JOURNAL OF VIROLOGY, June 1991, p. 2987–2992
0022-538X/91/062987-06$02.00/0

Vol. 65, No. 6

Characterization of Multiple mRNA Species of Simian Immunodeficiency Virus from Macaques in a CD4$^+$ Lymphoid Cell Line

IN-WOO PARK, ROBERT STEEN, AND YEN LI*

New England Regional Primate Research Center and Department of Microbiology and Molecular Genetics, Harvard Medical School, Southborough, Massachusetts 01772

Received 14 November 1990/Accepted 4 March 1991

Cytoplasmic poly(A)$^+$ RNA was isolated from CEMX721.174 cells 5 to 10 days after infection with molecularly cloned simian immunodeficiency virus SIVmac239. Expression of SIV RNA was analyzed by Northern (RNA) blot hybridization and by sequencing of cDNA clones. As expected, a splice donor site was demonstrated in the untranslated leader sequence outside the left long terminal repeat. The region between *pol* and *env* was found to contain at least two splice donor and six splice acceptor sites. Splice acceptor and donor sites in the intergenic region were suitably positioned for expression of *vpx*, *vpr*, *tat*, and *rev*. Splice acceptor sites at nucleotides 8802 and 8805 were demonstrated in singly and doubly spliced RNAs with the potential of expressing *nef* and the second exons of *tat* and *rev*. Our results demonstrate a complex pattern of alternative splicing of SIV mRNAs. The results are very similar to what has been observed in human immunodeficiency virus type 1-infected cells, suggesting that both human and simian immunodeficiency viruses are subject to multiple levels of regulation.

The simian immunodeficiency viruses (SIVs) are a diverse group of nonhuman primate lentiviruses that exhibit extensive similarity with the human immunodeficiency viruses (HIVs) in morphology, biological properties, and genome organization (7, 8, 22, 27, 29). Both viruses contain at least five or six accessory genes in addition to the standard retrovirus structural genes *gag*, *pol*, and *env* (5, 14). Some of the accessory genes such as *tat* and *rev* are absolutely essential for viral replication, and other accessory genes have been shown, in some cases, to be dispensable in vitro. Each of the eight or nine genes in the SIV genome overlaps with at least one of its neighboring genes. This type of gene structure is also observed in the HIV genome. It is unclear, however, what mechanisms regulate the expression of these overlapping genes with only a single promoter in the viral long terminal repeat (LTR).

In addition to transcriptional regulation, it appears that posttranscriptional and translational regulation may play a role in regulating the expression of the overlapping genes. Production of monocistronic mRNAs for each protein via splicing is one possible way to achieve the expression of the overlapping coding regions. There have been reports of alternative splicing of full-length transcripts to produce multiply spliced viral RNAs (1, 6, 16, 33, 34). Additionally, the selective transport of unspliced or singly spliced transcripts by Rev protein has been reported as another type of posttranscriptional regulation (12, 23, 24). Accumulation of these mRNAs in the cytoplasm is known to be mediated by *rev*-responsive element and Rev protein (9, 10, 17, 18, 23, 24). The latter is also indirectly involved in the down regulation of *tat*, *rev*, and *nef* gene expression (13). However, alternative splicing and selective transport by Rev protein may be only part of the complex regulatory network of the primate lentiviral expression. Translational regulation may also play a role in the expression of overlapping genes. For example, the overlapping reading frames of *gag* and *pol* are expressed by a ribosomal frameshift to produce a Gag-Pol fusion protein which is in turn processed to mature Pol proteins by the viral protease (20, 21).

In this study, a combination of Northern (RNA) blot analysis and cDNA cloning and sequencing was used to elucidate the viral mRNA structure and mode of splicing of the accessory genes. The mRNAs of SIVmac239 in a clonal human T × B somatic hybrid cell line, CEMX721.174, were analyzed by Northern blot analysis. The mRNAs were further characterized by cDNA cloning and sequencing. The cDNA clones were constructed by a traditional method instead of using the polymerase chain reaction to minimize possible artifacts (11). The results revealed a complex pattern of splicing of the viral RNA and suggested that both SIV and HIV-1 employ alternative splicing for the expression of overlapping open reading frames.

MATERIALS AND METHODS

Viruses and cells. The CD4$^+$ human T × B somatic hybrid cell line CEMX721.174 (19) was maintained in RPMI 1640 medium with 10% fetal bovine serum. SIVmac239 was obtained by transfection of a full-length molecular clone into macaque peripheral blood lymphocytes (28). An inoculum of the virus containing 10^5 cpm of reverse transcriptase activity was infected into 2×10^6 CEMX721.174 cells, and the infected cells were maintained in 40 ml of culture medium by fourfold dilution every 5 days.

