JOURNAL OF VIROLOGY, Nov. 1990, p. 5448–5456
0022-538X/90/115448-09$02.00/0

Vol. 64, No. 11

Env and Vpu Proteins of Human Immunodeficiency Virus Type 1 Are Produced from Multiple Bicistronic mRNAs

STEFAN SCHWARTZ,[1,2] BARBARA K. FELBER,[1] EVA-MARIA FENYÖ,[2] AND GEORGE N. PAVLAKIS[1]*

Basic Research Program, National Cancer Institute-Frederick Cancer Research and Development Center, Frederick, Maryland 21702-1201[1] and Department of Virology, School of Medicine, Karolinska Institute, Stockholm S-105 21, Sweden[2]

Received 18 May 1990/Accepted 19 August 1990

Three size classes of human immunodeficiency virus type 1 (HIV-1) mRNAs are produced in infected cells: full-length, intermediate, and small. Here we report that the intermediate-size class of viral mRNAs is heterogeneous, consisting of at least 12 differentially spliced species. This group contains nine bicistronic mRNAs producing Env and Vpu and three mRNAs expressing only the first exon of *tat*. In the latter mRNAs, Env and Vpu expression is blocked by the presence of the upstream *tat* open reading frame. We conclude that internal initiation of translation is not the mechanism for generation of the bicistronic *env* mRNAs. Translation of HIV-1 mRNAs is consistent with the scanning mechanism in which Env is produced by leaky scanning from mRNAs that contain *env* as the second or third reading frame. Env and Vpu proteins are expressed from the same mRNAs and are coordinately regulated by Rev. This arrangement may reflect a requirement for coordinate expression of Vpu and Env.

The genomic organization of lentiviruses, including human immunodeficiency virus type 1 (HIV-1), is more complex than that of other retroviruses. This organization allows the production of several additional viral proteins by alternative splicing. Three size classes of HIV-1 mRNAs are produced in infected cells: full-length, intermediate, and small. More than 20 small multiply spliced HIV-1 mRNA species have been identified. They are generated by alternative splicing in the middle part of the HIV-1 genome (2, 3, 31, 41, 42) and produce regulatory proteins. Two of these proteins, Tat and Rev, are essential for viral replication (8, 11, 16, 44). Tat increases the steady-state levels of all viral transcripts by acting on a *cis*-acting element named the Tat-responsive element (TAR), located in the R region of the HIV-1 long terminal repeat (LTR) (15, 17, 18, 22, 24, 32, 36, 43, 48). Posttranscriptional effects of Tat have also been reported, suggesting that TAR may also respond to the Tat protein at the RNA level (4, 7, 14, 23, 48). The second essential transacting protein of HIV-1, Rev, also acts through a *cis*-acting sequence termed the Rev-responsive element (RRE), located in the *env* region (9, 10, 13, 20, 21, 29, 37). Rev functions at the RNA level by promoting the export of RRE containing unspliced and partially spliced mRNAs to the cytoplasm (10, 13, 21, 29). In the absence of Rev, the levels of unspliced and partially spliced HIV-1 mRNAs in the cytoplasm are very low (13). The presence of Rev results in an increased ratio of these mRNAs versus the small, multiply spliced HIV-1 mRNAs and production of structural proteins. The regulation by Rev may also require the presence of other elements in the mRNA in addition to RRE (10, 13, 20, 37). Since the expression of the viral mRNAs is regulated during the viral life cycle, knowledge about their exact structure and coding potential is important in order to understand how the protein expression is regulated. However, the intermediate-size mRNAs producing Env and several other HIV-1 proteins had not been characterized.

To investigate the structures of the intermediate-size mRNAs containing the *env* region, we cloned and sequenced cDNAs amplified by the polymerase chain reaction (PCR). Sequence analysis showed that at least 12 intermediate-size mRNAs were present in the HIV-1-infected cells. Eucaryotic expression plasmids containing complete intermediate-size cDNAs were constructed, and protein expression was studied after transfection into human cells. The results revealed that mRNAs encoding Env were bicistronic, producing high levels of both Env and Vpu proteins in a Rev-dependent manner. mRNAs containing the *tat* open reading frame (ORF) upstream of the *env* and *vpu* ORFs did not express detectable amounts of Env or Vpu protein. The results indicate that translation of HIV-1 mRNAs follows the modified scanning model for initiation of translation (28).

MATERIALS AND METHODS

RNA preparation, PCR amplification, cloning, and sequencing. Total RNA was prepared from HIV-1-infected cells by the guanidine isothiocyanate method (5). Cytoplasmic RNA was prepared as described previously (13). The RNA was subjected to reverse transcription and PCR amplification (39, 40) as described previously (42) using primers NARS, 2608S, 534A, 2913A, and BAMA. Their exact locations and sequences are as follows: NARS, 5′-CTCTAG CAGTGGCGCCCGAACAGGG-3′ (nucleotide [nt] 173 to 197); 2608S, 5′-GGGTCTAGACCCGGGTCTCTCTGGT TAGACCAGATCTGAGCCTGGGAGC-3′ (nt 1 to 37; the underlined 12-base tail introducing an *Xba*I site was added at the 5′ end of the primer); 534A, 5′-CTATGATTACTATG GACCAC-3′ (nt 5673 to 5692); 2913A, 5′-GTGCTGATATT TCTCCTTCACTCTC-3′ (nt 5773 to 5797); and BAMA, 5′-GCCAAGGATCCGTTCACTAATCGAATGG-3′ (nt 8004to 8031). The amplification products were digested with *Bss*HII and cloned as *Bss*HII-cut blunt-end fragments into a modified Bluescript vector. Clones were screened by colony hybridization with a panel of exon-specific oligonucleotide probes, and representative clones were selected for sequencing. Double-stranded sequencing was performed on miniprep plasmid DNA, using a Sequenase kit (United States Biochemicals) as described previously (42).

* Corresponding author.

Southern blot analysis of PCR-amplified cDNA. Samples of each PCR amplification were analyzed on 5% nondenaturing polyacrylamide gels, and the cDNAs were visualized by UV fluorescence after staining of gels with ethidium bromide. Southern blots were performed after transfer of the cDNAs to nylon membranes by electroblotting, followed by hybridization to the ^{32}P-end-labeled oligonucleotide probe 565 or SACS as described previously (42). The sequences and exact locations of the oligonucleotide probes were as follows: 565, 5′-GGTTGCGTTACATGTTGGACCAC-3′ (nt 5596 to 5618); SACS, 5′-GAAGAAGCGGAGACAGCGACGAA GAGCTC-3′ (nt 5523 to 5551).

Recombinant plasmids. The partial cDNAs amplified with primer pair NARS-534A were cloned into a modified Bluescript KS$^-$ vector (Stratagene) and ligated to a proviral DNA *Sac*I fragment (nt 5546 in exon 5 to nt 9118 in the R region of the 3′ LTR). These cDNA constructs are missing the first 257 bases of exon 1. The resulting plasmids contain the cDNAs under the control of the T3 promoter. Two different infectious HIV-1 clones were used in the PCR amplification experiments, NL43 (1) and HXB2 (34). The cDNAs obtained from NL43 contain the complete *env* and *vpu* genes. cDNA 1.5EU$^-$ was constructed similarly but was obtained from the *vpu* mutant proviral clone HXB2. Eucaryotic expression plasmids were constructed by replacing the *Bss*HII-*Xho*I viral fragment of plasmid pNL43 (1) with the *Bss*HII-*Xho*I fragments of the in vitro expression cDNA clones. This provided the cDNAs with LTR promoter and polyadenylation signal and resulted in vectors expressing authentic mRNAs 1.5EU$^-$, 1.5E, 1.4BE, 1.4AE, and 1.4E (Fig. 1D) initiated at position +1. (The mRNAs were named according to the exons they contain.)

