JOURNAL OF VIROLOGY, June 1995, p. 3938–3944 Vol. 69, No. 6
0022-538X/95/$04.00+0

Human Immunodeficiency Virus Type 1 Preintegration Complexes Containing Discontinuous Plus Strands Are Competent To Integrate In Vitro

MICHAEL D. MILLER, BEIBEI WANG, AND FREDERIC D. BUSHMAN*

Infectious Disease Laboratory, The Salk Institute for Biological Studies, La Jolla, California 92037

Received 30 November 1994/Accepted 1 March 1995

Despite intensive study, the mechanism by which many retroviruses complete reverse transcription has remained unclear. Most retroviruses and all lentiviruses fail to synthesize a full-length second strand of the viral cDNA (plus strand) efficiently in infected cells. For human immunodeficiency virus type 1, we find in synchronous infection experiments that full-length plus strands are rare (<1% of products) at times when integration is likely taking place. Subviral nucleoprotein complexes containing such discontinuous cDNA can be extracted from infected cells and used to generate integration products in vitro. Analysis of such integration products using two-dimensional gel electrophoresis revealed that the discontinuous viral DNA was efficiently integrated into an added target DNA. These data support a model in which the discontinuities in the plus strand need not be sealed until after integration, potentially by the enzymes that are already thought to repair DNA gaps at the junctions between host and viral DNA.

Shortly after initiation of infection by a retrovirus, the viral RNA genome is reverse transcribed into a double-stranded DNA form, and this viral cDNA molecule is then covalently integrated into the host cell genome. The integrated viral DNA, or provirus, serves as the template for the generation of new viral genomic RNAs and mRNAs encoding viral proteins required to form new virions.

A widely accepted model of reverse transcription (reviewed in reference 30) holds that viral DNA synthesis is completed prior to integration, but for many retroviruses and all lentiviruses the mechanism of plus-strand synthesis differs. In addition to the primer that the RNase H activity of reverse transcriptase leaves at the normal polypurine tract, RNase H can also generate additional oligoribonucleotides that remain hybridized to the minus strand at other sites (28). Such oligoribonucleotides can be used by reverse transcriptase to prime plus-strand synthesis and so generate DNA molecules with discontinuous plus strands. For human immunodeficiency virus type 1 (HIV-1), a second site of plus-strand initiation, called the central polypurine tract (PPT_c), has been identified (11, 22). In HIV-1, for which this central plus-strand initiation site is used, reverse transcription products have been reported to contain two plus-strand fragments corresponding to the right and left halves of the viral genome. Other retroviruses, such as visna virus, human and simian foamy viruses, and avian sarcoma virus, also display multiple plus-strand initiation sites, and as a result such cDNAs contain discontinuous plus strands (2, 20, 23, 24; and below). Although it is often assumed that the plus strand is completed prior to integration (4, 19, 24, 31), it has also been suggested that cDNAs containing plus-strand discontinuities might be competent for integration (5, 22).

Several hours after infection, the linear viral cDNA can be extracted from infected cells as part of a nucleoprotein complex that is competent to integrate the viral DNA into an added target DNA in vitro (6, 12, 13). Such active complexes, operationally defined as preintegration complexes, contain the viral DNA, the viral IN protein, and possibly other viral proteins (8, 9, 15). In characterizing preintegration complexes from cells infected with HIV-1, we noticed that the viral DNA contained discontinuous plus-strand fragments of heterogeneous size. Despite the discontinuities, such DNA molecules were efficiently integrated into target DNA in vitro. These data suggest that HIV-1 DNA synthesis may in fact be completed after integration, potentially by host cell enzymes instead of by viral enzymes.

Viral plus strands are discontinuous and heterogeneous. DNA samples were prepared from SupT1 cells infected by coculture with the HIV-1-producing cell line Molt-IIIB (13). As described previously, linear viral DNA could be detected by 4 h after initiation of coculture and preintegration complexes isolated from infected cells at about this time were competent for integration (see below). To characterize the structure of viral DNA, cytoplasmic and nuclear DNA samples were isolated at various times after the initiation of coculture and analyzed by denaturing gel electrophoresis. Viral DNA strands were examined by Southern blotting and probing with strand-specific riboprobes [Fig. 1; (+) indicates a probe recognizing the plus strand, (−) indicates a probe recognizing the minus strand; probes are named by the genomic region recognized]. The locations of the probe sequences in the HIV-1 genome and the major reverse transcription products identified are summarized in Fig. 1A.

All of the probes detected DNA in the cytoplasmic fraction at earlier times than in the nuclear fraction, as expected (for example, compare cytoplasmic and nuclear DNA samples detected by probe U5(+) at 4 h). This observation confirms that nuclear fractions were not substantially contaminated by cytoplasmic DNA. A probe specific for the U3 region of the minus strand [U3(−)] revealed that virtually all of the molecules in the cytoplasm and the nucleus contained full-length (9.6-kb) or near full-length (8.8-kb) minus strands by 4 h postinfection. The smaller minus strands are probably reverse transcription intermediates that have not carried out strand displacement synthesis to complete the long terminal repeats (LTRs) (30).

* Corresponding author. Mailing address: The Salk Institute for Biological Studies, Infectious Disease Laboratory, 10010 N. Torrey Pines Rd., La Jolla, CA 92037.