Synthesis of oligonucleotides. Antisense (complementary to the mRNA) oligonucleotides were synthesized in a Cyclone DNA synthesizer (Milligen/Biosearch), using solid-phase phosphoramidite synthesis chemistry, and purified by affinity column chromatography by using oligonucleotide purification cartridges (Applied Biosystems). The sequences and nucleotide positions of each oligonucleotide primer and probe are based on the SIVmac239 proviral sequence and numbering system (31) and are as follows: *vif*-specific oligo-

* Corresponding author.

nucleotide probe, 5′-AAATGCCAATACCCTTGTACTTCT AAATGG-3′ (nucleotide positions 5563 to 5534); *vpx*-specific oligonucleotide probe, 5′-CATGTCCTTCCCCTAGAC ATCTACAGCCTT-3′ (nucleotide positions 6094 to 6065); *vpr*-specific oligonucleotide probe, 5′-GTTCCCTTTGTGGT CCTTCATTTTCTGGAG-3′ (nucleotide positions 6193 to 6164); *tat*-specific oligonucleotide probe, 5′-ACTGGCAA TGGTAGCAACACTTTTACAAT-3′ (nucleotide positions 6494 to 6465); *rev*-specific probe, 5′-ATTGGTTGGAGGAT CTGGTATACTCTCGAT-3′ (nucleotide positions 9005 to 8976); *nef*-specific oligonucleotide probe, 5′-TCTCTCTTCA GCTGGGTTTCTCCATGGAGT-3′ (nucleotide positions 9307 to 9378); *vif* first-strand cDNA primer, 5′-GTCTGCAT AGTTTGGTG-3′ (nucleotide positions 5657 to 5641); *vpx* and *vpr* first-strand cDNA primer, 5′-CAAGTGCAGTTA GCAAGCG-3′ (nucleotide positions 6280 to 6262); *tat*, *rev*, and *nef* first-strand cDNA primer, 5′-TGAATACAGAGC GAAATGCAG-3′ (nucleotide positions 9980 to 9960); and colony-screening oligonucleotide probe, 5′-GCAGGGAA CACCCAGGCTCTACCTGCTAG-3′ (nucleotide positions 600 to 572).

Analysis of viral mRNAs. Cytoplasmic RNA was isolated from the SIVmac239-infected CEMX721.174 cells at either 5 or 10 days postinfection. Nuclei were removed by low-speed centrifugation in 0.5% Nonidet P-40–50 mM Tris-HCl (pH 8.0)–100 mM NaCl–5 mM $MgCl_2$. Cytoplasmic mRNA was selected with poly(A) Quick columns (Stratagene). After denaturation in formaldehyde-formamide, the poly(A)$^+$ RNA was fractionated in an agarose gel containing formaldehyde (2). The RNA was transferred to nitrocellulose membranes (Schleicher & Schuell) and hybridized either with a nick-translated viral LTR or *env*-specific probe or with kinase-labeled gene-specific synthetic antisense oligonucleotide probes. Prehybridization and hybridization with nick-translated probes have been described elsewhere (2). For the kinase-labeled gene-specific synthetic oligonucleotide probes, prehybridization was performed at 42°C for 16 h in prehybridization buffer containing 6× SSC (1× SSC is 0.15 M NaCl and 0.015 M sodium citrate), 10× Denhardt's solution (1× Denhardt's solution is 0.02% bovine serum albumin, 0.02% Ficoll, and 0.02% polyvinylpyrrolidone), 0.5% sodium dodecyl sulfate (SDS), 50 mM Na_2HPO_4 (pH 6.5), and 0.1 mg of denatured salmon sperm DNA (Sigma) per ml. Hybridization was carried out at 42°C for 24 h in fresh prehybridization buffer containing the appropriate oligonucleotide probes. After hybridization, filters were washed with 6× SSC–0.05% sodium pyrophosphate at 42°C for 30 min and then washed once at 55°C for 30 min.

Construction of cDNA library. Double-stranded cDNA was synthesized by using virus-specific first-strand cDNA primer according to the protocol of the manufacturer (Bethesda Research Laboratories). Briefly, 4 μg of cytoplasmic poly(A)$^+$ RNA was mixed with the appropriate first-strand cDNA synthesis primer and Moloney murine leukemia virus reverse transcriptase. Second-strand synthesis was accomplished with *Escherichia coli* DNA polymerase I, deoxynucleoside triphosphate, and *E. coli* RNase H. The double-strand cDNA was tailed with dCTP by using terminal deoxynucleotidyltransferase (Pharmacia) (30) and annealed to the dG-tailed pUC9 vector (Pharmacia). The annealed cDNA was transformed into *E. coli* XL-1 Blue competent cells (Stratagene).

Colony hybridizations. Colonies were transferred to nitrocellulose filters and hybridized as described previously (15). The filters were prehybridized at 60°C for 1 h in 5× SSC–3× Denhardt's solution–0.1% SDS–200 μg of denatured salmon sperm DNA per ml. Hybridizations were performed at 60°C for 16 h in the fresh prehybridization buffer containing ^{32}P-end-labeled colony-screening oligonucleotide probe. The filters were washed twice at room temperature for 30 min and once at 60°C for 10 min in 2× SSC–0.2% SDS and then exposed to X-ray film with an intensifying screen at −70°C.

