High Rate of Genetic Recombination in Murine Leukemia Virus: Implications for Influencing Proviral Ploidy

Jianling Zhuang,^{1†} Sayandip Mukherjee,^{1,2†} Yacov Ron,¹ and Joseph P. Dougherty^{1*}

Department of Molecular Genetics, Microbiology and Immunology, Robert Wood Johnson Medical School, 675 Hoes Lane, Piscataway, New Jersey 08854,¹ and Graduate Program in Molecular Biosciences, Rutgers University, New Brunswick, New Jersey 08901²

Received 7 February 2006/Accepted 12 April 2006

A significant difference in the recombination rates between human immunodeficiency virus type 1 (HIV-1) and the gammaretroviruses was previously reported, with the former being 10 to 100 times more recombinogenic. It is possible that preferential copackaging of homodimers in the case of gammaretroviruses, like murine leukemia virus (MLV), led to the underestimation of their rates of recombination. To reexamine the recombination rates for MLV, experiments were performed to control for nonrandom copackaging of viral RNA, and it was found that MLV and HIV-1 exhibit similar crossover rates. The implications for control of proviral ploidy and preferential recombination during minus-strand DNA synthesis are discussed.

It is not surprising that retroviruses can efficiently undergo homologous recombination since the virions are diploid, providing the opportunity to recombine during reverse transcription. Moreover, two obligatory strand transfers occur during reverse transcription, so the type of template-switching ability that would be required for recombination is built into the viral life cycle. Results from our laboratory and from others also indicate that retroviral recombination occurs predominantly during minus-strand synthesis using viral RNA as a template (2, 10, 28–30).

Previously, we demonstrated that human immunodeficiency virus type 1 (HIV-1) recombines at a rate of at least three crossovers per cycle of replication in HeLaT4 cells (30). It was reported to be higher in other cell types, including primary cells and established cell lines. For macrophages, it was reported to be an astounding 30 crossovers per genome per cycle of replication (13), although this has been recently called into question (14). It was found through *in vivo* studies that the mutation rates of HIV-1, spleen necrosis virus, and murine leukemia virus (MLV) were quite similar (21, 24, 25), and the strand transfer activities of HIV-1 and MLV were found to be equivalent on the basis of measuring the rates of homologous recombination between direct repeats within the same genome (17). However, the intermolecular recombination frequencies between separate, copackaged viral RNAs were reported to be significantly higher for HIV-1 than for MLV (10). For MLV, the recombination rate was calculated to be 0.3 to 0.4 crossovers per cycle of replication, which is about 10-fold lower than what we and others found for HIV-1 (2, 10, 30). One of the assumptions previously made for measuring recombination rates was that the two genetically distinct viral RNAs were randomly copackaged into virions. However, this was recently reported not to be the case for MLV (7, 11). Instead, there

is preferential copackaging of homodimers, which likely led to an underestimation of the MLV recombination rate since the recombinant proviral progeny of homodimeric virions would not be scored (17). In this report, we sought to determine the MLV rate of crossover by exclusively examining the progeny of heterodimeric virions.

Strategy for analyzing proviral progeny of heterozygous virions. The MLV-derived vectors and the strategy to identify progeny proviruses of heterodimeric virions are depicted in Fig. 1. The concept is to utilize two vectors (Fig. 1A), one with a deletion of the 5' long terminal repeat (LTR) along with a deleted 3' U5 (CMVG₂IP) in combination with a second vector with intact LTRs (G₈INeo). The entire 5' LTR was deleted, and the CMV promoter forms a junction with the viral primer binding site. It should be noted here that the 3' U5 in CMVG₂IP was deleted to eliminate the possibility of restoring a functional 5' LTR via recombination during transfection. The bovine growth hormone polyadenylation [poly(A)] signal was inserted into CMVG₂IP because part of the viral poly(A) was lost due to deletion of the 3' U5. Homodimers of CMVG₂IP should not be able to replicate due to the lack of 5' R since it cannot produce a primer for successful minus-strand transfer (Fig. 1B), and homodimers of G₈INeo with the intact LTR lack the *puro* selection marker; therefore, they cannot survive *puro* selection. Identification of proviral progeny of heterodimeric virions can be obtained by selecting for puromycin resistance, which is linked to the defective vector. This would require an intermolecular minus-strand transfer, which has been shown to occur 50% of the time during HIV-1 reverse transcription (23, 28) and has been shown to occur during gammaretrovirus replication as well (9, 12, 20). The frequency of recombination is measured by restoration of wild-type *gfp*, as the two vectors contain complementary inactivating mutations in *gfp*. To provide a comparison, parallel experiments were conducted utilizing the vectors, both with intact LTRs, G₂IP, and G₈INeo, from which puromycin-resistant proviruses were obtained as a result of target cell infection with both homodimeric (G₂IP) and heterodimeric (G₂IP/G₈INeo) virions (Fig. 2). The only difference was that both the *neo* and *puro*

