JOURNAL OF VIROLOGY, Feb. 1994, p. 979–987
0022-538X/94/$04.00+0

Vol. 68, No. 2

Negative-Strand RNA Transcripts Are Produced in Human Immunodeficiency Virus Type 1-Infected Cells and Patients by a Novel Promoter Downregulated by Tat

NELSON L. MICHAEL,[1]* MARYANNE T. VAHEY,[1] LISA D'ARCY,[2] PHILIP K. EHRENBERG,[2] JOSEPH D. MOSCA,[2] JAY RAPPAPORT,[3] AND ROBERT R. REDFIELD[1]

Department of Retroviral Research, Walter Reed Army Institute of Research,[1] Henry M. Jackson Foundation,[2] Rockville, Maryland 20850 and Laboratory of Oral Medicine, National Institute of Dental Research, Bethesda, Maryland 20892[3]

Received 10 September 1993/Accepted 8 November 1993

Current understanding of human immunodeficiency virus type 1 (HIV-1) transcription is based on unidirectional expression of transcripts with positive-strand polarity from the 5′ long terminal repeat. We now report HIV-1 transcripts with negative-strand polarity obtained from acutely and chronically infected cell lines by use of a template orientation-specific reverse transcriptase-PCR assay. These findings were confirmed in natural infection by analysis of RNA derived from peripheral blood mononuclear cell samples from 15 HIV-1-infected patients. A cDNA derived from a 2.3-kb polyadenylated HIV-1 RNA with negative-strand polarity which encodes a highly conserved 189-amino-acid open reading frame antiparallel to the envelope gene was isolated from acutely infected A3.01 cells. Through use of reporter gene constructions, we further found that a novel negative-strand promoter functions within the negative response element of the 3′ long terminal repeat, which is downregulated by coexpression of Tat. Site-directed mutagenesis experiments demonstrated that NF-κB I and USF sites are crucial for negative-strand promoter activity. These data extend the coding capacity of HIV-1 and suggest a role for antisense regulation of the viral life cycle.

The positive-strand promoter (PSP) of human immunodeficiency virus type 1 (HIV-1), contained entirely in the U3 region of the long terminal repeat (LTR), is composed of three SP1-binding sites located at positions −78 through −47 and a TATA box at positions −28 to −24 (28, 33). Two NF-κB enhancer elements are located upstream, at positions −106 and −92 (29), and sequences with limited homology to AP-1 enhancer elements are found further upstream. A 59-bp region which confers transcriptional transactivating potential on the core promoter elements by the viral Tat protein is located within the R region of the LTR (3, 37). This transactivation response (TAR) region folds into alternate RNA stem-loop structures that are found in the 5′ termini of all HIV-1 transcripts. The TAR element mediates a substantial increase in transcriptional initiation and elongation through interactions with the viral Tat protein and other cellular factors (9, 19, 21, 35).

Bidirectional expression of transcription units is frequently observed in small genetic systems, such as mammalian mitochondria and animal viruses, where relatively large amounts of genetic information are encoded in relatively small genomes (1, 4, 24, 38, 39). Mammalian mitochondrial DNA contains a single structural gene and 8 of 22 transfer DNA genes encoded on the DNA strand complementary to the major coding strand. Both strands are expressed by a pair of promoters localized in close proximity to each other in the D-loop region (6, 7, 11). Herpes simplex virus type 1, a pathologic human DNA virus, possesses a gene, *lat-1*, which partially overlaps the α0 regulatory gene encoded on the major coding strand of herpes simplex virus type 1 and is thought to play a role in viral latency (38, 39). HIV-1 utilizes a panoply of genetic mechanisms to streamline a large amount of coding information into a 9.7-kb genome (8, 15, 27). These mechanisms include overlapping genes expressed by a complex pattern of alternative RNA splicing (2, 10, 16, 34), bicistronic transcripts employing ribosomal frameshifting (2, 18, 34), and posttranslational cleavage of polyproteins into products with a wide range of functions (40).

A limited report (5) has suggested HIV-1 negative-strand RNAs of 1.0, 1.1, and 1.6 kb in H9 cells acutely infected with $HIV\text{-}1_{111B}$. Negative-strand transcripts of 2.5 and 2.9 kb have also been identified in human T-cell leukemia virus type 1 (HTLV-1)-infected cells produced by a weak negative-strand promoter (NSP; 20). The HTLV-1 genome also encodes an open reading frame (ORF) of 265 amino acids 3′ to the envelope gene (20). We now present evidence that HIV-1 exhibits bidirectional transcription of coding regions mediated by a novel NSP in the negative response region of the 3′ LTR. Negative-strand RNA production may, therefore, be a general but underappreciated feature of lentiviral gene expression with important coding and regulatory functions.

MATERIALS AND METHODS

Detection of HIV-1 RNAs with positive- and negative-strand polarities by reverse transcriptase (RT)-PCR technology. The method used for sample preparation and quantitative RT-PCR and the *gag* and *nef* primer-probe combinations were previously described (25). We used primers specific for a 136-bp fragment of *gag* or a 304-bp fragment of *nef* that could amplify cDNA with either positive- or negative-strand orientation, depending upon the polarity of the primer chosen for the reverse transcription step. A single primer was used for the reverse transcription step prior to heat inactivation of the RT and subsequent PCR. For detection of positive-strand RNA, an oligodeoxynucleotide with negative-strand polarity was used

* Corresponding author. Mailing address: 13 Taft Court, Suite 200, Rockville, MD 20850. Phone: (301) 762-0089. Fax: (301) 762-4177. Electronic mail address: mich9009@hjfhiv.hjf.org.

to prime the RT reaction. For detection of negative-strand RNA, an oligodeoxynucleotide with negative-strand polarity was used to prime the RT reaction. Detection and determination of the sequence specificity of the amplified fragment were done by hybridization with an internal oligodeoxynucleotide probe. Copy number (see Fig. 1C) was assigned relative to parallel amplifications with cRNA transcribed in vitro from a cloned template. Average values from three iterations were used to assign both standard and experimental copy numbers. Signal detection and quantitation were performed by storage phosphor technology (Molecular Dynamics) (25).

Construction and screening of a cDNA library from acutely infected A3.01 cells. A 5-μg of sample poly(A)-enriched RNA from A3.01 cells acutely infected with HIV_{IIIB} was converted to cDNA, ligated into bacteriophage arms, and packaged by using the λZAPII system (Stratagene) in accordance with the manufacturer's instructions. This method, which employs a 3′ primer containing the *Xho*I recognition sequence followed by an oligo(dT) sequence, provides for the 3′ end of each cDNA to have an *Xho*I-site and the 5′ termini to end with an *Eco*RI site. Bacteriophage plaques were plaque purified with three successive hybridization screenings with the 8.9-kb insert from pBH10 (17) prior to in vivo phagemid rescue with helper phage R408. Sixty highly purified clones were obtained and subsequently oriented by DNA sequencing. The three negative-strand cDNAs were completely sequenced by using fluorescent-dye cycle sequencing and an Applied Biosystems 373A automated sequencer.

