Acta virologica **55**: 169 – 174, 2011 doi:10.4149/av_2011_02_169

# Minimal size of prototype foamy virus integrase for nuclear localization

U. HYUN, D.H. LEE, C.-G. SHIN*

Department of Biotechnology, Chung-Ang University, Ansung 456-756, South Korea

*Received September 20, 2010; accepted May 6, 2011*

**Summary. –** We have reported previously that the prototype foamy virus (PFV) integrase (IN) has a strong nuclear localization signal (NLS) in its C-terminal domain, in particular in a region of aa 306–334 including highly karyophilic arginines or lysines at positions 308, 313, 318, 324, and 329. In this study, we used various mutants of the C-terminal domain to further analyze its karyophilic determinants. Plasmids expressing these mutants fused to maltose binding protein (MBP) and enhanced green fluorescent protein (EGFP) were transfected to COS-1 cells and subcellular localization of these fluorescent fusion proteins was determined by fluorescent microscopy. The results revealed that a maximum karyophilicity was exhibited by a region longer than the previously described one of 29 aa (aa 306–334), in particular by a 64 aa region (aa 289–352) with Arg341 and Lys349 as critical determinants.

**Keywords:** integrase; karyophilic determinants; nuclear localization; prototype foamy virus

## Introduction

The family *Retroviridae* is taxonomically classified into subfamilies *Orthoretrovirinae* that consists of the six genera and *Spumaretrovirinae* that has only one genus *Spumavirus* containing foamy viruses (Fauquet *et al*., 2005). PFV previously referred to as the human foamy virus is one of the most well-known foamy viruses. PFV was initially isolated from the lymphoblastoid cells of a Kenyan patient with nasopharyngeal carcinoma (Achong *et al*., 1971). Multiple studies have since revealed that foamy viruses are different from the typical retroviruses in several ways. For example, their viral particles contain large amount of functionally relevant DNA and their life cycle includes intracellular movement of the viral particles (Rethwilm, 1996; Weiss, 1996; Linial, 1999).

Retroviral integrase is a viral enzyme that has two essential roles during the viral replication. First, it catalyzes covalent insertion of the DNA copy of viral genome into the host genomic DNA, a process called integration (Engelman *et al*., 1991; Pahl and Flugel, 1995). Second, it mediates the transport of preintegration complex (PIC) from cytoplasm to the nucleus of infected cell (Craigie, 2001). IN is a key component of PIC, which is composed of the viral DNA and proteins. The movement of PIC from cytoplasm to the nucleus is directed by NLS found in the viral proteins of PIC. The integration of human immunodeficiency virus 1 (HIV-1) DNA is influenced by NLS found in the IN (Gallay *et al*., 1995, 1996). In this process, IN directs the nuclear localization and insertion of viral DNA into the cellular DNA. HIV-1 IN has two NLS of the basic bipartite type extending from aa 186–188 and aa 211–219, respectively (Gallay *et al*., 1997). However, a further study found an additional NLS located in the central domain of HIV-1 IN at the region aa 161–173 (Bouyac-Bertoia *et al*., 2001). Avian sarcoma virus IN has been reported to have a functional NLS found between aa 206–235 (Kukolj *et al*., 1997). In addition, the karyophilic determinant of feline immunodeficiency virus IN was mapped to the highly conserved N-terminal zinc-binding HHCC motif (Woodward *et al*., 2003). However, little is known about the functional domains and NLS of PFV IN. We previously characterized the functional region and aa in PFV IN using chimeric proteins and domain-swapping

*Corresponding author. E-mail: cgshin@cau.ac.kr; fax: +82-31-675-0409.

**Abbreviations:** EGFP = enhanced green fluorescent protein; HIV-1 = human immunodeficiency virus 1; IN = integrase; MBP = maltose binding protein; NLS = nuclear localization signals; PFV = prototype foamy virus; PIC = preintegration complex; SV40 = simian virus 40

with HIV-1 IN (Lee *et al*., 2005). Using specific monoclonal antibodies, Imrich *et al*. (2000) reported that PFV IN has the karyophilic properties. Recently, we reported that the C-terminal domain of PFV IN has a strong NLS containing 5 basic aa (Arg308, Arg313, Lys318, Lys324, and Arg329) that are karyophilic (An *et al*., 2008). However, the short array of 29 aa located between aa 306-334 containing 5 basic aa is not karyophilic.