In vitro transcription and translation. The in vitro expression plasmids were linearized with *Bam*HI, and run-off transcripts were synthesized by using T3 RNA polymerase and an mRNA capping kit (Stratagene) according to the recommendations of the supplier. In vitro translation of the capped RNA was performed in a nuclease-treated rabbit reticulocyte lysate (Promega) containing 79 mM potassium acetate and 500 μM magnesium acetate, in the presence of 10 μCi of [^{35}S]methionine (New England Nuclear), according to the recommendations of the manufacturer.

Transfections and immunoprecipitations. Vectors expressing mRNA 1.4E, 1.4BE, 1.5E, or 1.5EU$^-$ were transfected into the HLtat cell line (12, 42), which constitutively produces Tat, using the calcium phosphate coprecipitation technique (19) as described previously (13). Rev was provided by cotransfection with the Rev-expressing plasmid pL3crev (13). The transfected HLtat cells were metabolically labeled with [^{35}S]methionine or [^{35}S]cysteine. Cells were harvested 2 days posttransfection in 0.5× RIPA buffer (50 mM Tris hydrochloride [pH 7.4], 150 mM NaCl, 1% Triton X-100, 1% sodium deoxycholate, 0.1% sodium dodecyl sulfate), and analyzed by immunoprecipitations with monospecific rabbit antisera against either Env, Vpu, Tat, or Nef as described previously (12, 42). In vitro translation products were immunoprecipitated with either rabbit anti-Tat antiserum or HIV-1 patient serum reacting with both Env and Vpu.

RESULTS

Identification of intermediate-size mRNAs encoding Env, Vpu, and one-exon Tat. To investigate the exact structure of the Env-encoding mRNA(s) produced by HIV-1, we analyzed PCR-amplified cDNA obtained from cells infected by the molecularly cloned HIV-1 isolate NL43 (1) or HXB2 (34) as described previously (42), using two different primers (534A and 2913A [Fig. 1C]) located 3′ of the splice donor at nt 5591. This splice donor is used by all the small multiply spliced mRNAs to join exon 4, 4A, 4B, or 5 to exon 7 (Fig. 1B) (42). The region to which these amplification primers hybridized was absent in all the small multiply spliced mRNAs; therefore, only unspliced and partially spliced (intermediate-size) mRNAs could be amplified. Primer 534A hybridized immediately upstream of the *env* AUG, while primer 2913A hybridized within the *env* gene. cDNA synthesis and PCR amplification using either of these primers in combination with primer NARS located in exon 1 (Fig. 1C), which is common to all HIV-1 transcripts, resulted in a large number of amplified products. Analysis of the amplification products on acrylamide gels revealed that a similar pattern of PCR products was obtained with primer pair 534A-NARS and 2913A-NARS, indicating that the same mRNA species were detected by these two primer pairs. It could be argued that not all of the PCR-amplified products represented cytoplasmic mRNAs. To address this point, we prepared cytoplasmic RNA from HeLa cells transfected with either of the infectious HIV-1 clones and performed PCR with primer pair 534A-NARS. The same cDNAs were amplified from the cytoplasmic RNA fraction as from total RNA (data not shown). Figure 2 shows that cDNAs amplified with primers NARS and 534A hybridized to either of the two oligonucleotide probes SACS (Fig. 2, lane 2) and 565 (Fig. 2, lane 4). SACS hybridizes to a region 5′ of the splice donor at position 5591 that is expected to be present in all small multiply spliced mRNAs. Probe 565 hybridizes to a region 3′ of the splice donor at position 5591 (Fig. 1B). This region is present only in the unspliced and partly spliced, intermediate-size mRNAs. cDNAs representing the small multiply spliced mRNAs were generated by PCR amplification using primer pair NARS-BAMA (42) and were hybridized to the same probes (Fig. 2, lanes 1 and 3). The similarities in hybridization patterns obtained using probe SACS indicated that the intermediate-size mRNAs utilized the same splice sites as the small multiply spliced mRNAs (42) (Fig. 2, lane 2 versus lane 1). Probe 565 hybridized to cDNAs amplified with primer pair NARS-534A (Fig. 2, lane 4) but not to those amplified with primers NARS and BAMA (Fig. 2, lane 3), demonstrating that primer 534A specifically detected intermediate-size mRNAs.

cDNA synthesis and PCR amplification of RNA from cells infected or transfected with HXB2 or NL43 resulted in similar patterns of amplification products. In addition to transfected HeLa cells, we also analyzed RNA from infected T-lymphoid cell lines (H9 and Molt-3) and infected peripheral blood mononuclear cells. The same pattern of amplification products was detected in all cell types. Cloning and sequencing of the amplified cDNAs from HXB2-infected cells obtained with primer pair NARS-534A revealed that the intermediate-size mRNAs utilize the previously known splice acceptors (2, 31, 42) (Fig. 1B). The structures of four intermediate mRNAs are shown in Fig. 1D. We named the newly identified exons 4E, 4AE, 4BE, and 5E by analogy to the exons starting at the same site in the small mRNAs. Exon 4E has been described previously as exon 6 (31). In addition to the four intermediate-size mRNAs shown, eight other mRNAs have been characterized, which contain either, but not both of the small noncoding exons 2 and 3 spliced after exon 1 (Fig. 1B). Previous results have shown that the presence or absence of these exons in other HIV-1 mRNAs did not significantly affect protein expression (42).

