Human Immunodeficiency Virus Integration in a Cell-Free System

VIOLA ELLISON,¹ HOLLY ABRAMS,² TAIYUN ROE,³ JEFF LIFSON,² AND PATRICK BROWN^{1,3,4*}

Departments of Pediatrics¹ and Biochemistry³ and Howard Hughes Medical Institute,⁴ Stanford University Medical Center, Stanford, California 94305-5428, and Genelabs Incorporated, Redwood City, California, 94063²

Received 17 January 1990/Accepted 6 March 1990

Integration of the viral genome into the nuclear DNA of a host cell plays a pivotal role in the replication of retroviruses. We have developed an in vitro method for studying the biochemistry of human immunodeficiency virus (HIV) integration by using extracts from HIV-infected cells. Analysis of the reaction products showed that HIV integration in vitro accurately reproduces the in vivo process. Integration occurred without apparent specificity for the target sequence, and the integrated provirus was directly flanked by a 5-base-pair duplication of DNA from the target site. HIV integration did not require a high-energy cofactor, and the enzymatic activities required for integration were recovered with the viral DNA when cell extracts were fractionated by gel exclusion chromatography.

The human immunodeficiency virus (HIV), a retrovirus, is an important cause of morbidity and mortality in many parts of the world (11). It has been estimated that in the United States alone at least a million individuals have already been infected with HIV (8). A substantial proportion, perhaps the majority, of those infected will ultimately develop acquired immunodeficiency syndrome or related conditions (4). Thus, there is a need for agents to prevent, reverse, or ameliorate the consequences of HIV infection. One approach to the identification of such agents is through biochemical characterization of the critical steps in the viral life cycle which provide potential targets for antiviral compounds.

Soon after a retrovirus enters the cytoplasm of its host cell, the viral reverse transcriptase synthesizes a double-stranded linear DNA copy of the viral genome. This viral DNA molecule enters the nucleus of the infected cell, where it is integrated into a host chromosome in a process that depends on sequences at the ends of the viral DNA molecule and on viral proteins brought into the cell by the virion. The integrated viral genome, or provirus, is thereafter transmitted as a stable element of the host genome. The provirus provides the sequences required in *cis* for its own expression and serves as the template for the next generation of viral RNA (18, 19). Integration, therefore, plays a pivotal role in retroviral replication and is thus an attractive target for the development of new antiviral agents. We describe here a method for studying HIV integration in vitro and the results of our preliminary characterization of the HIV integration reaction and its products.

MATERIALS AND METHODS

Viruses and cells. The CD4⁺ human T-lymphoblastoid cell line H9 (13) and H9 cells chronically infected with HIV type 1 (HIV-1_{HXB-2} [15]) were maintained in RPMI 1640 supplemented with 10% (vol/vol) heat-inactivated fetal calf serum and passaged three times weekly. More than 95% of the infected cells in the chronically infected population expressed viral p24 antigen as measured by immunohistochemistry. The cells used for experiments were in logarithmic growth.

Preparation of extracts from HIV-infected cells. Equal numbers of infected and uninfected H9 cells were coculti-

vated at a combined density of 1.0×10^6 /ml. After 8 to 9 h, characteristic HIV-induced cell fusion was evident and extensive, with up to 50% of cells involved in syncytia. At this time, cells were harvested by low-speed centrifugation ($175 \times g$) and then washed once with cold buffer A (10 mM Tris hydrochloride [pH 7.4], 150 mM KCl, 5 mM MgCl₂, 1 mM dithiothreitol, 20 μ g of aprotinin per ml [Sigma Chemical Co.]). All subsequent steps were performed at 4°C or on ice. Cells were then lysed in buffer A containing 0.5% (vol/vol) Triton X-100 by using $1.0 \text{ ml}/2.0 \times 10^7$ cells. After incubation in lysis buffer for 20 to 30 min, lysates were centrifuged at $1,000 \times g$ to pellet nuclei, and the supernatant was collected and clarified by centrifugation at $8,000 \times g$ for 10 min. The resulting supernatant, designated the cytoplasmic extract, was adjusted to 8% (wt/vol) sucrose, frozen in aliquots by immersion in liquid nitrogen, and stored at -80°C .