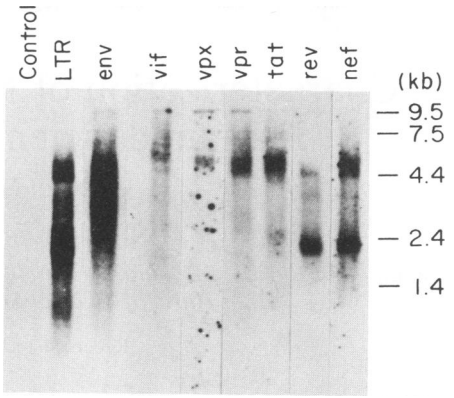

FIG. 1. Northern blot analysis of viral mRNAs. At 5 or 10 days postinfection, cytoplasmic poly(A)$^+$ mRNAs were denatured, transferred to a nitrocellulose filter, and hybridized with the gene-specific probes indicated above the lanes: LTR, LTR probe containing the U3 and R regions of the LTR; env, *env*-specific DNA fragment probe containing sequence not overlapping with any other known genes (from 6624 to 8304); vif, vpx, vpr, tat, rev, and nef, the gene-specific oligonucleotide probes used in the hybridization. Sequences of the oligonucleotides are given in Materials and Methods.

DNA sequencing. Double-stranded plasmid DNA was prepared from 2 ml of a stationary-growth-phase bacterial culture by a standard procedure as described previously (4) and denatured in 0.2 N NaOH–2 mM EDTA for 5 min at 22°C. Denatured DNA was precipitated with 2 volumes of absolute ethanol, washed once with 70% ethanol, and dissolved in 30 μl of H_2O. For each sequencing reaction, 10 ng of primer was annealed to ~3 μg (7 μl) of denatured DNA in Sequenase reaction buffer (U.S. Biochemical Corp.), and the reaction was carried out with Sequenase and [α-^{35}S]dATP according to the protocol of the manufacturer.

RESULTS

Northern blot analysis of viral mRNAs. Poly(A)$^+$ RNA was isolated from SIVmac239-infected CEMX721.174 cell cytoplasm by oligo(dT)-cellulose column chromatography at either 5 or 10 days postinfection. Approximately equivalent amounts of mRNAs were denatured, loaded on a denaturing agarose gel, and analyzed by Northern blot hybridization (2).

The LTR probe hybridization revealed multiple RNA species ranging from the full-length genomic 9.6-kb RNA to the subgenomic 5.4-, 5.0-, 4.5-, 3.1-, 2.8-, 1.7- to 2.0-, 0.7-, and 0.6-kb mRNAs (Fig. 1). The envelope-specific probe detected genomic and all intermediate-size RNA species, including 3.1- and 2.8-kb transcripts (Fig. 1). The detection of 0.7- and 0.6-kb RNAs by the LTR probe but not by any other gene-specific probes suggested that these RNAs contained primarily LTR sequences. The *vif*-specific probes detected a 5.4-kb mRNA (major) and 6.0- and 4.5-kb mRNAs (minor) but no smaller mRNA species that could be a possible monocistronic *vif* mRNA (Fig. 1), suggesting that the Vif protein might be encoded by at least one of the *env*-containing mRNAs. A 5.0-kb RNA was detected by the