* Corresponding author. Mailing address: UMDNJ, Robert W. Johnson Medical School, Dept. of Molecular Genetics, Microbiology and Immunology, 675 Hoes Lane, Piscataway, NJ 08854-5635. Phone: (732) 235-4588. Fax: (732) 235-5223. E-mail: doughejp@umdnj.edu.

† These authors contributed equally to the work.

FIG. 1. Strategy for analyzing the proviral progeny of heterozygous virions. (A) Constructs used for the study. GIP is an MLV-based vector with deleted *gag-pol* and *env* genes and an expression cassette consisting of wild-type *gfp* and a puromycin resistance gene (*puro*) separated by an encephalomyocarditis virus internal ribosome entry site (IRES). G₂IP is identical to GIP except it contains a *gfp* with a 5-nucleotide frameshift mutation at nucleotide position 65. CMVG₂IP is identical to G₂IP except that the 5' LTR has been replaced with a cytomegalovirus (CMV) immediate early gene promoter and the 3' U5 has been replaced with a bovine growth hormone-derived polyadenylation signal (pA). G₈INeo contains full-length LTRs and a *gfp* with a 5-nucleotide frameshift mutation at nucleotide position 352 along with a neomycin resistance gene (*neo*) separated from the *gfp* by an IRES. Asterisks indicate positions of complementary frameshift mutations. (B) *puro*-resistant proviruses are progeny of heterodimers. Without *r*, the CMVG₂IP RNA strong-stop DNA primer cannot pair with the 3' *r* and effectively strand transfer. An intermolecular primer strand transfer from G₈INeo RNA to CMVG₂IP RNA is shown, as only *puro*-resistant cells will be scored. A crossover is depicted to indicate an example of recombination. Thin lines represent viral RNAs, and thick lines represent DNA. pbs, primer binding site; ψ , viral packaging signal. u3, *r*, and u5 indicate sequences in the RNA; U3, R, and U5 indicate corresponding sequences in the DNA derived from the RNA.

titers were obtained to allow for correction using the Hardy-Weinberg equation.

To ensure that CMVG₂IP was totally incapable of replication on its own, we transfected the plasmid into the NIH 3T3-based GP+E-86 ecotropic packaging cell line followed by harvesting supernatant and inoculating NIH 3T3 target cells. We used the transfer of puromycin resistance to target cells as a test for the ability to produce vector virus. As anticipated, we obtained a puromycin resistance titer of 0.0 infectious units (IU)/ml in target cells using CMVG₂IP. For a positive control for vector virus passage, we transfected, in parallel, GIP into

GP+E-86 cells and obtained *puro* titers equal to $5.6 \pm 0.7 \times 10^6$ IU/ml on NIH 3T3 cells.

To implement the strategy outlined above (Fig. 1B), we established three GP+E-86-based producer cell clones, each of which stably harbors a single G₈INeo provirus (Fig. 2). This was done to eliminate the need to cotransfect G₈INeo plus CMVG₂IP simultaneously, which might result in restoring the CMVG₂IP LTR as well as the parental *gfp* before viral replication.

