

Circular DNA of Human Immunodeficiency Virus: Analysis of Circle Junction Nucleotide Sequences

TAO HONG, KARL DRLICA, ABRAHAM PINTER, AND ELLEN MURPHY*

Public Health Research Institute, 455 First Avenue, New York, New York 10016

Received 14 August 1990/Accepted 10 October 1990

During infection of cells by retroviruses, some of the nonintegrated viral DNA can be found as a circular form containing two tandem, directly repeated long terminal repeats. The nucleotide sequence at the point where the long terminal repeats join (the circle junction) can be used to deduce the terminal nucleotides of the linear form of the viral DNA. Comparison of the termini of linear viral DNA with sequences at the junctions between the integrated provirus and the host chromosome has revealed that for most retroviruses 2 bp are removed from each end of the linear viral DNA during integration. For human immunodeficiency virus type 1 (HIV-1), however, sequence considerations involving primer-binding sites had suggested that only 1 bp is removed during integration. We obtained the nucleotide sequences at the ends of HIV-1 DNA by using the polymerase chain reaction to amplify fragments corresponding to the HIV-1 circle junction. Of 17 clones containing amplified sequences, 10 had identical circle junctions that contained an additional 4 bp (GTAC) relative to the integrated provirus. This indicates that, as for other retroviruses, 2 bp are removed from each end of the linear HIV-1 viral DNA during integration. The remaining seven isolates contained insertions or deletions at the circle junction.

Cells infected with retroviruses contain integrated viral DNA and both linear and circular forms of nonintegrated DNA. The linear form is thought to be the substrate for integration (2, 8, 21). One of the circular species contains two tandemly repeated long terminal repeats (LTRs); in these molecules, the junction between the LTRs, the circle junction, corresponds to the ligated ends of linear viral DNA. Thus, the nucleotide sequence across the circle junction can be used to determine the sequences at the ends of the linear viral DNA. For a variety of murine and avian retroviruses, the linear DNA contains 2 bp at each end which are not present at the junctions between the integrated proviral DNA and the host chromosome (4, 11, 25, 26, 31). Removal of these 4 bp arises from a two-step reaction involving the viral integrase, the product of the 3' domain of the *pol* gene (5, 17, 18, 24). In the first step, integrase removes two nucleotides from the 3' ends of linear viral DNA. In the second step, the target DNA is cleaved and its 5' ends are ligated to the recessed 3' ends of the viral DNA (7, 8). The 2-nucleotide (nt) protrusions at the 5' ends of the viral DNA are presumably removed before ligation with the 3' ends of the host DNA.

The positioning of RNA primers used by reverse transcriptase causes linear viral DNA to extend slightly beyond U5 at one end and U3 at the other (Fig. 1). For initiation of minus-strand DNA synthesis, priming involves hybridization of the 3' end of a tRNA to about 18 nt within the unique portion of the genome just outside U5. For plus-strand DNA synthesis, priming involves hybridization of a short oligonucleotide to a nucleotide sequence in the unique region of the genome just outside U3 (10, 32). With avian and murine retroviruses, the primer-binding sites for both minus- and plus-strand DNA synthesis end 2 nt away from U5 and U3, respectively. Thus, the 2 bp removed from each end during integration are protrusions beyond the ends of the LTRs that are generated by the position of the primer. For human immunodeficiency virus type 1 (HIV-1), however,

only 1 nt separates the tRNA primer-binding site and U5 (30) (the plus-strand primer ends 2 nt from U3 [10]). Consequently, it has been suggested that HIV-1 integration involves asymmetric removal of 2 bp from U3 and 1 from U5 (28, 32). Below we show that this is not the case: 2 bp are removed from each end of viral DNA during integration of HIV-1.