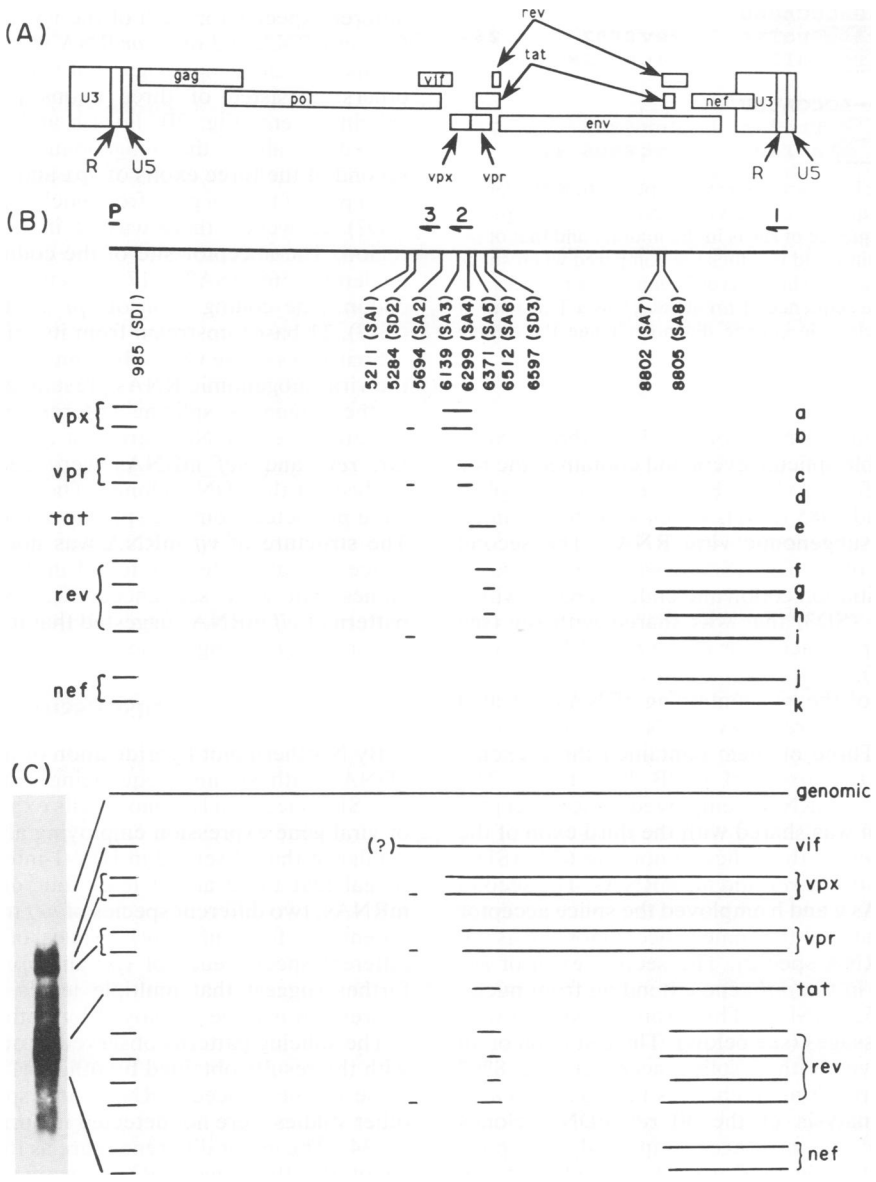

FIG. 2. (A) Genomic organization of SIVmac239. The arrows (1, 2, and 3) and bar (P) represent cDNA synthesis oligonucleotide primers and the colony-screening probe, respectively; see Materials and Methods for sequences. Primer 1 was used for *tat*, *rev*, and *nef* first-strand cDNA synthesis; primer 2 was used for *vpx* and *vpr* first-strand cDNA synthesis; primer 3 was used for *vif* first-strand cDNA synthesis. (B) Splicing pattern of viral messages. Exons are represented by solid lines. The splice donor (SD) and acceptor (SA) sites are all identified by cDNA sequencing. The corresponding nucleotide number of each splice site is indicated at the left of each splice junction. (C) Viral RNA structures. The viral RNA structures of *tat*, *rev*, and *nef* were determined by cDNA cloning and sequencing. The viral RNA structures of *vpx* and *vpr* are predicted from the splicing pattern and Northern hybridization pattern. The RNA structure of *vif* is not clear.

vpx probe, and a 4.5-kb RNA species was detected by the *vpr* probe (Fig. 1), indicating that both Vpx and Vpr might also be encoded by *env*-containing mRNAs. The *tat*-, *rev*-, and *nef*-specific probes hybridized to a cluster of 1.7- to 2.0-kb RNAs in addition to the genomic and intermediate-size (4.5- to 5.4-kb) mRNAs (Fig. 1). The 3.1- and 2.8-kb RNA species were detected by the LTR, *env*, *rev*, and *nef* probes but not by the *vif*, *vpx*, and *vpr* probes (Fig. 1).

Analysis of cDNA clones of *tat*, *rev*, and *nef*. To further characterize the viral RNAs, cDNA libraries were constructed by a traditional cDNA synthesis method, using cloned Moloney murine leukemia virus reverse transcriptase and an oligonucleotide primer complementary to nucleotide positions 9980 to 9960 at the end of the U3 region of the LTR. After transformation of competent bacterial cells, colonies were screened by filter hybridization by using a ^{32}P-end-labeled oligonucleotide probe that was complementary to nucleotide positions 572 to 600 in the R region of the LTR. Plasmid DNAs from positive colonies were prepared and sequenced. Of the 34 sequenced cDNA clones, 1 was *tat*, 3 were *nef*, and 30 were *rev* clones.

The fact that only one *tat* cDNA clone was obtained in these analyses was consistent with the results of the Northern blot hybridization in which *tat* mRNA was least abun-

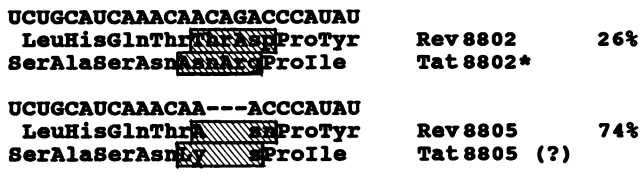

FIG. 3. Amino acid changes at splice acceptor site 8802 versus 8805. The nucleotide sequence of the viral RNA is at the top, the predicted amino acid sequence of *rev* is in the middle, and that of *tat* is at the bottom. The amino acid residues are compared when splice sites 8802 and 8805 are used. The different amino acid residues are boxed. *, The nucleotide sequence of *tat* at 8803 was a T instead of a C. However, this nucleotide change did not change the coding amino acid.

dant (Fig. 1). The *tat* cDNA consisted of three exons generated from a double splicing event and contained the *rev* and *nef* open reading frames (Fig. 2B, line e). The first splice donor was at nucleotide 985 (SD1), which was the common splice donor for all subgenomic viral RNAs. The second exon started at nucleotide 6299 (SA4), which was 2 bases upstream of the *tat* initiation codon and ended at 6597, which was the splice donor (SD3) that was shared with *rev* (see below). The second splice acceptor of the *tat* mRNA was at nucleotide 8802 (SA7).

