JOURNAL OF VIROLOGY, Sept. 1990, p. 4585–4588
0022-538X/90/094585-04$02.00/0

Vol. 64, No. 9

Characterization and Expression of Novel Singly Spliced RNA Species of Human Immunodeficiency Virus Type 1

SALVATORE J. ARRIGO,[1] STACY WEITSMAN,[1] JEROME A. ZACK,[1] AND IRVIN S. Y. CHEN[1,2]*

Department of Microbiology and Immunology,[1] and the Department of Medicine,[2] UCLA School of Medicine and Jonsson Comprehensive Cancer Center, Los Angeles, California 90024-1678

Received 10 January 1990/Accepted 5 June 1990

Human immunodeficiency virus type 1 (HIV-1) expresses the Vif, Vpr, Vpu, and Env proteins through complex differential splicing of a single full-length RNA precursor. We used HIV-1-specific oligonucleotide primer pairs in a quantitative polymerase chain reaction procedure on RNA from fresh peripheral blood lymphocytes infected with HIV-1_{JR-CSF} to detect and characterize the singly spliced RNA species which might encode these proteins. The nucleotide sequences at the junctions of splice donor and acceptor sites of these RNAs were determined. One of these RNAs, which has not been previously described, appears to be a novel HIV-1 RNA encoding Env and/or Vpu proteins.

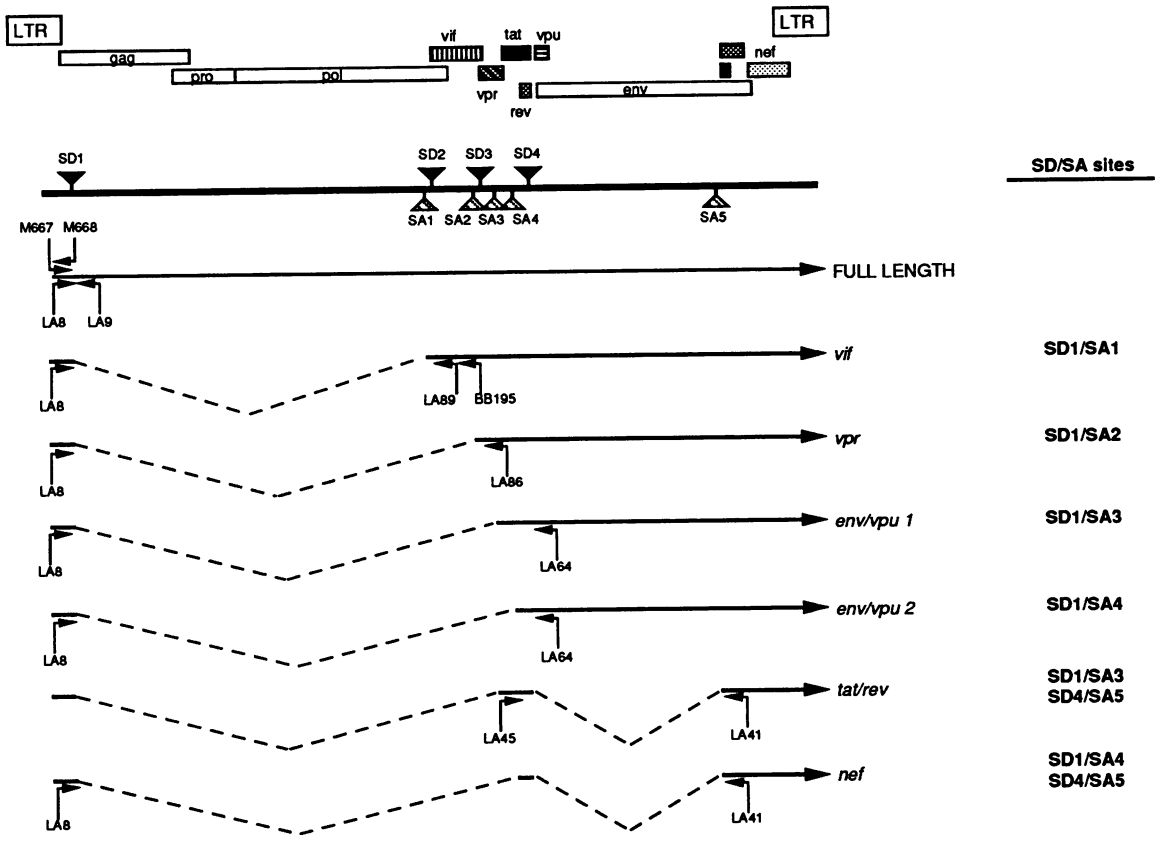
In addition to *gag*, *pol*, and *env*, the three genes found in all replication-competent retroviruses, human immunodeficiency virus type 1 (HIV-1) has at least seven additional genes, including *tat*, *rev*, *nef*, *vif*, *vpr*, *env*, and *vpu* (3, 5, 9, 11–16). Although the mRNAs with the potential to express some of these genes have been elucidated (1, 3, 7, 8, 10, 14), the RNAs encoding Vif, Vpr, Env, and Vpu proteins have not been clearly defined. Several 4.2- to 5.5-kilobase spliced RNA species that might encode these proteins have been detected by Northern (RNA) blot analysis (8). One of these RNAs with the potential to encode Env and Vpu proteins has been characterized by sequence analysis of cDNA clones (7). Another partial cDNA clone with the potential to encode Vif has also been characterized (14). Potential splice donors (SDs) and splice acceptors (SAs) which might be used in the splicing of these RNAs have been identified only in small doubly and triply spliced HIV-1 RNAs by cDNA cloning (1, 3, 7).

