Forms and Function of Intracellular HIV DNA

Andreas Meyerhans[1*], Tanja Breinig[1], Jean-Pierre Vartanian[2], and Simon Wain-Hobson[2]

[1]*Department of Virology, University of the Saarland, D-66421 Homburg, Germany*
[2]*Unité de Rétrovirologie Moléculaire, Institut Pasteur, F-75724 Paris cedex 15, France*
* to whom correspondence should be addressed

Abstract

Like all retroviruses HIV DNA exists in a variety of forms. When quantified, the provirus is clearly in a minority. The majority of viral DNA forms are dead-ends. *In vivo* probably upwards of 100 virions are necessary to infect a typically infected cell. Because of this, there is a strong stochastic component to determining which virions make it to the provirus. These findings help reconcile many basic features of HIV virology with general biology.

Productive HIV infection is dependent on provirus formation

Integration is an intimate part of the retrovirus life cycle. To accomplish integration, the infecting virion undergoes complex structural changes and gains access to the nucleus (recently reviewed in Greene and Peterlin, 2002). The first step of the infection process is specific binding to the surface receptor CD4 and chemokine receptors mainly CCR5 and CXCR4 on target cells. Receptor interactions trigger conformational changes in the virus envelope proteins which ultimately lead to the fusion of the lipid bilayers of the virus and the cell. Virus uncoating within the cell gives rise to the reverse transcription complex, a loose structure of viral and host proteins which is porous to dNTPs. This allows synthesis of a single DNA molecule from the two copies of plus stranded genomic RNA in the virion.

Completion of this reverse transcription step yields the so-called preintegration complex (PIC) which can be isolated from the cytoplasmic fraction of infected cells and is capable of integrating its dsDNA into target DNA *in vitro* (Miller *et al*., 1997). The PIC contains viral dsDNA, virus-derived proteins including integrase (IN), as well as host proteins *i.e.,* HMGa1 and BAF (Chen and Engelman, 1998; Farnet and Bushman, 1997; Lin and Engelman, 2003; Mansharamani *et al*., 2003; Miller *et al*., 1997). As PICs have an estimated diameter of 28 nm, too large to enter the nucleus merely by diffusion through the aqueous channel of the nuclear pore, they depend on active nuclear import (figure 1).

For oncoretroviruses it is known that PICs gain access to the host DNA when the nuclear membrane is disassembled during mitosis. Yet recent experiments have shown that HIV DNA nuclear import and integration are mitosis-independent even in proliferating cells (Katz *et al*., 2003), hence HIV transport follows another route. The signals needed for nuclear localization are found in the PIC-associated viral proteins as well as in the viral DNA in the form of a triple-helical structure called the DNA-flap (Greene and Peterlin, 2002; Zennou *et al*., 2000). Once translocated to the nucleus, the integration reaction can proceed (figure 1). For this, IN cleaves the 3' termini of the viral dsDNA to generate two nucleotide 5' overhangs at each end (Pauza, 1990). Subsequently IN triggers a transesterification reaction in which the 3' hydroxyl groups attack phosphodiester bonds on opposite strands of chromosomal DNA that are 5 bases apart (Brown, 1997; Vincent *et al*., 1990). This initial joining of the viral DNA to the host DNA results in a transient gapped structure with mismatched viral 5' ends that has to be modified by gap filling, trimming of the 5' ends and ligation to form the final integrated provirus (Brown, 1997; Hindmarsh and Leis, 1999).

HIV provirus formation is a prerequisite to efficient transcription and virus production (Englund *et al*., 1995; LaFemina *et al*., 1992; Sakai *et al*., 1993; Stevenson *et al*., 1990). In addition it ensures the stable maintenance of the retroviral genomic information within the host cell. After integrating into the host chromosome, the HIV DNA may remain silent (latent) or be actively transcribed. Both, the chro-

matin structure around the integration site and, probably more important, the metabolic state of the host cell determine the outcome (Adams *et al.*, 1994; Jordan *et al.*, 2001). As transcriptionally active chromosomal loci are favored in HIV integration, and completion of HIV reverse transcription depends on cellular activation at least for the CD4 T cells, the main targets of HIV *in vivo*, latent infection is of low frequency (Schroder *et al.*, 2002; Stevenson *et al.*, 1990; Zack *et al.*, 1992).