Propagation of heterodimeric virions. The GP+E-86 producer cell clones harboring the G₈INeo provirus were tran-

FIG. 2. Experimental protocol for measurement of MLV recombination rates. GP+E-86 producer cell clones harboring G₈INeo provirus were transfected separately with CMVG₂IP and G₂IP, and viral supernatant was harvested to infect NIH 3T3 cells in serial dilutions. The target cells were subjected to puromycin selection, and a manual count of puromycin-resistant green fluorescent protein (GFP)-positive clones was performed after selection was completed. Target cells were also subjected to G418 selection as indicated.

siently transfected with CMVG₂IP using a standard protocol, and viral supernatant was harvested within 48 h posttransfection and used to infect NIH 3T3 target cells (Fig. 2). This was followed by puromycin selection. All puromycin-resistant clones are progeny of heterozygous virions. Experiments were also done in parallel by transfecting G₂IP instead of CMVG₂IP and subjecting the target cells to puromycin or G418 selection. Superinfection of these producer cells is blocked as a result of superinfection interference (3, 16) and the relatively short period between transfection and virus harvest. Moreover, vector virus cannot spread in the target cells because of lack of helper plasmids, thereby restricting viral replication to a single cycle. To control for superinfection, we subjected the producer cell clones to puromycin selection after withdrawing the viral supernatant and could not detect any measurable puromycin resistance, indicating that, as expected, superinfection of the producer cells was not contributing to heterodimeric virus production. The green fluorescent protein (GFP) titer yielded the recombination titer and the *puro* titer the overall virus titer, which allows calculation of the recombination rate. The rate obtained by transfecting G₂IP into producer cells gives us the "apparent" rate of recombination since puromy-

cin-resistant clones are progenies of cells infected with both heterodimeric and homodimeric virions. In keeping with previous experimental protocols, we assumed random copackaging and applied the Hardy-Weinberg equation to correct for homodimeric virions contributing to puromycin resistance in the NIH 3T3 cells (Table 1). The average recombination rate obtained was 2.9×10^{-5} /base/replication cycle, which is comparable to the previously published MLV recombination frequency without controlling for nonrandom copackaging (1). On the other hand, the rate obtained by transfecting CMVG₂IP provides the "real" rate of recombination since all puromycin-resistant clones are progenies of heterodimeric virions (Table 2). The average rate obtained was 4.6×10^{-4} /base/replication cycle (equivalent to 3.8 crossovers/genome/replication cycle), which is quite similar to the rate of 3.05×10^{-4} /base/replication cycle, previously calculated for HIV-1 (10, 30). It is noteworthy that here, we are measuring recombination from a specific RNA template to the other strand. In the event that minus-strand priming and synthesis occurred on both viral RNAs, it would double the chance of homologous recombination occurring at least until the first crossover happened (Fig. 3). There was a >1-log-higher rate of recombination ("real rate" versus "apparent rate") in the three

TABLE 1. "Apparent" rates of recombination (transfected plasmid: G₂IP)^a

Clone no.	GFP titer (IU/ml)	<i>puro</i> titer (IU/ml)	<i>neo</i> titer (IU/ml)	Correction factor ^b	Recombination rate ^c	Extrapolated recombination rate ^d
1	1.7×10^3	2.3×10^5	1.5×10^6	0.87	3.0×10^{-5}	0.25
2	2.6×10^3	3.8×10^5	2.3×10^6	0.86	2.8×10^{-5}	0.23
3	3.1×10^3	4.2×10^5	2.6×10^6	0.86	3.0×10^{-5}	0.25
Avg					2.9×10^{-5}	0.24

^a All experiments were repeated at least twice, and infections were done in duplicates.

^b According to Hardy-Weinberg equilibrium and assuming random copackaging, the percentage of heterozygous virions contributing to *puro* resistance is calculated as $PN/P^2 + PN$, where P and N denote *puro* and *neo* titers, respectively. For example, for clone 1, the *puro/neo* titer was 1:6.5, and therefore, $PN/P^2 + PN = 6.5/7.5$, or 0.87.

^c The recombination rate is calculated as GFP titer/(*puro* titer $\times$ correction factor) per 287 bases per replication cycle.

^d Calculated by multiplying the recombination rate per base per replication cycle with the MLV genomic RNA size, 8,332 bases (GenBank accession no. AF033811).

different producer clones, confirming the hypothesis that, considering the progeny of only heterodimeric virions, the rate of recombination is significantly higher. Primers were used to amplify *gfp* from puromycin-resistant *gfp*-positive recombinant clones and subjected to sequencing and were found to have regenerated the parental *gfp* by recombination during reverse transcription (data not shown).