Oligonucleotides for the detection of HIV-1 RNAs. We generated HIV-1-specific oligonucleotide pairs to use in a quantitative RNA polymerase chain reaction (PCR) procedure for detection of singly spliced RNAs which might encode Vif, Vpr, Env, and Vpu proteins (Fig. 1A). We previously used this RNA PCR procedure to quantitatively detect HIV-1 *tat/rev*, *nef*, and full-length *gag/pol* RNAs, as well as total HIV-1 RNA (2). The oligonucleotide primer pairs to detect singly spliced RNAs were designed to flank potential SD-SA junctions. Vif, Vpr, Env, and Vpu proteins should be translated from RNAs which have utilized the 5′ SD (SD1) at nucleotide (nt) 743 of HIV-1_{JR-CSF}. On the basis of sequencing of cDNA clones, potential SAs for *vif* and *vpr* RNAs have been postulated to reside at nt 4924 (SA1) and nt 5401 (SA2) (7, 14). However, the original cDNA clones in which these SAs were detected were derived from RNAs which excised three introns, additionally utilizing the SDs at nt 4973 (SD2) and 5474 (SD3), and thus could not encode Vif or Vpr proteins. A partial cDNA clone which might encode Vif, utilizes SA1, and extends unspliced past SD2 and SD3 has been described previously (14). One cDNA clone derived from an RNA which might encode Vpu and Env has been described previously (7). This RNA uses an SA at nt 5788 (SA3). We generated 3′ oligonucleotide primers LA89 (nt 4995 to 4976), LA86 (nt 5495 to 5476), and LA64 (nt 6079 to 6060) to detect RNAs which utilize the known SAs of HIV-1 and which might encode the Vif, Vpr, Env, and Vpu proteins of HIV-1 (Fig. 1). These oligonucleotide primers, when used in PCR analysis with LA8 (nt 711 to 730) as the 5′ oligonucleotide primer, should detect HIV-1 RNAs which splice together the leader sequence of HIV-1 with any SA located upstream of the 3′ oligonucleotide primer. Any RNAs that utilize additional SDs and SAs which flank the 3′ oligonucleotide primer will not be detected. The sizes of PCR products that should be generated by these oligonucleotide primers on the various singly spliced RNAs are shown in Fig. 1B.