Non-integrated HIV DNA forms dominate

Although the linear DNA molecule is the precursor to the provirus, HIV infection of target cells generates a number of non-integrated DNA species (Barbosa *et al.*, 1994; Butler *et al.*, 2001; Kim *et al.*, 1989; Robinson and Zinkus, 1990; Vandegraaff *et al.*, 2001). For example, (i) the ends of the linear DNA may be joined to form a 2-LTR circle; (ii) homologous recombination between the two LTRs in a 2-LTR circle yields circles with a single LTR; (iii) autointegration of the viral DNA ends into a preexisting HIV DNA circles results in rearranged circular products. Depending on the site of integration and subsequent recombination a series of complex DNA structures may result (Brown, 1997; Green and Peterlin, 2002).

Non-integrated forms represent the largest fraction of HIV DNA in the nucleus. The relative abundance is non-integrated linear DNA > integrated provirus > 1-LTR circles > 2-LTR circles (Butler *et al.*, 2001; Kim *et al.*, 1989; Vandegraaff *et al.*, 2001). Kinetically they seem to appear in the same order. Three to fours hours after synchronous HIV infection, full-length viral DNA can be detected. Approximately
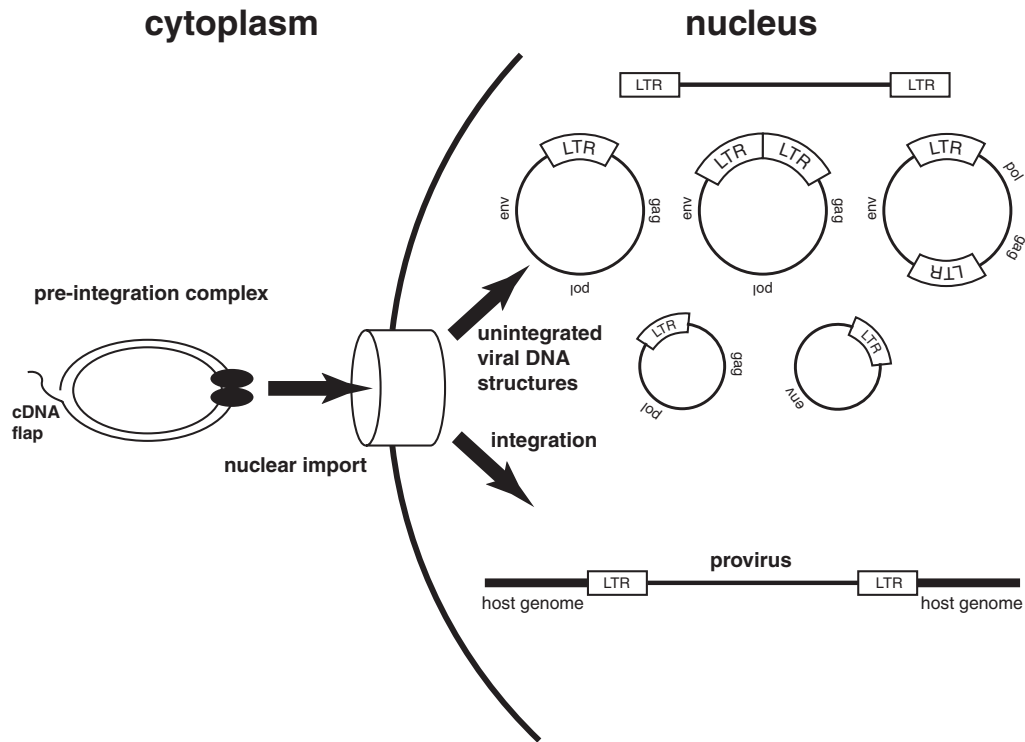