Fractionation of cytoplasmic extracts by gel exclusion chromatography. A 2-ml sample of cytoplasmic extract was loaded onto a 45-ml column of Bio-Gel A-5m (Bio-Rad Laboratories) previously equilibrated with buffer A. The A_{280} of the column effluent was monitored, and fractions were assayed for integration activity with and without ATP.

Integration reactions. Components were added to a 1.5-ml microcentrifuge tube on ice in the following order: 15 μ l of 10 $\times$ cocktail (200 mM HEPES [*N*-2-hydroxyethylpiperazine-*N'*-2-ethanesulfonic acid]-sodium [pH 7.5], 25 mM MgCl₂, 10 mM dithiothreitol), 10 μ l of a 0.1-mg/ml solution of ϕ X174 replicative-form I (RFI) DNA in H₂O, 100 μ l of crude or partially purified cytoplasmic extract. These components were mixed well, and then 25 μ l of a 30% solution of polyethylene glycol 8000 (Calbiochem-Behring) was added, giving a final volume of 150 μ l. The mixture was incubated at 22°C for 15 to 30 min. Reactions were stopped by addition of 300 μ l of freshly prepared stop solution (10 mM Tris hydrochloride [pH 8.0], 1 mM EDTA, 300 mM NaCl, 0.4% [wt/vol] sodium dodecyl sulfate, 400 μ g of proteinase K (Boehringer Mannheim Biochemicals) per ml. Following digestion at 45°C for several hours, the mixture was extracted once with CHCl₃, twice with phenol, and once with CHCl₃-phenol (50:50). The DNA was then recovered by ethanol precipitation and analyzed as described in the legends to Fig. 1 and 4.

Cloning and sequencing of junctions between target and proviral sequences. Integration reactions were performed as

* Corresponding author.

described above, except that 1 μ g of π AN13 DNA (10) was used instead of 1 μ g of ϕ X174 RFI DNA. Recovered DNA was digested with *Xho*I restriction endonuclease, and the predicted 10.7-kilobase recombination product was purified by preparative gel electrophoresis. The recovered DNA was treated with DNA polymerase I (16) to translate the putative gaps away from the target-provirus boundaries (3, 7). Polymerase chain reactions were then performed by using Taq DNA polymerase (The Perkin-Elmer Corp.) as directed by the manufacturer. Primers were HIVU3.4: 5'-GAGAT CTCGA GTCTT CGTTG GGAGT GAATT AGCCC, and HIVU5.3: 5'-GAGAT CTCGA GACCC TTTTA GTCAG TGTGG AAAAT CTC. The underlined portions of these primers hybridized to sequences present in the HIV-1_{HXB-2} long terminal repeat (LTR) such that their 3' ends in each case were 6 bases from the corresponding end of the LTR (14). The 5' ends of each primer were designed to introduce an *Xho*I restriction site at each end of the amplified DNA. The primers were used to prime amplification across the junctions between HIV and plasmid DNA and through the 894-base-pair plasmid itself. After 25 cycles of amplification, the amplified DNA (not visible by ethidium bromide staining) was purified on the basis of its predicted size by blind excision from a preparative agarose gel. The recovered DNA was then subjected to an additional 25 cycles of polymerase chain reaction amplification. A DNA band of the predicted size was visible when the products of this second round of amplification were analyzed by agarose gel electrophoresis. This DNA was then digested with *Xho*I endonuclease, circularized by ligation in a dilute solution, and used to transform *Escherichia coli* MC1061/p3 (10). Transformants were selected and restriction mapped. Three arbitrarily selected clones were sequenced by the dideoxy method by using a modified T7 polymerase (Sequenase version 2; U.S. Biochemical) in $MnCl_2$ as recommended by the manufacturer. The primers used for sequencing were HXB2L (5'-TGGAA GGGCT AATTC ACTCC CAACG) and HXB2R (5'-TGCTA GAGAT TTTCC AACT GACTA).