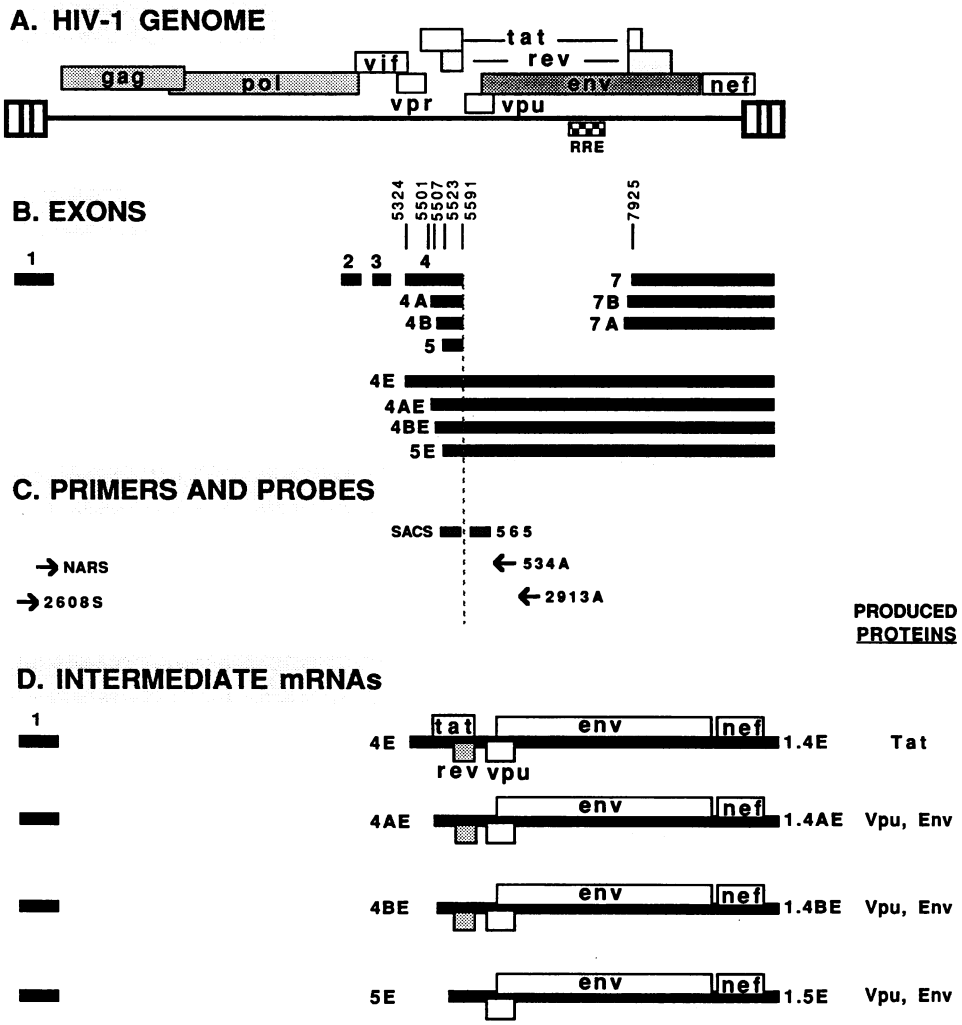

FIG. 1. (A) HIV-1 proviral genome. ORFs and the location of the RRE are indicated. (B) The exons present in the cloned HIV-1 mRNAs are indicated by solid bars. The exon numbering system follows that of Muesing et al. (31) and Schwartz et al. (42). Vertical numbers indicate the nucleotide positions of the splice acceptors of exons 4, 4A, 4B, 5, and 7 and the splice donor utilized by all small multiply spliced mRNAs at position 5591 (2, 3, 31, 42). In the numbering system of the genome used here, +1 is the mRNA start site. (C) The location and orientation of the oligonucleotide primers used for PCR is indicated by arrows, and the location of the oligonucleotide probes SACS and 565 is indicated by gray bars. For the sequences and exact locations of oligonucleotides, see Materials and Methods. (D) Structures of cloned intermediate-size HIV-1 mRNAs. These mRNAs were named according to the exons they contain, as indicated on the right. The proteins produced by these mRNAs are also indicated on the right. The structures have been verified by cloning and sequencing of the amplified cDNAs. Either of the small exons 2 and 3 (but not both) are found spliced to the intermediate mRNAs (not shown). Therefore, a total of 12 mRNAs are predicted on the basis of Southern blot analysis of PCR-amplified cDNAs, of which 10 have been cloned and sequenced.

All PCR-amplified cDNAs were repeatedly cloned and sequenced. The most frequently cloned amplification product was 1.5E cDNA. It remains to be verified by other methods that the cloning frequency of PCR-amplified cDNA reflects the relative abundance of the mRNA in the infected cell.

Expression of Env, Vpu, and one-exon Tat in vitro. To identify the proteins expressed by these cDNAs in vitro, we constructed RNA expression vectors by supplying the missing sequences downstream of primer 534A, as described in Materials and Methods. The cDNAs were inserted in the Bluescript KS$^-$ vector under the control of the T3 promoter, resulting in plasmids 1.4E, 1.4AE, 1.4BE, 1.5E, and 1.5EU$^-$. The 1.5EU$^-$ cDNA has the same structure as 1.5E but was cloned from HXB2, which contains ACG instead of the *vpu* AUG initiation and does not express Vpu. Immunoprecipitations of the in vitro translation products showed that all five cDNAs produced Env protein of the expected size (Fig. 3A). All cDNAs except 1.5EU$^-$ also directed the synthesis of the small 17-kDa Vpu protein (6, 30, 45, 46). However, smaller amounts of Vpu and Env proteins were produced from 1.4E than from the other transcripts. In addition, high levels of a 14-kDa protein precipitated with an anti-Tat antiserum were expressed by 1.4E RNA (Fig. 3B). This protein, one exon Tat, is the product of the first coding exon of *tat*, which terminates at a stop codon located at position 5593, immediately after the splice donor at position 5591.

Env and Vpu are expressed from bicistronic mRNAs, whereas one-exon Tat is expressed from monocistronic mRNAs in eucaryotic cells. To analyze the different cDNAs

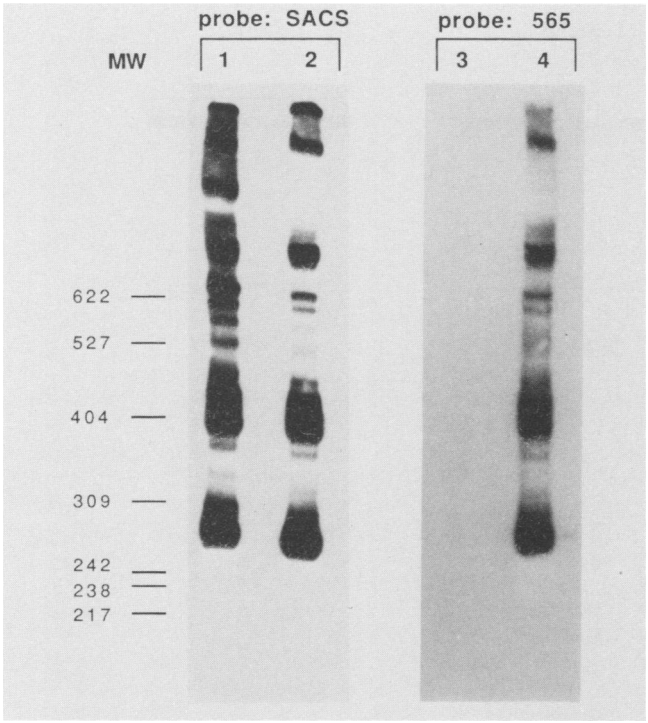

FIG. 2. Southern blots of PCR-amplified cDNA. Lane 1, cDNAs of small multiply spliced mRNAs that had been PCR amplified with primer pair NARS-BAMA and hybridized to probe SACS located 5′ of the splice donor at nt 5591; lane 2, cDNAs of intermediate-size mRNAs that had been PCR amplified with primer pair NARS-534A and hybridized to probe SACS; lane 3, cDNAs of small multiply spliced mRNAs that had been PCR amplified with primer pair NARS-BAMA hybridized to probe 565 located after the splice donor at nt 5591; lane 4, cDNAs of intermediate mRNAs that had been PCR amplified with primer pair NARS-534A and hybridized to probe 565. Size markers in nucleotides are shown to the left (MW).

