# Cellular and viral determinants of retroviral nuclear entry

<table>
<thead>
<tr>
<th>Journal:</th>
<th>Canadian Journal of Microbiology</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID</td>
<td>cjm-2015-0350.R2</td>
</tr>
<tr>
<td>Manuscript Type:</td>
<td>Review</td>
</tr>
<tr>
<td>Date Submitted by the Author:</td>
<td>04-Oct-2015</td>
</tr>
<tr>
<td>Complete List of Authors:</td>
<td>Hamid, Faysal Bin; Chung-Ang University, Department of Systems Biotechnology</td>
</tr>
<tr>
<td></td>
<td>Kim, Jinsun; Chung-Ang University, Department of Systems Biotechnology</td>
</tr>
<tr>
<td></td>
<td>Shin, Cha-Gyun; Chung Ang University, Department of Systems Biotechnology</td>
</tr>
<tr>
<td>Keyword:</td>
<td>HIV-1, PIC, nuclear import, TNPO3, 2-LTR</td>
</tr>
</tbody>
</table>
Title: Cellular and viral determinants of retroviral nuclear entry

Faysal Bin Hamid¹, Jinsun Kim¹, Cha-Gyun Shin¹*

¹Department of Systems Biotechnology, Chung-Ang University, Ansung 456-756, Republic of Korea

*Correspondence:
Cha-Gyun Shin, Ph.D.
Department of Systems Biotechnology
Chung-Ang University
Ansung 456-756
Republic of Korea
Tel: +82-31-670-3067
HP: +82-10-9047-6678
Fax: +82-31-675-0409
E-mail: cgshin@cau.ac.kr
Abstract

Retroviruses must integrate their cDNA into the host genome to generate proviruses. Viral DNA-protein complexes interact with cellular proteins and produce pre-integration complexes (PICs), which carry the viral genome and cross the nuclear pore channel to enter the nucleus and integrate viral DNA into host chromosomal DNA. If the reverse transcripts fail to integrate, linear or circular DNA species such as 1- and 2-long-terminal-repeats (1-LTRs and 2-LTRs, respectively) are generated. Such complexes encounter numerous cellular proteins in the cytoplasm, which restrict viral infection and protect the nucleus. To overcome host cell defenses, the pathogens have evolved several evasion strategies. Viral proteins often contain nuclear localization signals (NLSs), allowing entry into the nucleus. Among more than 1000 proteins identified as required for HIV infection by RNA interference screening, karyopherins, cleavage and polyadenylation specific factor 6, and nucleoporins have been predominantly studied. This review discusses current opinions about the synergistic relationship between the viral and cellular factors involved in nuclear import, with focus on the unveiled mysteries of the host-pathogen interaction, and highlights novel approaches to pinpoint therapeutic targets.

Keywords

HIV-1, PIC, nuclear import, TNPO3, 2-LTR.
Introduction

Like other viruses, retroviruses are obligatory intracellular parasites. However, the DNA copies are formed by reverse transcription and integrated into host DNA, which makes them a peculiar group in the microbial world. Retroviruses belong to the Retroviridae family. To meet the requirements of viral replication, from entry to egress, they orchestrate a complex relationship with cellular factors. Several viruses such as HIV-1, influenza A, hepatitis B, and herpes simplex viruses exploit nuclear transport machinery (Cohen et al. 2011). Recent studies on retroviruses, particularly HIV-1, have provided insights on these underlying processes. HIV-1 can invade CD4+ T cells and thus abate the host immune system (Walker and McMichael 2012). Moreover, integration of the viral genome into CD4+ memory cells and macrophages establishes a latent or chronic infection. To detect invaders, cells express several pattern-recognition receptors (PRRs) and other “sensor” proteins that can detect the pathogen-associated molecular patterns (PAMPs) of viral components, which in turn activate antiviral IFN and other responses (Iwasaki and Medzhitov 2010; Melchjorsen 2013). After overcoming obstacles from the cell membrane to the nucleus (Melchjorsen 2013), viral cDNA associated pre:integration complexes (PICs) enter the nucleus. Then, integrase (IN) facilitates viral DNA integration. If integration does not take place, the PIC self:ligates to form unintegrated DNA called 1- and 2-long terminal repeats (1-LTRs and 2-LTRs, respectively). Recent large-scale siRNA and shRNA screens revealed that over 1000 host proteins participate in a broad array of functions during the HIV-1 replication cycle (Brass et al. 2008; König et al. 2008; Bushman et al. 2009; Borner et al. 2010), although the majority of these hits were probably were false.