RESULTS

In vitro HIV integration assay. To detect HIV integration in vitro, we used an assay similar to one previously used to study murine leukemia virus integration (1-3). Uninfected human T-lymphoblastoid cells (H9) were infected by cocultivation with chronically HIV-1_{HXB-2}-infected, virus-producing H9 cells (13, 15). After several hours, the cells were collected by centrifugation and a cytoplasmic extract was prepared. This extract contained the bulk of the unintegrated viral DNA, as well as the enzymatic machinery required for its integration. The crude extract or partially purified fractions were mixed in a defined solution with ϕ X174 RFI DNA (Fig. 1, lanes 1 to 3) or ϕ X RFI DNA that had been linearized by digestion with *Xho*I (lanes 4 and 5) and incubated at 22°C for 15 to 30 min. The recovered DNA was then analyzed by restriction enzyme digestion, gel electrophoresis, and Southern blotting. A hybridization probe specific for HIV sequences detected a novel band at the position expected for the 15.2-kilobase product of HIV integration into ϕ X DNA only when the reaction was incubated at a temperature greater than 0°C and only when the ϕ X target DNA was present during incubation (Fig. 1, lanes 1 and 4). Bands corresponding to unintegrated HIV DNA were detected under the same conditions and also in control experiments in which either the incubation step (lanes 3 and 5) or target DNA (lane 2) was omitted. In addition to the

FIG. 1. Detection of HIV integration in vitro. Integration reactions using ϕ X174 RFI DNA as the target were performed as described in Materials and Methods, with the exceptions noted. Lanes: 1, standard reaction; 2, target DNA omitted from reaction mixture; 3, incubation on ice instead of at 22°C; 4, same as lane 1, except that ϕ X RFI DNA linearized by cleavage with *Xho*I was used in place of circular ϕ X DNA; 5, same as lane 4, except that incubation was on ice. DNA recovered from the reactions was digested with *Sall* and *Nco*I (lanes 1 to 3) before electrophoresis through a 0.8% agarose gel or loaded directly onto the gel (lanes 4 and 5). *Nco*I and *Sall* do not cut ϕ X174 DNA but cleave HIV-1_{HXB-2} DNA at sites 5.7 and 5.8 kilobases, respectively, from the left end of the molecule. Following electrophoresis, DNA was transferred by blotting to a nylon filter (Hybond N; Amersham Corp.) and HIV sequences were detected by hybridization by using a 32 P-labeled probe prepared from a complete molecular clone of the HIV-1_{HXB-2} provirus. The bands representing the unintegrated linear molecule, the products of *Nco*I and *Sall* digestion of the one-LTR and two-LTR circular forms of unintegrated HIV DNA, or the left-end fragment produced by cleavage of the linear molecule with *Sall* and *Nco*I are indicated. The arrowheads indicate the positions of bands representing putative recombinant molecules resulting from in vitro integration of HIV DNA into ϕ X RFI DNA.

linear form, circular forms of viral DNA containing either one or two copies of the LTR were present in the cell extracts. The putative integration product represented approximately 5% of the HIV DNA present at the start of the reaction.

Structure of HIV proviruses integrated in vitro. The products of authentic retroviral integration are characterized by a distinctive arrangement of sequences at the junction between proviral and target DNAs (18, 19). In HIV, proviral DNA is normally flanked by a direct repeat of a 5-base-pair sequence that was initially present in a single copy at the target site, and the boundaries of the provirus are defined by a CA dinucleotide found 1 base from each end of the LTR (12; Karen Vincent, unpublished data). To confirm the identity of the putative in vitro integration products, we sequenced the junctions from several recombinant molecules. The lack of a genetic marker in HIV suitable for direct selection in *E. coli* (2) and the low absolute yield of in vitro recombinants made direct cloning of the junctions between viral and target DNAs difficult. To clone the junctions for sequencing, we therefore used a small plasmid, π AN13 (10), as the target DNA for integration and used the polymerase chain reaction to amplify the junctions between the HIV

FIG. 2. Procedure for cloning of the junctions between target DNA and HIV proviruses produced by integration in vitro. Details are presented in Materials and Methods. PCR, Polymerase chain reaction.

provirus and target DNAs, along with the plasmid vector, as schematized in Fig. 2. The recovered products in each case had the same structural characteristics as proviruses integrated in vivo, suggesting that the in vitro reaction faithfully mimics in vivo HIV integration (Fig. 3). In one case, recombinant 3 in Fig. 3, the thymidine normally present at the extreme left end of the provirus was apparently replaced by a cytosine. We suspect that this mutation was introduced during polymerase chain reaction amplification, although we cannot exclude the possibility that it occurred before or during integration. The sites used for integration in vitro were scattered throughout the regions of the target plasmid that are dispensable for its replication and selection, and the sequences in the vicinity of the integration sites followed no clear pattern.