Figure 1. Schematic description of the generation of the multiple HIV DNA forms. The pre-integration complex (PIC) can be isolated from the cytoplasm of HIV-infected cells and is fully competent to integrate into dsDNA in vitro. It contains the viral dsDNA with its triple-helical region (flap), the virus-derived proteins including integrase (IN), and cellular proteins (not shown). After passing through the nuclear pore, the linear HIV DNA may be integrated into the host chromosome to generate the provirus. Alternatively, it may stay non-integrated linear or may undergo circularization to one of the various forms.

one hour thereafter, the integrated DNA form starts to accumulate but accounts for only about 10% of the total HIV DNA synthesized in a single infection round. Around 8–12 h post infection when the linear viral DNA reaches its maximum level, circular HIV DNA forms become detectable. The 1-LTR and the 2-LTR forms are minor products that may appear in the order of 10% and 1% of the total HIV DNA respectively. However, these values should be taken as rough estimates as the infection kinetics and the relative distribution of the retroviral DNA forms depend on many factors including the target cells and the virus used, as well as the particular infection conditions. One illustrative example is an early infection experiment performed with spleen necrosis virus. While chicken cells accumulate about 200 copies of non-integrated retroviral DNA shortly after infection, susceptible rat cells accumulate just 2 copies (Keshet and Temin, 1979). Likewise the presence or absence of regulatory elements such as the central polypurine tract in HIV-based vectors may decrease the kinetics and the level of provirus formation (Butler *et al*., 2001; Zennou *et al*., 2000).

As linear HIV DNA is the substrate for integration, the circular forms are side products that do not give rise to infectious progeny. They are solely located in the nucleus where they are generated by non-homologous end joining (2-LTR circles), homologous recombination (1-LTR circle) or intramolecular integration of the linear DNA form (auto-integration circles) (Brown, 1997; Kim *et al*., 1989). With the exception of the latter, their generation depends on cellular enzymes rather then the viral integrase. Since the free ends of linear viral DNA mimic double strand breaks of the chromosome and thus may provide a signal for apoptosis, circularization might be considered as a repair process to reduce such cellular danger signals (Li *et al*., 2001). A further complication is the observation of 2-LTR circles with non-complete LTRs, the region around the ligation point lacking numerous bases. The interpretation here is that the linear form of retroviral DNA underwent exonuclease digestion followed by aberrant blunt end ligation (Jurriaans *et al*., 1992).

Multiple HIV DNA forms in vivo

How are the multiple HIV DNA forms distributed in infected cells *in vivo* and what does their existence tell us about the infection process within infected individuals? The first demonstration of high levels of non-integrated HIV DNA *in vivo* was made by Pang and colleagues (Pang *et al*., 1990). By combining DNA separation via gel electrophoresis and quantitative PCR they showed (i) that non-integrated linear, non-integrated circular and integrated HIV DNA forms are present in blood and brain tissue of AIDS patients and (ii) that the ratio of non-integrated to integrated HIV may vary by 2 orders of magnitude between patients with a maximum of ~80:1. Using PCR adapted to identify and distinguish between these various forms several groups have analyzed the relative proportions of HIV DNA forms in patient material and have addressed specific questions about HIV infection *in vivo*. Most notable is the extensive study of Chun and colleagues who characterized HIV DNA forms and replication competence in activated/resting CD4 T cells from lymph nodes and blood of 14 asymptomatic HIV carriers (Chun *et al*., 1997). They estimated that around 10^9 CD4 T cells per person might be HIV-infected of which 1.2×10^7 contain latent provirus and 1.4×10^6 may produce infectious virus after appropriate T cell stimulation. Non-integrated HIV DNA was about 100 fold more frequent than the provirus form. Thus, like the situation *ex vivo*, the majority of HIV DNA *in vivo* is non-integrated.