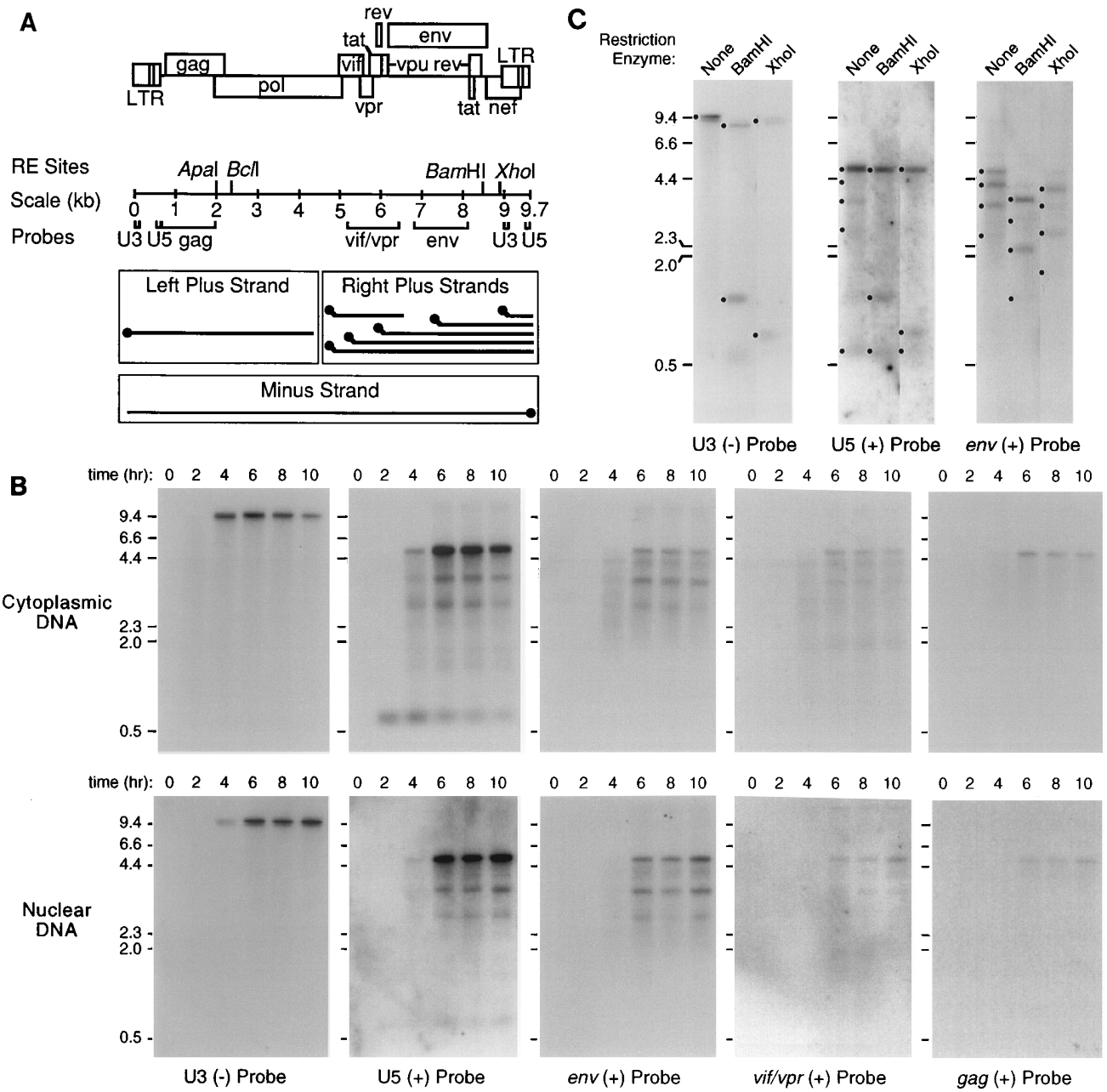


FIG. 1. Most unintegrated HIV-1 cDNA molecules contain discontinuous plus strands. (A) Summary of the major HIV-1 reverse transcription products identified in panels B and C. A physical map of HIV-1 genes is shown at the top, and the locations of various probes are shown in the middle. The bottom part of the panel shows the locations of individual DNA fragments mapped in panels B and C. The minus strand is shown as a thin line, and the plus-strand fragments are shown as thick lines. Filled circles represent 5′ ends of DNA strands. (B) Time course of viral DNA synthesis. Infections were initiated by culturing SupT1 cells (2×10^6/ml) with phorbol myristate acetate-stimulated chronically HIV-1-infected Molt-IIIB cells at a ratio of approximately 50:1. At various times after infection, cytoplasmic extracts were prepared by using a modification of a previously described protocol (13). In brief, cells were harvested from the coculture, washed twice with buffer K (20 mM HEPES [*N*-2-hydroxyethylpiperazine-*N*′-2-ethanesulfonic acid] pH 7.3, 150 mM KCl, 5 mM $MgCl_2$, 1 mM dithiothreitol, 10 U of aprotinin per ml) at 4°C, and resuspended at 4×10^7 cells per ml in buffer K containing 0.025% Brij-96 (Sigma Chemical Co., St. Louis, Mo.). Cells were left at room temperature (RT) for 10 min and then centrifuged at $1{,}000 \times g$ for 3 min at RT. The supernatant was clarified by centrifugation at $8{,}000 \times g$, for 3 min at RT, snap-frozen in liquid nitrogen, and stored at −75°C (cytoplasmic extract). Cytoplasmic DNAs were prepared from extracts by digestion with proteinase K in the presence of 0.5% sodium dodecyl sulfate (SDS)–10 mM EDTA. For preparation of unintegrated nuclear viral DNA (21), the pellet obtained after lysis with buffer K–0.025% Brij-96 was resuspended at 8×10^7 cell equivalents per ml in 0.6% SDS–10 mM Tris-HCl, pH 8.0–10 mM EDTA. Lysates were incubated at RT for 10 min, and then NaCl was added to a final concentration of 1 M and the lysates were placed at 4°C overnight. Precipitated material was removed by centrifugation at $14{,}000 \times g$ at 4°C for 60 min, and then the supernatant was extracted once with phenol-chloroform-isoamyl alcohol (25:24:1) and ethanol precipitated with tRNA carrier. DNA samples prepared from cytoplasmic or nuclear fractions at the indicated times were separated by electrophoresis through 1.2% alkaline agarose gels and analyzed by Southern blotting (29) using Hybond-N membranes (Amersham). DNA fragments were detected with ^{32}P-labeled RNA probes synthesized with T3 or T7 RNA polymerase (29). Each probe is named for the genomic region and DNA strand it detects. The positions and sizes (in kilobases) of fragments of lambda phage DNA generated by digestion with *Hin*dIII are indicated beside each panel. DNA in cytoplasmic fractions from 6 to 10 h was competent for integration (data not shown). Since no steps were taken to block continued infection during the coculture (1), some or all of the viral cDNA present in the cytoplasmic fractions at later times may represent newly reverse-transcribed viral genomes. (C) Restriction mapping of the viral plus-strand fragments. DNA prepared from cytoplasmic extracts harvested 5 h after infection were digested with the indicated restriction endonucleases (see restriction map in panel A). Alkaline agarose gel electrophoresis and hybridization with the indicated probes were performed as described in the legend to panel B. Size markers are described in the legend to panel B. Expected bands (see text) are marked with filled circles.