Quantitative detection of novel HIV-1 RNAs. RNA was prepared from phytohemagglutinin-stimulated peripheral blood lymphocytes (PBL) 8 days subsequent to infection with HIV-1_{JR-CSF} (4). This RNA was subjected to treatment with RNase-free DNase to remove all contaminating DNA and was used in RNA PCR analysis with oligonucleotide primer pairs to detect *vif*, *vpr*, and *env/vpu* RNAs. Oligonucleotide primers LA89 (nt 4995 to 4976), LA86 (nt 5495 to 5476), and LA64 (nt 6079 to 6060) were used in combination with oligonucleotide primer LA8 (nt 711 to 730) for detection of *vif*, *vpr*, and *env/vpu* RNAs, respectively. RNA was sequentially diluted threefold from 10^5 to 4×10^2 cell equivalents prior to reverse transcription and PCR amplification. Figure 2 shows that HIV-1-specific *vif*, *vpr*, and *env/vpu-2* RNAs can be detected quantitatively in HIV-1-infected PBL. The sizes of major bands generated by PCR amplification with these oligonucleotide primer pairs correspond directly to the predicted sizes of the products expected from the joining of the SDs and SAs shown in Fig. 1. These bands were not detected in uninfected PBL (data not shown). By using oligonucleotide pairs LA8 (nt 711 to 730) and LA64 (nt 6079 to 6060) to detect RNAs which could encode Env or Vpu, we detected a novel spliced RNA (*env/vpu-2*), which utilizes the joining of SD1 (nt 743) with an SA at nt 5987 (SA4). This RNA does not utilize SD4 (nt 6055), which is used in further splicing to *nef* RNA. The first and second methionines encountered within this novel RNA would be the initiation codons for Vpu and Env, respectively. This is in contrast to the previously identified *env/vpu-1* RNA, in which *vpu* and *env* utilize the third and fourth

* Corresponding author.

A



B

RNA Species	PCR Products						
	M667/M668	LA8/LA9	LA8/LA89	LA8/LA86	LA8/LA64	LA45/LA41	LA8/LA41
Full length	161	95	—	—	—	—	—
vif	161	—	105	—	—	—	—
vpr	161	—	—	128	—	—	—
env/vpu 1	161	—	—	—	325	—	—
env/vpu 2	161	—	—	—	126	—	—
tat/rev	161	—	—	—	—	123	—
nef	161	—	—	—	—	—	133

FIG. 1. Detection of HIV-1 RNAs by PCR. (A) Schematic representation of the HIV-1 genome showing the putative RNAs and the location and orientation of oligonucleotide primer pairs used to detect each RNA. The locations of potential SDs and SAs are indicated. Oligonucleotide primer LA8 (5′-GCGCGCACAGCAAGAGGCGA-3′, nt 711 to 730) is homologous to HIV-1 sequences just upstream of the 5′ SD site (SD1 at nt 743). It can be used in PCR analysis in combination with oligonucleotide primer LA89 (5′-TGTATTACTACTGC CCCTTC-3′, nt 4995 to 4976) on the cDNA products of reverse transcription reactions primed by oligonucleotide primer BB195 (5′-TATGTCACTATTATCTTGTA-3′, nt 5011 to 4992) to detect a potential *vif* RNA which utilizes the joining of SD1 (nt 743) with SA1 (nt

methionines as initiation codons. Minor bands were observed with several of the oligonucleotide pairs tested. DNA sequence analysis of several of these bands has shown that some of these bands are due to nonspecific hybridization to both cellular sequences and other sites within HIV during reverse transcription (data not shown). Some of these bands may represent minor species of RNA produced during infection of PBL.

We have not successfully detected the singly spliced RNA joining SD1 to SA3 (*env/vpu-1*) previously observed by cDNA cloning. Since the *tat/rev* and *env/vpu-1* RNAs both employ SA3, the 3′ oligonucleotide primer that discriminates *env/vpu-1* from *tat/rev* RNAs must be positioned 3′ to SD4 (used only by *tat/rev* RNAs); thus, oligonucleotide primer LA64 (nt 6079 to 6060) was used. Both *env/vpu-1* and *env/vpu-2* RNAs would be predicted to be detected by the LA8-LA64 oligonucleotide primer pair (nt 711 to 730 and nt 6079 to 6060); however, perhaps due to the large size (325 base pairs) of the anticipated PCR product for *env/vpu-1* RNA, only the smaller PCR product for *env/vpu-2* RNA is detected. The inability to detect this larger *env/vpu* RNA does not suggest that this RNA is not produced, since we have previously shown that when oligonucleotide primer pairs are used in this RNA PCR procedure, which should detect more than one RNA species, only the smaller PCR product is readily detected (2). We have also been unable to detect a distinct RNA species encoding Rev (10) in infected PBL by using various oligonucleotide primer pairs (2; data not shown).