An interesting attempt was made to use the episomal 2-LTR form of HIV DNA as an indicator of recent infection events *in vivo* (Sharkey *et al*., 2000). The underlying logic was as follows: HIV infection results in the generation of various DNA forms including episomal forms; since the virus has no mechanism to maintain such forms in infected cells, they should be labile; thus the detection of an episomal form should imply a recent infection event. With HIV episome-specific PCR that amplifies across the LTR circle junction, Sharkey and colleagues analyzed HIV carriers having undetectable HIV RNA loads due to highly active anti-retroviral therapy. Two-LTR circles were detectable in the peripheral blood mononuclear cells of 48 of the 63 patients tested. This was interpreted as a result of ongoing replication within aviremic patients, a conclusion in line with studies using RNA transcripts or HIV sequence evolution as an indicator. More recent work however questioned the lability of episomal HIV DNA (Butler *et al*., 2001; Pierson *et al*., 2002). In fact, the circular HIV DNA forms were found to be rather stable and decreased in concentration only as a function of dilution by cell division. In addition,

a cohort study of structured treatment interruption showed that the level of episomal HIV DNA does not correlate with new cycles of HIV infection *in vivo* (Fischer *et al*., 2003), again questioning the notion that episomal forms are short lived.

HIV DNA forms within single infected cells

All the above studies are basically population-based studies while the phenomena are occurring at the single cell level. By limiting dilution of peripheral blood mononuclear cells followed by virus isolation and HIV-specific PCR, the HIV provirus copy number was estimated to be close to 1 per infected cell (Brinchmann *et al*., 1991; Simmonds *et al*., 1990). As infection is occurring mainly in lymphoid organs, it is possible that the situation might be somewhat different. More recent analysis using HIV-specific fluorescence in situ hybridization (FISH) on splenocytes from splenectomized HIV patients showed that the majority of infected cells carries more than one provirus with a mean of 3–4 proviruses per infected cell (Jung *et al*., 2002). Subsequent isolation of HIV-positive nuclei by laser microdissection and transfer to PCR tubes allowed nested amplification of nuclear DNA capable of detecting ~1 copy. The segment studied were the hypervariable regions V1 and V2 of the gp120 coding sequences. This was possible without interference from the probe because the 7 kb FISH probe was deleted in *env*.

Sequence analysis of cloned DNA indicated an extraordinary genetic diversity on the single cell level—up to 29% amino acid variation in the V1V2 regions! Interestingly, the number of HIV variants per infected cell were much greater than the number of proviruses. This can readily be explained considering that the nested PCR amplified all HIV DNA forms while FISH detected the integrated forms only. An analogy with the Birnboim-Doly rapid preparation of plasmid DNA helps understand this difference—the two strands of circular DNA remain intertwined despite denaturation and hence can rapidly rehybridize on themselves when neutralized.

As (i) the number of HIV DNA molecules per nucleus (N) is finite and (ii) the sequence heterogeneity is so extensive that most sequences per nucleus are unique, extensive sequencing of cloned DNA amplified from a single nucleus should asymptote towards N. While such a study is ongoing it is already clear that N equals ~10–20 times the number of proviruses per cell (Vartanian *et al*., unpublished). As one virion gives rise to a single DNA molecule and on average a splenocyte harbours 3–4 proviruses per nucleus, a nucleus is invaded by between 30–60 PICs. Schwartz *et al* showed that the proteosome degrades up to 70% of incoming virions (Schwartz *et al*., 1998). Therefore the number of virions necessary to infect a cell *in vivo* may be in excess of (30–60)/0.3 or 100–200.

Indeed, given the capacity of an infected lymphoblast to produce hundreds of progeny as evidenced by electron microscopic studies *ex vivo* (Fais *et al*., 1995), in the dense confines of lymphoid tissue it is hard to imagine that only a single virion would infect an adjacent lymphoblast. In line with the above observations are *in vitro* infection experiments with two recombinant HIV that carry different marker genes. Both, direct and cell-mediated infection of T cells showed a preference for multi-infection providing mechanistic support for the *in vivo* observations and suggesting that individual target cells might differ in their susceptibility towards infection (Dang *et al*., 2004). Indeed, this latter notion fits well the striking diversity of proviral copy numbers in transduced CD34+ hematopoietic stem cell clones after infection with retroviral gene therapy vectors (Kustikova *et al*., 2003).