We used the polymerase chain reaction (PCR) to amplify a fragment of approximately 250 bp corresponding to the circle junction sequence of DNA extracted from H9 cells infected with HIV-1 (Fig. 2). Separate experiments were carried out with two strains of HIV-1. One was a viral stock of HIVNL43 obtained by previous transfection with molecular clone pNL4-3, obtained from M. Martin (1). The other was ARV2 (23), obtained from D. Dina of Chiron Corp. Infected cells were grown in RPMI 1640 medium containing 10% fetal bovine serum. When >90% of the cells stained positive for HIV-1 by immunofluorescence with positive human anti-HIV-1 serum (3 to 5 days), cells were harvested by centrifugation and washed and DNA was extracted by a modified Hirt procedure (9). PCR primers which allowed efficient amplification of circle junction sequences were selected; alternative full-length products are not expected to be efficiently amplified or cloned because of their large sizes (Fig. 2). For each reaction, 1 µg of the template, 125 nmol of each primer (described in the legend to Fig. 2), and 5 U of *Taq* polymerase were used. Reaction mixtures were incubated at 94°C for 1 min, 37°C for 2 min, and 72°C for 3 min for 35 cycles. The amplified DNA was extracted with phenol-chloroform, precipitated, cleaved with *Hind*III and *Bam*HI, ligated to M13mp18, and transformed into *Escherichia coli* JM109 (15, 35). Single-stranded DNA was sequenced by using [³⁵S]dATP and T7 DNA polymerase (Sequenase; US Biochemicals).

Seventeen M13mp18 clones containing HIV-1 LTR-specific sequences were identified by sequence analysis (Fig. 3). Of the 17 clones, 10 (N25, N81, N51, N70, A1, A5, A3, A8, A46, and A47) had circle junction sequences identical to those at the ends of the provirus (19), except for the addition of 4 nt, GTAC, at the centers of the circle junctions.

* Corresponding author.

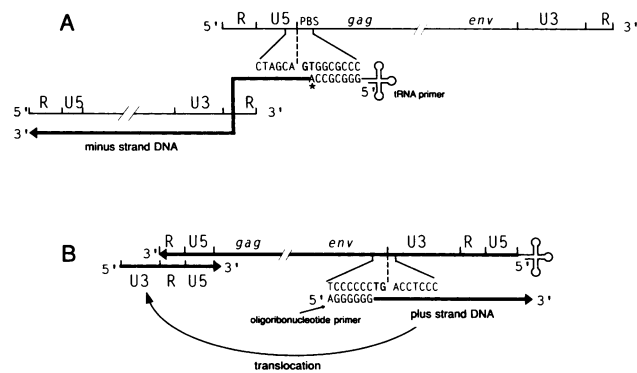


FIG. 1. Initiation of HIV-1 DNA synthesis with possible genesis of deletion and insertion mutations. The sequence of events involved in production of double-stranded viral DNA is diagrammed in simplified form. For details, see reference 33. Thin lines represent RNA, and thick lines represent DNA; a few base pairs of the sequences surrounding the U5-unique and U3-unique borders are shown. The two nucleotides bordering U5 (A) or U3 (B) that are copied to generate the 2-nt extensions on the linear DNA are shown in boldface. (A) Initiation of minus-strand DNA synthesis. Reverse transcription is initiated by binding of the tRNA^{lys} primer on the viral RNA template at the primer-binding site (PBS). Translocation to a second template occurs within R. The asterisk indicates the terminal adenosine residue of the primer tRNA which must be removed before initiation of DNA synthesis to generate the predominant form of circle junction clone. If the tRNA primer contains additional bases at its 3' end or if synthesis is primed by a random oligonucleotide hybridizing within U5, deletions of U5 will result. (B) Initiation of plus-strand DNA synthesis. An oligoribonucleotide primer 19 nts long, generated from the genomic RNA by RNase H, hybridizes 2 nt upstream from U3 for initiation of plus-strand synthesis (10), which continues until the first modified base of the tRNA is encountered (plus short-stop DNA). The tRNA is removed by RNase H. Use of an incorrect primer, hybridizing to sequences within U3, will generate U3 deletions; positioning of the primer further into the unique portion of the viral genome would generate insertions at the circle junction containing portions of the unique region. Translocation of the plus short-stop DNA to the 3' end of the minus strand allows extension of the minus strand through U3 and extension of the plus strong-stop DNA through the full length of the genome. Failure of RNase H to remove the tRNA primer or failure of the strong-stop DNA to translocate could result in copying of tRNA sequences into the plus strand at the 3' end of the viral DNA; subsequent removal of the ribonucleotides, ligation of the plus strand, and filling in of the minus strand would generate a circle junction containing an insertion of tRNA sequences.