Sequencing of RNA PCR products. In order to verify the identity of the PCR products generated with each set of oligonucleotide primer pairs, the major bands were subjected directly to DNA sequence analysis (6). The results of this DNA sequence analysis are shown in Table 1. We were able to demonstrate the correct joining of SD1 with SA1 for *vif* RNA, SD1 with SA2 for *vpr* RNA, and SD1 with SA4 for *env/vpu-2* RNA. DNA sequence analysis additionally demonstrated that the *vif*, *vpr*, and *env/vpu-2* RNAs detected by PCR analysis do not utilize the SDs, SD2, SD3, and SD4, which have previously been associated with doubly and triply spliced RNA species (1, 3, 7). We conclude that the RNAs detected have the potential to encode Vif, Vpr, Env, and Vpu proteins. These RNAs most likely correspond to the 4.2- to 5.5-kilobase HIV-1 RNAs detected by Northern blot analyses (8).

We generated oligonucleotide primer pairs to be employed in a quantitative RNA PCR procedure to detect *vif*, *vpr*, and *env/vpu-2* HIV-1-specific RNA species. We identified RNAs which join the 5′ HIV-1 SD (SD1) with previously detected SAs (SA1, SA2, and SA4), which have the potential to encode Vif, Vpr, Env, and Vpu proteins. Although usage of these SAs has been previously demonstrated by sequencing of cDNA clones (7), the RNAs which had been detected were, in all cases, additionally spliced at SD2, SD3, and SD4. A partial cDNA clone which utilizes SA1 but not SD2