Sequence analysis of the *rev*-containing cDNAs revealed that there were four different *rev* mRNAs generated by alternative splicing. Three of them contained three exons, and one consisted of four exons (Fig. 2B, lines f to i). The second exon of the *rev* mRNA f employed splice acceptor site 6371 (SA5), which was shared with the third exon of the *rev* mRNA i and ended at the splice donor site 6597 (SD3) that was shared by all of the *rev* and *tat* mRNAs. The second exon of the *rev* mRNAs g and h employed the splice acceptor site 6512 (SA6) and ended at the same splice donor site as did all of the other *rev* RNA species. The second exon of *rev* mRNA i was located in the *pol* gene extending from nucleotides 5211 (SA1) to 5284 (SD2). This exon was shared with some *vpx* and *vpr* messages (see below). The last exon of all *rev* messages employed either splice acceptor site 8802 (SA7) or 8805 (SA8), resulting in changes in two amino acid residues (Fig. 3). Analysis of the 30 *rev* cDNA clones revealed that 26% of the messages employed the splice acceptor site at nucleotide 8802 (SA7) and 74% of the messages spliced at nucleotide 8805 (SA8). Whether *tat* might also employ nucleotide 8805 as the splice acceptor site for its last exon remains to be determined. It should be noted that all four *rev* mRNAs were *nef*-containing RNA species.

Two different species of *nef* mRNAs from three cDNA clones were detected (Fig. 2B, lines j and k). Both species were generated from a single splicing event, but the splice acceptor sites were different: one acceptor site was at nucleotide 8802 (Fig. 2B, line j), and the other was at nucleotide 8805 (Fig. 2B, line k). There was no AUG triplet between these splice sites and the initiation codon of the *nef* RNA, suggesting that the sequences between them were part of the untranslated leader sequence of *nef*.

Analysis of cDNA clones of *vpx* and *vpr*. To obtain cDNA clones representing *vif*, *vpx*, and *vpr* mRNAs, cDNA was synthesized by using a primer complementary to the sequence from nucleotides 6280 to 6262 in the *vpr* region. Positive colonies were identified with the same colony-screening oligonucleotide probe (nucleotides 572 to 600) that was used to identify *tat*-, *rev*-, and *nef*-positive colonies. Sequence analysis of the positive cDNA clones revealed two different species for each of the *vpx* and *vpr* RNAs (Fig. 2). One *vpx* RNA and one *vpr* RNA were each made up of two exons by a single splicing event (Fig. 2B, lines a and c), the others consisted of three exons produced by a double splicing event (Fig. 2B, lines b and d). The first exon was shared by all of the subgenomic RNA species, and the second of the three exons of *vpx* and *vpr* RNAs was derived from part of the *pol* gene from nucleotides 5211 (SA1) to 5284 (SD2). However, there was no initiation codon within this region. The acceptor site of the coding exon of *vpx* was at nucleotide 5694 (SA2), 117 bases upstream from its initiation codon. The coding exon of *vpr* started at nucleotide 6139 (SA3), 11 bases upstream from its initiation codon.

Structures of the viral subgenomic RNAs. The structures of the viral subgenomic RNAs presented in Fig. 2C were based on the sequences, splicing patterns, and sizes of the mRNAs as determined by Northern blot analysis. The structures of *tat*, *rev*, and *nef* mRNAs were deduced from sequence analysis of the cDNA clones. The structures of *vpx* and *vpr* were predicted from the splicing and hybridization patterns. The structure of *vif* mRNA was not determined, since no splice acceptor site was found in more than 20 *vif*-specific clones that were sequenced. The Northern hybridization pattern of *vif* mRNA suggested that it might contain *tat*, *rev*, *env*, and *nef* coding sequences.

DISCUSSION

By Northern blot hybridization of the viral mRNA and by cDNA synthesis and sequencing, our results demonstrate that SIVmac239 in lymphoid cells exhibits a complex pattern of viral gene expression employing alternative splicing very similar to that observed in HIV-1-infected cells. The results reveal that there are at least four different species of *rev* mRNAs, two different species of *nef* mRNAs, one (however, potentially four different) species of *tat* mRNA, and two different species each of *vpx* and *vpr* mRNAs. Our results further suggest that multiple levels of regulation of gene expression may be employed by both SIV and HIV-1.