During infection, retroviruses bind to specific receptors and coreceptors on the cell surface to penetrate into the cytoplasm and either fuse with these cell surface proteins or are endocytosed...
(Overbaugh et al. 2001) (Fig. 1). Once the virions penetrate the host cell, the poorly understood process of uncoating occurs. Uncoating or disassembly, an important early phase of HIV-1 replication, is mainly dependent on capsid stability and determines the nuclear entry of the mature PIC (Moore and Stevenson 2000; Matreyek and Engelman 2013a). The process results in the release of viral RNA from the capsid core into the cytoplasm, where the RNA is reverse transcribed within the reverse transcription complex (RTC) (Fig. 1) (Moore and Stevenson 2000). Upon uncoating, the viral capsid and nucleocapsid, along with the viral reverse transcripts, reverse transcriptase (RT), IN, protease (PR), and viral accessory proteins Vif, Nef, and Vpr, collectively called the reverse transcription complex (RTC), are released. These elements trigger the formation of the PIC, which traverses along the cytoskeleton and enters the nucleus through nuclear pores (Fig. 1). Due to the difficulty of isolating PICs from infected cells, all components of the PIC have not been identified. Additionally, the process of uncoating is still under investigation (Arhel 2010). HIV-1 PICs formed in the cytoplasm are the viral cDNA integration-competent complexes: they transit into the nucleus to form the integrated provirus by utilizing several cellular factors (Stewart 2007).

**Canonical nuclear import mechanisms**

Retroviruses generally contain two copies of positive single-stranded RNA and three structural polyproteins (Gag, Pol, Env), which are proteolytically cleaved into smaller proteins (Coffin et al. 1997). After reverse transcription, viral PICs need to be transported to the nucleus. In eukaryotic cells, the cytoplasm and nucleus are interconnected through the nuclear pore complex (NPC), which selectively filters histones, transcription factors, tRNA, rRNA, and mRNA (Stewart 2007). During interphase, the NPC strictly regulates the entry of proteins about 9 nm in diameter and 45 nm long by passive diffusion (Pante and Kann 2002). Uncoating is thought to be
maintained to limit exposure of the viral DNA to cytosolic DNA sensors. Both cytoplasmic dynein and kinesin 1 heavy chain KIF5B silencing delay uncoating, which does not affect reverse transcription but hampers the nuclear import process (Lukic et al. 2014; Fernandez et al. 2015). Interestingly, several imaging-based experiments have shown that a minor amount of CA is associated with the RTC, indicating disassembly of some viral cores in the cytoplasm (Xu et al. 2013; Campbell and Hope 2015). Uncoating may occur when the viral cores pass through the NPC (Lahaye et al. 2013; Campbell and Hope 2015). However, macromolecules (>40 kD) are transported actively by binding to several soluble transport factors called karyopherins, which are specific in their direction of movement. Karyopherin β1, also known as importin β, can recognize and bind to the nuclear localization signals (NLSs) on the cargo via its HEAT repeats (Cingolani et al. 2002). Most of the cytosolic cargoes are imported after binding to the adaptor protein importin α. Some proteins such as HIV Rev (Henderson and Percipalle 1997), Tat (Truant and Cullen 1999), and HTLV REX protein (Palmeri and Malim 1999) bind directly to only importin β. After binding, importin β triggers translocation of the trimeric complex into the nucleus through interaction with the cytoplasmic filaments of the NPC (Moroianu et al. 1996; Ball and Ullman 2005; Woodward et al. 2009). Importin β can also bind GTP-bound Ran proteins, which supply energy for transport by hydrolysis of GTP into GDP (Stewart 2007; Matreyek and Engelman 2013a). In the nucleus, Ran-GDP is dissociated from the complex and importin β and the NLS-containing protein from the adaptor molecule are imported. Afterward, the Ran-GTP concentration gradient generated between the nucleus and cytoplasm causes discharge of both subtypes of importin to the cytoplasm using several exportins, such as Crm1. However, this nuclear import mechanism is subverted by some viruses for import of their PIC,
allowing cDNA to be successfully incorporated into the host genome (Stewart 2007; Hossain et al. 2014; Ali et al. 2015).

