

Integration of Human Immunodeficiency Virus Types 1 and 2 DNA In Vitro by Cytoplasmic Extracts of Moloney Murine Leukemia Virus-Infected Mouse NIH 3T3 Cells

CORNELIS VINK, DIK C. VAN GENT, AND RONALD H. A. PLASTERK*

*Division of Molecular Biology, The Netherlands Cancer Institute,
Plesmanlaan 121, 1066 CX Amsterdam, The Netherlands*

Received 30 April 1990/Accepted 10 July 1990

An essential step in the life cycle of the human immunodeficiency virus (HIV) is integration of a DNA copy of the viral RNA into the genome of the infected cell. We show here that this step can be faithfully accomplished in vitro by the enzymatic machinery of another retrovirus, Moloney murine leukemia virus (MoMLV). Mini-HIV substrates, which are linearized plasmids with long terminal repeat sequences at their ends, were incubated with cytoplasmic extracts of MoMLV-infected NIH 3T3 cells and target DNA. The MoMLV integration apparatus carried out integration of the mini-HIV substrates correctly; the terminal nucleotides of the viral substrate were removed, and a 4-base-pair duplication of the target DNA flanked the inserted viral DNA (C. Shoemaker, S. P. Goff, E. Gilboa, M. Paskind, S. W. Mitra, and D. Baltimore, Proc. Natl. Acad. Sci. USA 77:3932-3936, 1980). Our experiments show that the substrate sequence requirements for integration in vitro were limited to a few nucleotides, as the similarity between HIV and MoMLV long terminal repeat ends is minimal.

Integration of a double-stranded DNA copy of the viral RNA into the genome of the infected cell is an essential step in the life cycle of all retroviruses. Although much is still unknown about the process of integration, the development of cell-free integration assays has greatly contributed to the understanding of this process. Cytoplasmic extracts from NIH 3T3 cells infected with a mouse retrovirus, Moloney murine leukemia virus (MoMLV), were shown to contain MoMLV DNA that could be integrated into an exogenous target DNA (1). By using this in vitro integration system, the structure of the integration intermediate was elucidated (6).

The basic features of retroviral integration are the following. After reverse transcription and second-strand DNA synthesis, a linear, double-stranded DNA molecule with blunt ends is formed. This DNA copy of the viral genome is flanked by direct repeat sequences called long terminal repeats (LTRs). At the outer edges of the two LTRs, short, imperfect inverted repeats are present (for a review, see reference 19). These sequences are sufficient for correct integration (3, 4). The terminal nucleotides at each 3' end of the blunt-ended DNA are then cleaved off, resulting in the putative immediate precursor for integration (6, 13). This cleavage reaction is probably executed by the viral integrase protein, which is the only viral protein known to be required for integration (2, 9, 13). The integrase is encoded by the 3' part of the viral *pol* gene. A random sequence in the target DNA is cut in a staggered pattern, giving rise to a 5' overhang. The length of this staggered cut is virus specific (20). The 5' ends of the target DNA are then joined to the viral 3' OH ends. The last step in proviral integration is repair of the gaps present in the intermediate. It is possibly in this step that the terminal nucleotides at each proviral 5' end are lost. The resulting provirus is now flanked by a short direct duplication of the target DNA and terminates with 5'-TG --/--CA-3'.

Recently, an assay was described in which cytoplasmic extracts from MoMLV-infected NIH 3T3 cells (or disrupted MoMLV particles) were capable of integrating exogenously added mini-MoMLV DNA into target DNA (5). This assay makes it possible to investigate whether the MoMLV integration machinery is also able to integrate other mini-retroviral DNAs, such as mini-human immunodeficiency virus (HIV) DNA. Although the general features of HIV and MoMLV integration are probably similar, there are some clear differences in the actual integration reactions. First, the integrase proteins of HIV and MoMLV are not very similar; there is less than 20% similarity on the amino acid level (7, 8, 17, 21). Second, the imperfect inverted repeats at the outer edges of the LTRs, the only sequences necessary in *cis* for correct integration, differ significantly between the two retroviruses. The terminal 18 base pairs (bp) of MoMLV are aligned with the corresponding sequences from HIV-1 and HIV-2. Only 7 of 18 bp at the HIV-1 U3 end match the MoMLV sequence; at the HIV-2 U3 end, the match is 6 of 18 (Fig. 1B). At the U5 terminus, the match is 6 of 18 bp for HIV-1 and 5 of 18 bp for HIV-2. Third, integrated MoMLV proviruses are flanked by 4-bp directly repeated target DNA duplications (18), whereas HIV proviruses are flanked by 5-bp target DNA duplications (12; C. Vink, M. Groenink, Y. Elgersma, R. A. M. Fouchier, M. Tersmette, and R. H. A. Plasterk, submitted for publication).