The splicing patterns observed in our study are consistent with the results obtained by others (1, 16, 33, 34). However some of the spliced mRNAs and splice sites observed in other studies were not detected in ours and vice versa (1, 16, 33, 34). The major differences are as follows. (i) In our study, all of the three *nef* mRNAs were derived from a single splicing event, which is consistent with results of another SIVmac splicing study (34), whereas HIV-1 *nef* mRNAs were either double or triple splicing products (33). (ii) In contrast to what has been found in HIV-1 infection, none of the SIVmac *tat*, *rev*, and *nef* mRNAs contained sequences derived from the *vif* region (33, 34). (iii) SIVmac *tat*, *rev*, and *nef* mRNAs employed at least two different acceptor sites (SA7 and SA8 or their equivalents in another study) for the second coding exon. However, the frequencies of usage of these two splicing acceptor sites were different in the two studies: 74% of the *rev* mRNA utilized 8805 as the splice acceptor site in our study, whereas 35 to 40% of the *rev* mRNA utilized 8805 equivalent in the other study (34). (iv) The first coding exon of *rev* mRNA utilized SA5 (6371) and SA6 (6512) with equal frequency, whereas Viglianti et al. (34) found that all of their *rev* mRNAs utilized only one splice acceptor site (SA6 equivalent). The differences among these different studies are most likely attributed to the following factors: first, different strains of virus and host cells have been used in these studies (HIV-1 and SIVmac251 versus SIV mac239; CEM and HuT-78 versus CEMX721.174) and there-

Splice Site	Position	Nucleotide Sequence	Notes
SD1	985	UUGCAG/GUAAGUGCA	5′ splice donor for all subgenomic messages
SD2	5284	UUAAAG/GUAGGGACA	vpx, vpr, rev
SD3	6597	AAACAA/GUAAGUAUG	rev, tat
		---A/CAG/GUA/GAGU---	consensus splice donor sequence
SA1	5211	UCUAUUACAG/AGAAGG	vpx, vpr, rev
SA2	5694	UGCUUUACAG/CGGGAG	vpx
SA3	6139	CCCCCUCCAG/GACUAG	vpr
SA4	6299	UCUAUAAUAG/ACAUGG	tat
SA5	6371	UGCAUUUCAG/AGGCGG	rev
SA6	6512	CUUAAAAAAG/GCUUGG	rev
SA7	8802	UUAUUUCCAG/CAGACC	rev, tat*, nef
SA8	8805	UUUCCAGCAG/ACCCAU	rev, nef
		UUUUUU-C/CCCCCC U AG/------	consensus splice acceptor sequence

FIG. 4. Splice donor and acceptor sequences. The splice donor and acceptor sites are determined by nucleotide sequence analysis of the cDNAs. Some splice donor and acceptor sites are utilized by multiple genes, and some splice acceptor sites are probably gene specific. The splice donor and acceptor consensus sequences were obtained from Mount (25). *, The nucleotide sequence of *tat* at 8803 was a T instead of a C. This nucleotide change most likely represents a mutation in our *tat* cloning, since all *rev* cDNAs have a C in that position. The original SIVmac239 also contains a C at 8803. The C→T change will not affect the GU-AG rule.

fore may lead to different spliced RNA products; second, different approaches have been used in constructing cDNA clones. The polymerase chain reaction was used in other studies. The ability of this technique to amplify sequences that are present in very low copy number may explain why low-abundance mRNAs may not have been detected in our studies. Other contributing factors may be the time and multiplicity of infection and whether the infection is chronic or acute. For example, RNA was isolated from 5- or 10-day-postinfection cultures in our studies, whereas Viglianti et al. isolated their RNA from 31-day-postinfection cultures (34). All of these factors may contribute to the differences observed between our and other studies. As shown in Fig. 4, the splice donor and acceptor consensus sequences (25) suggest that no spurious splicing has occurred.

As previously mentioned, mRNAs present in low copy number may be underrepresented in our study. For example, the *tev* message detected in HIV-1-infected cells (3, 32) was not observed in our cDNA cloning. Our Northern blot analysis using LTR and gene-specific probes reveals that the 3.1- and 2.8-kb mRNAs hybridize with the LTR, *rev*, *nef*, and *env* probes, indicating that these two mRNA species are the most likely candidates for *tev* even though *tat* hybridization is not distinct. Further study is needed to determine the identity of these RNA species. The splicing patterns of *env* and *vif* have not been identified in this report. The primers chosen in our studies will not allow us to detect *env*-specific splicing patterns. However, we suspect that *env* expression may employ any one of the five splice acceptor sites (SA2, SA3, SA4, SA5, and SA6 in Fig. 2B) which are all properly positioned for *env* utilization. Experiments are under way to determine which one (or ones) is actually used in *env* RNA synthesis. We have attempted to obtain splicing information for *vif* by using *vif*-specific primers. Several cDNA libraries have been constructed and screened. Sequence analysis of more than 20 *vif*-containing clones did not reveal splice acceptor sites near the *vif* coding exon. The inability to obtain *vif* splicing information may be due to the low abundance of *vif* mRNA in infected cells at the time of the RNA isolation or due to the fact that the *vif* splice acceptor site is further upstream than our sequence showed. The third possibility is that *vif* may be translated from the genomic mRNA by the ribosome frameshift mechanism. However, sequence analysis and secondary structure search revealed neither the putative ribosomal frameshift sequences nor potential secondary structures which are thought to be important for ribosome frameshift, indicating that this possibility is remote.