for protein expression in human cells, they were transferred into eucaryotic expression vectors containing the HIV-1 LTRs, which provided the authentic HIV-1 promoter and polyadenylation signals, as described in Materials and Methods. This resulted in expression plasmids that produced mRNAs identical to those expressed after viral infection and allowed the independent analysis of the different mRNAs. We expressed four cDNAs (1.4E, 1.4BE, 1.5E, and 1.5EU$^-$) that differed in the number of ORFs preceding the *env* initiator AUG (Fig. 4A). These plasmids produced both intermediate-size and (after splicing) small viral mRNAs when transfected into eucaryotic cells. Expression of unspliced and intermediate HIV-1 mRNAs has been shown to be dependent on Rev (9, 11, 13, 20, 26, 29). To maximize expression of the unspliced mRNA, the cDNA vectors were cotransfected with the Rev expression plasmid pL3crev (13) into a cell line constitutively producing Tat protein (HLtat) (12, 42). Analysis of [^{35}S]methionine-labeled cellular lysates by immunoprecipitations showed that Env protein was expressed from all cDNAs except 1.4E and that Vpu was expressed from all cDNAs except 1.4E and 1.5EU$^-$ (Fig. 4A). Transfection of the 1.4E cDNA into HeLa cells and analysis of protein production with anti-Tat antiserum revealed that large amounts of one-exon Tat protein were produced from the 1.4E mRNA (Fig. 4B). To exclude the possibility that a mutation abolishing *env* and *vpu* expression had occurred in the 1.4E cDNA during cloning into the

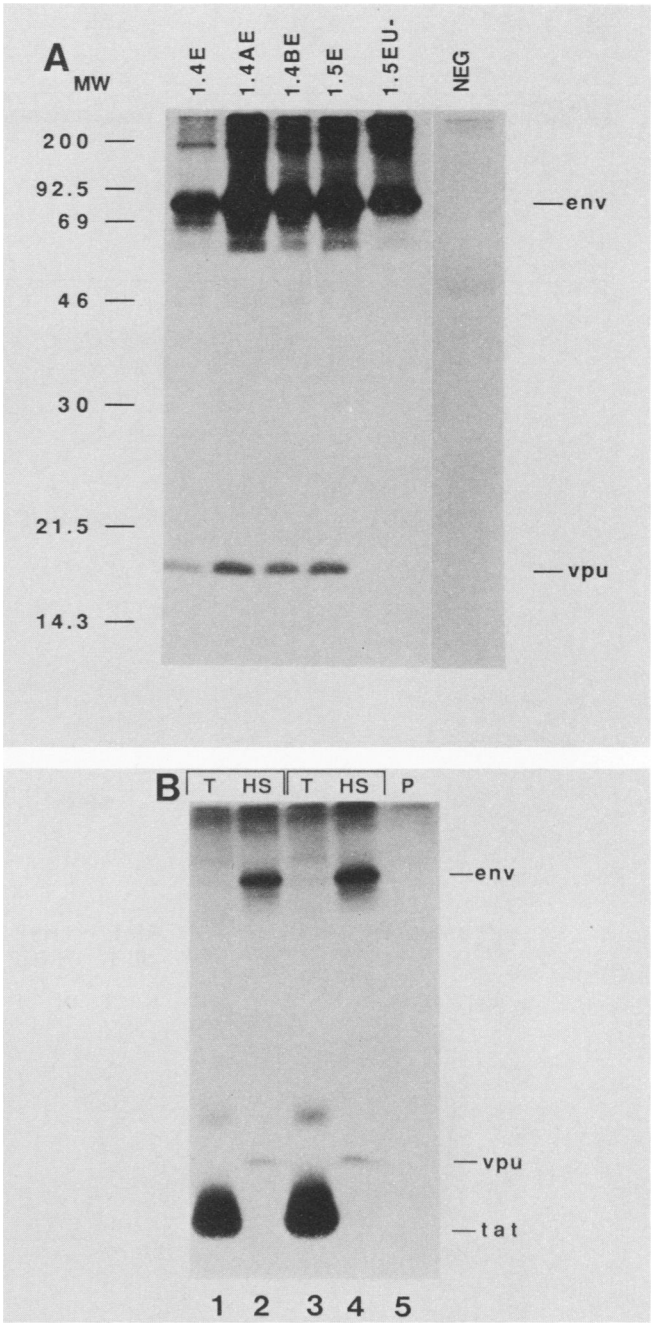

FIG. 3. Translational analysis of intermediate-size HIV-1 mRNAs in vitro. (A) Immunoprecipitation of in vitro translation products. cDNA clones 1.4E, 1.4AE, 1.4BE, 1.5E, and 1.5EU$^-$ were placed downstream of the T3 promoter and were transcribed in vitro by using T3 RNA polymerase, and 1 μg of RNA was subsequently translated by using the rabbit reticulocyte lysate system. The translation products were immunoprecipitated with HIV-1 patient serum reacting with both Env and Vpu proteins. Positions of molecular weight (MW) standards are shown to the left. (B) Immunoprecipitation of cDNA 1.4E translation products in vitro (lanes 1 and 2). cDNA 1.4E was also transferred back from the eucaryotic expression plasmid into a Bluescript vector, and the produced RNA was translated in vitro. In vitro translation products were immunoprecipitated (lanes 3 to 5) with anti-Tat antiserum (T), HIV-1 patient serum (HS) reacting with both Env and Vpu proteins, or preimmune serum (P).