Significance. We have shown that the rate of intermolecular crossovers for the gammaretrovirus MLV is quite high and similar to that of HIV-1. Previous attempts to examine recombination rates in MLV did not control for the preferential copackaging of viral RNAs, which appears to have resulted in the underestimation of its crossover rate. Similarly, it is possible that the recombination frequencies of other retroviruses reported might have been experimentally underestimated due to nonrandom copackaging. It is important to note that we have calculated the minimum rate of recombination for MLV since double crossovers between the frameshift mutations will not be scored and since we are examining crossovers only from the *puro*-containing template to the *neo*-containing template and not the other way around.

TABLE 2. "Real" rates of recombination (transfected plasmid: CMVG₂IP)^a

Clone no.	GFP titer (IU/ml)	<i>puro</i> titer (IU/ml)	Recombination rate ^b	Extrapolated recombination rate ^c
1	12	1.3×10^2	4.6×10^{-4}	3.8
2	14	1.4×10^2	5.0×10^{-4}	4.1
3	16	1.9×10^2	4.2×10^{-4}	3.5
Avg			4.6×10^{-4}	3.8

^a All experiments were repeated at least twice, and infections were done in duplicates.

^b The recombination frequency is calculated as GFP-positive titer/*puro* titer, which averages 9.1%. To determine a more accurate rate of recombination, the denominator should be slightly adjusted because some of the *puro* progeny would have undergone recombination before reaching the position of the 3' *gfp* frameshift mutation, so these progeny would not be able to contribute to the population that can score for recombination since even if they crossed over between the frameshift mutations, they would result in a GFP-negative phenotype. To obtain a minimum correction, the denominator should be decreased by the recombination frequency times the distance between *puro* and the 3' frameshift mutation (974 bases) divided by the distance between the frameshift mutations (287 bases), or by $9.1\% \times 974/287$, or 30.8%. Thus, recombination rate is calculated as GFP-positive titer/(*puro* titer $\times$ 0.692) per 287 bases per replication cycle.

^c Calculated by multiplying the recombination rate per base per replication cycle with the MLV genomic RNA size, 8,332 bases (GenBank accession no. AF033811).

The high rate of crossover in MLV has important mechanistic implications (Fig. 3). It would help to explain why there seems to be a preference for recombination during minus-strand synthesis since once a crossover happens, it blocks synthesis of the second minus-strand, so that there is only one template for the synthesis of one plus-strand, thereby preventing plus-strand recombination (8, 28). The high rate of crossover also implies that recombination might influence MLV proviral ploidy. Since there are 50 to 100 reverse transcriptase molecules and 50 to 100 tRNAs found in a virion (5, 18, 19), it should be possible to produce two proviruses from a single virion, which is supported by the finding that two tRNAs can prime synthesis in a single virion (26, 27). However, it has been reported that only a single provirus is derived from an individual virion (9), leading to the question of what controls this elemental facet of retroviral replication. First, dual initiation occurs less than 30% of the time during reverse transcription in both MLV- and HIV-1-based vectors containing two primer binding sites on the same strand (26, 27). Second, such a high crossover rate suggests that essentially all proviruses are subjected to crossover during DNA synthesis, so even if dual initiation did occur, it would be possible to synthesize only one provirus most of the time because after the first recombination crossover, the template for a second DNA would be destroyed (Fig. 3). In order to synthesize two proviruses in an instance in which dual priming transpired, it would require that no crossovers occur, where the average number of crossovers that the virus would have to forego is 7.6 per genome (2×3.8). It is noteworthy that when considering the probability of whether two proviruses can be formed, the rate needs to be doubled because the virus would have to forego almost four crossovers from each template, as they would both need to remain intact during replication. Given a Poisson distribution, the chance of this occurring would be 0.05%. Thus, it is possible that the relative inefficiency of dual priming coupled with the high crossover rate might, at least in part, account for the pseudodiploid nature of the virus. However, one cannot rule out the possibility that other factors might also contribute to the control of proviral ploidy.