Viral RNAs generated by alternative splicing have two structural differences. One of the structural differences of each RNA species of the same gene is the length of the leader sequence preceding the coding exon. Muesing et al. (26) reported that *tat* cDNA with a longer leader sequence transactivated the HIV-1 LTR five times better than did the *tat* cDNA with a shorter leader sequence. However, Schwartz et al. (33) suggested that there were no significant differences between the two cDNAs with different lengths of 5′ leader sequences. Therefore, whether the length of the leader sequence has any effect on the function of the mRNA remains to be determined. The other structural difference among *rev* RNAs generated by alternative splicing is the change of two amino acid residues, depending on whether the splice occurred at nucleotide position 8802 or 8805 (Fig. 3). At present, it is not clear what the effect may be on the function of the Rev (and possibly Tat) protein by the amino acid changes at the splice junction. Structural analysis of the spliced viral RNAs reveals that all but *nef* mRNAs have more than one open reading frame. For example, *tat* mRNA contains *rev* and *nef* coding sequences, and *rev* mRNA contains a *nef* coding sequence. Northern blot hybridization analysis suggests that *vpx* and *vpr* mRNAs may contain all of the downstream coding sequences. Determination of whether these viral RNAs are functionally bicistronic or monocistronic must await further analysis.

Our studies and those of others demonstrate that both SIVmac and HIV-1 exhibit a complex pattern of mRNA expression by complicated alternative splicing, suggesting that expression of the primate lentiviruses may be controlled at many different levels. The possibility of translational regulation of SIVmac expression is under study.

ACKNOWLEDGMENTS

We thank James Hoxie for the human T × B hybrid cell line CEMX721.174, Dean Regier for assistance in computer analysis and helpful information regarding sequencing, Nancy Adams and Joanne Newton for preparing the manuscript, Beverly Blake for editorial

assistance, and Ronald Desrosiers and Gregory Viglianti for helpful discussions and critical reading of the manuscript.

This work was supported by grants from the National Institutes of Health to Y.L. (AI25108 and CA50146) and from the Division of Research Resources of NIH to the N.E.R.P.R.C.