Unlike the minus strand, less than 1% of the plus strands were full length after 10 h, by which time integrated HIV-1 genomes can be detected (1). The pattern of plus-strand DNA fragments in the nucleus was very similar to that seen in the cytoplasm (Fig. 1B). These fragments were also the predominant plus-strand forms present in the nuclear fraction 24 h after initiation of coculture, although this fraction contained an increased amount (20 to 25%) of possible full-length plus strands (data not shown). Restriction mapping and hybridization with various plus-strand-specific riboprobes identified the following plus-strand fragments (Fig. 1B and C, summarized in Fig. 1A).

(i) Plus-strand strong-stop DNA. The earliest plus-strand DNA species expected is the plus-strand strong-stop DNA, which is initiated from an RNA primer generated by RNase H at the polypurine tract and extended to the right end of the genome by using the minus strand as template (30). As early as 2 h after infection, probe U5(+) detected a 0.6-kb fragment in cytoplasmic DNA that did not hybridize to any other plus-strand probe, thus displaying the size and sequence content expected for the plus-strand strong-stop DNA (Fig. 1B). A similar fragment was detected in nuclear DNA by 6 h after infection, though in lesser abundance.

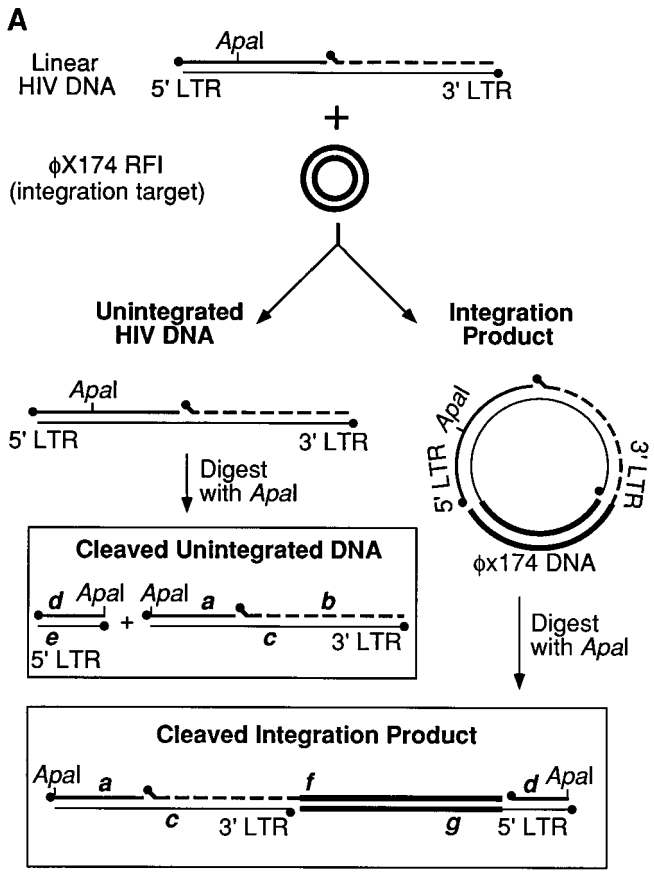
(ii) Left half of the plus strand. A *gag*(+) probe detected a single fragment of 4.8 kb in all cytoplasmic DNA samples beginning at 4 h after infection and in nuclear DNA fractions beginning at 6 h after infection (Fig. 1B). This fragment also hybridized to probe U5(+). The size and hybridization pattern of this fragment are as expected for plus-strand DNA extended from plus-strand strong-stop DNA following the second template switch (Fig. 1) and terminated at a discrete site in the center of the genome, as described previously (Fig. 1A, left plus strand) (11, 22).

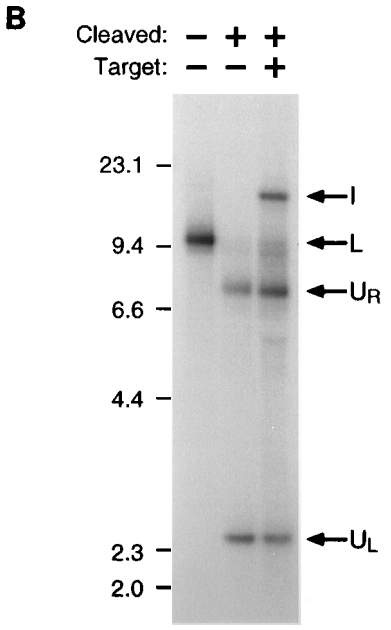
(iii) Right half of the plus strand. A second fragment of 4.8 kb hybridized to the *env*(+) and *vif/vpr*(+) probes from the right side of the plus strand and also to the U5(+) probe. This fragment comigrated with the 4.8-kb left half of the fragment just described but is derived from the right half of the plus strand on the basis of restriction mapping and reactivity with different plus-strand probes (Fig. 1B and C). For example, the 4.8-kb left plus strand detected by probe U5(+) was not cleaved by *Bam*HI or *Xho*I, which cleave in the right half of the genome (Fig. 1C). The comigrating right half fragment detected by probe *env*(+), however, was cleaved by all of these enzymes (Fig. 1C). Evidently this strand was generated by initiation of reverse transcription near the center of the genome and extension to the right end (11, 22).