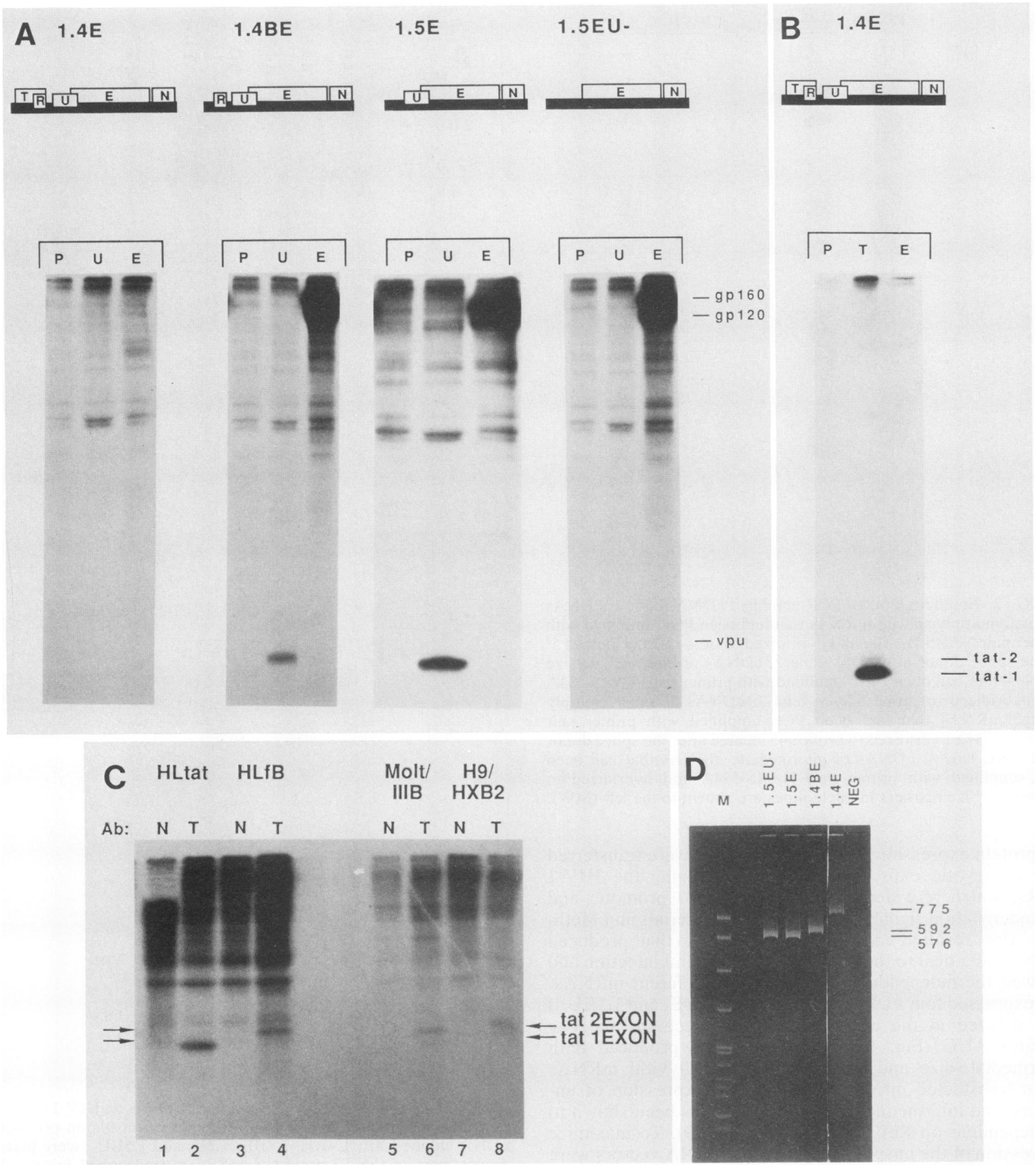

FIG. 4. (A) Immunoprecipitation of Env and Vpu expressed in human cells. cDNA plasmids expressing mRNA 1.4E, 1.4BE, 1.5E, or 1.5EU$^-$ were cotransfected with the Rev-expressing plasmid pL3crev (12) into HLtat cells (14, 42), which express Tat constitutively. The structures of the mRNAs are shown at the top. The ORFs for *tat* (T), *rev* (R), *vpu* (U), *env* (E), and *nef* (N) are shown as open boxes. (T and R represent the first exons of *tat* and *rev* ORFs present on the unspliced mRNAs. These ORFs encode truncated forms of Tat [72-amino-acid] and Rev [27-amino-acid] proteins, since stop codons are present in all reading frames immediately after the splice donor at nt 5591). Immunoprecipitations with pre-immune serum (P), anti-Vpu antiserum (U), and anti-Env antiserum (E) are shown. (B) Immunoprecipitation of one-exon Tat protein produced in human cells. cDNA plasmid expressing mRNA 1.4E was transfected into HeLa cells with the Rev-expressing plasmid pL3crev. Cells were metabolically labeled with [^{35}S]cysteine, and the cell lysate was immunoprecipitated with preimmune serum (P), anti-Env antiserum (E) and anti-Tat antiserum (T). Tat-1 indicates the one-exon Tat protein produced from the first exon of the *tat* ORF, which terminates immediately after the splice donor at nt 5591 on the unspliced 1.4E mRNA. Tat-2 is the complete Tat protein produced from a small amount of spliced mRNA, utilizing the splice donor at nt 5591 and the splice acceptor at nt 7925 (see Fig. 1). (C) Immunoprecipitation of one-exon Tat protein produced in HIV-1-infected cells. HLtat and HLfB are HeLa-derived cell lines containing stably integrated copies of the pL3tat (42) and the *rev* proviral mutant fB (20), respectively. pL3tat contains the first exon of *tat* and produces

eucaryotic expression plasmid, we transferred the cDNA back into the Bluescript T3 vector and expressed RNA in vitro as described above. Immunoprecipitation of the proteins produced by this RNA in the reticulocyte lysate revealed the presence of Env, Vpu, and Tat proteins of the expected sizes (Fig. 3B). This demonstrated that the lack of Env and Vpu production by this cDNA in vivo was not due to a mutation in the *env* or *vpu* gene. The results indicated that the presence of the *tat* ORF upstream of the *vpu* and *env* AUGs prevented Vpu and Env expression from 1.4E mRNA, whereas the presence of the *rev* or *vpu* ORFs on mRNAs 1.4BE or 1.5E allowed Env expression in vivo. The discrepancy between the in vivo and in vitro data for Env expression by 1.4E mRNA may be explained by relaxed specificity of the rabbit reticulocyte lysate or RNA degradation in vitro.

To verify that one-exon Tat protein is produced in HIV-1-infected cells, H9 cells infected with the molecular clone HXB2 or Molt4 cells infected with the viral isolate $HTLV_{IIIB}$ were metabolically labeled with [^{35}S]cysteine and proteins were immunoprecipitated with anti-Tat antiserum (Fig. 4C). These experiments demonstrated that both one-exon Tat and two-exon Tat were present in HIV-1-infected cells. The levels of one-exon Tat protein were lower than two-exon Tat. Immunoprecipitations of Tat from HLtat cells, which produce exclusively one-exon Tat protein, verified that the smaller Tat protein detected in infected cells is one-exon Tat (Fig. 4C). Immunoprecipitations of Tat from HLfB cells containing a *rev* mutant viral clone of HXB2, fB (20), revealed that one-exon Tat protein is not produced in the absence of Rev, verifying that expression of one-exon Tat is dependent on Rev and is produced from mRNA 1.4E.

Expression of Env and Vpu proteins from bicistronic mRNAs is coordinately regulated. The results of the in vivo expression of the cDNAs indicated that *env* mRNAs containing exon 4AE, 4BE, or 5E were bicistronic and produced both Env and Vpu. Alternatively, Env and Vpu might be expressed from two different mRNAs generated by further splicing of the cDNA transcript. To verify that Vpu and Env are indeed produced from bicistronic mRNAs, we analyzed the structures of the RNAs produced from the cDNA expression plasmids in HeLa cells by reverse transcription and PCR amplification (Fig. 4D). A primer located 3′ of the *env* AUG (2913A) was used for cDNA synthesis and PCR amplification in combination with primer 2608S containing the first 37 bases of the HIV-1 mRNAs (Fig. 1C). These primers should detect any mRNA that would use a splice acceptor upstream of the *env* AUG. After cDNA synthesis and PCR amplification, single bands of the correct sizes were detected (Fig. 4D), verifying that no splicing upstream of the *env* AUG occurred in these mRNAs. Therefore, all mRNAs expressing Env were bicistronic and produced both Env and Vpu. To verify that both Env and Vpu production was dependent on Rev, we transfected the 1.5E cDNA in the presence or absence of the Rev-expressing plasmid pL3crev into HLtat cells and measured the expression of Env, Vpu, and Nef. Immunoprecipitations revealed that only Nef was produced in the absence of Rev, whereas very little Nef was produced in its presence (Fig. 5B). In the absence of Rev, the 1.5E mRNA is spliced to mRNA 1.5.7 (Fig. 5A) that has *nef* as the first ORF and has been shown to express high levels of Nef protein (42). Expression of both Vpu and Env was dependent on the presence of Rev, indicating that Vpu and Env are coordinately regulated (Fig. 5B).