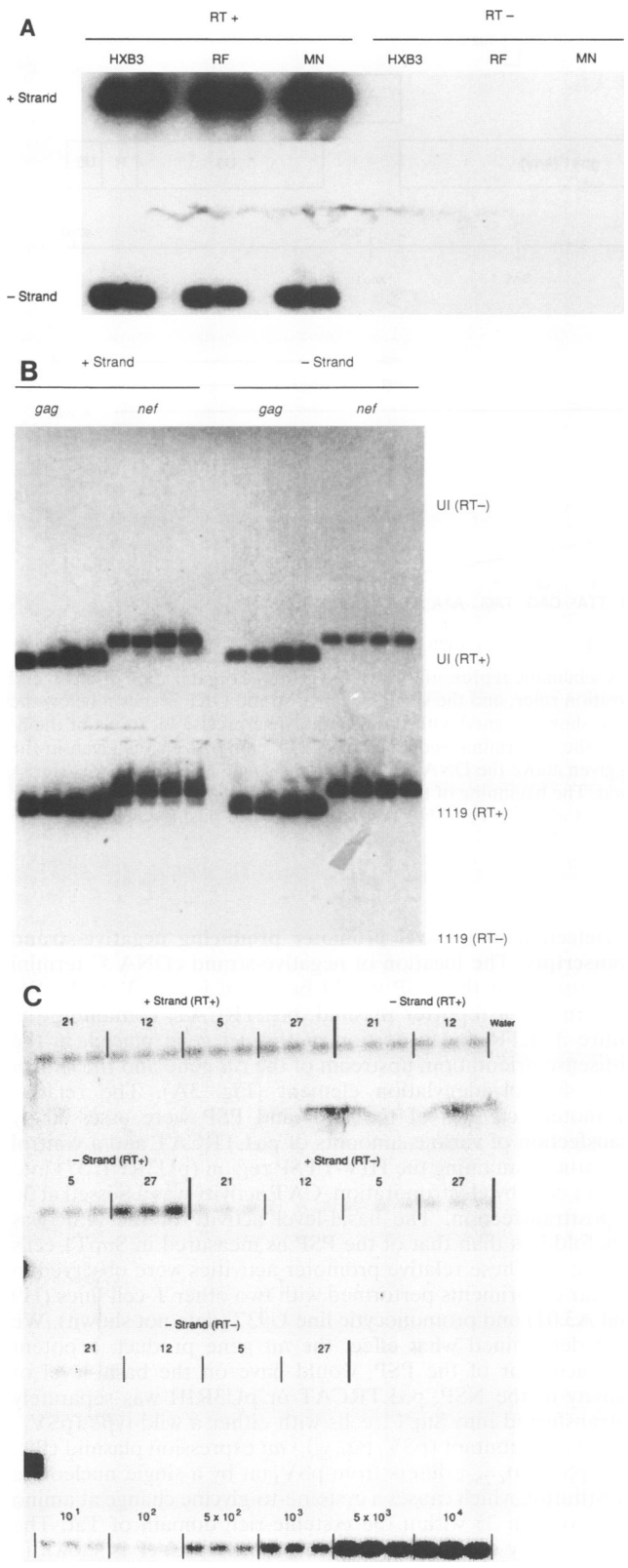
Reporter gene constructions. pαLTRCAT was constructed by amplifying a 719-bp fragment of pHXB2 representing sequence positions (32) 9718 (U5) to 9009 (*nef*) by PCR. An *Xba*I site was added to the U5 primer, and the *nef* primer was phosphorylated prior to amplification. pSV_2CAT was digested with *Hin*dIII, treated with T4 DNA polymerase, and then digested with *Bam*HI, and the resulting 1.6-kb fragment containing the *cat*-poly(A) region was gel purified. The HIV-1 fragment was ligated in a tripartite reaction with the *cat*-poly(A) fragment and an *Xba*I-*Bam*HI double-digested pBluescript KSII plasmid. Deletion mutations of the 3′ LTR were prepared by PCR and then recloned upstream of the *cat* gene in vector pSKCAT. pSKCAT was constructed by cloning the *cat*-poly(A) fragment from pSV_2CAT into pBluescript II SK− at the *Hin*dIII and *Bam*HI sites. The 3′ deletion constructions share a 5′ terminus at position 9,718 at the end of U5. The 3′ ends are as follows: 14-2, 9,650; 14-4, 9,584; 14-6, 9,538; 14-8, 9,503; 14-10, 9,482; 14-12, 9,471; 14-14, 9,458; 14-16, 9,444; 14-18, 9,430, 14-20, 9,366; 14-22, 9,288; 14-24, 9,225; 14-26, 9,122. The 5′ deletion constructions share a 3′ end at position 9,009. The 5′ ends are as follows: 1-3′, 9,679; 3-3′, 9,649; 7-3′, 9,604; 9-3′, 9,558. Construction 5-3′-ΔUSF spans positions 9,639 to 9,009. It contained a spontaneous G-to-A transition in the USF site conserved hexanucleotide core (HXB2 position 9,372). Constructions 11-26 and 13-26 span positions 9,495 to 9,122 and 9,460 to 9,122, respectively. Constructions 11-26-NF-κBI, 11-26-NF-κB II, and 11-26-NF-κB I+II have triple mutations in the first, second, or both NF-κB elements, respectively (see Fig. 4A). Constructions were recovered by transformation of *Escherichia coli* DH5α cells (GIBCO/BRL) and subsequent restriction enzyme screening. All constructions were purified by double banding in ethidium bromide-cesium chloride gradients and confirmed by nucleotide sequencing.

Reporter gene assays. For each transfection, 1.5×10^7 SupT1 cells were grown to the mid-logarithmic phase in complete RPMI 1640 supplemented with 10% fetal calf serum, resuspended in 0.3 ml of complete RPMI 1640 supplemented with 20% fetal calf serum, and mixed with appropriate amounts of plasmid DNA. The total amount of DNA per transfection was adjusted with pBluescript KSII− to 20 μg. Cell-DNA mixtures were kept on ice for 20 min, transferred to Gene-Pulser (Bio-Rad) cuvettes, subjected to 960 μF and 0.25 kV of electricity, kept on ice for 20 min, transferred into six-well plates with an additional 3.5 ml of complete RPMI 1640 supplemented with 20% fetal calf serum, and placed at 37°C with a 5% CO_2 atmosphere. Cells were harvested 36 h posttransfection, washed once with phosphate-buffered saline, resuspended in 85 μl of 0.25 M Tris-HCl (pH 7.6), and subjected to three successive freeze-thaw cycles with baths held at −80 and 37°C. Total protein concentrations of the lysates were determined by a commercial dye-binding method (Pierce), and the same amount of protein from each lysate was used in the standard chloramphenicol acetyltransferase (CAT) assay (14). Chromatograms were quantitated by using a storage phosphor system (Molecular Dynamics). Percent conversion was determined by dividing the amount of radioactivity in acetylated forms of [^{14}C]chloramphenicol by the amount of radioactivity in the acetylated and nonacetylated forms of [^{14}C]chloramphenicol multiplied by 100.

Electrophoretic mobility shift assays. Two pairs of oligodeoxynucleotides corresponding to both strands of the region of U3 containing the USF site were synthesized and gel purified by using denaturing acrylamide gels. One set was the wild type, and the other contained a G-to-A transition in the core hexanucleotide of the USF element. One member of each set was 5′ end labeled with ^{32}P prior to annealing in stoichiometric amounts with its cognate complementary strand at 70°C for 30 min. For noncompetitive reactions, 10^4 cpm of a double-stranded probe was mixed with 4.0 μl of 5× binding buffer [1.0 mg/ml of poly(dA-dT), 100 mM *N*-2-hydroxyethylpiperazine-*N*′-2-ethanesulfonic acid (HEPES; pH 7.3), 250 mM NaCl, 25% glycerol, 25 mM $MgCl_2$, 10 mM dithiothreitol, 20 mM spermidine, 1.0 mM EDTA) and 5.0 μg of HeLa cell nuclear extract (Promega) for 10 min at 22°C in a final volume of 20 μl. For competitive reactions, a 50-, 100-, 200-, or 400-fold molar excess of unlabeled double-stranded DNA was preincubated with the HeLa cell extract and binding buffer for 10 min at 22°C prior to the addition of labeled probes. A 2.0-μl volume of loading buffer (48% glycerol, 0.05% bromophenol blue) was then added prior to electrophoresis through a 4% acrylamide gel (acrylamide-to-bisacrylamide ratio, 50:1) at 14.4 V/cm for 1 h. The gel was cast and run with a buffer containing 40 mM Tris base, 307 mM glycine, and 0.16% Nonidet P-40. The gel was dried prior to storage phosphor imaging and standard autoradiography. The wild-type oligodeoxynucleotides used were 5′-TCTCGGGCCACGTGATGAAATGCTAGGC and 5′-GCCTAGCATTTCATCACGTGGCCCGAGA. The mutant oligodeoxynucleotides used were 5′ TCTCGGGCCACGTAATGAAATGCTAGGC and 5′ GCCTAGCATTTCATTACGTGGCCCGAGA.