Infection of a single cell by many virions could well be a general phenomenon in virology. However, in order to quantify the number of incoming virions some trait is needed to distinguish them from *de novo* production of progeny viruses. For retroviruses the incoming genome is RNA which is reverse transcribed into DNA providing a simple difference. For the Epstein-Barr herpesvirus it has been known for awhile that cells can be multiply infected (Lindahl *et al*., 1976; Meerbach *et al*., 2001). The distinguishing trait was the circular episomal form that serves as template for linear genomes via a rolling circle mechanism. For RNA viruses, such as poliovirus, there is no such trait distinguishing parent and daughter genomes. However, recombinants are a telltale sign. Following vaccination with the three attenuated polioviruses, recombinants among them have been described, as have recombinants between vaccinating strains and wild type poliovirus and other enteroviruses (Cuervo *et al*., 2001; Guillot *et al*., 2000). Along with the HIV data, such examples indicate that multi-infection is probably commonplace, an inevitable consequence of the capacity of a cell to produce hundreds of virions in a very small space.

Ramifications of massive infection

Massive infection provides a *raison d'être* for the large burst sizes typical of productively infected cells. Secondly, they throw new light on the meaning of particle to plaque forming unit (PFU) ratios. If indeed the majority of incoming virions are degraded or sidelined in the cytoplasm and nucleus, then this will show up as a lower PFU count than expected. As the prevailing hypothesis does not incorporate massive infection of target cells, the inevitable conclusion was that there are large numbers of defective virions. The data now indicate an alternative explanation for low PFU to particle ratios. Third, if non-integrated forms dominate over proviruses than there should be numerous infected cells that harbor circular or degraded DNA and no provirus. As the non-integrated forms are, to a good first approximation, transcriptionally silent these cells should not be recognized by HIV specific cellular immunity and persist for some time becoming diluted by mitosis. Fourth, the excess of non-integrated DNA show that when simply performing PCR on total HIV DNA of any sample, the read out reflects dead-end forms far more than proviral DNA. Equating a single DNA copy to a single cell is clearly erroneous—a sample with say 600 copies per 150 000 cells for example may reflect as few as 10 cells! Fifth, as a minority of incoming viruses attains the nucleus there is a strong stochastic component to those virions that "make it". This suggests that genetic drift will be very important in the radiation into sequence space. Sixth, as the majority (~80%) of infected cells harbor more than one provirus then, while there could be transcriptional biases, there will be a great deal of heterokaryon formation and hence recombination. This is born out by analysis (Wain-Hobson *et al.*, 2003).

In conclusion, not only is the extracellular space a highly hostile environment for the virus with millions of viruses being destroyed per day (Ho *et al.*, 1995; Wei *et al.*, 1995) but also the intracellular space decimates virions. There is a striking analogy with the high carrier capacity of plants, fish and some other life forms where the majority of progeny die before reproducing themselves.

Acknowledgements

References

Adams, M., Sharmeen, L., Kimpton, J., Romeo, J. M., Garcia, J. V., Peterlin, B. M., Groudine, M., and Emerman, M. (1994). Cellular latency in human immunodeficiency virus-infected individuals with high CD4 levels can be detected by the presence of promoter-proximal transcripts. *Proc. Natl. Acad. Sci. USA* **91**, 3862–3866.

Barbosa, P., Charneau, P., Dumey, N., and Clavel, F. (1994). Kinetic analysis of HIV-1 early replicative steps in a coculture system. *AIDS Res Hum Retroviruses* **10**, 53–59.