Removal of these 4 bp unambiguously generated the sequences of the LTRs found in integrated proviral DNA. Thus, integration of HIV-1, as for other retroviruses, involves removal of 2 bp from each end of the linear viral DNA. This conclusion has also been reached in two other studies reported recently (14, 34).

A dilemma arises from the observation that for HIV-1, the primer-binding site for minus-strand synthesis is only 1 nt from the 3' end of U5 (Fig. 1A). The two most likely mechanisms for generating the observed circle junction sequences are (i) removal of the 3' adenosine from the primer tRNA before reverse transcription and (ii) failure of RNase H to remove the terminal riboadenine nucleotide of the primer from the completed DNA. Although there are precedents for the use of truncated tRNA primers (6, 13, 22) or incorrectly processed tRNAs (16), Whitcomb et al. (34) have argued that the tRNA primer used for HIV-1 replication is intact and that the 5'-terminal A of the minus strand of

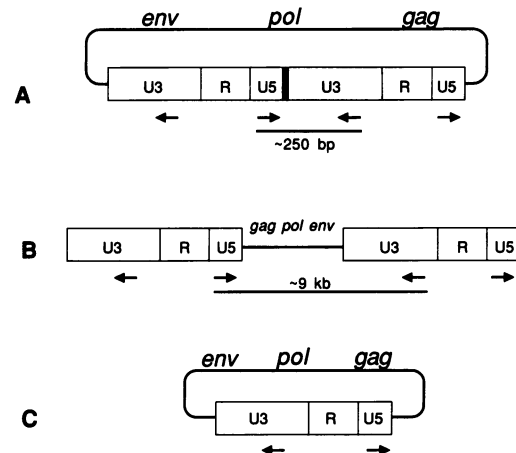


FIG. 2. Strategy for PCR amplification of circle junctions. (A) An approximately 250-bp fragment was amplified across the circle junction (thick black bar) of a two-LTR circle. Theoretically, amplification with the same primers would also yield a nearly full-length (approximately 9-kb) viral DNA molecule from a two-LTR circle (A), a linear template (B), or a single-LTR circle (C); no such clones were recovered. The primers were 5'-GCCTCAATA AAGCTTGCTTGTGAGTGC-3' (U5) and 5'-CAGGATCCAAAGGTC AGTGGATATCTG-3' (U3), which correspond to nts 522 to 547 of U5 and nts 132 to 110 of U3 of the HIVNL43 sequence (GenBank accession number M19921) (2a). Nucleotides added or altered to produce the *Bam*HI restriction site in the U3 primer are shown in lowercase letters. The *Hind*III site in the U5 primer occurs at nt 531.

U5 is a ribonucleotide derived from the primer. Occasional removal of this ribonucleotide by RNase H would generate a linear viral DNA molecule lacking a T at the 3' end of U5. We found one clone (A20) that could be explained in this way.

The GTAC at the circle junction forms the central tetranucleotide of an *Sca*I restriction site (AGTACT); cleavage occurs precisely at the circle junction. Since cleavage by *Sca*I produces flush-ended DNA, digestion of circle junction clones with this enzyme generates a linear DNA molecule that has at its ends the precise sequence found in the presumed integration substrate, linear viral DNA. This DNA may be a useful substrate for analyzing biochemical properties of integrase, such as its expected 3' recessing activity (2). The circular form itself may be a useful substrate for analysis of certain interactions between DNA and integrase, although it is probably not a productive intermediate for viral integration.

Two clones contained insertions similar in origin to those that have been reported for other retroviruses. One, N72, had what appeared to be a 342-bp insertion at the circle junction; the insertion corresponded to a portion of the 5' unique region of the virus, starting at the primer-binding site (nt 637) and ending at nt 978 in *gag*. Similar murine leukemia virus and *copia* clones have been interpreted as arising from integration of one viral DNA into another (6, 27). The insert in N72 contained the sequences located between the 5' LTR of the target molecule and the 5' LTR of the integrating viral DNA; the U3 boundary represents the 5' end of the integrated virus. N72 lacked the 2 nt (AC) at the outer border of U3 of the integrating virus, supporting the conclusion that integration involves removal of 2 bp from U3. Another clone, A34, contained a remnant of the tRNA primer, a 16-bp insertion precisely at the circle junction that corre-