RESULTS

Detection of HIV-1 RNAs with negative-strand polarity by PCR. The presence of negative-strand HIV-1 transcripts produced by H9 cells acutely infected with the HXB3, RF, or MN strain of HIV-1 (Fig. 1A) was determined by using an RT-PCR approach with primers specific for a 136-bp fragment of *gag* that could amplify cDNA with either positive- or negative-strand orientation, depending upon the polarity of the primer chosen for the reverse transcription step. The presence and sequence specificity of the amplified fragment were determined by hybridization with an internal oligodeoxynucleotide probe. The dependence on RNA as the initial template was



shown by parallel reactions performed in the absence of RT (right side of Fig. 1A). RT-PCRs performed in the absence of a specific primer during the reverse transcription step to control for the possibility of nonspecific priming by small DNAs contained in the RNA samples were uniformly negative (data not shown). Relative signal intensities in these experiments could not be interpreted, since the assay was performed in a strictly qualitative fashion. Similar analysis done with primer sets in both the *gag* and *nef* regions revealed positive- and negative-strand transcripts in poly(A)-enriched RNA samples derived from two persistently infected lines, U1 (promonocytic derivative; 12) and 1119 (SupT1/HIV_{LAI}); reference 26a) (Fig. 1B). We next sought to establish the physiologic relevance of these findings to natural infection. Thus, negative-strand RNA expression was assessed in RNA extracted from 15 uncultured peripheral blood mononuclear cell samples from 15 different early-stage asymptomatic HIV-1-infected patients. Negative-strand RNA was detected in all of the samples with the *gag* primer set and a quantitative RT-PCR assay (Table 1). The proportion of HIV-1 RNA containing these *gag* sequences with negative-strand orientation ranged from 1.3 to 94%, with a mean of 33%. Results of a representative experiment with four patients are shown in Fig. 1C. Patient 21.1.027 expressed nearly four times as much negative-strand as positive-strand *gag* RNA.

Isolation and characterization of cDNAs derived from negative-strand RNAs. To characterize the structure of negative-strand transcripts further, a cDNA library was constructed from A3.01 cells acutely infected with $HIV\text{-}1_{IIIB}$ by using the λZAPII directional bacteriophage cloning vector. Fifteen percent of the HIV-1 *gag* RNA from this source had negative-strand polarity as shown by quantitative RT-PCR experiments (data not shown). This cloning system adds *Eco*RI restriction endonuclease sites to the 5′ ends and *Xho*I sites to the 3′ ends of the cDNA generated. This allows for the orientation of the original polyadenylated mRNA template for each cDNA recovered. We screened 3×10^5 plaques with an HIV-1-specific

FIG. 1. HIV-1 RNA with negative-strand polarity as evidenced by RT-PCR. (A) Duplicate 6-ng aliquots of total cellular RNA from H9 cells acutely infected with either HXB3, RF, or MN for 3 days were subjected to RT-PCR analysis (25) either with (RT+) or without (RT−) RT. The upper row of signals resulted from use of an NSP (antiparallel to positive-strand viral RNA) during the RT step, while the lower row of signals resulted from the use of a PSP (parallel to positive-strand viral RNA) during the RT step. The polarity of the amplified cDNA is indicated at the left. The *gag* primers encompass HXB2 sequence positions 1403 to 1538. The signals were obtained by hybridization with an oligodeoxynucleotide representing an internal sequence within the fragment. (B) Quadruplicate 330-pg aliquots of poly(A)-enriched RNA from persistently infected cell lines U1 and 1119 were analyzed with the same *gag* region primers as described in panel A. An additional *nef* region primer set (HXB2 positions 9144 to 9449) and a cognate internal oligonucleotide probe were also utilized in this experiment. (C) Total RNA from three aliquots of 10^5 cryopreserved cells from four early-stage asymptomatic HIV-1-infected patients was analyzed by using a quantitative *gag* region RT-PCR strategy (25) modified for template orientation specificity as described for panel A. All signals are shown in three iterations of 10^5 cells. The patients are identified numerically. RT+ reactions performed with RT. RT−, reactions performed without RT. The polarity of the amplified signals is indicated above each group of 12 signals. Triplicate copy number standard reactions are shown at the bottom. Panels A and B were produced by digital scanning of original autoradiographs, and panel C is an image produced by a storage phosphor imaging device.