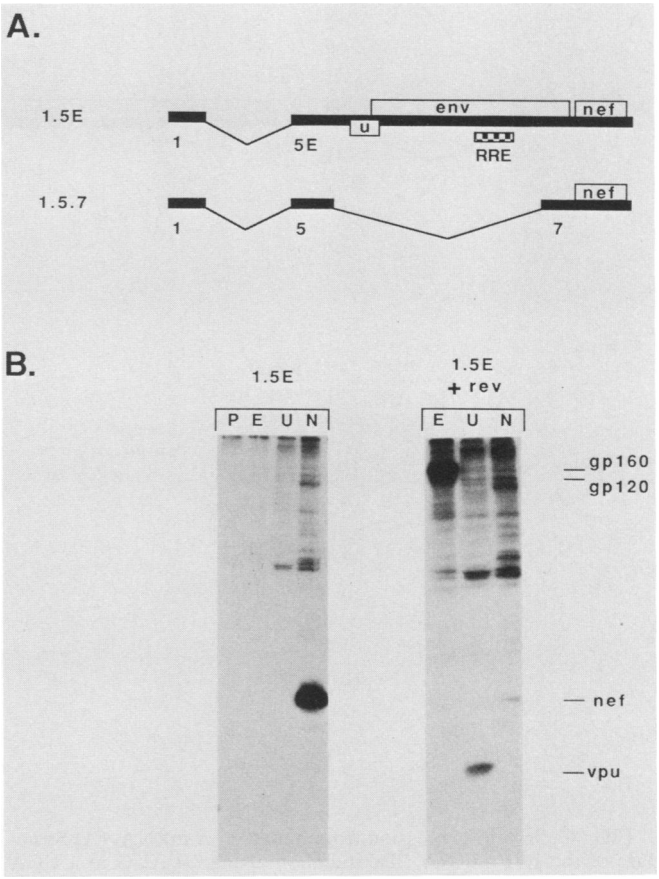

FIG. 5. Expression of Env and Vpu is coordinately regulated by Rev. The expression vector producing mRNA 1.5E was transfected into HLtat cells (14, 42) in the presence or absence of Rev. This mRNA remains unspliced in the presence of Rev but is spliced to a small mRNA designated 1.5.7 (42) in the absence of Rev. (A) Structures of the unspliced 1.5E mRNA and the spliced 1.5.7 mRNA. The position of the RRE in the 1.5E mRNA is indicated. (B) Protein expression was analyzed by immunoprecipitations with preimmune serum (P), anti-Env (E), anti-Vpu (U), and anti-Nef (N) antisera.

Expression of one-exon Tat protein prevents expression of the downstream ORFs *env* and *vpu*. In contrast to 1.4AE, 1.4BE, and 1.5E mRNAs, the 1.4E mRNA contains the first exon of *tat* (Fig. 1D). This mRNA produced only one-exon

only one-exon Tat protein. Molt/IIIB is a T-cell line (Molt 4) infected by HIV-1 ($HTLV_{IIIB}$ isolate). H9/HXB2 is a T-cell line (H9) infected by the molecular clone HIV_{HXB2}. Lanes 1, 3, 5, and 7, immunoprecipitations of [^{35}S]cysteine-labeled cell extracts with nonimmune rabbit serum (N). Lanes 2, 4, 6, and 8, immunoprecipitations of the same cell extracts with anti-Tat antiserum (T). The arrows indicate the position of the one-exon Tat protein and the two-exon Tat protein. (D) Analysis of RNA expressed from the cDNA clones in HLtat cells (14, 42). Expression vectors producing mRNA 1.5EU$^-$, 1.5E, 1.4BE, or 1.4E were transfected into HLtat cells in the presence of Rev. cDNA was synthesized and PCR amplified using primers 2913A and 2608S. The expected sizes of the amplification products were 576 bp for 1.5E and 1.5EU$^-$, 592 bp for 1.4BE, and 775 bp for 1.4E, as indicated on the right. The negative control (NEG) shows amplification of RNA from sample 1.5E after omitting the reverse transcription step. Size marker (M) is *Msp*I-digested pBR322 plasmid.

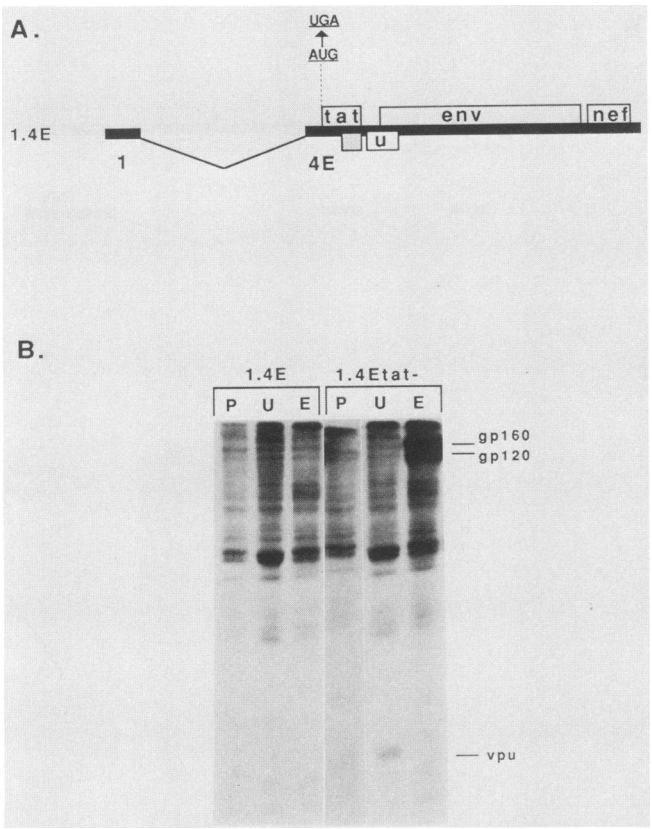

FIG. 6. Protein expression from mutated *tat* mRNA. (A) Structure of the 1.4E mRNA. The mutation of the *tat* AUG to a UGA which generates the *tat* mutant form of the authentic mRNA (named 1.4Etat$^-$) is indicated. (B) cDNA expression plasmids producing 1.4E mRNA or 1.4Etat$^-$ mRNA were cotransfected with the Rev-expressing plasmid pL3crev (12) into HLtat cells (14, 42). Expression of Env and Vpu were analyzed by immunoprecipitation with preimmune serum (P), anti-Vpu (U), and anti-Env (E) antisera.

Tat protein. To prove that translation of the *tat* ORF was responsible for the inability of the 1.4E mRNA to express Env and Vpu proteins, a mutation converting the *tat* AUG to a termination codon (UGA) was introduced in the 1.4E cDNA expression vector (Fig. 6A). This mutation has been shown to inactivate Tat expression from the intact provirus (38) and from the 1.4E cDNA (data not shown). If expression of downstream ORFs were inhibited by *tat* translation, the *tat* mutant construct should produce Env and Vpu after transfection of the 1.4Etat$^-$ cDNA into HLtat cells. The results showed that both Env and Vpu were produced from the 1.4Etat$^-$ cDNA, while neither was expressed from the 1.4E cDNA containing the *tat* AUG (Fig. 6B). These data indicated that the HIV-1 mRNAs show polarity and that the tat ORF affects expression of downstream ORFs to a high extent while other ORFs (i.e., *rev* and *vpu*) do not.