Brinchmann, J. E., Albert, J., and Vartdal, F. (1991). Few infected CD4+ T cells but a high proportion of replication-competent provirus copies in asymptomatic human immunodeficiency virus type 1 infection. *J Virol* **65**, 2019–2023.

Brown, O. P. (1997). Integration. In Retroviruses, S. H. H. J. M. Coffin, H. E. Varmus, ed. (Cold Spring Harbour Laboratory Press), pp. 161–204.

Butler, S. L., Hansen, M. S., and Bushman, F. D. (2001). A quantitative assay for HIV DNA integration *in vivo*. *Nat Med* **7**, 631–634.

Chen, H., and Engelman, A. (1998). The barrier-to-autointegration protein is a host factor for HIV type 1 integration. *Proc. Natl. Acad. Sci. USA* **95**, 15270–15274.

Chun, T. W., Carruth, L., Finzi, D., Shen, X., DiGiuseppe, J. A., Taylor, H., Hermankova, M., Chadwick, K., Margolick, J., Quinn, T. C., et al. (1997). Quantification of latent tissue reservoirs and total body viral load in HIV-1 infection. *Nature* **387**, 183–188.

Cuervo, N. S., Guillot, S., Romanenkova, N., Combiescu, M., Aubert-Combiescu, A., Seghier, M., Caro, V., Crainic, R., and Delpeyroux, F. (2001). Genomic features of intertypic recombinant sabin poliovirus strains excreted by primary vaccinees. *J Virol* **75**, 5740–5751.

Dang, Q., Chen, J., Unutmaz, D., Coffin, J. M., Pathak, V. K., Powell, D., KewalRamani, V. N., Maldarelli, F., and Hu, W. S. (2004). Nonrandom HIV-1 infection and double infection via direct and cell-mediated pathways. *Proc. Natl. Acad. Sci. USA* **101**, 632–637.

Englund, G., Theodore, T. S., Freed, E. O., Engleman, A., and Martin, M. A. (1995). Integration is required for productive infection of monocyte-derived macrophages by human immunodeficiency virus type 1. *J Virol* **69**, 3216–3219.

Fais, S., Capobianchi, M. R., Abbate, I., Castilletti, C., Gentile, M., Cordiali Fei, P., Ameglio, F., and Dianzani, F. (1995). Unidirectional budding of HIV-1 at the site of cell-to-cell contact is associated with co-polarization of intercellular adhesion molecules and HIV-1 viral matrix protein. *AIDS* **9**, 329–335.

Farnet, C. M., and Bushman, F. D. (1997). HIV-1 cDNA integration: requirement of HMG I(Y) protein for function of preintegration complexes *in vitro*. *Cell* **88**, 483–492.

Fischer, M., Trkola, A., Joos, B., Hafner, R., Joller, H., Muesing, M. A., Kaufman, D. R., Berli, E., Hirschel, B., Weber, R., and Gunthard, H. F. (2003). Shifts in cell-associated HIV-1 RNA but not in episomal HIV-1 DNA correlate with new cycles of HIV-1 infection *in vivo*. *Antivir Ther* 8**,** 97–104.

Greene, W. C., and Peterlin, B. M. (2002). Charting HIV's remarkable voyage through the cell: Basic science as a passport to future therapy. *Nat Med* 8**,** 673–680.

Guillot, S., Caro, V., Cuervo, N., Korotkova, E., Combiescu, M., Persu, A., Aubert-Combiescu, A., Delpeyroux, F., and Crainic, R. (2000). Natural genetic exchanges between vaccine and wild poliovirus strains in humans. *J Virol* **74**, 8434–8443.

Hindmarsh, P., and Leis, J. (1999). Retroviral DNA integration. *Microbiol Mol Biol Rev* **63**, 836–843, table of contents.

Ho, D. D., Neumann, A. U., Perelson, A. S., Chen, W., Leonard, J. M., and Markowitz, M. (1995). Rapid turnover of plasma virions and CD4 lymphocytes in HIV-1 infection. *Nature* **373**, 123–126.