Cofactor requirements for HIV integration. To determine whether any nucleotides or other small molecules were required as cofactors for HIV integration, the crude cytoplasmic extract from infected cells was fractionated by gel exclusion chromatography by using a matrix with an exclu-

FIG. 3. Structure of the junctions between π AN13 DNA and HIV proviruses integrated in vitro. The sequences shown correspond to the viral plus strand. The probable sequence of the unsequenced linear precursor is shown at the top for comparison. The ends of this precursor were inferred from the positions of the putative primers for synthesis of the plus and minus strands of viral DNA but were not determined directly. The 5-base-pair repeats flanking proviruses 1, 2, and 3 correspond to bases 237 to 241, 644 to 648, and 129 to 133, respectively, in the sequence of π AN13, as depicted in reference 16.

sion limit of 5×10^6 daltons (Fig. 4A). As expected on the basis of studies of the murine leukemia virus system (2), the viral DNA was recovered in the excluded volume fractions, well separated from ATP and other small molecules (as determined in parallel experiments). Although separated from most cellular proteins and from ATP, these partially purified fractions retained integration activity and did not depend on addition of ATP for their activity (Fig. 4B). These results are consistent with a model in which a nucleoprotein complex mediates HIV integration and the DNA breaking and joining reactions involved in HIV integration are energetically coupled (2).

DISCUSSION

We have shown that HIV integration can occur in a cell-free system. The products of the in vitro reaction have all the distinctive characteristics of authentic HIV proviruses integrated in vivo. Previous published data have left unsettled the question of the length of the flanking host sequence duplication (cf. references 12 and 17). Our in vitro results and unpublished sequences of the host-provirus junctions of proviruses integrated in vivo (K. Vincent et al., unpublished results) show that a 5-base stretch of host DNA is duplicated flanking the integrated provirus. The cell-free system constitutes a practical approach to definition of many of the properties of the HIV integration process (1-3, 7). Our results show that naked plasmid or bacteriophage DNA can serve as a target for HIV integration, that many sites in these molecules can serve as targets, and that the reaction can occur in the absence of a high-energy cofactor. Preliminary results from fractionation of infected cell extracts suggest also that the enzymatic machinery for integration is in a complex with its viral DNA substrate. In all of these respects, the HIV integration reaction resembles murine leukemia virus and Ty1 integration (1-3, 5). However, in view of their differences in structure, genetic organization, and regulation, further biochemical investigation may yet reveal important differences among these mobile genetic elements in the molecular details of the integration process.

In the current work, an extract from HIV-infected cells served as the source of both viral DNA and the proteins required to carry out integration. Because of the efficiency with which viral DNA molecules are integrated in such a system, it provides the best available tool for characterization of the structures of the viral DNA precursor, the active nucleoprotein complex, and DNA intermediates in integration (1-3, 6, 7). Nevertheless, complete characterization of

FIG. 4. Fractionation of integration activity by gel filtration and assessment of the ATP requirement. Cytoplasmic extract was fractionated by gel exclusion chromatography as described in Materials and Methods. (A) Elution profile of the optical density at 280 nm. Fractions were collected as shown and assayed for integration activity with or without added ATP (final concentration, 1 mM). Fractions 25 to 27 were shown to be free of measurable ATP (<2.5 nM) in parallel experiments using a luciferase assay (data not shown). (B) Results of *in vitro* assays for integration activities of column fractions (lanes 1 to 6) or unfractionated cytoplasmic extract (lanes 7 to 9) assayed with (lanes 1, 3, 5, 7, and 9) or without (lanes 2, 4, 6, and 8) ATP. DNA products were analyzed as described in the legend to Fig. 1, except that the reaction products were digested before gel electrophoresis with *EcoRI* (which cleaves twice near the middle of the HIV-1_{HXB-2} genome) rather than with *Sall* and *NcoI*. The positions of the band representing recombinants (*) and bands representing the products of *EcoRI* cleavage of one-LTR circles, two-LTR circles, and unintegrated linear DNA (left and right ends) are indicated. Lanes: 1 and 2, fraction 25; 3 and 4, fraction 26; 5 and 6, fraction 27; 7, standard reaction using crude cytoplasmic extract plus ATP; 8, same as lane 7 but without ATP; 9, same as lane 7 but incubation was on ice. Note that integration activity with ATP appears indistinguishable from that without added ATP.

the enzymology of HIV integration will depend on the ability to reconstitute the activity by using isolated and purified components. Steps toward the development of such a reconstituted system for integration of murine leukemia virus (6)

and avian retrovirus (9) DNAs have recently been reported. Similar modifications to the *in vitro* HIV integration system reported here should allow us to identify and purify the proteins that carry out each step in this important reaction.