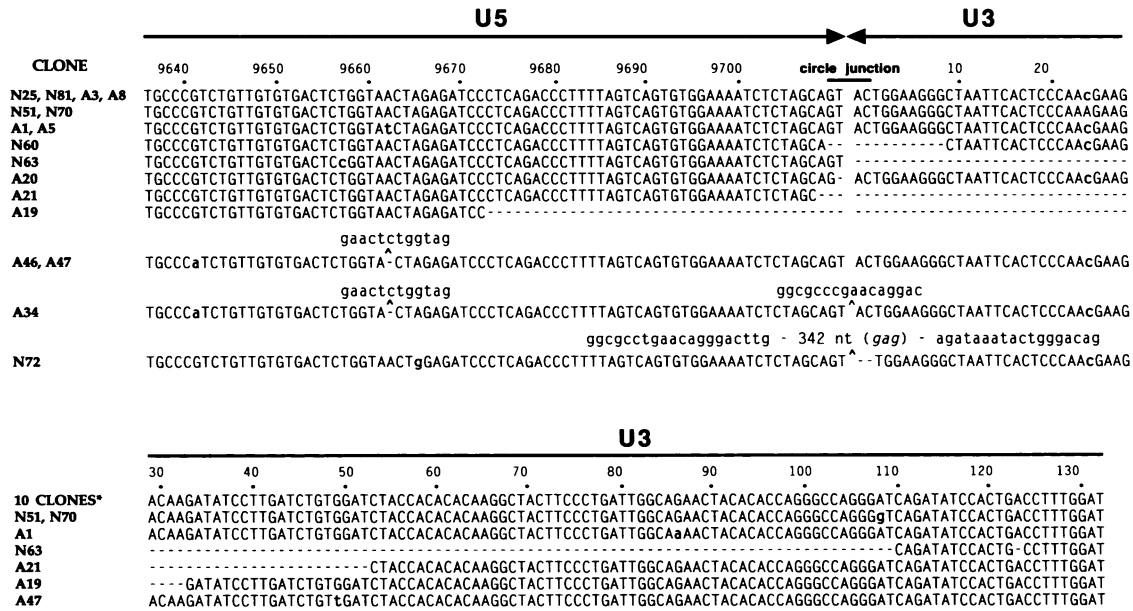


FIG. 3. Nucleotide sequences of cloned circle junctions. Circle junctions amplified by PCR were cloned, and the nucleotide sequences were determined (see the text). Each line represents a unique sequence; clones containing that sequence are indicated at the left. In the first line of the lower set of sequences, the 10 clones with identical sequences are A3, A8, A5, A20, A34, A46, N25, N60, N72, and N81. U3 is numbered starting from nt 1, since it is derived from the 5' LTR; U5 is numbered from nt 9636, since it is derived from the 3' LTR. The additional 4 bp at the circle junction are not included in the numbering scheme. Only part of the U5 portion of the amplified fragment is shown; the entire fragment was sequenced in each case; and no deviations from the published sequence were found in the region of U5 not shown. The prefix A indicates clones from cells infected with ARV2, and the prefix N indicates clones from HIVNL43 infection. Dashed lines indicate deleted sequences. The sequence shown for clones N25 and N81 is identical to that of the HIVNL43 GenBank entry (2a), except for the substitution of C for A at nt 24 in U3. This and other substitutions relative to the HIVNL43 sequence are shown in lowercase letters.

sponded to 17 bp of the tRNA primer-binding site (nts 637 to 653), less one internal nucleotide. Such clones have been postulated to arise from the failure of RNase H to remove the entire tRNA primer from minus-strand DNA, which results in copying of the tRNA sequence into the 3' end of the plus-strand DNA (3,29). An alternative explanation for the origin of clone A34 is failure of the plus strong-stop DNA to translocate to the 3' end of the minus strand (Fig. 1). Clone A34 also contained a second insertion: an adenine at nt 9662 in U5 was replaced by a 12-bp imperfect tandem duplication of the adjacent sequence (nts 9652 to 9662). An identical duplication was found at the same location in two additional clones, A46 and A47, both of which contained wild-type circle junction sequences. Thus, this duplication probably represents a variant present in the infecting virus population.