In this work, we attempted to define the minimal size and critical amino acids of NLS of PFV IN in relation to the nuclear localization of PFV IN. For this purpose, we prepared plasmids expressing various mutants of a 64 aa region (aa 289–352) of NLS fused to MBP and EGFP, transfected them to the COS-1 cells and determined subcellular localization of the expressed fluorescent fusion proteins. The obtained results showed that the analyzed 64-aa region can be regarded as a karyophilic core of PFV IN and Arg341 and Lys349 as its critical determinants.

## Materials and Methods

*Cells and transfection.* COS-1 cells were cultured in DMEM supplemented with 10% fetal bovine serum (Gibco), 100 U/ml penicillin, and 100 μg/ml streptomycin at 37°C in a humidified atmosphere containing 5% $CO_2$. Trypsinized cells $1.6 \times 10^4$ cells/ml were resuspended in 300 μl of medium and plated in an 8-well chambered coverglass. Next morning, the cells were transfected with DNA (about 1.0 μg) coding for the fusion proteins using the SuperFect reagent (Qiagen). Subcellular localization of the fusion proteins was analyzed 24 hrs post-transfection.

*Vector constructs.* Expression vectors for single aa substitution and deletion mutants were constructed by the ligation of aa 306–334, 306–344, 306–352, 306–371, 289–334, 289–344, 289–352, and 289–371 fragments of PFV IN into the HindIII/KpnI site of the pEGFP-C3 (Clontech, An *et al*., 2008). Mutant fragments were created via overlapping PCR with the mutagenic primers described in Table 1 (Lee *et al*., 2005). Briefly, PCR produced the half fragments of the 292G, 305S, 341S, and 349T mutants by using the PFIN241S forward primer with the appropriate reverse primer (PFIN292A, PFIN305A, PFIN341A, and PFIN349A, respectively). The other half fragments were amplified with the appropriate forward primer (PFIN292S, PFIN305S, PFIN341S, and PFIN349S, respectively) and the FC6A reverse primer. Next, PCRs were carried out using PFIN289S or PFIN289S292 (only for the mutant 292G) and PFIN352A, PFIN352A349 (only for the mutant C289/352-349), or FC6A, respectively, in the presence of appropriate template fragments generated from the first PCRs. The mutant 369F (C289/371-369) was directly constructed by the amplification with HFIN289S and FC6A369.

The deletion mutants in C-terminal region were constructed by using PFIN289S or PFIN306S and PFIN334A, PFIN344A, PFIN352A, or FC6A, respectively. We inserted MBP of 43 K between the EGFP and PFV IN fragment to increase the size of fusion

**Table 1. PCR primers**

| Primers | Sequences (5′–3′) | Restriction or mutagenic sites |
|---|---|---|
| MBP-3S | 5′-GCATCTCGAGATCGAAGAAGGTAAACTGGTAATCTGG-3′ | Xho I |
| MBP-367A | 5′-TTGTTAAGCTTTCCTCCAGTCTGCGCGTCTTTCAGGGCTTC-3′ | Hind III |
| PFIN241S | 5′-AAATATACTCCACATCAACTC-3′ | |
| PFIN289S | 5′-CATCCAAAGCTTGCCTCCTCTCGTTCCTGG3′ | Hind III |
| PFIN289S292 | 5′-TCCACCAAGCTTGCCTCCTCT*GGA*TCCTGGTCTC-3′ | Hind III/292G |
| PFIN306S | 5′-CAGAAGCTTGTGGCTAGGCCTGCTTCTTTG-3′ | Hind III |
| PFIN334A | 5′-AAGGGTACCCAAAATAACAACAGTCCTTGG-3′ | Kpn I |
| PFIN344A | 5′-ATTGGTACCACTTACAGTTCTGTTGTTGCC-3′ | Kpn I |
| PFIN352A | 5′-ATTGGTACCAGAAGTAGGTTTTAAATTATCTATAC-3′ | Kpn I |
| PFIN352A349 | 5′-ATTGGTACCAGAAGTAGG*TGT*TAAATTATCTATAC-3′ | Kpn I |
| FC6A | 5′-TGCGGGTACCTCATTTTTTTCCAAATGATCC-3′ | Kpn I |
| PFIN292S | 5′-CCAGCCTCCTCT*GGA*TCCTGGTCTCCTGTTGTTGGCCAATTG-3′ | 292G |
| PFIN292A | 5′-CAGGAGACCAGGATCCAGAGGAGGCTGGAGGGGTGGATGG-3′ | 2 92G |
| PFIN305S | 5′-CAATTGGTCCAGGAG*ACG*GTGGCTACGCCTGCTTCTTTG-3′ | 305T |
| PFIN305A | 5′-CAAAGAAGCAGGCGTAGCCAC*CGT*CTCCTGGACCAATTG-3′ | 305T |
| PFIN341S | 5′-CTTGGCAACAAC*AGT*ACTGTAAGTATAGATAATTTAAAACC-3′ | 341S |
| PFIN341A | 5′-CTATACTTACAGT*ACT*GTTGTTGCCAAGATGGTCCAAAATAAC-3′ | 341S |
| PFIN349S | 5′-GTATAGATAATTTA*ACA*CCTACTTCTCATCAGAATGGCACC-3′ | 349T |
| PFIN349A | 5′-GATGAGAAGTAGG*TGT*TAAATTATCTATACTTACAGTTCTGTTG-3′ | 349T |
| FC6A369 | 5′-TGCGGGTACCTCATT*GAA*TTCCAAATGATCCATTGTTGC-3′ | Kpn I/369F |