Besides having mechanistic implications, the higher rate of recombination in MLV also affects viral fitness since it allows for the transfer of multiple nucleotide changes in a new allelic combination and occurs at a frequency higher than those of point mutations (4, 6). Also, the emergence of new viral strains is governed by the rates of appearance and stabilization of

FIG. 3. Hypothetical model for recombination controlling proviral ploidy and inhibiting plus-strand recombination. (i) Two tRNAs prime minus-strand DNA synthesis. (ii) r-u5 digested by RNase H activity of reverse transcriptase; the minus-strand DNA primer transfers to the 3' end of the RNA genome (only intramolecular strand transfer is shown). (iii) Synthesis of minus-strand DNA is not synchronized for the two strands with one ahead of the other. (iv) Once strand transfer occurs for the nascent DNA strand that is leading, the RNA template for the lagging strand is lost. (v) This prevents two proviruses from being formed and would inhibit recombination during plus-strand synthesis by preventing synthesis of two complete minus-strand DNAs. Thin lines represent RNA. Thick lines represent DNA. Dotted lines represent RNase H-digested RNA. The X represents abortive minus-strand DNA synthesis. u3, r, and u5 indicate sequences in the RNA; U3, R, and U5 indicate corresponding sequences in the DNA derived from the RNA.

point mutations, and recombination plays a significant role in how these mutations spread through the viral population (22). Recent observations have highlighted the role of genetic diversity during the progression of HIV-1 infection (15), and a similarly high rate of recombination in MLV, coupled with its propensity for forming homodimers, could be crucial in determining the evolutionary adaptiveness of the virus under severe selection pressure since even mild fluctuations in the local frequency of recombination could bias the generation of specific recombinant forms.

We sincerely thank Annmarie L. Pacchia, Hui-ling R. Lee, and

Bradley D. Phelan for helpful suggestions and critically reviewing the manuscript.

This work was supported by NIH grants AI51910 and CA50777.

REFERENCES

1. Anderson, J. A., E. H. Bowman, and W. S. Hu. 1998. Retroviral recombination rates do not increase linearly with marker distance and are limited by the size of the recombining subpopulation. *J. Virol.* **72**:1195-1202.
2. Anderson, J. A., R. J. Teufel II, P. D. Yin, and W. S. Hu. 1998. Correlated template-switching events during minus-strand DNA synthesis: a mechanism for high negative interference during retroviral recombination. *J. Virol.* **72**:1186-1194.
3. Berkowitz, R. D., and S. P. Goff. 1993. Point mutations in Moloney murine leukemia virus envelope protein: effects on infectivity, virion association, and superinfection resistance. *Virology* **196**:748-757.