(iv) Other fragments from the right half of the genome. An array of bands smaller than 4.8 kb but larger than strong-stop DNA was detected by probes U5(+), *vif/vpr*(+), and *env*(+) (Fig. 1B). These fragments map to the right side of the genome since they did not hybridize detectably to probe *gag*(+) (Fig. 1B). Some of the major DNA species observed were mapped by Southern blot analysis of DNA cleaved with restriction enzymes. The results of one such experiment are shown in Fig. 1C. As expected, minus-strand probe U3(−) detected full-length and near-full-length bands (9.6 and 8.8 kb) in uncut DNA but two fragments of the expected sizes after cleavage with *Bam*HI or *Xho*I, which cleave at unique sites 1.2 and 0.8 kb, respectively, from the right end of the HIV-1 genome. In addition to detecting the expected fragments of 8.4 and 1.2 kb in DNA cleaved with *Bam*HI, probe U3(−) also detected a 0.6-kb band. The origin of this band is unknown, but since it is reproducibly generated by *Bam*HI digestion of this DNA (not shown) it may be a product of *Bam*HI star activity or of a second *Bam*HI site present in a fraction of the proviruses in Molt-IIIB cells.

As expected, the plus-strand probe U5(+) detected heterogeneous plus-strand fragments, including major species of 4.8, 4.2, 3.5, 2.5, and 0.6 kb in uncut DNA (the 4.2-, 2.5-, and 0.6-kb bands are diffuse and faint). After digestion of DNA with *Bam*HI, probe U5(+) detected the 4.8- and 0.6-kb bands plus a faint new band of 1.2 kb representing the expected fragment of *Bam*HI cleavage (Fig. 1C). Similarly, after cleavage of the DNA with *Xho*I, probe U5(+) detected the 4.8- and 0.6-kb bands plus a faint new band of 0.8 kb, the expected size of the right end of the HIV-1 genome after *Xho*I digestion. The *env*(+) probe also detected heterogeneous plus-strand fragments, including major species of 4.8, 4.2, 3.5, and 2.5 kb in uncut DNA. After cleavage with *Bam*HI, these fragments were all reduced in size by ~1.2 kb to yield fragments of 3.6, 3.0, 2.3, and 1.3 kb (the 3.0- and 1.3-kb bands are diffuse and faint). The major species in uncut DNA that hybridized to the *env*(+) probe were all reduced in size by ~0.8 kb after cleavage with *Xho*I, yielding fragments of 4.0, 3.5, 2.7, and 1.6 kb (the 3.5- and 1.6-kb bands are diffuse and faint). Taken together, these results suggest that the major heterogeneous right plus strands share a common 3′ end that maps to the right end of the HIV-1 genome. This conclusion is supported by the fact that the right plus-strand fragments identified by U5(+) must extend to within 100 nucleotides (nt) of the right end of the genome to be reactive with this probe. The 5′ ends of these fragments must therefore map to distinct sites in the right half of the genome. Smaller bands that hybridized to *vif/vpr*(+) but not to *env*(+) or U5(+) were also apparent (Fig. 1B) and probably represent DNA strands that initiated near the center of the genome but failed to extend to the 3′ end. The smear detected with the U5(+), *env*(+), and *vif/vpr*(+) probes (Fig. 1B) reflects additional heterogeneity in the right half of the plus strand and suggests the presence of myriad minor plus-strand initiation sites, although some of this heterogeneity could also result from termination of DNA synthesis at different sites. The initiation sites of the major plus-strand fragments found in the right half of the HIV-1 genome have been mapped only roughly (Fig. 1A) but are predicted to start near nt 5500 (4.2-kb fragment), nt 6200 (3.5-kb fragment), and nt 7200 (2.5-kb fragment). For each fragment, inspection of the nucleotide sequence of HIV-1 clone HXB2 (derived from the IIIB strain used in these studies) revealed a reasonable candidate for a purine-rich initiation site (10) within 100 bp of the predicted start site (nt 5580 to 5594, 6253 to 6269, and 7152 to 7168).

Integration of discontinuous viral DNA in vitro. The observation that full-length plus strands are so rare leads to the question of whether discontinuous reverse transcripts are actually competent for integration. To address this issue, cytoplasmic extracts containing preintegration complexes were prepared as described above and incubated with target DNA in vitro. The structures of the integration products were then analyzed. The predicted structures of the integrated and unintegrated DNAs produced by this reaction are shown in Fig. 2A. To facilitate the analysis, integration products were cleaved with a restriction endonuclease that cuts once in the HIV-1 genome but not in the ϕX174 target DNA. Joining of the heterogeneous gapped plus-strand fragments (*b* fragments in Fig. 2A, shown as a dashed line) to the 5.4-kb target DNA in the integration reaction produces larger heterogeneous fragments in the integration product (*f* fragments). Two additional gaps are predicted to exist in the integration product, since only the 3′ end of each viral DNA strand is joined to the target DNA in the in vitro integration reaction (7, 16).