Our *in vitro* assay for HIV integration can be adapted for use in screening for agents that specifically block integration of HIV DNA. Drugs currently available for treatment of HIV infection possess significant and often dose-limiting toxicity. Retroviral integration depends specifically on virus-encoded functions (18), and no closely related activity appears to be essential to the economy of normal cells. It is therefore reasonable to hope that inhibitors of HIV integration that might be identified with our *in vitro* assay would include highly specific antiviral agents with low toxicity.

ACKNOWLEDGMENTS

We thank Harold Varmus, Mike Bishop, and Bruce Bowerman for helpful discussions; Ka-Cheung Luk for DNA analysis useful in optimizing extract preparation; Brenda Ross for technical assistance; and Sue Klapholz and Harold Varmus for helpful comments on the manuscript.

This work was supported by Public Health Service grant AI27205 from the National Institutes of Health to P.B., a National Science Foundation predoctoral fellowship to V.E., and the Howard Hughes Medical Institute. P.B. is an assistant investigator of the Howard Hughes Medical Institute.

LITERATURE CITED

1. Bowerman, B., P. O. Brown, J. M. Bishop, and H. E. Varmus. 1989. A nucleoprotein complex mediates the integration of retroviral DNA. *Genes Dev.* 3:469-478.
2. Brown, P. O., B. Bowerman, H. E. Varmus, and J. M. Bishop. 1987. Correct integration of retroviral DNA *in vitro*. *Cell* 49:347-356.
3. Brown, P. O., 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.
4. Curran, J. W., H. W. Jaffe, A. M. Hardy, W. M. Morgan, R. M. Selik, and T. J. Dondero. 1988. Epidemiology of HIV infection and AIDS in the United States. *Science* 239:610-616.
5. Eichinger, D. J., and J. D. Boeke. 1988. The DNA intermediate in yeast Ty1 element transposition copurifies with virus-like particles: cell-free Ty1 transposition. *Cell* 54:955-966.
6. 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.
7. Fujiwara, T., and K. Mizuuchi. 1988. Retroviral DNA integration: structure of an integration intermediate. *Cell* 54:497-504.
8. Heyward, W. L., and J. W. Curran. 1988. The epidemiology of AIDS in the U.S. *Sci. Am.* 259:72-81.
9. Katzman, M., R. A. Katz, A. M. Skalka, and J. Leis. 1989. The avian retroviral integration protein cleaves the terminal sequences of linear viral DNA at the *in vivo* sites of integration. *J. Virol.* 63:5319-5327.
10. Lutz, C. T., W. C. Hollifield, B. Seed, J. M. Davie, and H. V. Huang. 1987. Syrinx 2A: an improved λ phage vector designed for screening DNA libraries by recombination *in vivo*. *Proc. Natl. Acad. Sci. USA* 84:4379-4383.
11. Mann, J. M., J. Chin, P. Piot, and T. Quinn. 1988. The international epidemiology of AIDS. *Sci. Am.* 259:82-89.
12. Muesing, M. A., D. H. Smith, C. D. Cabradilla, C. V. Benson, L. A. Lasky, and D. J. Capon. 1985. Nucleic acid structure and expression of the human AIDS/lymphadenopathy retrovirus. *Nature (London)* 313:450-458.
13. Popovic, M., M. Sarngadharan, E. Read, and R. C. Gallo. 1984. Frequent detection and isolation of cytopathic retroviruses (HTLV-III) from patients with AIDS and at risk for AIDS. *Science* 224:497-500.
14. 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. Ghayeb, 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.
15. Robert-Guroff, M., M. S. Reitz, W. G. Robey, and R. C. Gallo. 1986. In vitro generation of an HTLV-III variant by neutralizing antibody. *J. Immunol.* **137**:3306-3311.
16. Sambrook, J., E. F. Fritsch, and T. Maniatis. 1989. *Molecular cloning: a laboratory manual*, 2nd ed. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
17. 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.
18. 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.
19. Weiss, R., N. Teich, H. E. Varmus, and J. Coffin (ed.). 1984. *RNA tumor viruses*. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.