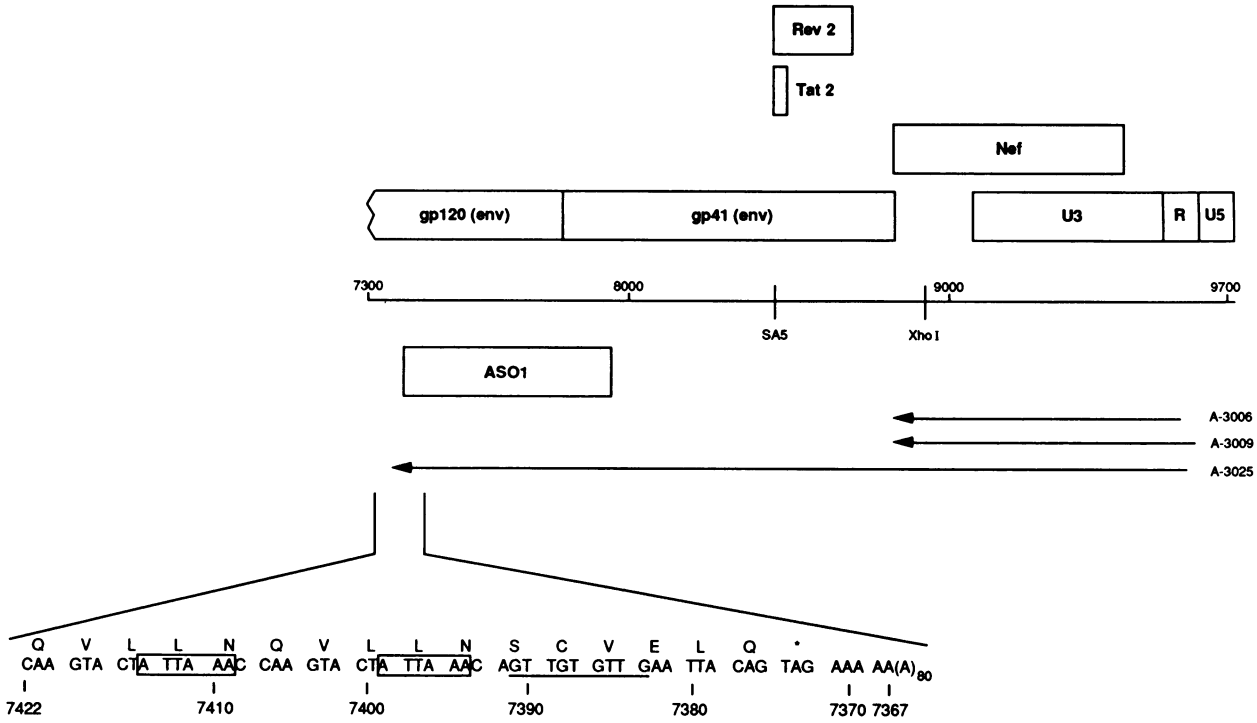


FIG. 2. Characterization of HIV-1 negative-strand RNA-derived cDNAs. A schematic representation of the genomic organization of the 3′ end of the HXB2 provirus is shown. Positive-strand genes are shown above the position ruler, and the single negative-strand ORF is shown below the ruler. The positions and polarity of three negative-strand cDNAs (arrows) are shown aligned with the genome. The precise locations of the 5′ termini of each of the cDNAs are indicated in the text. An expanded view of the 3′-terminal sequence of cDNA clone A-3025 is given in the negative-strand orientation. A translation of the carboxyl terminus of *ASO1* is given above the DNA sequence in the single-letter code. An asterisk connotes the termination codon for the gene. Putative poly(A) signals are boxed. The beginning of the posttranscriptionally synthesized 80-residue poly(A) tail is given [$(A)_{80}$]. The sequence numbering scheme is that of Ratner et al. (32). The *Xho*I restriction site is at position 8896. The SA5 major 3′ splice acceptor site is at position 8377.

probe; this resulted in 60 highly purified clones, of which 3 had negative-strand orientation upon subsequent sequence analysis across the *Xho*I sites. Complete sequencing of these cDNAs revealed a cluster of 5′ ends mapping within the R region of the 3′ LTR (Fig. 2). The 5′ ends of two of the cDNAs, A-3006 and A-3009, mapped to HXB2 positions (32) 9581 and 9618, respectively, in the R region. These two cDNAs were truncated at the *Xho*I site at position 8,896 because of incomplete protection of internal *Xho*I sites by methylated dCTP residues during first-strand synthesis of the cDNA library. However, the 5′ end of cDNA A-3025 mapped to position 9608, extended 2,242 nucleotides to position 7367, and then terminated in a poly(A) tract of approximately 80 residues. This cDNA encodes a previously identified, highly conserved 189-amino-acid negative-strand ORF located across the gp120-gp41 boundary (positions 7941 to 7372; 26) (Fig. 2). The protein predicted by this ORF is remarkable for its hydrophobicity and number of cysteine and proline residues. A-3025 acquired a poly(A) tail 5 nucleotides downstream of the TAG termination codon for this ORF. Two common variant polyadenylation signals with the sequence ATTAAA (30) were found in an unusual position, within the coding sequences, for this ORF at 48 and 33 nucleotides upstream of the poly(A) addition site. Poly(A) signals are normally found in the 3′ untranslated region of structural genes (30). A GT-rich domain, associated with efficient polyadenylation sites (30), is present 21 bp downstream of the poly(A) addition site. We propose that this ORF be named *ASO1* (antisense ORF1).

Detection of a novel promoter producing negative-strand transcripts. The location of negative-strand cDNA 5′ termini suggested that the NSP would be found in the 3′ LTR. We constructed a reporter plasmid, pαLTRCAT, containing the entire 3′ LTR and a portion of the *nef* gene placed, in the antisense orientation, upstream of the *cat* gene and the simian virus 40 polyadenylation element (Fig. 3A). The relative promoter activities of the NSP and PSP were assessed by transfection of various amounts of pαLTRCAT and a control construct containing the HIV-1 PSP region (pU3RIII; 37) into SupT1 cells by electroporation. CAT activity was assessed at 36 h posttransfection. The basal-level activity of the NSP was ninefold less than that of the PSP as measured in SupT1 cells (Table 2). These relative promoter activities were observed in similar experiments performed with two other T-cell lines (H9 and A3.01) and promonocytic line U937 (data not shown). We next determined what effect the *tat* gene product, a potent transactivator of the PSP, would have on the basal-level of activity of the NSP. pαLTRCAT or pU3RIII was separately cotransfected into SupT1 cells with either a wild-type (pSV_L-tat; 31) or a mutant ($pSV_L tat_{C25G}$) *tat* expression plasmid (Fig. 3B). $pSV_L tat_{C25G}$ differs from pSV_Ltat by a single nucleotide substitution which causes a cysteine-to-glycine change at amino acid position 25 within the cysteine-rich domain of Tat. The basal activity of the NSP carried on pαLTRCAT is shown in lane 1. An 18-fold fall in the basal activity of the NSP was caused by cotransfection of 1.0 μg of pSV_Ltat and pαLTRCAT (lane 2). This suppression was minimal with cotransfection of