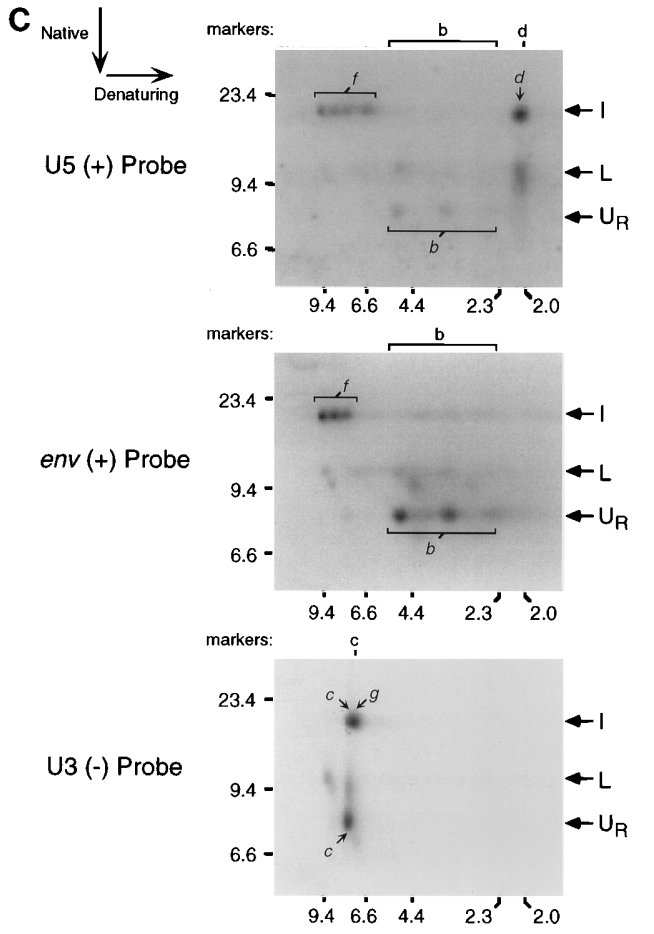


FIG. 2. Discontinuous HIV-1 reverse transcription products are competent for integration. (A) Experimental strategy. Thin lines represent the HIV-1 DNA minus strand, thick lines represent phage φX174 target DNA. Solid and dashed medium-width lines represent the left plus strand and heterogeneous right plus strands, respectively, of HIV-1 DNA (see Fig. 3A) Filled circles mark the 5′ ends of DNA strands. Linear unintegrated HIV-1 DNA and circular φX174 target DNA are shown at the top. After the integration reaction, some of the HIV-1 DNA remains unintegrated (middle left), whereas some of the HIV-1 DNA becomes integrated into the target DNA (middle right). The predicted structures of unintegrated HIV-1 DNA and of the integration product after cleavage with *Apa*I are shown at the bottom, with individual strands labeled *a* to *g*. The heterogeneous *b* fragments and the heterogeneous parts of the *f* fragments are represented as dashed lines. The predicted sizes of the fragments are as follows: *a*, 2.8 kb; *b*, heterogeneous, 0.6 to 4.8 kb; *c*, 7.7 kb; *d*, 2.0 kb; *e*, 2.0 kb; *f*, heterogeneous, 6.0 to 10.1 kb; *g*, 7.4 kb. (B) Integration in vitro directed by preintegration complexes: assay of products by native gel electrophoresis. Cytoplasmic extracts (150 μl) were incubated in the presence or absence of 3 μg of φX174 replicative-form I DNA per ml at 37°C for 45 min and then deproteinized. Where indicated, DNA products were cleaved with the restriction endonuclease *Bcl*I, which cleaves once in the viral DNA 2.4 kb from the left end. Products were resolved by electrophoresis through a 0.75% agarose gel in 1× TAE, transferred to Hybond-N, and detected by hybridization with a ^{32}P-labeled HIV-1 LTR probe (13). Size markers are as described in the legend to Fig. 1B. Symbols: I, integration product; L, size of linear viral genome; U_R and U_L, right and left fragments, respectively, of the unintegrated genome. (C) Two-dimensional electrophoretic analysis of reaction products. Integration reactions were performed by incubating 800 μl of cytoplasmic extract with 6 μg of φX174 replicative-form intermediate DNA per ml followed by digestion with *Apa*I. Reaction products were first resolved by electrophoresis through a 0.75% agarose gel in 1× TAE (29), then the lane was cut out and cast into a new 1.2% agarose gel and alkaline electrophoresis, blotting, and probing were performed as described in the legend to Fig. 1B. Letters at the top of each panel correspond to positions of viral DNA fragments (as shown in panel A) that were electrophoresed in parallel as size markers. Other markers and symbols are as described for panel B.

The connectivity of the DNA strands in integration products was analyzed by two-dimensional gel electrophoresis. After the integration reaction and cleavage of the DNA with a restriction enzyme, integration products and unintegrated DNAs were separated in the first dimension on a native agarose gel. Figure 2B shows products resolved in the first dimension analyzed by Southern blotting. The DNA species marked I, seen only in the presence of the target DNA, has the structure expected for covalently linked viral and target DNA (reference 13 and data not shown). The unintegrated linear viral DNA was cleaved once, yielding two fragments (U_R and U_L). Other DNA forms, including some migrating with the size of uncleaved linear viral DNA (L), can also be seen (discussed below). For the second