Mutagenic sequences are indicated in italics. Restriction sites are underlined.

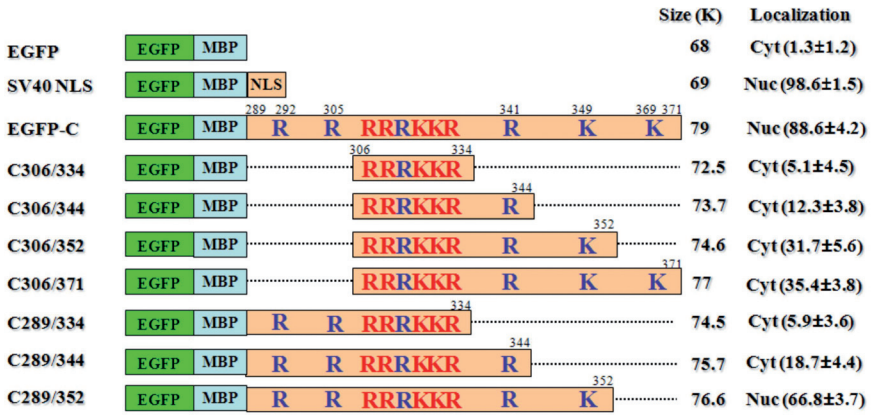


**Fig. 1**

**Karyophilic properties of deletion mutants of the C-terminal domain**

The numbers in parentheses indicate the percentage of nuclear localization efficiency. Cytoplasm (Cyt), nucleus (Nuc).

protein. The template for amplification of MBP DNA was pMal-c2 (New England Biolabs). Larger size of the fusion protein could prevent a non-specific passive diffusion across the nuclear envelope. Plasmid DNA for transfection was purified using a Jetstar plasmid purification kit (Genomed). The mutations were verified by DNA sequencing conducted by Macrogen, Korea.

*Fluorescent microscopy*. Subcellular localization of the fusion proteins in transfected cells was determined using Nikon fluorescence microscope TE 2000. Nuclear localization efficiency is defined as a ratio of the number of cells showing nuclear fluorescence to the number of cells showing nuclear or cytoplasmic fluorescence (Woodward *et al*., 2003). Transfected cells were considered as the cells exhibiting primarily cytoplasmic localization, when nuclear localization efficiency was below 50% or as the cells exhibiting nuclear localization, when its efficiency was greater than 50% (Woodward *et al*., 2003). At least 100–150 cells were measured for one representative in an independent experiment. All of the data represent the average values (the mean ± standard deviation) from three independent experiments with at least three microscopic fields counted each time.

## Results

*Deletion mutants of the C-terminal domain*

To determine a minimal karyophilic region of PFV IN, we generated seven plasmid DNA constructs that encode various deletion mutants of the PFV IN C-terminal domain fused to the C-terminus of EGFP-MBP (Fig. 1). Subcellular localization was determined by the nuclear localization efficiency expressed as a percentage of number of cells showing nuclear fluorescence to the number of cells showing nuclear or cytoplasmic fluorescence.