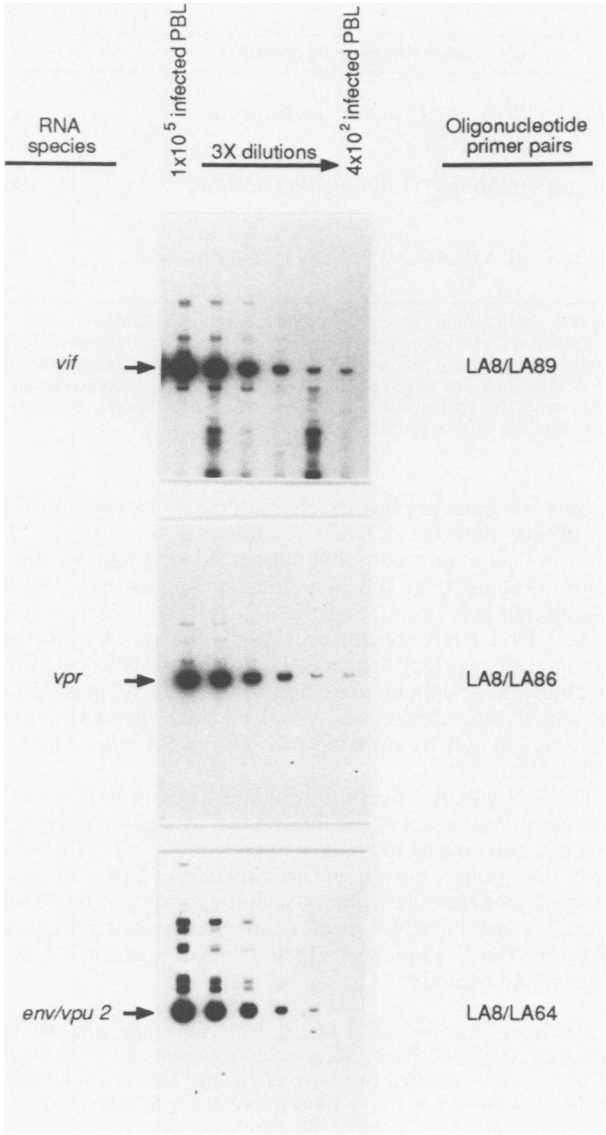


FIG. 2. Quantitative detection of HIV-1 RNAs. RNA PCR analysis was performed on cell equivalents of RNA from the indicated number of PHA-stimulated PBL harvested on day 8 after infection with HIV-1_{JR-CSF}. RNA was prepared by the method of Arrigo et al. (2). The same conditions for RNA PCR analysis as those described by Arrigo et al. were also used, with the exception that the PCR denaturation temperature was changed to 94°C. The PCR products were resolved on a 6% polyacrylamide gel and visualized by autoradiography. Oligonucleotide primer pairs used to detect each RNA species are shown. The sections of the gels shown span approximately 70 to 350 base pairs, as determined by molecular weight markers.

4924). It will not detect an RNA which additionally utilizes SD2 (nt 4973). The use of different oligonucleotides in the reverse transcription and PCR amplifications helped to reduce the background with this oligonucleotide set. Oligonucleotide primer LA8 can also be used in RNA PCR analysis in combination with oligonucleotide primer LA86 (5′-CAAGTACTGTAGAGATCCTA-3′, nt 5495 to 5476) to detect a potential *vpr* RNA that utilizes the joining of SD1 (nt 743) with SA2 (nt 5401), but it will not detect an RNA which additionally utilizes SD3 (nt 5474). Oligonucleotide primer LA8 can also be used in an RNA PCR in combination with oligonucleotide primer LA64 (5′-GGTTGCATTACAT GCACTAC-3′, nt 6079 to 6060) to detect potential *env/vpu* RNAs which utilize the joining of SD1 (nt 743) with SA3 (nt 5788) or SA4 (nt 5987) but do not utilize SD4 (nt 6055). SD4 (nt 6055) is used by both *tat/rev* and *nef* RNAs. Oligonucleotide primers M667, M668, LA9, LA45, and LA41 have been previously described (2). The SD and SA sites used to generate each RNA are shown. LTR, long terminal repeat. (B) Sizes of amplified PCR products which should be generated by RNA PCR analysis with pairs of oligonucleotide primers specific for the HIV-1 RNAs shown in panel A.

TABLE 1. Sequences of SDs and SAs[a]

RNAs (oligonucleotide primer pairs)	Sequence		
	SD1		SA1
vif (LA8/LA89; nt 711 to 730, 4995 to 4976)	GGGGCGGCGACTG (743)	GTGAG.....TTTATTACAG	GGACAACA (4924)
	SD1		SA2
vpr (LA8/LA86; nt 711 to 730, 5495 to 5476)	GGGGCGGCGACTG (743)	GTGAG.....TGTTTTTCAG	AATCTGCT (5401)
	SD1		SA4
env/vpu-2 (LA8/LA64; nt 711 to 730, 6079 to 6060)	GGGGCGGCGACTG (743)	GTGAG.....CCTATGGCAG	GAAGAAGC (5987)

[a] PCR amplifications were resolved on a 6% polyacrylamide gel, and the major band was excised. DNA was eluted in 500 mM ammonium acetate and 1 mM EDTA and was precipitated with ethanol. A fraction of this eluent was reamplified for an additional 25 cycles of PCR. PCR conditions for both amplifications were identical, although the second amplification included a fourfold increase in all reagents. Since the PCR products were end labeled with [γ-^{32}P]ATP, a portion of the PCR product was subjected directly to DNA sequence analysis by the method of Maxam and Gilbert (6). The sequences include SDs, introns, and SAs, from left to right. The vertical line indicates the exon-intron borders. Numbers in parenthesis shown under nucleotide sequences correspond to the nucleotides to the immediate left (SD) or right (SA) of the vertical lines.

or SD3 has been previously characterized (14). It is a formal possibility that the RNAs we detected for *vif*, *vpr*, and *env/vpu-2* contain additional splices downstream of the regions homologous to our oligonucleotides; however, we feel this is unlikely, because of the 4.2- to 5.5-kilobase sizes of these HIV-1 RNA species previously detected by Northern blot analyses with probes specific for these RNAs (8). Taken together, these data indicate that Vif, Vpr, Env, and Vpu are most likely encoded on singly spliced RNA species 4.2 to 5.5 kilobases in size by utilizing the SD and SA sites which we define here.