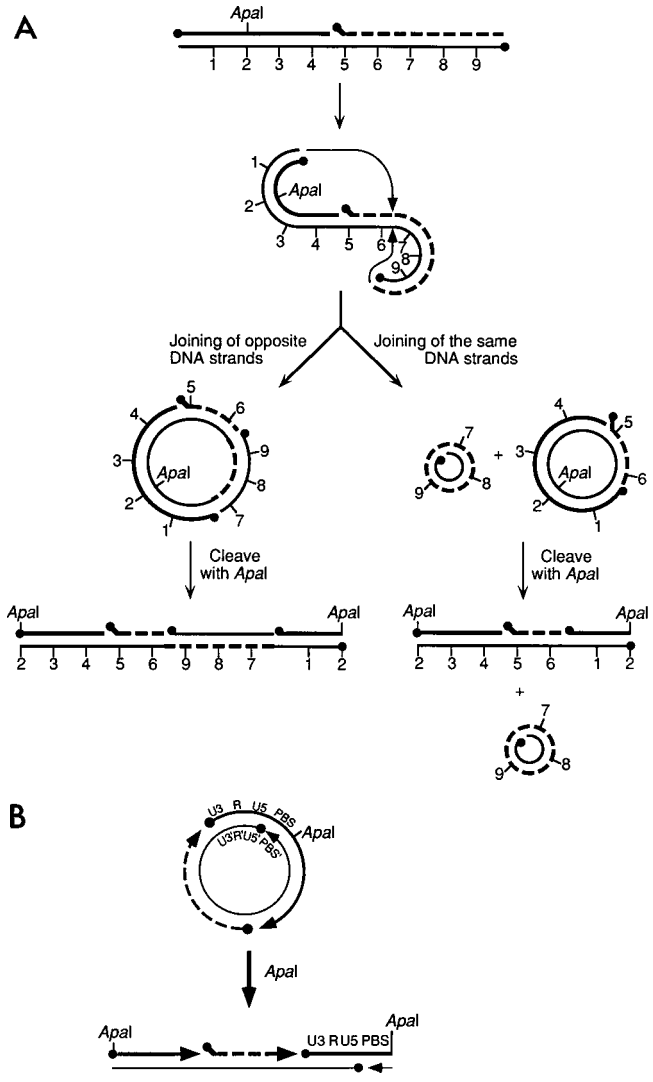


FIG. 3. Additional HIV-1 DNA forms observed by two-dimensional gel electrophoresis. (A) HIV-1 DNA autointegration. The figure shows the two possible outcomes of an integration event in which both ends of the viral DNA are integrated into a site in the right half of the same genome. In the example on the left, each viral DNA 3′ end integrates into the strand of opposite polarity, yielding an internally inverted circle (14). Cleavage of such circles with *Apa*I generates linear molecules the size of unintegrated linear viral DNA. Thus, a heterogeneous population of inverted circles cleaved with *Apa*I would migrate as a single species (L) in native gels (Fig. 2B and C). In the example on the right, each viral DNA 3′ end integrates into the strand of the same polarity to yield two circular molecules. Cleavage of these products yields one linear molecule and one circular molecule, the sizes of which depend on the specific site of integration. As in intermolecular integration, the 5′ ends of the viral DNA are not joined to target DNA during autointegration, resulting in the introduction of two additional gaps in the autointegration product(s). Thus, after *Apa*I cleavage, fragment *d* (Fig. 2C) is present in both of the examples shown and indeed would be present in any autointegrant in which integration occurred at a site in the HIV-1 DNA 3′ of the *Apa*I site. Similarly, autointegration into any site 5′ of the *Apa*I site in HIV-1 DNA (not shown) would result in a product containing the minus-strand fragment *c* (Fig. 2A). As can be seen in Fig. 2C, fragments *c* and *d* are both present in the species migrating with the size of linear DNA in the first dimension (L). Evidently, some of these species arose from autointegration and not from the well-known 2-LTR circles (14), which would not be expected to contain fragments *c* and *d*. In addition, some of the spots in this region (e.g., spots with the mobilities of *b* fragments) could be derived from uncut linear viral DNA. However, these spots cannot all be from uncut linear genomes, because fragments *d* and *c* are only present in cleaved genomes. Two prominent smears were detected in Fig. 2C. One smear, detected by probe U5(+), had the mobility of fragment *d* in the second dimension and heterogeneous mobility in the first dimension. The other smear, detected by probe U3(−), had the mobility of fragment *c* in the second dimension and heterogeneous mobility in the first dimension, such a gel lane was cut out and cast into a denaturing agarose gel, and electrophoresis was carried out perpendicular to the first dimension. After electrophoresis, the DNA fragments were transferred to a nylon membrane and identified by using strand-specific riboprobes (Fig. 2C).

DNA fragments with the mobility of the larger unintegrated fragment (U_R) in the first dimension displayed a series of spots in the second dimension of the sizes expected for the heterogeneous plus strands (Fig. 2C, probes U5(+) and *env*(+), labeled *b*). The U_L fragment is not present on these blots because of its small size. By contrast, the integrated DNA (I) yielded a series of spots of higher molecular size in both dimensions that hybridized to the U5(+) and *env*(+) probes (labeled *f*). These spots have the mobilities expected for the *b* fragments integrated into ϕX174 target DNA to yield *f* fragments (Fig. 2A). The U5(+) probe hybridized to four prominent *f* fragment spots, the largest three of which also hybridized to the *env*(+) probe. The smallest of the four *f* spots detected by U5(+) presumably corresponds to a target DNA molecule ligated to a short plus strand that does not extend into the *env* sequences recognized by the *env*(+) probe. In the integrated DNA, probe U5(+) also hybridized to an additional spot of about 2 kb, representing the expected fragment *d*.

The minus-strand probe, U3(−), detected 7.7-kb spots in both integrated and unintegrated DNAs corresponding to fragment *c*, the larger minus-strand fragment generated by *Apa*I cleavage. In addition, this probe hybridized to a 7.4-kb spot in the integration product corresponding to the minus strand of the 5′ LTR ligated to a target DNA (fragment *g*). Fragments *c* and *g* were not resolved in this gel system.

Among the products of the in vitro integration reaction (Fig. 2B and C) were DNA species that migrated in the first dimension with the same or similar mobility as that of uncleaved linear viral DNA (L). Analysis of these species by two-dimensional electrophoresis indicated that some of these forms are likely derived from two circular DNA species. One is probably derived from viral cDNA molecules that integrated into themselves (autointegration) to yield internally inverted circles, (Fig. 3A, left side), and another may be derived from a reverse transcription strand transfer intermediate shown in Fig. 3B as described previously (30). The origins of these and other forms are discussed in the legend to Fig. 3.