EGFP-MBP fusion protein of 68 K served as a control for the nuclear exclusion of fusion proteins that were incapable of nuclear import. The cells transfected with the EGFP-MBP plasmid showed the expected cytoplasmic fluorescence. In contrast, a fusion protein containing the highly karyophilic simian virus 40 (SV40) NLS localized exclusively in the nucleus and served as a positive control for the nuclear import (Fig. 1). The PFV IN C-terminal domain (83 aa region) fused to the EGFP-MBP, signed as EGFP-C had a strong NLS and localized primarily to the nucleus (Fig. 1). Six of the seven deletion mutants of the PFV IN C-terminal domain were less karyophilic and exhibited nuclear localization efficiencies of 5.1 to 35.4%. Only C289/352 deletion mutant was more karyophilic showing the nuclear localization efficiency of 66.8 ± 3.7% (Fig. 1). Therefore, we concluded that a 64 aa region (aa 289–352) of the PFV IN C-terminal domain including 10 basic aa could be the minimal region able to act as a NLS.

*Single amino acid substitution mutants of the 289/352 region*

When the minimal size of PFV IN C-terminal domain for the nuclear localization had been identified, we were interested to identify amino acids playing a critical role in the localization of IN protein to the nucleus. There are 10 basic aa (lysine or arginine) in the 289/352 region. The six basic aa – Arg308, Arg313, Arg315, Lys318, Lys324, and Arg329 – were tested in our previous study (An *et al*., 2008). In this study, four other basic aa – Arg292, Arg305, Arg341, and Lys349 – were examined for their karyophilic properties by the single aa substitution (Fig. 2). We found that the mutant fusion proteins C289/352-292 and C289/352-305

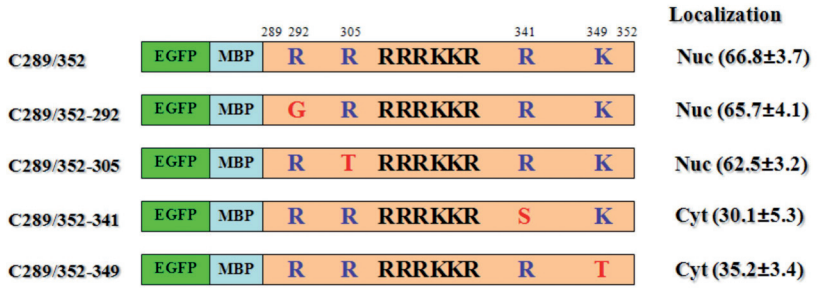


**Fig. 2**

**Karyophilic properties of single amino acid substitution mutants in the 289/352 region**

See legend in Fig. 1.

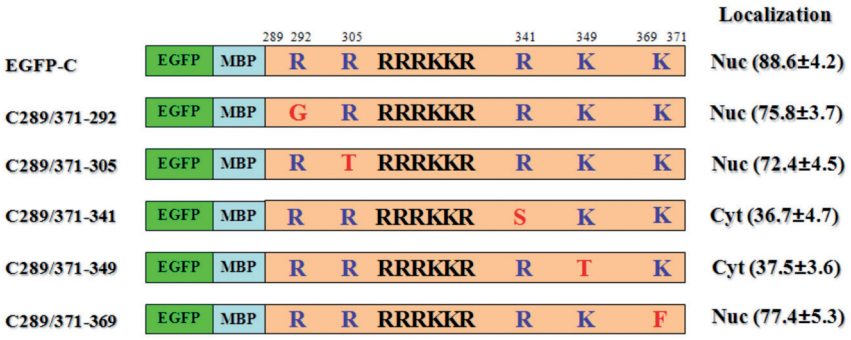


**Fig. 3**

**Karyophilic properties of single amino acid substitution mutants in the C-terminal domain**

See legend in Fig. 1.

primarily localized to the nuclei of transfected COS-1 cells with the nuclear localization efficiencies of 65.7±4.1 and 62.5±3.2%, respectively. In contrast, the mutant fusion proteins C289/352-341 and C289/352-349 primarily localized to the cytoplasm of transfected COS-1 cells with the nuclear localization efficiencies of 30.1 ± 5.3 and 35.2 ± 3.4 %, respectively (Fig. 2). These data indicated that not all of the basic aa present in the region contributed equally to the karyophilicity. In addition, the karyophilic strength of basic aa varied depending on their location in the region. Lys341 and Arg349 were more important for the karyophilicity of fusion protein than Arg292 and Arg305 (Fig. 2).

