirus expression in mouse embryonal carcinoma cells mediated by an intragenic domain. *J. Virol.* 62:4086-4095.
- Maxam, A. M., and W. Gilbert. 1980. Sequencing end-labeled DNA with base-specific chemical cleavages. *Methods Enzymol.* 65:499-559.
- Mergia, A. 1994. Simian foamy virus type 1 contains a second promoter at the 3' end of the *env* gene. *Virology* 199:219-222.
- Mergia, A., E. Pratt-Lowe, K. E. S. Shaw, L. W. Renshaw-Gegg, and P. A. Luciw. 1992. *cis*-acting regulatory regions in the long terminal repeat of simian foamy virus type 1. *J. Virol.* 66:251-257.
- Mergia, A., K. E. S. Shaw, J. E. Lackner, and P. A. Luciw. 1990. Relationship of the *env* genes and the endonuclease domain of the *pol* genes of simian foamy virus type 1 and the human foamy virus. *J. Virol.* 64:406-410.
- Mergia, A., K. E. S. Shaw, E. Pratt-Lowe, P. A. Barry, and P. A. Luciw. 1991. Identification of the simian foamy virus transcriptional transactivator gene (*taf*). *J. Virol.* 65:2903-2909.
- Miller, C. L., R. Garner, and V. Paetkau. 1992. An activation-dependent, T-lymphocyte-specific transcriptional activator in the mouse mammary tumor virus *env* gene. *Mol. Cell. Biol.* 12:3262-3272.
- Nordeen, S. K., P. I. Green, and D. M. Fowlkes. 1987. A rapid, sensitive, and inexpensive assay for chloramphenicol acetyltransferase. *DNA* 6:173-178.
- Nosaka, T., Y. Ariumi, M. Sakurai, R. Takeuchi, and M. Hatanaka. 1993. Novel internal promoter/enhancer of HTLV-I tax expression. *Nucleic Acids Res.* 21:5124-5129.
- Rethwilm, A. 1995. Regulation of foamy virus gene expression. *Curr. Top. Microbiol. Immunol.* 193:1-24.
- Van Lint, C., A. Burny, and E. Verdin. 1991. The intragenic enhancer of human immunodeficiency virus contains functional AP-1-binding sites. *J. Virol.* 65:7066-7072.
- Verdin, E., N. Becker, F. Bex, L. Droogmans, and A. Burny. 1990. Identification and characterization of an enhancer in the coding region of the genome of human immunodeficiency virus type 1. *Proc. Natl. Acad. Sci. USA* 87:4874-4878.
- Verdin, E., and C. Van Lint. 1995. Internal transcriptional regulatory elements in HIV-1 and other retroviruses. *Cell. Mol. Biol.* 41:365-369.
- Willems, L., R. Kettmann, and A. Burny. 1995. Bovine leukemia virus has an internal promoter, p. 175. *In* *Proceedings of the 1995 Cold Spring Harbor Meeting on Retroviruses*.
- Yu, S. F., D. N. Baldwin, S. R. Gwynn, S. Yenapilli, and M. Linial. 1996. Human foamy virus replication: a pathway distinct from that of retroviruses and hepadnaviruses. *Science* 271:1579-1582.
- Zou, J. X., and P. A. Luciw. 1996. The transcriptional transactivator of simian foamy virus 1 binds to a DNA target in the viral internal promoter. *Proc. Natl. Acad. Sci. USA* 93:326-330.