dimension. These smears may have originated from autointegrants such as those shown on the right side of panel A. Autointegration into different sites would create a heterogeneous population of molecules that would migrate as a smear in the first dimension. Such molecules would nevertheless contain fragments *c* or *d* (depending on the integration site) that would migrate with characteristic mobility in the second dimension. In the example shown on the left of panel A, one of the products is a full-length continuous strand containing LTR sequences from both plus and minus strands as well as *env* sequences from the plus strand. In the two-dimensional gel analysis shown in Fig. 2C, a spot of 9.7 kb in both dimensions hybridized to the U3(−), U5(+), and *env*(+) probes. This spot is likely the product of an autointegration reaction such as the one shown in panel A. (B) Reverse transcription intermediate. The DNA form migrating slightly faster than the full-length linear DNA in Fig. 2B has the size expected for a linearized 1-LTR circle. In the denaturing dimension (Fig. 2C), this form yielded spots corresponding to apparently continuous 1-LTR-length plus and minus strands, consistent with the structure of 1-LTR circles. In addition, however, fragments *b*, *c*, and *d* were also visible (Fig. 2A and C). The well-known 1-LTR circles should not contain discontinuities and therefore would not be expected to contain any of these fragments. The data are consistent with the idea that these fragments originated from *Apa*I digestion of a circular structure such as the one shown here. As described previously, the second-strand transfer (plus-strand strong-stop) product might proceed through such a circular intermediate (30), and ligation of the nicks in this form may also account for formation of 1-LTR circles (25). Symbols are described in the legend to Fig. 2A. PBS, primer-binding site.

In summary, the two-dimensional gel analysis of in vitro integration products showed that heterogeneous plus-strand fragments in the unintegrated DNA (*b*) gave rise to heterogeneous integration products (*f*). All of the DNA fragments predicted to react with our probes as diagrammed in Fig. 2A were detected. Evidently, viral DNA molecules containing discontinuous plus strands are competent to integrate into target DNA molecules, and these discontinuities are not repaired during integration.

Implications. Textbook models of retroviral reverse transcription include as a final step the synthesis of a full-length plus strand prior to integration (18, 30). Even for retroviruses which fail to synthesize detectable levels of full-length plus strands either in vitro or in vivo, it has been stated explicitly that some mechanism must exist to generate mature cDNA molecules containing a complete plus strand prior to integration (4, 24). For the large number of retroviruses displaying discontinuous plus strands, completion of the plus strand has been postulated to occur by extension of the most upstream plus strand with displacement of downstream fragments or by filling of DNA gaps and ligation of plus-strand fragments (5). However, our data indicate that there is no strict requirement to synthesize a full-length plus strand prior to integration. Up to 30% of linear viral cDNA products synthesized in a single round of HIV-1 infection become integrated, and integration of HIV-1 proviruses in vivo is first detectable beginning 8 h after initiation of infection by coculture (1). Our analysis showed that >99% of the plus strands in unintegrated linear HIV-1 genomes were discontinuous during the first 10 h of infection by coculture. Furthermore, such genomes were competent for integration in vitro as demonstrated by two-dimensional electrophoretic analysis of integration products. Although the possibility that plus strands are repaired and then quickly integrated in vivo is not ruled out, the simplest model to explain our observations is that discontinuous reverse transcripts first integrate and then are completed.

In fact, previously published data on simple retroviruses are consistent with the idea that discontinuous reverse transcripts may be the predominant substrates for integration in these systems also. Discontinuous plus strands are the major products of avian sarcoma virus reverse transcription in vivo (3, 4, 23, 26). Mouse mammary tumor virus and Moloney murine leukemia virus also display short heterogeneous plus strands at early times after infection (4 to 10 h) (17, 23), but in these cases they display primarily full-length plus strands at later times (24 h). This latter observation led investigators to conclude that the short plus strands apparent at early times were precursors to full-length plus strands, possibly the result of pausing by reverse transcriptase during synthesis (23). The eventual appearance of full-length plus strands in unintegrated viral cDNAs led to the view that a continuous plus strand is completed prior to integration. However, since it was recently found that integration of newly synthesized Moloney murine leukemia virus reverse transcripts into target DNA in vivo begins as early as 4 h after infection and peaks at about 8 h after infection (27), the early reverse transcription products containing heterogeneous plus strands may well represent the major cDNA forms that are destined to integrate.

Although integration of reverse transcripts containing discontinuous plus strands would result in single-stranded gaps in the integrated viral DNA, similar gaps are already thought to be left at each junction between host and viral DNA during the integration reaction. These gaps are then repaired after integration by viral or cellular enzymes to generate two continuous DNA strands. Thus, both internal gaps in the plus strand and gaps at the junctions between host and viral DNA could be repaired by the same or similar mechanisms.

We thank Janet Bushman, Amy Espeseth, and members of the Infectious Disease Laboratory for comments on the manuscript.

This work was supported by National Institutes of Health grant RO1 AI34786 (F.D.B.) and NRSA fellowship 1F32 AI09068-01 (M.D.M.) F.D.B. is a special fellow of The Leukemia Society of America.