Nearly one-third of our clones contained deletions that either spanned the circle junction or terminated precisely at the U3-U5 boundary. Clone N63 had a 109-bp deletion starting at the junction and extending only into U3. Three other clones contained deletions at 12 (N60), 55 (A21), and 73 (A19) bp that spanned the circle junction. Similar deletions have been found at high frequencies in unintegrated murine or avian retroviral DNAs (11, 12, 16, 26, 27). In these systems, a bias toward deletion of U3 is evident; our sample size was too small to support an equivalent conclusion. Deletions spanning the circle junction or having one endpoint at the circle junction are most easily explained as arising from incorrect priming or from exonucleolytic digestion of one or both ends of the viral DNA before ligation. Both explanations are consistent with the absence of deletions outside the junction sequences, although mispriming

might more easily explain any bias in deletion frequency of U3 over U5, since the mechanisms of priming of U3 and U5 differ.

None of our deletions lacked only the GT nucleotides of U5. Such deletions, suggested to arise via the activity of the viral integrase, represented half of the clones found by Kulkosky et al. (14). However, deletions extending more than 2 bp into U3 and U5, as described above, were obtained at a much higher frequency here than in two other studies. In those studies, the relative recovery of full-length clones was increased by gel purification of the amplified DNA before cloning (34) or by screening of clones for full-length inserts before sequencing (14). Thus, our results probably reflect the situation in vivo more accurately, since there was no preselection based on insert size.

To determine whether PCR provided a representative sample of the population of circle junction sequences, we performed several reconstruction amplifications. Cloned DNAs from M13 clones containing wild-type circle junction (clone A1), plus DNA from either clone A19 (73-bp deletion at the circle junction) or N72 (342-bp insertion), were added to 4 µg of DNA from uninfected H9 cells to approximately 0.4 to 4 viral molecules per cell. The molar ratio of wild-type to mutant clones was varied over a 10-fold range, and the mixtures were sampled after 25, 30, and 35 cycles of amplification. In all cases, amplified DNAs were obtained in amounts that were roughly proportional to the concentrations of template DNA clones (data not shown). Proportional results were also obtained when DNAs from all three clones were mixed in equal ratios or in a 10:10:1 ratio. The results were unaffected by omission of H9 genomic DNA or addition of 10-fold excess *ScaI*-digested clone A1 DNA to

simulate the presence of linear viral DNA. These experiments indicate that preexisting deletions are not preferentially amplified relative to wild-type circle junctions. Thus, it is likely that our results do reflect a high proportion of mutant circle junction-containing DNA in infected cells.

The clones we examined exhibited a small amount of sequence variation in regions distant from the circle junction. None of our clones had a sequence identical to that of HIVNL43; they differed by up to 3 nt (lowercase letters in Fig. 3). For the seven clones obtained after infection with molecular clone HIVNL43, the error rate was 0.4% (seven point mutations in 1,638 nts sequenced). This is considerably higher than the misincorporation rate we have experienced with PCR in other systems (approximately one error per 3,000 nts sequenced; 15a) and is within the range of reported misincorporation rates for HIV-1 reverse transcriptase (20). Thus, these differences are likely to reflect real variation occurring during virus replication.

There are seven positions within the amplified fragment at which the ARV2 and HIVNL43 sequences differ. All 17 clones contained the HIVNL43 sequence, even though 10 were obtained after infection with ARV2. It is likely that the ARV2 nucleotide sequence, which was obtained from a molecular clone, represents only a small fraction of the sequences present in the uncloned ARV2 virus stock. The major variant present in this stock is a virus more closely related to HIVNL43 in this region.

It is now clear that 2 bp are removed from each end of HIV-1 during integration. It is also evident that a large proportion of the circular DNA molecules in infected cells carry mutations at the circle junction. It may be that circular viral DNA is derived preferentially from defective linear molecules. If not, our observations could indicate that a large proportion of the linear, integrative form of HIV-1 viral DNA normally has incorrect termini, making it unable to integrate. A better understanding of mutant formation may lead to ways to increase this fraction, thereby blocking productive infection.