REFERENCES

1. **Arrigo, S. J., S. Weitsman, J. A. Zack, and I. S. Y. Chen.** 1990. Characterization and expression of novel singly spliced RNA species of human immunodeficiency virus type 1. J. Virol. **64**:4585–4588.
2. **Ausubel, F. M., R. Brent, R. E. Kingston, D. D. Moore, J. G. Seidman, J. A. Smith, and K. Struhl (ed.).** 1987. Current protocols in molecular biology, vol. 2. John Wiley & Sons, Inc., New York.
3. **Benko, D. M., S. Schwartz, G. N. Pavlakis, and B. K. Felber.** 1990. A novel human immunodeficiency virus type 1 protein, Tev, shares sequences with *tat*, *env*, and *rev* proteins. J. Virol. **64**:2505–2518.
4. **Birnboim, H. C.** 1983. A rapid alkaline extraction method for the isolation of plasmid DNA. Methods Enzymol. **100**:243–255.
5. **Chakrabarti, L., M. Guyader, M. Alizon, M. D. Daniel, R. C. Desrosiers, P. Tiollais, and P. Sonigo.** 1987. Sequence of simian immunodeficiency virus from macaque and its relationship to other human and simian retroviruses. Nature (London) **328**:543–547.
6. **Colombini, S., S. K. Arya, M. S. Reitz, L. Jagodzinski, B. Beaver, and F. Wong-Staal.** 1989. Structure of simian immunodeficiency virus regulatory genes. Proc. Natl. Acad. Sci. USA **86**:4813–4817.
7. **Daniel, M. D., N. L. Letvin, N. W. King, M. Kannagi, P. K. Sehgal, R. D. Hunt, P. J. Kanki, M. Essex, and R. C. Desrosiers.** 1985. Isolation of T-cell tropic HTLV-III-like retrovirus from macaques. Science **228**:1201–1204.
8. **Daniel, M. D., N. L. Letvin, P. K. Sehgal, D. K. Schmidt, D. P. Silva, K. R. Solomon, F. S. Hodi, Jr, D. J. Ringler, R. D. Hunt, N. W. King, and R. C. Desrosiers.** 1988. Prevalence of antibodies to 3 retroviruses in a captive colony of macaque monkeys. Int. J. Cancer **41**:601–608.
9. **Dayton, A. I., E. F. Terwilliger, J. Potz, M. Kowalski, J. G. Sodroski, and W. A. Haseltine.** 1988. Cis-acting sequences responsive to the *rev* gene product of the human immunodeficiency virus. J. AIDS **1**:441–452.
10. **Emerman, M., R. Vazeux, and K. Peden.** 1989. The *rev* gene product of human immunodeficiency virus affects envelope-specific RNA localization. Cell **57**:1155–1165.
11. **Ennis, P. D., J. Zemmour, R. D. Salter, and P. Parham.** 1990. Rapid cloning of HLA-A, B cDNA by using the polymerase chain reaction: frequency and nature of errors produced in amplification. Proc. Natl. Acad. Sci. USA **87**:2833–2837.
12. **Felber, B. K., C. M. Drysdale, and G. N. Pavlakis.** 1990. Feedback regulation of human immunodeficiency virus type 1 expression by the *rev* protein. J. Virol. **64**:3734–3741.
13. **Felber, B. K., M. Hadzopoulou-Cladaras, C. Cladaras, T. Copeland, and G. N. Pavlakis.** 1989. The *rev* protein of human immunodeficiency virus type 1 affects the stability and transport of the viral mRNA. Proc. Natl. Acad. Sci. USA **86**:1495–1499.
14. **Fukasawa, M., T. Miura, A. Hasegawa, S. Morikawa, H. Tsujimoto, K. Miki, T. Kitamura, and M. Hayami.** 1988. Sequence of simian immunodeficiency virus from African green monkey, a new member of the HIV/SIV group. Nature (London) **333**:457–461.
15. **Grunstein, M., and D. S. Hogness.** 1975. Colony hybridization: a method for the isolation of cloned DNAs that contain a specific gene. Proc. Natl. Acad. Sci. USA **72**:3961–3965.
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. **Hadzopoulou-Cladaras, M., B. K. Felber, C. Cladaras, A. Athanassopoulos, A. Tse, and G. N. Pavlakis.** 1989. The *rev* (*trs/art*) protein of human immunodeficiency virus type 1 affects viral mRNA and protein expression via a *cis*-acting sequence in the *env* region. J. Virol. **63**:1265–1274.
18. **Hammarskjold, M.-L., J. Helmer, B. Hammarskjold, I. Sangwan, L. Albert, and D. Rekosh.** 1989. Regulation of human immunodeficiency virus *env* expression by the *rev* gene product. J. Virol. **63**:1959–1966.
19. **Hoxie, J. A., B. S. Haggarty, S. E. Bonser, J. L. Rackowski, H. Shan, and P. J. Kanki.** 1988. Biological characterization of a simian immunodeficiency virus-like retrovirus (HTLV-IV): evidence for CD4-associated molecules required for infection. J. Virol. **62**:2557–2568.
20. **Jacks, T., H. D. Madhani, F. R. Masiarz, and H. E. Varmus.** 1988. Signals for ribosomal frameshifting in the Rous sarcoma virus *gag-pol* region. Cell **55**:447–458.
21. **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.
22. **Kannagi, M., J. M. Yetz, and N. L. Letvin.** 1985. In vitro growth characteristics of simian T-lymphotropic virus type III. Proc. Natl. Acad. Sci. USA **82**:7053–7057.
23. **Malim, M. H., J. Hauber, S.-Y. Le, J. V. Maizel, and B. R. Cullen.** 1989. The HIV-1 *rev* trans-activator acts through a structured target sequence to activate nuclear export of unspliced viral mRNA. Nature (London) **338**:254–257.
24. **Malim, M. H., L. S. Tiley, D. F. McCarn, J. R. Rusche, J. Hauber, and B. R. Cullen.** 1990. HIV-1 structural gene expression requires binding of the *rev* trans-activator to its RNA target sequence. Cell **60**:675–683.
25. **Mount, S. M.** 1982. A catalogue of splice junction sequences. Nucleic Acids Res. **10**:459–472.
26. **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.
27. **Murphey-Corb, M., L. N. Martin, S. R. S. Rangan, G. B. Baskin, B. J. Gormus, R. H. Wolf, W. A. Andes, M. West, and R. C. Montelaro.** 1986. Isolation of an HTLV-III-related retrovirus from macaques with simian AIDS and its possible origin in asymptomatic mangabeys. Nature (London) **321**:435–437.
28. **Naidu, Y. M., H. W. Kestler III, Y. Li, C. V. Butler, D. P. Silva, D. K. Schmidt, C. D. Troup, P. K. Sehgal, P. Sonigo, M. D. Daniel, and R. C. Desrosiers.** 1988. Characterization of infectious molecular clones of simian immunodeficiency virus (SIVmac) and human immunodeficiency virus type 2: persistent infection of rhesus monkeys with molecularly cloned SIVmac. J. Virol. **62**:4691–4696.
29. **Ohta, Y., T. Masuda, H. Tsujimoto, K. Ishikawa, T. Kodama, S. Morikawa, M. Nakai, S. Honjo, and M. Hayami.** 1988. Isolation of simian immunodeficiency virus from African green monkeys and seroepidemiologic survey of the virus in various non-human primates. Int. J. Cancer **41**:115–122.
30. **Okayama, H., and P. Berg.** 1982. High-efficiency cloning of full-length cDNA. Mol. Cell. Biol. **2**:161–170.
31. **Regier, D., and R. C. Desrosiers.** 1990. The complete nucleotide sequence of a pathogenic molecular clone of simian immunodeficiency virus. AIDS Res. Hum. Retroviruses **6**:1221–1231.
32. **Salfeld, J., H. G. Gottlinger, R. A. Sia, R. E. Park, J. G. Sodroski, and W. A. Haseltine.** 1990. A tripartite HIV-1 *tat-env-rev* fusion protein. EMBO J. **9**:965–970.
33. **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.
34. **Viglianti, G. A., P. L. Sharma, and J. I. Mullins.** 1990. Simian immunodeficiency virus displays complex patterns of RNA splicing. J. Virol. **64**:4207–4216.