REFERENCES

1. **Barbosa, P., P. Charneau, N. Dumey, and F. Clavel.** 1994. Kinetic analysis of HIV-1 early replicative steps in a coculture system. AIDS Res. Hum. Retroviruses **10:**53–59.
2. **Blum, H. E., J. D. Harris, P. Ventura, D. Walker, K. Staskus, E. Retzel, and A. T. Haase.** 1985. Synthesis in cell culture of the gapped linear duplex DNA of the slow virus visna. Virology **142:**270–277.
3. **Boone, L. R., and A. M. Skalka.** 1981. Viral DNA synthesized in vitro by avian retrovirus particles permeablized with melittin. I. Kinetics of synthesis and size of minus- and plus-strand transcripts. J. Virol. **37:**109–116.
4. **Boone, L. R., and A. M. Skalka.** 1981. Viral DNA synthesized in vitro by avian retrovirus particles permeablized with melittin. II. Evidence for a strand displacement mechanism in plus-strand synthesis. J. Virol. **37:**117–126.
5. **Boone, L. R., and A. M. Skalka.** 1993. Strand displacement synthesis by reverse transcriptase, 119–133. *In* A. M. Skalka and S. P. Goff (ed.), Reverse transcriptase. Cold Spring Harbor Laboratory Press, Plainview, N.Y.
6. **Brown, P. O., B. Bowerman, H. E. Varmus, and J. M. Bishop.** 1987. Correct integration of retroviral DNA in vitro. Cell **49:**347–356.
7. **Brown, P. O., B. Bowerman, H. E. Varmus, and J. M. Bishop.** 1989. Retroviral integration: structure of the initial covalent complex and its precursor, and a role for the viral IN protein. Proc. Natl. Acad. Sci. USA **86:**2525–2529.
8. **Bukrinsky, M. I., S. Haggerty, M. P. Dempsey, N. Sharova, A. Adzhubel, L. Spitz, P. Lewis, D. Goldfarb, M. Emerman, and M. Stevenson.** 1993. A nuclear localization signal within HIV-1 matrix protein that governs infection of non-dividing cells. Nature (London) **365:**666–669.
9. **Bukrinsky, M. I., N. Sharova, T. L. McDonald, T. Pushkarskaya, G. W. Tarpley, and M. Stevenson.** 1993. Association of integrase, matrix, and reverse transcriptase antigens of human immunodeficiency virus type 1 with viral nucleic acids following acute infection. Proc. Natl. Acad. Sci. USA **90:**6125–6129.
10. **Champoux, J. J.** 1993. Roles of ribonuclease H in reverse transcription, p. 103–117. *In* A. M. Skalka and S. P. Goff (ed.), Reverse transcriptase. Cold Spring Harbor Laboratory Press, Plainview, N.Y.
11. **Charneau, P., and F. Clavel.** 1991. A single-stranded gap in human immunodeficiency virus unintegrated linear DNA defined by a central copy of the polypurine tract. J. Virol. **65:**2415–2421.
12. **Ellison, V. H., H. Abrams, T. Roe, J. Lifson, and P. O. Brown.** 1990. Human immunodeficiency virus integration in a cell-free system. J. Virol. **64:**2711–2715.
13. **Farnet, C. M., and W. A. Haseltine.** 1990. Integration of human immunodeficiency virus type 1 DNA in vitro. Proc. Natl. Acad. Sci USA **87:**4164–4168.
14. **Farnet, C. M., and W. A. Haseltine.** 1991. Circularization of human immunodeficiency virus type 1 DNA in vitro. J. Virol. **65:**6942–6952.
15. **Farnet, C. M., and W. A. Haseltine.** 1991. Determination of viral proteins present in the human immunodeficiency virus type 1 preintegration complex. J. Virol. **65:**1910–1915.
16. **Fujiwara, T., and K. Mizuuchi.** 1988. Retroviral DNA integration: structure of an integration intermediate. Cell **54:**497–504.
17. **Gianni, A. M., and R. A. Weinberg.** 1975. Partially single-stranded form of free Moloney viral DNA. Nature (London) **225:**646–648.
18. **Gilboa, E., W. M. Sudha, S. Goff, and D. Baltimore.** 1979. A detailed model of reverse transcription and tests of crucial aspects. Cell **18:**93–100.
19. **Goff, S. P.** 1990. Retroviral reverse transcriptase: synthesis, structure, and function. J Acquired Immune Defic. Syndr. **3:**817–831.
20. **Harris, J. D., J. V. Scott, B. Traynor, M. Brahic, L. Stowring, P. Ventura, A. T. Haase, and R. Peluso.** 1981. Visna virus DNA: discovery of a novel gapped structure. Virology **113:**573–583.
21. **Hirt, B.** 1967. Selective extraction of polyoma DNA from infected mouse cell cultures. J. Mol. Biol. **26:**365–371.
22. **Hungnes, O., E. Tjotta, and B. Grinde.** 1991. The plus strand is discontinuous in a subpopulation of unintegrated HIV-1 DNA. Arch. Virol. **116:**133–141.
23. **Kung, H. J., Y. K. Fung, J. E. Majors, J. M. Bishop, and H. E. Varmus.** 1981. Synthesis of plus strands of retroviral DNA in cells infected with avian sarcoma virus and mouse mammary tumor virus. J. Virol. **37:**127–138.
24. **Kupiec, J.-J., J. Tobaly-Tapiero, M. Canivet, M. Santillana-Hayat, R. M. Flugel, J. Peries, and R. Emanoil-Ravier.** 1988. Evidence for a gapped linear duplex DNA intermediate in the replicative cycle of human and simian spumaviruses. Nucleic Acids Res. **16:**9557–9565.

25. **Lee, Y. M. H., and J. M. Coffin.** 1990. Efficient autointegration of avian retrovirus DNA in vitro. J. Virol. **64:**5958–5965.
26. **Lee, Y. M. H., and J. M. Coffin.** 1991. Relationship of avian retrovirus DNA synthesis to integration in vitro. Mol. Cell. Biol. **11:**1419–1430.
27. **Pryciak, P., H.-P. Muller, and H. E. Varmus.** 1992. Simian virus 40 minichromosomes as targets for retroviral integration in vivo. Proc. Natl. Acad. Sci. USA **89:**9237–9241.
28. **Rattrey, A. J., and J. J. Champoux.** 1987. The role of Moloney murine leukemia virus RNase H activity in the formation of plus-strand primers. J. Virol. **61:**2843–2851.
29. **Sambrook, J., E. F. Fritsch, and T. Maniatis.** 1989. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
30. **Telenitsky, A., and S. P. Goff.** 1993. Stong-stop strand transfer during reverse transcription, p. 49–83. *In* A. M. Skalka and S. P. Goff (ed.), Reverse transcriptase. Cold Spring Harbor Laboratory Press, Plainview, N.Y.
31. **Whitcomb, J. M., and S. H. Hughes.** 1992. Retroviral reverse transcription and integration: progress and problems. Annu. Rev. Cell Biol. **8:**275–306.