We tested the abilities of cytoplasmic extracts of MoMLV-infected NIH 3T3 cells to use HIV-1 and HIV-2 DNA as substrates for correct integration. Extracts were prepared as previously described (5). As an exogenous substrate for integration, we first used a mini-MoMLV plasmid in the in vitro system. The mini-MoMLV substrate is a plasmid that can be digested with restriction endonuclease *Nde*I, after which the ends of the plasmid are almost identical to the termini of unintegrated viral DNAs within the cytoplasm of MoMLV-infected cells. The only difference between the mini-MoMLV DNA and wild-type MoMLV DNA is the presence of a T instead of an A nucleotide at the

* Corresponding author.

TABLE 1. Integration^a of mini-retroviral DNAs

Substrate	No. of integrants	Total plaques	Integration frequency (%)
MoMLV(pRP250, <i>NdeI</i>)	133	1.6×10^7	8.4×10^{-6} (100)
HIV-1(pRP131, <i>NdeI</i>)	75	1.6×10^8	4.7×10^{-7} (5.6)
HIV-2(pRP255, <i>NdeI</i>)	7	3.0×10^7	2.3×10^{-7} (2.7)
HIV-1(pRP137, <i>PvuII</i>)	0	1.8×10^8	$<0.6 \times 10^{-8}$ (<0.07)
Control(pRP268, <i>NdeI</i>)	0	9.2×10^7	$<1.0 \times 10^{-8}$ (<0.13)
HIV-1(pRP131, <i>NdeI</i>) ^b	0	3.4×10^7	$<3.0 \times 10^{-8}$ (<0.35)

^a Integration reactions were carried out as previously described (5). Clone 4, a wild-type MoMLV-producing cell line derived from murine NIH 3T3 cells, was originally made by S. P. Goff (Columbia University, New York). This cell line was obtained from R. Craigie (National Institutes of Health, Bethesda, Md.). Clone 4 and uninfected NIH 3T3 cells were grown in Dulbecco modified Eagle medium with 10% newborn calf serum. The preparation of cytoplasmic extracts from uninfected and MoMLV-infected NIH 3T3 cells was done as previously described (5). In each reaction, 0.26 pmol of *NdeI*-linearized (pRP250, pRP255, pRP131, and pRP268) or *PvuII*-linearized (pRP137) plasmid and 1.5 μ g of lambda *Aam Bam Sam* concatemers (obtained from L. Stubbs, Imperial Cancer Research Fund, London, England) were used. Integration reactions were terminated by addition of 8 μ l of 0.5 M EDTA (pH 8.0), 8 μ l of 10-mg/ml Proteinase K (Merck & Co., Inc.), 10 μ l of 20% sodium dodecyl sulfate, and 224 μ l of H₂O. After incubation at 45°C for 16 h, 100 μ l of 1.5 M NaOAc (pH 7.0) was added and the samples were extracted with buffer-saturated phenol and phenol-chloroform-isoamyl alcohol (25:24:1, vol/vol/vol). The DNA was precipitated with ethanol, washed with 70% ethanol, and suspended in 6 μ l of 1 mM Tris hydrochloride (pH 8.0)–0.05 mM EDTA. The recovered DNA was packaged in an in vitro packaging system (Amersham Corp.). Plating of the resulting phages on *E. coli* CA274 [*sup*⁰ *lacZ*(*Am*)] was done as described by the manufacturers of the packaging system. To the plating mixture, 0.4 mg of 5-bromo-4-chloro-3-indolyl- β -D-galactopyranoside and 0.1 mg of isopropyl- β -D-thiogalactopyranoside per ml was added, and the mixture was plated onto Luria-Bertani agar plates. Blue plaques were counted after about 16 h of incubation at 37°C. Total phage titers were determined by plating phage dilutions onto lawns of *E. coli* NM538 (*supF*). The columns show the total numbers of plaques on the selective strain (from phage DNAs that have an integrated miniretrovirus) and on the nonselective strain, the ratio between these two (the integration frequency), and the relative integration frequency (relative to the frequency for the mini-MoMLV substrate, expressed as a percentage). Data were pooled from a duplicate set of experiments.

^b Extract from uninfected NIH 3T3 cells.