Jordan, A., Defechereux, P., and Verdin, E. (2001). The site of HIV-1 integration in the human genome determines basal transcriptional activity and response to Tat transactivation. *EMBO J* **20**, 1726–1738.

Jung, A., Maier, R., Vartanian, J. P., Bocharov, G., Jung, V., Fischer, U., Meese, E., Wain-Hobson, S., and Meyerhans, A. (2002). Multiply infected spleen cells in HIV patients. *Nature* **418**, 144.

Jurriaans, S., de Ronde, A., Dekker, J., Goudsmit, J., and Cornelissen, M. (1992). Analysis of human immunodeficiency virus type 1 LTR-LTR junctions in peripheral blood mononuclear cells of infected individuals. *J Gen Virol* **73** (Pt 6), 1537–1541.

Katz, R. A., Greger, J. G., Boimel, P., and Skalka, A. M. (2003). Human immunodeficiency virus type 1 DNA nuclear import and integration are mitosis independent in cycling cells. *J Virol* **77**, 13412–13417.

Keshet, E., and Temin, H. M. (1979). Cell killing by spleen necrosis virus is correlated with a transient accumulation of spleen necrosis virus DNA. *J Virol* **31**, 376–388.

Kim, S. Y., Byrn, R., Groopman, J., and Baltimore, D. (1989). Temporal aspects of DNA and RNA synthesis during human immunodeficiency virus infection: evidence for differential gene expression. *J Virol* **63**, 3708–3713.

Kustikova, O. S., Wahlers, A., Kuhlcke, K., Stahle, B., Zander, A. R., Baum, C., and Fehse, B. (2003). Dose finding with retroviral vectors: correlation of retroviral vector copy numbers in single cells with gene transfer efficiency in a cell population. *Blood* **102**, 3934–3937.

LaFemina, R. L., Schneider, C. L., Robbins, H. L., Callahan, P. L., LeGrow, K., Roth, E., Schleif, W. A., and Emini, E. A. (1992). Requirement of active human immunodeficiency virus type 1 integrase enzyme for productive infection of human T-lymphoid cells. *J Virol* **66**, 7414–7419.

Li, L., Olvera, J. M., Yoder, K. E., Mitchell, R. S., Butler, S. L., Lieber, M., Martin, S. L., and Bushman, F. D. (2001). Role of the non-homologous DNA end joining pathway in the early steps of retroviral infection. *EMBO J* **20**, 3272–3281.

Lin, C. W., and Engelman, A. (2003). The barrier-to-autointegration factor is a component of functional human immunodeficiency virus type 1 preintegration complexes. *J Virol* **77**, 5030–5036.

Lindahl, T., Adams, A., Bjursell, G., Bornkamm, G. W., Kaschka-Dierich, C., and Jehn, U. (1976). Covalently closed circular duplex DNA of Epstein-Barr virus in a human lymphoid cell line. *J Mol Biol* **102**, 511–530.

Mansharamani, M., Graham, D. R., Monie, D., Lee, K. K., Hildreth, J. E., Siliciano, R. F., and Wilson, K. L. (2003). Barrier-to-autointegration factor BAF binds p55 Gag and matrix and is a host component of human immunodeficiency virus type 1 virions. *J Virol* **77**, 13084–13092.

Meerbach, A., Gruhn, B., Egerer, R., Reischl, U., Zintl, F., and Wutzler, P. (2001). Semiquantitative PCR analysis of Epstein-Barr virus DNA in clinical samples of patients with EBV-associated diseases. *J Med Virol* **65**, 348–357.

Miller, M. D., Farnet, C. M., and Bushman, F. D. (1997). Human immunodeficiency virus type 1 preintegration complexes: studies of organization and composition. *J Virol* **71**, 5382–5390.

Pang, S., Koyanagi, Y., Miles, S., Wiley, C., Vinters, H. V., and Chen, I. S. (1990). High levels of unintegrated HIV-1 DNA in brain tissue of AIDS dementia patients. *Nature* **343**, 85–89.