DISCUSSION

We have identified 12 intermediate-size mRNAs encoding one-exon Tat, Vpu, and Env. Evidence exists for additional intermediate mRNAs expressing Vif and Vpr proteins (unpublished data). The presence of several of the PCR-identified mRNAs in HIV-1-infected cells has been verified by S1 nuclease protection assays and by the identification of their protein products (12). A comparison between PCR-amplified cDNAs obtained from cytoplasmic RNA and cDNAs obtained from total RNA revealed that all the mRNAs identified here were present in the cytoplasm of infected cells. Repetitive cloning and sequencing of the different cDNAs revealed that the most frequently cloned cDNA was that of mRNA 1.5E. Since PCR amplification is nonlinear and shows preferences for certain target DNAs, further quantitative analysis is necessary to verify that mRNA 1.5E is the most abundant Env- and Vpu-producing mRNA in HIV-infected cells. The redundancy of RNAs might be caused by the organizational principles of the lentiviral genome. It may also allow the rapid change of the virus by mutations affecting some splice sites. The biological variability attributed to HIV-1 may also involve splicing, since different HIV-1 strains may produce quantitatively or qualitatively different mRNAs. We have also shown that HIV-1 produces both monocistronic and bicistronic mRNAs. Our data show that the presence of the *tat* ORF blocks expression of downstream ORFs, resulting in monocistronic mRNAs expressing only one-exon Tat protein. This finding differs from previously published data (21, 35) that reported expression of Env protein from a vector containing the *tat*, *rev*, and *vpu* ORFs upstream of the *env* reading frame. This discrepancy may be explained by the simian virus 40 expression vectors used by these investigators, which may cause aberrant initiation or splicing. The results presented here show that one-exon Tat is produced from an RRE-containing mRNA in infected cells and is dependent on Rev for its expression. This could secure Tat production at a late stage of the viral life cycle. However, the levels of one-exon Tat protein in HIV-1-infected cells are much lower than the levels of the two-exon Tat protein (Fig. 4C), indicating that the two-exon Tat is primarily responsible for transactivation at all times.

All mRNAs expressing Vpu were bicistronic and also expressed Env, since the *vpu* AUG allowed bypassing of ribosomes, resulting in expression of the downstream *env* ORF. Therefore, the results presented here suggest that translation of the HIV-1 mRNAs follows the scanning model (27, 28) for the selection of the AUG initiation site. This model predicts efficient initiation of translation when the AUG is in a favorable context, resulting in no or very low expression of downstream ORFs. In contrast, AUGs in an unfavorable context are used less frequently and allow passing of ribosomes and downstream initiation (leaky scanning). We compared the sequences surrounding the HIV-1 AUGs with the consensus sequence for efficient initiation of translation (28). In contrast to the *rev* and *vpu* AUGs, the *tat* AUG followed the consensus sequence at positions -3 and $+4$. The presence of a purine at position -3 and of a guanine at position $+4$ has been shown to be important for efficient initiation of translation (28). Expression of Env from the bicistronic mRNAs 1.5E, 1.4AE, and 1.4BE can therefore be explained by leaky scanning. Since *vpu* and *env* ORFs are overlapping, Env can only be expressed by ribosomes that bypass the *vpu* AUG. Although the scanning of capped mRNAs appears to be a general mechanism for translation initiation, it has been shown that in the case of picornaviruses internal initiation sites may exist (25, 33). Internal initiation is not responsible for Vpu and Env production from the bicistronic mRNAs in vivo. If internal initiation would occur, one should expect Env and Vpu production also from the 1.4E cDNA, which was not observed. Since the one-exon *tat* terminates upstream of both *env* and *vpu* ORFs, our data also exclude reinitiation of translation at the *vpu* or *env* AUG by ribosomes that have expressed the one-exon *tat* ORF on the 1.4E mRNA.

Both Env and Vpu expression were Rev dependent, indicating that production of these proteins occurs at a later stage in the viral life cycle. Functional studies of the Vpu protein have shown that Vpu increases the amount of viral particles released from infected cells (45–47). The expression of Env and Vpu from the same mRNA may reflect a need for coordinate regulation of Vpu and Env or a need to decrease the levels of Env production. This implies an additional mechanism for the regulation of expression of the viral proteins by modulation of translation of upstream ORFs.

ACKNOWLEDGMENTS

We thank M. Powers for oligonucleotide synthesis; L. Arthur for anti-Env antiserum; K. Strebel for anti-Vpu antiserum; J. Ghrayeb for anti-Nef antiserum; and D. Benko, D. Court, and S. Hughes for critical comments. We also thank M. Campbell, J. Harrison, and D. Vanderbur for technical assistance and A. Arthur and K. Toms for editorial assistance.

This research was sponsored in part by the National Cancer Institute under contract N01-CO-74101 with ABL, by the Swedish Cancer Society, and by the Swedish Medical Research Council.

ADDENDUM IN PROOF

After submission of the manuscript, two groups (J. C. Guatelli, T. R. Gingeras, and D. D. Richman, J. Virol. **64:** 4093–4098, 1990; and S. J. Arrigo, S. Weitsman, J. A. Zack, and I. S. Y. Chen, J. Virol. **64:**4585–4588, 1990) reported the identification of one of the mRNAs described here. Their data also suggest the presence of additional intermediate-size mRNAs, as described here. G. A. Viglianti, P. L. Sharma, and J. I. Mullins (J. Virol. **64:**4207–4216, 1990) reported that SIV also displays complex patterns of RNA splicing. Comparisons between HIV-1 and SIV show that the splicing characteristics of the two viruses are remarkably similar.