MoMLV (pRP250–*NdeI*) 1. (15,868) GTGG**AAACT**GGA – – // – – TC**AAAC**CACC
 2. (19,477) TATA**CCTG**TGA – – // – – TC**ACCTG**CACT

HIV-1 (pRP131–*NdeI*) 1. (33,608) TTCG**CTACT**GGA – – // – – GC**ACTAC**CATC
 2. (30,374) CCTG**AGAG**TGG – – // – – GCA**AGAG**TTAA

HIV-2 (pRP255–*NdeI*) 1. (27,972) CGCC**TACCT**GGA – – // – – GC**TACCT**TTTC
 2. (29,392) AAAA**CATAT**GGA – – // – – GC**CATAT**AGA

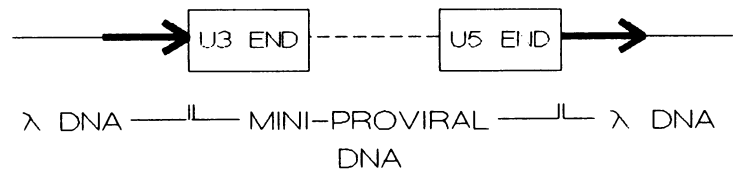


FIG. 2. Sequences of the junctions between lambda DNA and integrated mini-proviral DNA. From two blue plaques from each integration reaction, lambda DNAs were isolated by the method of Manfioletti and Schneider (10). The junctions between lambda DNAs and mini-retroviral DNAs were directly sequenced by the dideoxy sequencing method (15) with the Sequenase system (United States Biochemical Corp.). The sequencing primers used were 5'-CCACAGATCAAGGATAT-3' and 5'-ACTCTGGTAACTAGAGA-3' for mini-HIV-1 recombinants and 5'-TTACGTTGAGAAAGA-3' and 5'-AAACGCCAGCAACGC-3' for the other recombinants. The numbers in parentheses indicate the positions of the first nucleotide of the sequences in the lambda DNA.

lambda DNA and integrated mini-MoMLV DNA were sequenced. The sequences are shown in Fig. 2. In both recombinant phages, the mini-proviral DNAs terminate correctly with the sequence 5'-TG--/--CA-3' and are flanked by a 4-bp directly repeated duplication of the target DNA, which is characteristic of authentic MoMLV integration (18). Target DNA duplication was checked by comparing the duplicated sequence with the sequence of wild-type lambda DNA (14).

For mini-HIV-1 and mini-HIV-2 substrates, we used plasmids (pRP131 and pRP255) that comprise ends that, after cleavage by *NdeI*, are identical to the termini of the putative immediate precursors for integration of HIV-1 and HIV-2, except for the presence of the dinucleotide AT at the 5' ends, which is different in wild-type HIV-1 and HIV-2 DNAs. As for the mini-MoMLV substrate, we assume that the crucial aspect is the 3' terminal CA sequence at both ends that ligates to the 5' overhanging target and that minor differences in the 5' overhanging termini of the mini-retroviral DNAs are irrelevant. The results of the integration experiments in which the mini-HIV substrates were used are listed in Table 1. Both the mini-HIV-1 and the mini-HIV-2 plasmids that contain *NdeI* cohesive ends were integrated. The frequency of integration of the mini-HIV-1 plasmid is about 20 times lower than the frequency of integration of the mini-MoMLV substrate. The mini-HIV-2 plasmid was integrated at a frequency that is approximately 40 times lower than the integration frequency of the mini-MoMLV plasmid. To check whether the blue plaques were the product of correct retroviral integration, we isolated lambda DNAs from two blue plaques from each integration reaction and sequenced the junctions between the mini-HIV DNA and lambda DNA. The HIV-1 and HIV-2 substrates were correctly integrated (Fig. 2); the terminal mini-HIV sequences are 5'-TG--/--CA-3'. The integrated mini-HIV plasmids are

flanked by 4-bp directly repeated target DNA duplications, which is characteristic for MoMLV integration, and not by 5-bp target DNA duplications, which is characteristic for HIV integration. This suggests that the length of the target DNA duplication is determined by one or more virally encoded proteins (in this study, from MoMLV) and not by the DNA that is integrated (in this study, from HIV). The integration of mini-retroviral DNAs in our assay is dependent on MoMLV, since extracts from uninfected NIH 3T3 cells do not promote integration of the mini-HIV-1 plasmid pRP131 (Table 1). Although the mini-HIV DNAs are integrated at a lower frequency relative to the frequency of integration of the mini-MoMLV substrate, it is clear from our experiments that not all the nucleotides present at the LTR termini of MoMLV are required for correct integration (Fig. 1B). It is possible that only the nucleotides that are similar among MoMLV and HIV-1 and HIV-2 (these nucleotides are boxed in Fig. 1B) are necessary for correct integration. It cannot be excluded that for some of the positions the conservation in sequence requirement is partial (e.g., only purines are allowed) and that therefore some of the positions not boxed in Fig. 1B may still be of importance.