We thank D. Dina and M. Martin for providing strains, Jim Groarke and Joseph Hughes for stimulating discussions, and Samuel Kayman for critical comments on the manuscript.

Support for this work was provided by the Aaron Diamond Foundation, The American Foundation for AIDS Research, and the U.S. Army Medical Research and Development Command under contract DAMD17-88-C-8126. T.H. is an NIH predoctoral trainee of the Department of Microbiology, New York University School of Medicine.

REFERENCES

- Adachi, A., H. Gendelman, S. Koenig, T. Folks, R. Willey, A. Rabson, and M. Martin. 1986. Production of acquired immunodeficiency syndrome-associated retrovirus in human and non-human cells transfected with an infectious molecular clone. *J. Virol.* **59**:284-291.
- Brown, P., B. Bowerman, H. E. Varmus, and J. M. Bishop. 1989. Retroviral integration: structure of the initial covalent product and its precursor, and a role for the viral IN protein. *Proc. Natl. Acad. Sci. USA* **86**:2525-2529.
- Buckler, C. Personal communication.
- Colicelli, J., and S. P. Goff. 1986. Structure of a cloned circular retroviral DNA containing a tRNA sequence between the terminal repeats. *J. Virol.* **57**:674-677.
- Dhar, R., W. L. McClements, L. W. Enquist, and G. F. Vande Woude. 1980. Nucleotide sequences of integrated Moloney sarcoma provirus long terminal repeats and their host and viral junctions. *Proc. Natl. Acad. Sci. USA* **77**:3937-3941.
- Donehower, L. A., and H. E. Varmus. 1984. A mutant murine leukemia virus with a single missense codon in *pol* is defective in a function affecting integration. *Proc. Natl. Acad. Sci. USA* **81**:6461-6465.
- Flavell, A. J., and D. Ish-Horowicz. 1983. The origin of extra-chromosomal circular *copia* elements. *Cell* **34**:415-419.
- Fujiwara, T., and R. Craigie. 1989. Integration of mini-retroviral DNA: a cell-free reaction for biochemical analysis of retroviral integration. *Proc. Natl. Acad. Sci. USA* **86**:3065-3069.
- Fujiwara, T., and K. Mizuuchi. 1988. Retroviral DNA integration: structure of an integration intermediate. *Cell* **54**:497-504.
- Hirt, B. 1967. Selective extraction of polyoma DNA from infected mouse cell cultures. *J. Mol. Biol.* **26**:362-371.
- Huber, H. E., and C. C. Richardson. 1990. Processing of the primer for plus strand DNA synthesis by human immunodeficiency virus 1 reverse transcriptase. *J. Biol. Chem.* **265**:10565-10573.
- Ju, G., and A. M. Skalka. 1980. Nucleotide sequence analysis of the long terminal repeat (LTR) of avian retroviruses: structural similarities with transposable elements. *Cell* **22**:379-386.
- Katz, R. A., C. A. Omer, J. H. Weis, S. A. Mitsialis, A. J. Faras, and R. V. Guntaka. 1982. Restriction endonuclease and nucleotide sequence analyses of molecularly cloned unintegrated avian tumor virus DNA: structure of large terminal repeats in circle junctions. *J. Virol.* **42**:346-351.
- Kikuchi, Y., Y. Ando, and T. Shiba. 1986. Unusual priming mechanism of RNA-directed DNA synthesis in *copia* retrovirus-like particles in *Drosophila*. *Nature (London)* **323**:824-826.
- Kulkosky, J., R. A. Katz, and A. M. Skalka. 1990. Terminal nucleotides of the preintegrative linear form of HIV-1 DNA deduced from the sequence of circular DNA junctions. *J. Acquired Immune Defic. Syndr.* **3**:852-858.
- Messing, J. 1983. New M13 vectors for cloning. *Methods Enzymol.* **101**:10-89.
- Murphy, E. Unpublished data.
- Olsen, J., and R. Swanstrom. 1985. A new pathway in the generation of defective retrovirus DNA. *J. Virol.* **56**:779-789.
- Panganiban, A. T., and H. M. Temin. 1984. Circles with two tandem LTRs are precursors to integrated retrovirus DNA. *Cell* **36**:673-679.
- Quinn, T., and D. Grandgenett. 1988. Genetic evidence that the avian retrovirus DNA endonuclease domain of *pol* is necessary for viral integration. *J. Virol.* **62**:2307-2312.
- 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, I. 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.
- Roberts, J. D., K. Bebenek, and T. A. Kunkel. 1988. The accuracy of reverse transcriptase from HIV-1. *Science* **242**:1171-1173.
- Roth, M., P. Schwartzberg, and S. Goff. 1989. Structure of the termini of DNA intermediates in the integration of retroviral DNA: dependence on IN function and terminal DNA sequence. *Cell* **58**:47-54.
- Saigo, K. 1986. A potential primer for reverse transcription of *mdg3*, a *Drosophila copia*-like element, is a leucine tRNA lacking its 3' terminal 5 bases. *Nucleic Acids Res.* **14**:4370-4371.
- Sanchez-Pescador, R., M. D. Power, P. J. Barr, K. S. Steimer, M. M. Stempien, S. L. Brown-Shimer, W. W. Gee, A. Renard, A. Randolph, J. A. Levy, D. Dina, and P. A. Luciw. 1985. Nucleotide sequence and expression of an AIDS-associated retrovirus (ARV-2). *Science* **227**:484-492.
- Schwartzberg, P., J. Colicelli, and S. P. Goff. 1984. Construction and analysis of deletion mutations in the *pol* gene of Moloney murine leukemia virus: a new viral function required for productive infection. *Cell* **37**:1043-1052.
- Shimotohno, K., S. Mizutani, and H. M. Temin. 1980. Sequence of retrovirus provirus resembles that of bacterial transposable elements. *Nature (London)* **285**:550-554.
- Shoemaker, C., S. Goff, E. Gilboa, M. Paskind, S. W. Mitra, and D. Baltimore. 1980. Structure of a cloned circular Moloney murine leukemia virus DNA molecule containing an inverted segment: implications for retrovirus integration. *Proc. Natl.*