4. **Chao, L.** 1990. Fitness of RNA virus decreased by Muller's ratchet. *Nature* **348**:454–455.
5. **Coffin, J. M., S. H. Hughes, and H. E. Varmus.** 1997. *Retroviruses*. Cold Spring Harbor Laboratory Press, Plainview, N.Y.
6. **Domingo, E.** 2000. Viruses at the edge of adaptation. *Virology* **270**:251–253.
7. **Flynn, J. A., W. An, S. R. King, and A. Telesnitsky.** 2004. Nonrandom dimerization of murine leukemia virus genomic RNAs. *J. Virol.* **78**:12129–12139.
8. **Hu, W. S., E. H. Bowman, K. A. Delviks, and V. K. Pathak.** 1997. Homologous recombination occurs in a distinct retroviral subpopulation and exhibits high negative interference. *J. Virol.* **71**:6028–6036.
9. **Hu, W. S., and H. M. Temin.** 1990. Retroviral recombination and reverse transcription. *Science* **250**:1227–1233.
10. **Jetzt, A. E., H. Yu, G. J. Klarmann, Y. Ron, B. D. Preston, and J. P. Dougherty.** 2000. High rate of recombination throughout the human immunodeficiency virus type 1 genome. *J. Virol.* **74**:1234–1240.
11. **Kharytonchyk, S. A., A. I. Kireyeva, A. B. Osipovich, and I. K. Fomin.** 2005. Evidence for preferential copackaging of Moloney murine leukemia virus genomic RNAs transcribed in the same chromosomal site. *Retrovirology* **2**:3.
12. **Kulpa, D., R. Topping, and A. Telesnitsky.** 1997. Determination of the site of first strand transfer during Moloney murine leukemia virus reverse transcription and identification of strand transfer-associated reverse transcriptase errors. *EMBO J.* **16**:856–865.
13. **Levy, D. N., G. M. Aldrovandi, O. Kutsch, and G. M. Shaw.** 2004. Dynamics of HIV-1 recombination in its natural target cells. *Proc. Natl. Acad. Sci. USA* **101**:4204–4209.
14. **Lewinski, M. K., D. Bisgrove, P. Shinn, H. Chen, C. Hoffmann, S. Hannehalli, E. Verdin, C. C. Berry, J. R. Ecker, and F. D. Bushman.** 2005. Genome-wide analysis of chromosomal features repressing human immunodeficiency virus transcription. *J. Virol.* **79**:6610–6619.
15. **Malim, M. H., and M. Emerman.** 2001. HIV-1 sequence variation: drift, shift, and attenuation. *Cell* **104**:469–472.
16. **Nethe, M., B. Berkhout, and A. C. van der Kuyl.** 2005. Retroviral superinfection resistance. *Retrovirology* **2**:52.
17. **Onafuwa, A., W. An, N. D. Robson, and A. Telesnitsky.** 2003. Human immunodeficiency virus type 1 genetic recombination is more frequent than that of Moloney murine leukemia virus despite similar template switching rates. *J. Virol.* **77**:4577–4587.
18. **Panet, A., D. Baltimore, and T. Hanafusa.** 1975. Quantitation of avian RNA tumor virus reverse transcriptase by radioimmunoassay. *J. Virol.* **16**:146–152.
19. **Panet, A., and Z. Kra-Oz.** 1978. A competition immunoassay for characterizing the reverse transcriptase of mammalian RNA tumor viruses. *Virology* **89**:95–101.
20. **Panganiban, A. T., and D. Fiore.** 1988. Ordered interstrand and intrastrand DNA transfer during reverse transcription. *Science* **241**:1064–1069.
21. **Pathak, V. K., and H. M. Temin.** 1990. Broad spectrum of in vivo forward mutations, hypermutations, and mutational hotspots in a retroviral shuttle vector after a single replication cycle: substitutions, frameshifts, and hypermutations. *Proc. Natl. Acad. Sci. USA* **87**:6019–6023.
22. **Schierup, M. H., and J. Hein.** 2000. Consequences of recombination on traditional phylogenetic analysis. *Genetics* **156**:879–891.
23. **van Wamel, J. L., and B. Berkhout.** 1998. The first strand transfer during HIV-1 reverse transcription can occur either intramolecularly or intermolecularly. *Virology* **244**:245–251.
24. **Varela-Echavarria, A., N. Garvey, B. D. Preston, and J. P. Dougherty.** 1992. Comparison of Moloney murine leukemia virus mutation rate with the fidelity of its reverse transcriptase in vitro. *J. Biol. Chem.* **267**:24681–24688.
25. **Varela-Echavarria, A., C. M. Prorock, Y. Ron, and J. P. Dougherty.** 1993. High rate of genetic rearrangement during replication of a Moloney murine leukemia virus-based vector. *J. Virol.* **67**:6357–6364.
26. **Voronin, Y. A., and V. K. Pathak.** 2003. Frequent dual initiation of reverse transcription in murine leukemia virus-based vectors containing two primer-binding sites. *Virology* **312**:281–294.
27. **Voronin, Y. A., and V. K. Pathak.** 2004. Frequent dual initiation in human immunodeficiency virus-based vectors containing two primer-binding sites: a quantitative in vivo assay for function of initiation complexes. *J. Virol.* **78**:5402–5413.
28. **Yu, H., A. E. Jetzt, Y. Ron, B. D. Preston, and J. P. Dougherty.** 1998. The nature of human immunodeficiency virus type 1 strand transfers. *J. Biol. Chem.* **273**:28384–28391.
29. **Zhang, J., L. Y. Tang, T. Li, Y. Ma, and C. M. Sapp.** 2000. Most retroviral recombinations occur during minus-strand DNA synthesis. *J. Virol.* **74**:2313–2322.
30. **Zhuang, J., A. E. Jetzt, G. Sun, H. Yu, G. Klarmann, Y. Ron, B. D. Preston, and J. P. Dougherty.** 2002. Human immunodeficiency virus type 1 recombination: rate, fidelity, and putative hot spots. *J. Virol.* **76**:11273–11282.