Pauza, C. D. (1990). Two bases are deleted from the termini of HIV-1 linear DNA during integrative recombination. *Virology* **179**, 886–889.

Pierson, T. C., Kieffer, T. L., Ruff, C. T., Buck, C., Gange, S. J., and Siliciano, R. F. (2002). Intrinsic stability of episomal circles formed during human immunodeficiency virus type 1 replication. *J Virol* **76**, 4138–4144.

Robinson, H. L., and Zinkus, D. M. (1990). Accumulation of human immunodeficiency virus type 1 DNA in T cells: results of multiple infection events. *J Virol* **64**, 4836–4841.

Sakai, H., Kawamura, M., Sakuragi, J., Sakuragi, S., Shibata, R., Ishimoto, A., Ono, N., Ueda, S., and Adachi, A. (1993). Integration is essential for efficient gene expression of human immunodeficiency virus type 1. *J Virol* **67**, 1169–1174.

Schroder, A. R., Shinn, P., Chen, H., Berry, C., Ecker, J. R., and Bushman, F. (2002). HIV-1 integration in the human genome favors active genes and local hotspots. *Cell* **110**, 521–529.

Schwartz, O., Marechal, V., Friguet, B., Arenzana-Seisdedos, F., and Heard, J. M. (1998). Antiviral activity of the proteasome on incoming human immunodeficiency virus type 1. *J Virol* **72**, 3845–3850.

Sharkey, M. E., Teo, I., Greenough, T., Sharova, N., Luzuriaga, K., Sullivan, J. L., Bucy, R. P., Kostrikis, L. G., Haase, A., Veryard, C., et al. (2000). Persistence of episomal HIV-1 infection intermediates in patients on highly active anti-retroviral therapy. *Nat Med* **6**, 76–81.

Simmonds, P., Balfe, P., Peutherer, J. F., Ludlam, C. A., Bishop, J. O., and Brown, A. J. (1990). Human immunodeficiency virus-infected individuals contain provirus in small numbers of peripheral mononuclear cells and at low copy numbers. *J Virol* **64**, 864–872.

Stevenson, M., Stanwick, T. L., Dempsey, M. P., and Lamonica, C. A. (1990). HIV-1 replication is controlled at the level of T cell activation and proviral integration. *EMBO J* **9**, 1551–1560.

Vandegraaff, N., Kumar, R., Burrell, C. J., and Li, P. (2001). Kinetics of human immunodeficiency virus type 1 (HIV) DNA integration in acutely infected cells as determined using a novel assay for detection of integrated HIV DNA. *J Virol* **75**, 11253–11260.

Vincent, K. A., York-Higgins, D., Quiroga, M., and Brown, P. O. (1990). Host sequences flanking the HIV provirus. *Nucleic Acids Res* **18**, 6045–6047.

Wain-Hobson, S., Renoux-Elbé, C., Vartanian, J.-P., and Meyerhans, A. (2003). Network analysis of human and simian immunodeficiency virus sequence sets reveals massive recombination resulting in shorter pathways. *J Gen Virol* **84**, 885–895.

Wei, X., Ghosh, S. K., Taylor, M. E., Johnson, V. A., Emini, E. A., Deutsch, P., Lifson, J. D., Bonhoeffer, S., Nowak, M. A., Hahn, B. H., and et al. (1995). Viral dynamics in human immunodeficiency virus type 1 infection. *Nature* **373**, 117–122.

Zack, J. A., Haislip, A. M., Krogstad, P., and Chen, I. S. (1992). Incompletely reverse-transcribed human immunodeficiency virus type 1 genomes in quiescent cells can function as intermediates in the retroviral life cycle. *J Virol* **66**, 1717–1725.

Zennou, V., Petit, C., Guetard, D., Nerhbass, U., Montagnier, L., and Charneau, P. (2000). HIV-1 genome nuclear import is mediated by a central DNA flap. *Cell* **101**, 173–185.