- Acad. Sci. USA 77:3932-3936.
27. **Shoemaker, C., J. Hoffmann, S. P. Goff, and D. Baltimore.** 1981. Intramolecular integration within Moloney murine leukemia virus. *J. Virol.* 40:164-172.
 28. **Skalka, A. M.** 1988. Integrative recombination in retroviruses, p. 701-724. *In* R. Kucherlapati and G. R. Smith (ed.), Genetic recombination. American Society for Microbiology, Washington, D.C.
 29. **Sonigo, P., C. Barker, E. Hunter, and S. Wain-Hobson.** 1986. Nucleotide sequence of Mason-Pfizer monkey virus: an immunosuppressive D-type retrovirus. *Cell* 45:375-385.
 30. **Starcich, B., L. Ratner, S. F. Josephs, T. Okamoto, R. C. Gallo, and F. Wong-Staal.** 1985. Characterization of long terminal repeat sequences of HTLV-III. *Science* 227:538-540.
 31. **Van Beveren, C., E. Rands, S. K. Chattopadhyay, D. R. Lowy, and I. M. Verma.** 1982. Long terminal repeat of murine retroviral DNAs: sequence analysis, host-proviral junctions, and preintegration site. *J. Virol.* 41:542-556.
 32. **Varmus, H., and P. Brown.** 1989. Retroviruses, p. 53-108. *In* D. E. Berg and M. M. Howe (ed.), Mobile DNA. American Society for Microbiology, Washington, D.C.
 33. **Varmus, H., and R. Swanstrom.** 1985. Replication of retroviruses, p. 75-134. *In* R. Weiss, N. Teich, H. Varmus, and J. Coffin (ed.), RNA tumor viruses, vol. 2. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
 34. **Whitcomb, J. M., R. Kumar, and S. H. Hughes.** 1990. Sequence of the circle junction of human immunodeficiency virus type 1: implications for reverse transcription and integration. *J. Virol.* 64:4903-4906.
 35. **Yanisch-Perron, C., J. Vieira, and J. Messing.** 1985. Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC19 vectors. *Gene* 33:103-119.