*Single amino acid substitution mutants of the C-terminal domain*

To determine whether Lys341 and Arg349 contributed more significant to the nuclear localization of fusion proteins than Arg292 and Arg305, the same mutants were constructed in the PFV IN C-terminal domain (Fig. 3). The Lys369 was also tested by mutation to the Phe369. In the transfected cells, these mutant proteins C289/371-292, C289/371-305, and C289/371-369 localized primarily to the nucleus with nuclear localization efficiencies of 75.8 ± 3.7, 72.4 ± 4.5, and 77.4 ± 5.3%, respectively (Fig. 3). In contrast, the mutant proteins C289/371-341 and C289/371-349 were mainly observed in the cytoplasm with nuclear localization efficiencies of 36.7 ± 4.7 and 37.5 ± 3.6%, respectively (Fig. 3). As expected, the same pattern was observed for both regions (C289/352 and C289/371) highlighting the importance of Arg341 and Lys349 for the nuclear localization of fusion proteins. Therefore, we found that the 64 aa region of PFV IN C-terminal domain extending from aa 289 to 352 was the minimal region essential for karyophilicity and the basic aa Arg341 and Lys349 were required for PFV IN nuclear localization activity.

## Discussion

The nuclear transport of proteins occurs through nuclear pores embedded in the nuclear membrane. Proteins smaller than 50 K are able to diffuse freely into and out of the nucleus,

whereas proteins larger than 50 K need NLS to enter the nucleus (Gorlich and Kutay, 1999). Several pathways for the nuclear import of karyophilic proteins have been described (Nakielny and Dreyfuss, 1999). To prevent the free diffusion of fusion proteins between the cytoplasm and nucleus, we fused MBP of 43 K to the truncated PFV IN and EGFP. MBP was previously shown to increase successfully the size of fusion proteins without changing the karyophilic properties of the original proteins (Fassati *et al*., 2003).

In the nuclear translocation process, cellular chaperone proteins recognize the NLS of karyophilic proteins as the signal for import into the nucleus. Although there are no consensus sequences for the nuclear localization, many karyophilic determinants have been mapped to the protein regions enriched in basic aa, what is a characteristic feature of the NLS (Dingwall and Laskey, 1991; Gorlich and Kutay, 1999). For instance, the NLS for SV40 T-antigen is PKKKRKV, for adenovirus E1A is KRPRP, and for nucleoplasmin KRPAAIKKAGQAKKKK (Silver, 1991; Boulikas, 1993; Nakielny and Dreyfuss, 1999).

One of the important functions of retroviral INs is to mediate the nuclear localization of PIC during the viral life cycle. NLSs have been described for several retroviral INs (Gallay *et al*., 1997; Kukolj *et al*., 1997; Woodward *et al*., 2003). Recently, we demonstrated that PFV IN has a strong NLS in its C-terminal domain, which contains 5 arginines and lysines between aa 308–329 (An *et al*., 2008). In the present study, we extended our finding that the short array of 29 aa region (aa 306–334) containing those 5 arginines and lysines was not karyophilic. We defined instead a 64 aa region extending from aa 289 to 352 including the short array and two basic aa in positions aa 341 and 349, respectively. This 64 aa region was highly karyophilic and mediated a nuclear localization of the fusion protein. Although there were no continuous sequences of basic aa in PFV IN, we were able to study its karyophilic properties by focusing on the arginine and lysine residues present in C-terminal domain. The especially karyophilic characteristics of the region extending from aa 289 to 352 was well preserved and we found that Arg341 and Lys349 were more important for the karyophilicity than Arg292 and Arg305.

Hence, the research on NLS of retroviral INs should focus not only on the IN, but also on the components of PIC. More research is needed to determine possible interaction PFV IN with other cellular factors and to understand the interactions between the sub-domains of PFV IN. Retroviral IN is essential for the viral replication and therefore, the thorough knowledge of IN characteristics is critical for the protection of humans from severe diseases caused by the retroviruses. The knowledge about the NLS of PFV IN elucidated in this study could contribute to the research conducted on other retroviral INs.

**Acknowledgments.** This research was supported by the grant 2009-0074910 from the National Research Foundation of Korea funded by Korean government. The authors thank Dr. Samson Chow and Dr. Cora Woodward for the generous gift of SV40 NLS plasmid and Da-Young Youk and Dog Gn An for the assistance with cell culture and DNA experiments.

## References

AAchong BG, Mansell PW, Epstein MA, Clifford P (1971): An unusual virus in cultures from a human nasopharyngeal carcinoma. J. Natl. Cancer Inst. 46, 299–307.