To determine whether the cohesive ends that are generated by digestion with *NdeI* are possibly sufficient substrates for integration, we tested a plasmid that does not contain any viral sequences. We first cloned the *supF* gene into pUC18, and the resulting plasmid, pRP268, was digested with *NdeI* and used in the assay. No integration of this substrate was observed (Table 1), indicating that *NdeI* cohesive ends are not sufficient substrates for integration.

Another mini-retroviral substrate that we tested was a mini-HIV-1 plasmid containing a *PvuII* restriction site at the boundary between the U3 and U5 terminal sequences (pRP137) (Fig. 1). When digested with *PvuII*, a G nucleotide is present at the 3' ends of this plasmid. Prior to correct integration, this nucleotide will have to be removed by the integration machinery, thereby generating the 3' terminal sequence CA-3'. No integration of substrate pRP137 was detected. This is in agreement with the finding that the MoMLV integration machinery is not able to integrate blunt-ended mini-MoMLV substrates in vitro (5). A possible explanation might be that the step in vivo preceding the actual integration, the cutting off of the nucleotides beyond the CA-3', is not carried out with sufficient efficiency on exogenous DNA in this cell-free system and might depend on the precise architecture of the endogenous integration complex.

We thank R. Craigie (National Institutes of Health, Bethesda, Md.) for communicating his results prior to publication and for making available a MoMLV-producing cell line. The help of Henri van Luenen in growing these cells is greatly appreciated.

This work was supported by a grant from the Dutch Organization for Scientific Research, a grant from the Dutch Ministry of Health, and a CCH grant from the Dutch Organization for Scientific Research to R.H.A.P.

LITERATURE CITED

1. Brown, P. O., B. Bowerman, H. E. Varmus, and J. M. Bishop. 1987. Correct integration of retroviral DNA in vitro. *Cell* **49**:347-356.
2. 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.
3. Colicelli, J., and S. P. Goff. 1985. Mutants and pseudorevertants of Moloney murine leukemia virus with alterations at the integration site. *Cell* **42**:573-580.
4. Colicelli, J., and S. P. Goff. 1988. Sequence and spacing requirements of a retrovirus integration site. *J. Mol. Biol.* **199**:47-59.
5. 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.
6. Fujiwara, T., and K. Mizuuchi. 1988. Retroviral DNA integration: structure of an integration intermediate. *Cell* **54**:497-504.
7. Guyader, M., M. Emerman, P. Sonigo, F. Clavel, L. Montagnier, and M. Alizon. 1987. Genome organization and transactivation of the human immunodeficiency virus type 2. *Nature (London)* **326**:662-669.
8. Johnson, M. S., M. A. McClure, D.-F. Feng, J. Gray, and R. F. Doolittle. 1986. Computer analysis of retroviral *pol* genes: assignment of enzymatic functions to specific sequences and homologies with nonviral enzymes. *Proc. Natl. Acad. Sci. USA* **83**:7648-7652.
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. Manfioletti, G., and C. Schneider. 1988. A new and fast method for preparing high quality lambda DNA suitable for sequencing. *Nucleic Acids Res.* **16**:2873-2884.
11. Messing, J., R. Crea, and P. H. Seeberg. 1981. A system for shotgun DNA sequencing. *Nucleic Acids Res.* **9**:309-321.
12. 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.
13. Roth, M. J., P. L. Schwartzberg, and S. P. 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.
14. Sanger, F., A. R. Coulson, G. F. Hong, D. F. Hill, and G. B. Petersen. 1982. Nucleotide sequence of bacteriophage lambda DNA. *J. Mol. Biol.* **162**:729-773.
15. Sanger, F., S. Nicklen, and A. R. Coulson. 1977. DNA sequencing with chain terminating inhibitors. *Proc. Natl. Acad. Sci. USA* **74**:5463-5467.
16. Seed, B. 1983. Purification of genomic sequences from bacteriophage libraries by recombination and selection in vivo. *Nucleic Acids Res.* **11**:2427-2447.
17. Shinnick, T. M., R. A. Lerner, and J. G. Sutcliffe. 1981. Nucleotide sequence of Moloney murine leukaemia virus. *Nature (London)* **293**:543-548.
18. Shoemaker, C., S. P. 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.
19. Varmus, H., and P. Brown. 1989. Retroviruses, p. 53-108. In M. M. Howe and D. E. Berg (ed.), *Mobile DNA*. American Society for Microbiology, Washington, D.C.
20. Varmus, H. E. 1983. Retroviruses, p. 411-503. In J. Shapiro (ed.), *Mobile genetic elements*. Academic Press, Inc., New York.
21. Wain-Hobson, S., P. Sonigo, O. Danos, S. Cole, and M. Alizon. 1985. Nucleotide sequence of the AIDS virus, LAV. *Cell* **40**:9-17.