This RNA PCR procedure provides an extremely sensitive and quantitative assay for analysis of low abundance RNA species, allowing us to analyze expression of HIV-1 RNAs at early time points subsequent to infection of PBL. By generating oligonucleotide primers which span potential SD and SA pairs, specific RNA species can be examined individually and the DNA sequence of the splice junctions can be determined directly.

This work was supported by the University of California Task Force on AIDS and Public Health Service grants AI 27221 and AI 29107 from the National Institutes of Health. S.J.A. is a Leukemia Society Fellow, I.S.Y.C. is a Leukemia Society Scholar, and J.A.Z. is a recipient of a National Research Service award from the National Institutes of Health and a grant from the California Institute for Cancer Research.

We thank P. Green, W. O'Brien, K. Arrigo, S. Pang, and M. Yip for helpful comments and W. Aft for preparation of the manuscript.

LITERATURE CITED

1. **Ahmad, N., and S. Venkatesan.** 1988. *nef* protein of HIV-1 is a transcriptional repressor of HIV-1 LTR. Science **241**:1481–1485.
2. **Arrigo, S. J., S. Weitsman, J. D. Rosenblatt, and I. S. Y. Chen.** 1989. Analysis of *rev* gene function on human immunodeficiency virus type 1 replication in lymphoid cells by using a quantitative polymerase chain reaction method. J. Virol. **63**:4875–4881.
3. **Arya, S. K., C. Guo, S. J. Josephs, and F. Wong-Staal.** 1985. Transactivator gene of human T-lymphotropic virus type III. Science **229**:69–73.
4. **Koyanagi, Y., S. Miles, R. T. Mitsuyasu, J. E. Merrill, H. V. Vinters, and I. S. Y. Chen.** 1987. Dual infection of the central nervous system by AIDS viruses with distinct cellular tropisms. Science **236**:819–822.
5. **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.
6. **Maxam, A. M., and W. Gilbert.** 1977. A new method for sequencing DNA. Proc. Natl. Acad. Sci. USA **74**:560–564.
7. **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.
8. **Rabson, A. B., D. F. Daugherty, S. Venkatesan, K. E. Boulukos, S. I. Benn, T. M. Folks, P. Fiorino, and M. A. Martin.** 1985. Transcription of novel open reading frames of AIDS retrovirus during infection of lymphocytes. Science **229**:1388–1390.
9. **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, Jr., 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**:277–284.
10. **Sadie, M. R., J. Rappaport, T. Benter, S. F. Josephs, R. Willis, and F. Wong-Staal.** 1988. Missense mutations in an infectious human immunodeficiency viral genome: functional mapping of *tat* and identification of the *rev* splice acceptor. Proc. Natl. Acad. Sci. USA **85**:9224–9228.
11. **Sodroski, J., W. C. Goh, C. Rosen, A. Dayton, E. Terwillinger, and W. Haseltine.** 1986. A second post-transcriptional *trans*-activator gene required for HTLV-III replication. Nature (London) **321**:412–417.
12. **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.
13. **Sonigo, P., M. Alizon, K. Staskus, D. Klatzmann, S. Cole, O. Danos, E. Retzel, P. Tiollais, A. Haase, and S. Wain-Hobson.** 1985. Nucleotide sequence of the visna lentivirus: relationship to the AIDS virus. Cell **42**:369–382.
14. **Strebel, K., D. Daugherty, K. Clouse, D. Cohen, T. Folks, and M. A. Martin.** 1987. The HIV 'A' (*sor*) gene product is essential for virus infectivity. Nature (London) **328**:728–731.
15. **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.
16. **Wain-Hobson, S., P. Sonigo, O. Danos, S. Cole, and M. Alizon.** 1985. Nucleotide sequence of the AIDS virus, LAV. Cell **40**:9–17.