An DG, Hyun U, Shin C-G (2008): Characterization of nuclear localization signals of the prototype foamy virus integrase. J. Gen. Virol. 89, 1680–1684. doi:10.1099/vir.0.83689-0

Boulikas T (1993): Nuclear localization signals (NLS). Crit. Rev. Eukarytic Gene Expr. 3, 193–227.

Bouyac-Bertoia M, Dvorin JD, Fouchier RA, Jenkins Y, Meyer BE, Wu LI, Emerman M, Malim MH (2001): HIV-1 infection requires a functional integrase NLS. Mol. Cell 7, 1025–1035. doi:10.1016/S1097-2765(01)00240-4

Craigie R (2001): HIV integrase, a brief overview from chemistry to therapeutics. J. Biol. Chem. 276, 23213–23216. doi:10.1074/jbc.R100027200

Dingwall C, Laskey RA (1991): Nuclear targeting sequences a consensus? Trends Biochem. Sci. 16, 478–481. doi:10.1016/0968-0004(91)90184-W

Engelman A, Mizuuchi K, Craigie R (1991): HIV-1 DNA integration: mechanism of viral DNA cleavage and DNA strand transfer. Cell 67, 1211–1221. doi:10.1016/0092-8674(91)90297-C

Fassati A, Gorlich D, Harrison I, Zaytseva L, Mingot J-M (2003): Nuclear import of HIV-1 intracellular reverse transcription complexes is mediated by importin 7. EMBO J. 22, 3675–3685. doi:10.1093/emboj/cdg357

Fauquet CM, Mayo MA, Maniloff J, Desselberger U, Ball LA (2005): Virus taxonomy. Eighth Report of the International Committee on Taxonomy of Viruses. Elsevier-Academic Press, Amsterdam, pp 421-440.

Gallay P, Swinger S, Song J, Bushman F, Trono D (1995): HIV nuclear import is governed by the phosphotyrosine-mediated binding of matrix to the core domain of integrase. Cell 83, 569–576. doi:10.1016/0092-8674(95)90097-7

Gallay P, Stitt V, Mundy C, Oettinger M, Trono D (1996): Role of the karyopherin pathway in human immunodeficiency virus type 1 nuclear import. J. Virol. 70, 1027–1032.

Gallay P, Hope T, Chin D, Trono D (1997): HIV-1 infection of nondividing cells through the recognition of integration of integrase by the importin/karyopherin pathway. Proc. Natl. Acad. Sci. USA 94, 9825–9830. doi:10.1073/pnas.94.18.9825

Gorlich D, Kutay U (1999): Transport between the cell nucleus and the cytoplasm. Ann. Rev. Cell. Dev. Biol. 15, 607–660.

Imrich H, Heinkelin M, Herchenroder O, Rethwilm A (2000): Primate foamy virus pol protein are imported into the nucleus. J. Gen. Virol. 81, 2941–2947.

Kukolj G, Jones KS, Skalka AM (1997): Subcellular localization of avian sarcoma virus and human immunodeficiency virus type 1 integrases. J. Virol. 71, 843–847.

Lee HS, Kang SY, Shin C-G (2005): Characterization of the functional domains of human foamy virus integrase using chimeric integrases. Mol. Cells 19, 246–255.

Linial ML (1999): Foamy viruses are unconventional retroviruses. J. Virol. 73, 1747–1755.

Nakielny S, Dreyfuss G (1999): Transport of proteins and RNAs in and out of the nucleus. Cell 99, 677–690. doi:10.1016/S0092-8674(00)81666-9

Pahl A, Flugel RM (1995): Characterization of the human spuma retrovirus integrase by site-directed mutagenesis, complementation analysis, and by swapping the zinc finger domain of HIV-1. J. Biol. Chem. 270, 2957–2966. doi:10.1074/jbc.270.7.2957

Rethwilm A (1996): Unexpected replication pathways of foamy viruses. J. Acquir. Immune defic. Syndr. Hum. Retrovirol. 13 (suppl. 1), S248–S253. doi:10.1097/00042560-199600001-00037

Silver PA (1991): How proteins enter the nucleus. Cell 64, 489–497. doi:10.1016/0092-8674(91)90233-O

Weiss RA (1996): Foamy viruses bubble on. Nature 380, 201. doi:10.1038/380201a0

Woodward CL, Wang Y, Dixon WJ, Htun H, Chow SA (2003): Subcellular localization of feline immunodeficiency virus integrase and mapping of its karyophilic determinant. J. Virol, 77, 4516–4527. doi:10.1128/JVI.77.8.4516-4527.2003