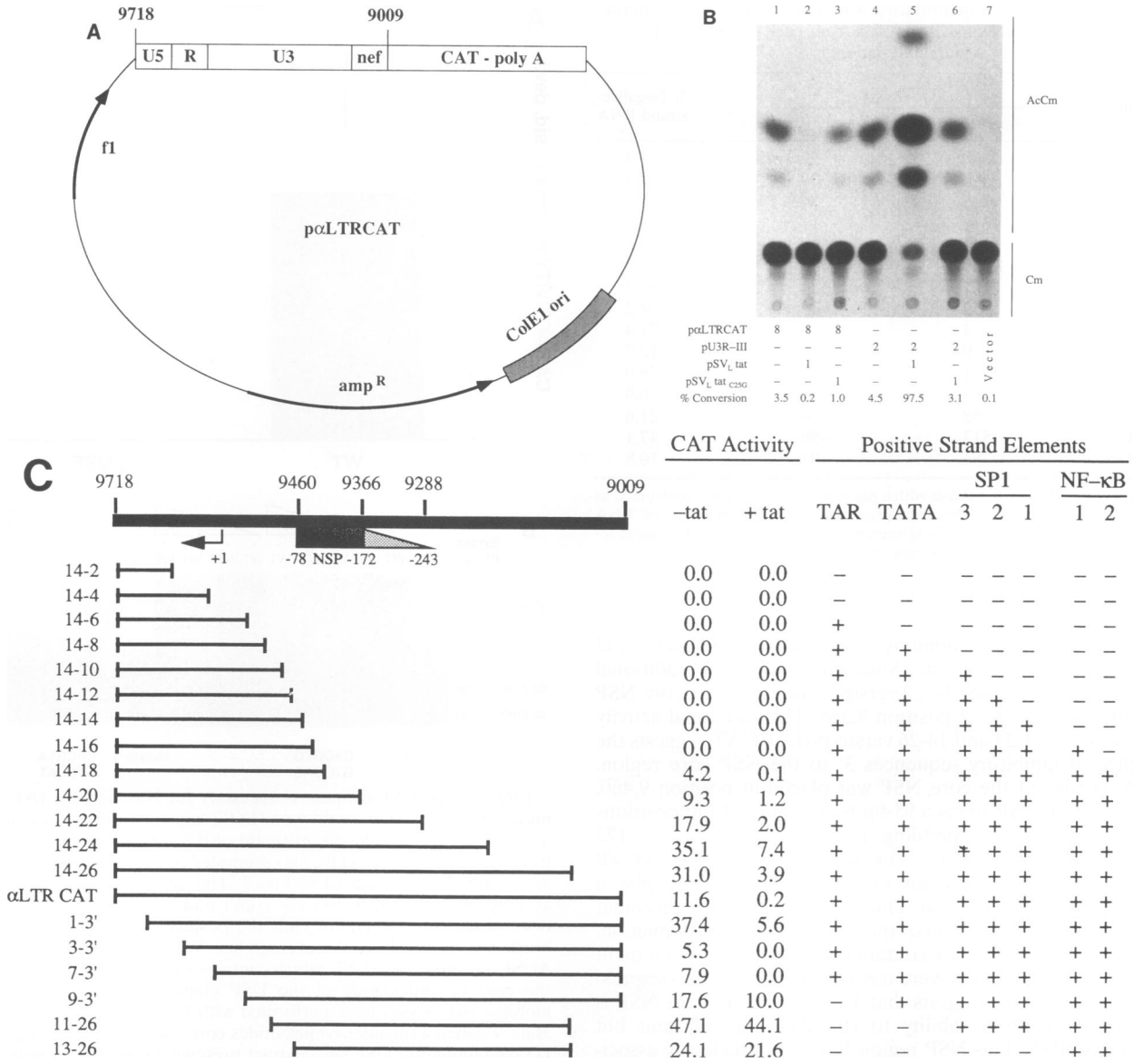


FIG. 3. Localization of the HIV-1 NSP. (A) Circular representation of the construct pαLTRCAT. The sequence positions of the HIV-1 portions of the plasmid are demarcated above the genomic organization of the LTR-*nef* region. CAT-poly A is the *cat*-poly(A) region from pSV_2CAT (14), f1 is the filamentous phage origin of replication, ColE1 ori is the plasmid origin of replication, and amp^R is the ampicillin resistance-encoding gene. (B) The effect of wild-type and mutant *tat* coexpression on pαLTRCAT was studied with a wild-type *tat* expression plasmid (pSV_Ltat) and a single-nucleotide mutation of pSV_Ltat which causes a cysteine-to-glycine change at amino acid position 25 in the transactivation domain (pSV_Ltat_{C25G}). Results of a control transfection with 20 μg of the *cat* vector without HIV-1 sequences (pSKCAT) are shown in lane 7. Cm, unacetylated chloramphenicol. AcCm, acetylated chloramphenicol. (C) Deletion mapping of the NSP performed with a nested series of 5′ and 3′ deletions of the HIV-1 portion of pαLTRCAT. The 3′ LTR region of HXB2 is shown in the negative-strand orientation aligned with the deletion clones. CAT activities, expressed as percent conversion, are shown for each clone to the right. The presence of certain elements encoded on the positive strand (TAR element, TATA box, SP1 sites, and NF-κB sites) is shown for each clone. The core NSP is shown by the solid bar, while less critical upstream sequences are shown by the faded bar. The cap site for the PSP is shown by the arrow. The positions of NSP sequences relative to the cap site for the PSP are given by the negative numbers below the line, while HXB2 sequence positions are shown above the line. −tat, without *tat* cotransfection. +tat, with *tat* cotransfection. Panel B was produced by digital scanning of the original autoradiograph.

pSV_Ltat_{C25G} (lane 3). Repeat cotransfections of pSV_Ltat_{C25G} with pαLTRCAT demonstrated that this minimal suppression was not statistically significant (data not shown). As expected, pSV_Ltat was capable of transactivating the PSP carried on pU3RIII but pSV_Ltat_{C25G} was not (lanes 4 to 6). Transfection of SupT1 cells with pSKCAT, the parental reporter construction which lacks HIV-1 sequences, resulted in no measurable promoter activity (lane 7).

The NSP maps to the negative response element. The NSP was more precisely localized by extensive 5′ and 3′ deletion mapping of the HIV-1 portion of pαLTRCAT (Fig. 3C). The 3′ border of the NSP was found to be between HXB2 positions

TABLE 1. Results of quantitative RT-PCR analysis of peripheral blood mononuclear cell samples from early-stage HIV-1-infected patients[a]

Patient	Copy no.		% Negative-strand RNA
	+*gag* RNA	−*gag* RNA	
21.1.001	1,650	22	1.3
21.1.002	390	720	64.9
21.1.003	115	4	3.4
21.1.004	684	11	1.6
21.1.005	140	69	33.0
21.1.007	1,950	130	6.3
21.1.010	21	350	94.3
21.1.012	52	27	34.2
21.1.015	4	10	71.4
21.1.021	69	10	12.7
21.1.027	51	191	78.9
21.1.031	68	5	6.9
21.1.035	58	16	21.6
21.2.A	513	456	47.1
21.2.B	754	91	10.8

[a] Quantitative RT-PCR analysis with a *gag* region primer set was performed as previously described (25). The data presented are copy numbers of RNA molecules per 10^5 peripheral blood mononuclear cells. +*gag* RNA, positive-strand determination; −*gag* RNA, negative-strand determination.

9,366 and 9,288, corresponding to positions −172 and −243 relative to the PSP cap site. Since there was little additional contribution from clone 14-22 versus clone 14-20, the core NSP 3′ border was placed at position 9,366. The additional activity seen in clones 14-24 and 14-26 versus pαLTRCAT suggests the presence of inhibitory sequences 3′ to the NSP core region. The 5′ border of the core NSP was placed at position 9,460. Thus, the core NSP spans a 95-bp region in U3 from positions 9,460 to 9,366, corresponding to positions −78 to −172 relative to the PSP cap site. This region includes both NF-κB sites but not the TATA box or the SP1 sites, which play a critical role in PSP function. Thus, the NSP has a requirement for U3 sequences upstream of those required for PSP function. Furthermore, only clones containing both the TAR element and the NSP core region were downregulated by Tat coexpression (Fig. 3C). This suggests that Tat repression of the NSP is mediated through its ability to stimulate the adjacent but antiparallel PSP. This NSP region has been previously associated with the negative response element of the LTR, which has

TABLE 2. Relative promoter strengths of the HIV-1 PSP and NSP[a]

Reporter plasmid (promoter) and amt of DNA transfected (μg)	% Conversion	% Conversion/μg of DNA
pαLTRCAT (NSP)		
2	1.1	0.6
4	2.8	0.7
pU3RIII (PSP)		
2	13.0	6.5
4	20.2	5.1

[a] The indicated amounts of plasmid DNAs were transfected into SupT1 cells by electroporation in triplicate reactions (see the legend to Fig. 3). CAT activity was determined at 48 h posttransfection. Similar relative promoter activities were determined with H9, A3.01, and U937 cells (data not shown). The PSP/NPN ratios of values for percent conversion per microgram of DNA were 10.8 and 7.2 (average, 9.0), respectively, when 2.0 and 4.0 μg of DNA was transfected.