LITERATURE CITED

1. **Adachi, A., H. E. Gendelman, S. Koenig, T. Folks, R. Willey, A. Rabson, and M. A. Martin.** 1986. Production of acquired immunodeficiency syndrome-associated retrovirus in human and nonhuman cells transfected with an infectious molecular clone. J. Virol. **59:**284–291.
2. **Arya, S. K., and R. C. Gallo.** 1986. Three novel genes of human T-lymphotropic virus type III: immune reactivity of their products with sera from acquired immune deficiency syndrome patients. Proc. Natl. Acad. Sci. USA **83:**2209–2213.
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. **Braddock, M., A. Chambers, W. Wilson, M. P. Esnouf, S. E. Adams, A. J. Kingsman, and S. M. Kingsman.** 1989. HIV-1 tat 'activates' presynthesized RNA in the nucleus. Cell **58:**269–279.
5. **Chirgwin, J. M., A. E. Przybyla, R. J. MacDonald, and W. J. Rutter.** 1979. Isolation of biologically active ribonucleic acid from sources enriched in ribonuclease. Biochemistry **18:**5294–5299.
6. **Cohen, E. A., E. F. Terwilliger, J. G. Sodroski, and W. A. Haseltine.** 1988. Identification of a protein encoded by the vpu gene of HIV-1. Nature (London) **334:**532–534.
7. **Cullen, B. R.** 1986. Trans-activation of human immunodeficiency virus occurs via a bimodal mechanism. Cell **46:**973–982.
8. **Dayton, A. I., J. G. Sodroski, C. A. Rosen, W. C. Goh, and W. A. Haseltine.** 1986. The trans-activator gene of the human T-cell lymphotrophic virus type III is required for replication. Cell **44:**941–947.
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 the human immunodeficiency virus affects envelope-specific RNA localization. Cell **57:**1155–1165.
11. **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.
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 HIV-1 affects the stability and transport of the viral mRNA. Proc. Natl. Acad. Sci. USA **86:**1495–1499.
14. **Felber, B. K., and G. N. Pavlakis.** 1988. A quantitative bioassay for HIV-1 based on trans-activation. Science **239:**184–187.
15. **Feng, S., and E. C. Holland.** 1988. HIV-1 tat trans-activation requires the loop sequence within tar. Nature (London) **334:** 165–167.
16. **Fisher, A. G., M. B. Feinberg, S. F. Josephs, M. E. Harper, L. M. Marselle, G. Reyes, M. A. Gonda, A. Aldovini, C. Debouk, R. C. Gallo, and F. Wong-Staal.** 1986. The transactivator gene of HTLV-III is essential for virus replication. Nature (London) **320:**367–371.
17. **Garcia, J. A., D. Harrich, L. Pearson, R. Mitsuyasu, and R. B. Gaynor.** 1988. Functional domains required for tat-induced transcriptional activation of the HIV-1 long terminal repeat. EMBO J. **7:**3143–3147.
18. **Garcia, J. A., D. Harrich, E. Soultanakis, F. Wu, R. Mitsuyasu, and R. B. Gaynor.** 1989. Human immunodeficiency virus type 1 LTR TATA and TAR region sequences required for transcriptional regulation. EMBO J. **8:**765–778.
19. **Graham, F. J., and A. J. Van der Eb.** 1973. A new technique for the assay of infectivity of human adenovirus 5 DNA. Virology **52:**456–460.
20. **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.
21. **Hammarskjöld, M. L., J. Heimer, B. Hammarskjöld, 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.
22. **Hauber, J., and B. R. Cullen.** 1988. Mutational analysis of the *trans*-activation-responsive region of the human immunodeficiency virus type 1 long terminal repeat. J. Virol. **62:**673–679.
23. **Hauber, J., A. Perkins, E. P. Heimer, and B. R. Cullen.** 1987. Transactivation of human immunodeficiency virus gene expression is mediated by nuclear events. Proc. Natl. Acad. Sci. USA **84:**6364–6368.
24. **Jakobovits, A., D. H. Smith, E. B. Jakobovits, and D. J. Capon.** 1988. A discrete element 3′ of human immunodeficiency virus 1 (HIV-1) and HIV-2 mRNA initiation sites mediates transcriptional activation by an HIV *trans* activator. Mol. Cell. Biol. **8:**2555–2561.
25. **Jang, S., M. Davies, R. Kaufman, and E. Wimmer.** 1989. Initiation of protein synthesis by internal entry of ribosomes into the 5′ nontranslated region of encephalomyocarditis virus RNA in vivo. J. Virol. **63:**1651–1660.
26. **Knight, D. M., F. A. Flomerfelt, and J. Ghrayeb.** 1987. Expression of the art/trs protein of HIV and study of its role in viral envelope synthesis. Science **236:**837–840.
27. **Kozak, M.** 1978. How do eucaryotic ribosomes select initiation regions in messenger RNA? Cell **15:**1109–1123.
28. **Kozak, M.** 1989. The scanning model for translation: an update. J. Cell Biol. **108:**229–241.
29. **Malim, M. H., J. Hauber, S. Le, J. V. Maizel, and B. R. Cullen.** 1989. The HIV-1 rev transactivator acts through a structured target sequence to activate nuclear export of unspliced viral mRNA. Nature (London) **338:**254–257.

30. **Matsuda, Z., M. J. Chou, M. Matsuda, J. H. Huang, Y. M. Chen, R. Redfield, K. Mayer, M. Essex, and T. H. Lee.** 1988. Human immunodeficiency virus type 1 has an additional coding sequence in the central region of the genome. Proc. Natl. Acad. Sci. USA **85**:6968–6972.
31. **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.
32. **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.
33. **Pelletier, J., and N. Sonenberg.** 1988. Internal initiation of translation of eukaryotic mRNA directed by a sequence derived from poliovirus mRNA. Nature (London) **334**:320–324.
34. **Ratner, L., W. A. Haseltine, R. Patarca, K. L. 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. Chang, R. C. Gallo, and F. Wong-Staal.** 1985. Complete nucleotide sequence of the AIDs virus, HTLV-III. Nature (London) **313**:277–283.
35. **Rekosh, D., A. Nygren, P. Flodby, M. L. Hammarskjöld, and H. Wigzell.** 1988. Coexpression of human immunodeficiency virus envelope proteins and tat from a single simian virus 40 late replacement vector. Proc. Natl. Acad. Sci. USA **85**:334–338.
36. **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.
37. **Rosen, C. A., E. Terwilliger, A. Dayton, J. G. Sodroski, and W. A. Haseltine.** 1988. Intragenic cis-acting art gene-responsive sequences of the human immunodeficiency virus. Proc. Natl. Acad. Sci. USA **85**:2071–2075.
38. **Sadaie, M. B., T. Benter, and F. Wong-Staal.** 1988. Site-directed mutagenesis of two trans-regulatory genes (tat-III, trs) of HIV-1. Science **239**:910–913.
39. **Saiki, R. K., D. H. Gelfand, S. Stoffel, S. J. Scharf, R. Higuchi, G. T. Horn, K. B. Mullis, and H. A. Ehrlich.** 1988. Primer-directed enzymatic amplification of DNA with a thermostabile DNA polymerase. Science **239**:487–491.
40. **Saiki, R. K., S. Scharf, F. Faloona, K. B. Mullis, G. T. Horn, H. A. Erlich, and N. Arnheim.** 1985. Enzymatic amplification of β-globin genomic sequences and restriction site analysis for diagnosis of sickle cell anemia. Science **230**:1350–1354.
41. **Salfeld, J., H. Gottlinger, R. Sia, R. Park, J. Sodroski, and W. Haseltine.** 1990. A tripartite HIV-1 tat-env-rev fusion protein. EMBO J. **9**:965–970.
42. **Schwartz, S., B. K. Felber, D. M. Benko, E.-M. Fenyö, 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.
43. **Selby, M. J., E. S. Bain, P. A. Luciw, and B. M. Peterlin.** 1989. Structure, sequence, and position of the stem-loop structure in tar determine transcriptional elongation by tat through the HIV-1 long terminal repeat. Genes Dev. **3**:547–558.
44. **Sodroski, J., W. C. Goh, C. Rosen, A. Dayton, E. Terwilliger, and W. A. Haseltine.** 1986. A second post-transcriptional trans-activator gene required for the HTLV-III replication. Nature (London) **321**:412–417.
45. **Strebel, K., T. Klimkait, F. Maldarelli, and M. Martin.** 1989. Molecular and biochemical analyses of human immunodeficiency virus type 1 *vpu* protein. J. Virol. **63**:3784–3791.
46. **Strebel, K., T. Klimkait, and M. A. Martin.** 1988. A novel gene of HIV-1, vpu, and its 16-kilodalton product. Science **241**:1221–1223.
47. **Terwilliger, E., E. Cohen, Y. Lu, J. Sodroski, and W. Haseltine.** 1989. Functional role of human immunodeficiency virus type 1 vpu. Proc. Natl. Acad. Sci. USA **86**:5163–5167.
48. **Wright, C. M., B. K. Felber, H. Paskalis, and G. N. Pavlakis.** 1986. Expression and characterization of the trans-activator of HTLV-III/LAV virus. Science **234**:988–992.