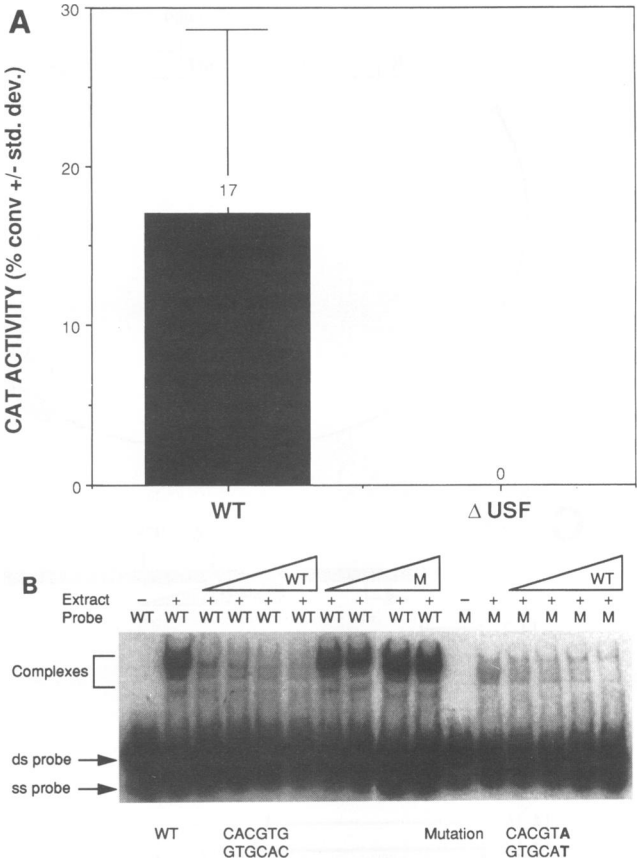


FIG. 4. The USF element is necessary for NSP activity. (A) Plasmids containing either a wild-type HXB2 sequence and one containing a mutated USF site cloned upstream of the *cat* gene were separately transfected into SupT1 cells, and promoter activities were determined as described in the legend to Fig. 3. The data are expressed as the average percent conversion (% conv) and standard deviation (std. dev.) of three transfections. Only the positive deflection of the error bracket is visible above the average bar. WT, wild-type plasmid 3-3′. ΔUSF, plasmid 5-3′-ΔUSF, which contains a single G-to-A mutation in the core hexanucleotide of the USF element. (B) Electrophoretic mobility shift assays were performed with HeLa cell nuclear extracts and ^{32}P-labeled oligodeoxynucleotides corresponding to the region in U3 containing the USF site. Extract presence or absence is indicated by a plus sign or a minus sign, respectively, at the top. The use of a wild-type (WT) or mutant (M) probe is indicated at the top. The migration of DNA-protein complexes is given to the left. The migration positions of labeled double-stranded (ds) and single-stranded (ss) probes are shown. Competition reactions are indicated by wedge-shaped symbols over contiguous groups of four lanes with competitor excesses, left to right, of 50-, 100-, 200-, and 400-fold. The symbol WT or M within a wedge symbol refers to the use of a wild-type or mutant competitor template, respectively. The USF core hexanucleotide sequences are given below both wild-type and mutant oligodeoxynucleotides. The mutated base pair is in boldface.

an inhibitory effect on both the PSP and virion production (13, 22, 23).

The USF and NF-κB I sites are crucial for NSP activity. Further delineation of the *cis*-acting sequence elements involved in NSP activity was provided through analysis of point mutations in the USF and NF-κB sites. A wild-type *cat* reporter construction and one containing a single-point mutation in the conserved hexanucleotide core sequence of the USF site (CACGTG to CACGTA) were independently transfected

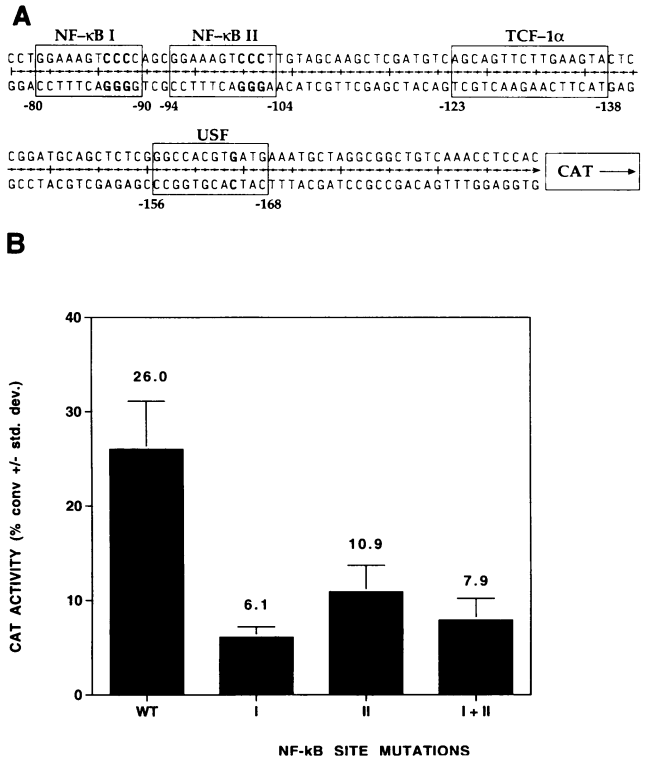


FIG. 5. Effect of NF-κB site mutations on NSP activity. (A) Nucleotide sequence of clone 11-26 in the negative-strand orientation. The two NF-κB sites, the TCF-1α site, and the USF site are boxed. The numbering scheme reflects the distance from the cap site of the PSP at the first nucleotide of the R region. The position and orientation of the *cat* gene, immediately downstream of this sequence, are shown. Regions in boldface within the NF-κB sites show the location of the triple mutation in each site. (B) Reporter plasmids containing either the wild-type HXB2 sequence contained in clone 11-26 or those containing NF-κB site mutations cloned upstream of the *cat* gene were separately transfected into SupT1 cells. Promoter activities were determined as described in the legend to Fig. 3. The data are expressed as the average percent conversion (% conv) and standard deviation (std. dev.) of three transfections. WT, wild-type plasmid 11-26. I, construction 11-26-NF-κB I containing a site I triple mutation. II, construction 11-26-NF-κB II containing a site II triple mutation. I+II, construction 11-26-NF-κB I+II containing both site I and site II triple mutations.

into SupT1 cells in triplicate, and promoter activities were assessed as shown in Fig. 4A. This single-point mutation in the USF site diminished NSP activity to undetectable levels. This activity was restored by reversion mutation (data not shown). On the basis of this finding, oligodeoxynucleotides representing both strands were synthesized to reflect either the wild-type or mutant USF site on these constructions. These DNAs were then used as probes in electrophoretic mobility shift assays with HeLa cell nuclear extract as shown in Fig. 4B. Wild-type probes shifted up into one major and two minor complexes which were eliminated by competition from excess unlabeled wild-type probe but not excess unlabeled mutant probe. Conversely, mutant probe showed minimal formation of DNA-protein complexes, and these were readily eliminated by excess unlabeled wild-type probe. These data substantiate the role of the USF binding site sequences in NSP activity. Whether or not USF polypeptides themselves are involved in these complexes awaits binding studies with purified USF or supershift experiments with antisera to this transcription factor.

The role of the two NF-κB sites in NSP activity was investigated by site-directed mutagenesis of reporter construction 11-26, whose sequence is shown in the negative-strand orientation in Fig. 5A. We made separate constructions containing mutations of each individual element, as well as both sites together. For any given NF-κB site, the CCC sequence was changed to GAG (shown in boldface). Results of triplicate CAT assays of SupT1 cells with this series of wild-type and mutant NSP sequences are shown in Fig. 5B. Mutation of NF-κB site I alone reduced NSP activity 4.3-fold, whereas mutation of site II alone reduced NSP activity by only 2.4-fold. The combined effect of mutation of both NF-κB sites was no different from the reduction in NSP activity associated with the site I mutation alone, as the standard deviation of the former assay fully overlapped with that of the latter.

DISCUSSION

The data presented here confirm the presence of HIV-1 RNA transcripts with negative-strand polarity in tissue culture models of acute and persistently infected cells. We feel that the physiologic relevance of these findings is greatly strengthened by the direct demonstration of negative-strand HIV-1 RNAs in the peripheral blood mononuclear cells of a large number of infected patients. What role do these transcripts play in the life cycle of the virus? We feel that negative-strand transcription of HIV-1 provides for expansion of the coding capacity of HIV-1 and suggests a role for antisense regulation of viral replication. These two implications of our findings will be discussed in turn.

The extended coding capacity afforded by negative-strand transcription of HIV-1 is underscored by the isolation of a cDNA that encodes the highly conserved ORF *ASO1* first described by Miller (26). The deduced amino acid sequence of *ASO1* predicts a very hydrophobic protein of 20.3 kDa that initiates with an AUG codon and is rich in cysteine, proline, and leucine residues. The isoelectric point of the predicted *ASO1*-encoded protein would be 6.993, with a charge of −0.017 at pH 7.0. Additional HIV-1 negative-strand ORFs can be found in HIV-1, but none is as well conserved or nearly as long as *ASO1*. The poly(A) signal sequences found in *ASO1*, with the sequence ATTAAA, represent the second most common variant of such elements (30). We feel that the second of the two signal sequences found in cDNA A-3025 is likely the functional element, as it, and not the first element, is conserved in HIV-1 sequence data bases. The sequence compactness of the HIV-1 proviral genome is evidenced by the location of this polyadenylation signal within the coding sequences of *ASO1* and not in the typical position within the 3′ untranslated region of the transcription unit. We are intensively searching for evidence of an *ASO1* translation product in infected cells by using both anti-peptide and anti-protein reagents (the latter were generated against bacterially synthesized material), as well as by screening for reactivity to this putative protein in a large number of patient sera. We are also searching for a phenotype for this gene by using isogenic recombinant proviruses that interrupt the *ASO1* reading frame but do not disturb the antiparallel envelope gene or RRE structure.

It is intriguing that poly(A) signal and addition sites were found to operate in the *env* region despite the fact that negative-strand RNAs could be detected in the *gag* region by RT-PCR techniques. It is possible that either other, unidentified, negative-strand promoters exist downstream of these poly(A) signals or that the *env* region poly(A) signals are leaky owing to their variant nature. We are attempting to answer this

question by building a more accurate transcription map for the HIV-1 negative strand through the use of RNA blotting, RNase protection, and 5′ rapid amplification of cDNA end techniques.

The sequence requirements of the NSP, although overlapping with those for the PSP, were found to be quite distinct. This argued strongly for the presence of other *cis*-acting elements critical for NSP activity within the minimal NSP sequences defined in this work. Mutagenesis of the conserved hexanucleotide core sequence of the USF site in reporter constructions containing both the core NSP region and regions downstream of the core implicates the USF binding site as an important *cis*-dependent region for NSP function. This was confirmed by the ability of wild-type, but not mutant, oligodeoxynucleotides representing this site to display protein-binding activity in electrophoretic mobility shift assays. Further work with purified USF and antisera to this factor is required to substantiate that it is the USF protein itself that is involved in these sequence-specific binding interactions. Coexpression of USF with NSP reporter constructions would also clarify the role of this transcription factor in NSP function. The presence of both NF-κB sites, encoded on the antiparallel positive strand, was shown to be critical to NSP function in the set of nested 3′ deletion clones. Site-directed mutagenesis of these sites, especially NF-κB site I, confirmed the role of these sequences for NSP activity. If NF-κB polypeptides were shown to partially mediate NSP activity, then this would substantiate the polarity-independent nature of this enhancer element. Confirmation of the role of NF-κB polypeptides for NSP activity will also require binding assays with purified protein and specific antisera, as well as overexpression of NF-κB in cells transfected with NSP reporter constructs. Cell lines with various levels of NF-κB expression would also be useful tools toward this end.

Analysis of the sequences contained by the NSP reveal no well-conserved TATA box. This suggests that transcriptional initiation of negative-strand transcripts is not precise, as has been shown for other promoters lacking a TATA box (36). Indeed, preliminary mapping experiments of 5′ RNA termini by RNase protection suggests that the initiation of negative-strand RNAs occurs in a ragged fashion (data not shown).

The finding that Tat downregulates the NSP must be interpreted cautiously, since the data were generated with circular reporter plasmids and not in the context of an integrated provirus. It is logical to assume that Tat mediates this effect through upregulation of the PSP, since Tat-mediated downregulation of the NSP is dependent upon an intact TAR element present in *cis*. It is possible that either helical unwinding or the presence of transcription factor complexes mobilized to the PSP sterically interferes with NSP initiation. Alternatively, PSP-initiated transcripts could conceivably interfere with NSP activity on these circular plasmids by a full cycle of procession around the vector. If this downregulatory effect on the NSP by Tat is relevant to an integrated provirus, then it is important to note that this effect was not absolute in our reporter gene experiments. Thus, the disparity in the amounts of negative- and positive-strand transcripts may be a consequence of both the weaker activity of the negative-strand promoter (Table 2) and the inhibitory effect of Tat on this promoter. The suppression of the NSP by Tat, taken together with the localization of the NSP to the negative response element, suggests a mechanism for antisense gene regulation in HIV-1. Giacca and coworkers have recently shown that the USF site is the critical domain contributing to the activity of the negative response element (13). We have shown that the core NSP contains both this region and other sequences previously shown to be associated with a negative regulatory effect on the PSP and on virion production. We have further shown that mutation of the core hexanucleotide of the USF site greatly diminishes the activity of the NSP. It is possible that the mechanism of action of the negative response element is mediated by the activity of the opposing promoter described in this report. As such, this new regulatory pathway may provide another target for therapeutic intervention in HIV disease. Elucidation of the transcriptional factors that contribute to the function of NSP and its interaction with the PSP will be critical to the dissection of the molecular mechanisms of NSP regulation and its role in the HIV-1 life cycle.

ACKNOWLEDGMENTS

We thank Donald S. Burke for support and encouragement of this work. G. Chang, J. Cooley, P. Morrow, and J. Ruderman provided expert technical assistance in the early phase of this project. We thank J. Kim for critical review of the manuscript. V. Hunter, D. Joynes, K. Hayes, and Liz Harris provided graphic design assistance.

This work was partially supported by a grant from the NIH (J.M.).

REFERENCES

1. **Anderson, S., A. T. Bankier, B. G. Barrell, M. de Bruijn, A. R. Coulson, J. Drouin, I. C. Eperon, D. P. Nierlich, B. A. Roe, F. Sanger, P. H. Schreier, A. J. Smith, R. Staden, and I. G. Young.** 1981. Sequence and organization of the human mitochondrial genome. Nature (London) **290**:457–465.
2. **Arrigo, S. J., S. Weitsman, J. A. Zack, and I. S. Chen.** 1990. Characterization and expression of novel singly spliced RNA species of human immunodeficiency virus type 1. J. Virol. **64**:4585–4588.
3. **Arya, S. K., C. Guo, S. F. Josephs, and F. Wong-Staal.** 1985. Trans-activator gene of human T-lymphotropic virus type III (HTLV-III). Science **229**:69–73.
4. **Bibb, M. J., R. A. Van Etten, C. T. Wright, M. W. Walberg, and D. A. Clayton.** 1981. Sequence and gene organization of mouse mitochondrial DNA. Cell **26**:167–180.
5. **Bukrinsky, M. I., and A. F. Etkin.** 1990. Plus strand of the HIV provirus DNA is expressed at early stages of infection AIDS Res. Hum. Retroviruses **6**:425–426.
6. **Chang, D. D., and D. A. Clayton.** 1986. Precise assignment of the light-strand promoter of mouse mitochondrial DNA: a functional promoter consists of multiple upstream domains. Mol. Cell. Biol. **6**:3253–3261.
7. **Chang, D. D., and D. A. Clayton.** 1986. Precise assignment of the heavy-strand promoter of mouse mitochondrial DNA: cognate start sites are not required for transcriptional initiation. Mol. Cell. Biol. **6**:3262–3267.
8. **Cullen, B. R., and W. C. Greene.** 1989. Regulatory pathways governing HIV-1 replication. Cell **58**:423–426.
9. **Feinberg, M. B., D. Baltimore, and A. D. Frankel.** 1991. The role of Tat in the human immunodeficiency virus life cycle indicates a primary effect on transcriptional elongation. Proc. Natl. Acad. Sci. USA **88**:4045–4049.
10. **Feinberg, M. B., R. F. Jarrett, A. Aldovini, R. C. Gallo, and F. Wong-Staal.** 1986. HTLV-III expression and production involve complex regulation at the levels of splicing and translation of viral RNA. Cell **46**:807–817.
11. **Fisher, R. P., J. N. Topper, and D. A. Clayton.** 1987. Promoter selection in human mitochondria involves binding of a transcription factor to orientation-independent upstream regulatory elements. Cell **50**:247–258.
12. **Folks, T. M., J. Justement, A. Kinter, S. Schnittman, J. Orenstein, G. Poli, and A. S. Fauci.** 1988. Characterization of a promonocyte clone chronically infected with HIV and inducible by 13-phorbol-12-myristate acetate. J. Immunol. **140**:1117–1122.
13. **Giacca, M., M. I. Gutierrez, S. Menzo, F. F. Di, and A. Falaschi.** 1992. A human binding site for transcription factor USF/MLTF mimics the negative regulatory element of human immunodeficiency virus type 1. Virology **186**:133–147.
14. **Gorman, C. M., L. F. Moffat, and B. H. Howard.** 1982. Recombi-

nant genomes which express chloramphenicol acetyltransferase in mammalian cells. Mol. Cell. Biol. **2**:1044–1051.
15. **Greene, W. C.** 1990. Regulation of HIV-1 gene expression. Annu. Rev. Immunol. **8**:453–475.
16. **Guatelli, J. C., T. R. Gingeras, and D. D. Richman.** 1990. Alternative splice acceptor utilization during human immunodeficiency virus type 1 infection of cultured cells. J. Virol. **64**:4093–4098.
17. **Hahn, B. H., G. M. Shaw, S. K. Arya, M. Popovic, R. C. Gallo, and F. Wong-Staal.** 1984. Molecular cloning and characterization of the HTLV-III virus associated with AIDS. Nature (London) **312**:166–169.
18. **Jacks, T., M. D. Power, F. R. Masiarz, P. A. Luciw, P. J. Barr, and H. E. Varmus.** 1988. Characterization of ribosomal frameshifting in HIV-1 gag-pol expression. Nature (London) **331**:280–283.
19. **Kao, S. Y., A. F. Calman, P. A. Luciw, and B. M. Peterlin.** 1987. Anti-termination of transcription within the long terminal repeat of HIV-1 by tat gene product. Nature (London) **330**:489–493.
20. **Larocca, D., L. A. Chao, M. H. Seto, and T. K. Brunck.** 1989. Human T-cell leukemia virus minus strand transcription in infected T-cells. Biochem. Biophys. Res. Commun. **163**:1006–1013.
21. **Laspia, M. F., A. P. Rice, and M. B. Mathews.** 1989. HIV-1 Tat protein increases transcriptional initiation and stabilizes elongation. Cell **59**:283–292.
22. **Lu, Y., M. Stenzel, J. G. Sodroski, and W. A. Haseltine.** 1989. Effects of long terminal repeat mutations on human immunodeficiency virus type 1 replication. J. Virol. **63**:4115–4119.
23. **Lu, Y. C., N. Touzjian, M. Stenzel, T. Dorfman, J. G. Sodroski, and W. A. Haseltine.** 1990. Identification of *cis*-acting repressive sequences within the negative regulatory element of human immunodeficiency virus type 1. J. Virol. **64**:5226–5229.
24. **Michael, N. L., J. B. Rothbard, R. A. Shiurba, H. K. Linke, G. K. Schoolnik, and D. A. Clayton.** 1984. All eight unassigned reading frames of mouse mitochondrial DNA are expressed. EMBO J. **3**:3165–3175.
25. **Michael, N. L., M. Vahey, D. S. Burke, and R. R. Redfield.** 1992. Viral DNA and mRNA expression correlate with the stage of human immunodeficiency virus (HIV) type 1 infection in humans: evidence for viral replication in all stages of HIV disease. J. Virol. **66**:310–316.
26. **Miller, R. H.** 1988. Human immunodeficiency virus may encode a novel protein on the genomic DNA plus strand. Science **239**:1420–1422.
26a. **Mosca, J.** Personal communication.
27. **Muesing, M. A., D. H. Smith, C. D. Cabradilla, C. V. Benton, L. A. Lasky, and D. J. Capon.** 1985. Nucleic acid structure and expression of the human AIDS/lymphadenopathy retrovirus. Nature (London) **313**:450–458.
28. **Muesing, M. A., D. H. Smith, and D. J. Capon.** 1987. Regulation of mRNA accumulation by a human immunodeficiency virus trans-activator protein. Cell **48**:691–701.
29. **Nabel, G., and D. Baltimore.** 1987. An inducible transcription factor activates expression of human immunodeficiency virus in T cells. Nature (London) **326**:711–713.
30. **Nevins, J. R.** 1983. The pathway of eukaryotic mRNA formation. Annu. Rev. Biochem. **52**:441–466.
31. **Rappaport, J., S. J. Lee, K. Khalili, and F. Wong-Staal.** 1989. The acidic amino-terminal region of the HIV-1 Tat protein constitutes an essential activating domain. New Biol. **1**:101–110.
32. **Ratner, L., W. Haseltine, R. Patarca, K. J. Livak, B. Starcich, S. F. Josephs, E. R. Doran, J. A. Rafalski, E. A. Whitehorn, K. Baumeister, L. Ivanoff, S. R. Petteway, M. L. Pearson, J. A. Lautenberger, T. S. Papas, J. Ghrayeb, N. T. Change, R. C. Gallo, and F. Wong-Staal.** 1985. Complete nucleotide sequence of the AIDS virus, HTLV-III. Nature (London) **313**:277–284.
33. **Rosen, C. A., J. G. Sodroski, and W. A. Haseltine.** 1985. The location of cis-acting regulatory sequences in the human T cell lymphotropic virus type III (HTLV-III/LAV) long terminal repeat. Cell **41**:813–823.
34. **Schwartz, S., B. K. Felber, D. M. Benko, E. M. Fenyo, and G. N. Pavlakis.** 1990. Cloning and functional analysis of multiply spliced mRNA species of human immunodeficiency virus type 1. J. Virol. **64**:2519–2529.
35. **Selby, M. J., E. S. Bain, P. A. Luciw, and B. M. Peterlin.** 1989. Structure, sequence, and position of the stem-loop in tar determine transcriptional elongation by tat through the HIV-1 long terminal repeat. Genes Dev. **3**:547–558.
36. **Smale, S. T., and D. Baltimore.** 1989. The "initiator" as a transcription control element. Cell **57**:103–113.
37. **Sodroski, J., R. Patarca, C. Rosen, F. Wong-Staal, and W. Haseltine.** 1985. Location of the trans-activating region on the genome of human T-cell lymphotropic virus type III. Science **229**:74–77.
38. **Stevens, J. G., L. Haarr, D. D. Porter, M. L. Cook, and E. K. Wagner.** 1988. Prominence of the herpes simplex virus latency-associated transcript in trigeminal ganglia from seropositive humans. J. Infect. Dis. **158**:117–123.
39. **Stevens, J. G., E. K. Wagner, G. B. Devi-Rao, M. L. Cook, and L. T. Feldman.** RNA complementary to a herpesvirus alpha gene mRNA is prominent in latently infected neurons. Science **235**: 1056–1059.
40. **Varmus, H.** 1988. Retroviruses. Science **240**:1427–1435.