Interestingly, production of lentiviruses in the infected cells is independent of cell division. Although HIV-1 can produce progeny in both dividing and non-dividing cells (Fig. 2A), simple viruses, including gammaviruses such as murine leukemia virus (MLV) and foamy viruses, cannot infect non-dividing and arrested cells (Patton et al. 2004; Randow et al. 2013). This limitation in infection might be because the viruses require the cell membrane to split during mitosis for incorporation into the nucleus of the daughter cells (Fig. 2B). Alternatively, viral proteins such as MLV CA may produce an unknown factor that blocks infection of non-dividing cells (Goff 2007).

HIV-1 evades detection by intrinsic immune factors in the cytoplasm such as apolipoprotein B mRNA editing enzyme, catalytic polypeptide-like 3G (APOBEC3G), tripartite motif-containing protein 5 (TRIM5α), tetherin, SAM domain and HD domain-containing protein 1 (SAMHD1), and adenosine deaminase acting on RNA 1 (ADAR1). HIV-1 enters the nucleus through the NPC (Cullen 2006; Huthoff and Towers 2008; Biswas et al. 2012; Chan et al. 2014). If the viral cDNA is arrested in the cytoplasm or is unable to undergo integration either in dividing or non-dividing cells, then different types of unintegrated DNA species (dead end products of reverse transcription) are formed, such as 1-LTR, 2-LTR, and autointegrants (Delelis et al. 2003; Sloan and Wainberg 2011). Among them, 2-LTR circles result from non-homologous end joining and are a nuclear import marker, while 1-LTR results from homologous recombination of one copy of the viral long terminal repeat. In contrast, autointegrants are produced from the processed 3′-end and any site within the reverse transcribed transcripts (Sloan and Wainberg 2011). Interestingly, MLV and PFV produce circular 2-LTRs in the cytoplasm at the early stage of
infection (Delelis et al. 2003; Serhanv et al. 2004). Moreover, 2-LTR is directly cleaved at the palindromic sequences at the LTR-LTR junction by IN to yield linear cDNA, which contributes to viral replication (Delelis et al. 2005). Using real time PCR, these different forms can be identified and their localization can be determined in “real time” by checking their presence in both cytoplasmic and nuclear fractions.

HIV and other retroviruses utilize the cellular import system for the transport of their own proteins. Different research groups have succeeded in detecting the cellular import systems by genome-wide or yeast two-hybrid screening. However, the roles of their components have not been determined by single or double mutations because viral infection is not inhibited (Cribier et al. 2011; De Iaco et al. 2013; De Houwer et al. 2014; Ali et al. 2015).

**Cellular proteins that promote viral protein import**

**Karyopherins and TNPO3.** The main function of karyopherins is to shuttle host mRNA splicing factors, particularly phosphorylated serine/arginine rich proteins (SR proteins) (Lai et al. 2001). In HIV-1 infected cells, karyopherins, especially transportin 3 (TNPO3 or TRN-SR2), act as transporters of HIV-1 proteins. Since the identification of TNPO3 as a binding partner of HIV-1 IN through Y2H and pull down assays, the role of TNPO3 in the retrovirus life cycle has been investigated (Christ et al. 2008). Interactions of TNPO3 with the HIV-1 PIC components IN (Christ et al. 2008; De Iaco and Luban, 2011; Larue et al. 2012; Ali et al. 2015) and CA (Valle-Casuso et al. 2012) were also identified by co-immunoprecipitation and pull-down assays.

Different retroviruses show different levels of TNPO3 dependency. For example, simian immunodeficiency virus (SIV), HIV-1, and HIV-2 production declines approximately 15-fold in TNPO3 knockdown (KD) HeLa cells, but FIV and MMLV do not show any sensitivity (Valle-
The infection of prototype foamy virus (PFV) is reduced by only 2-fold in TNPO3 KD baby hamster kidney (BHK-21) cells with no effect on late stages, including viral assembly and release (Ali et al. 2015), indicating that replication of PFV is less dependent on TNPO3. Consistently, beta-retroviruses, such as mouse mammary tumor virus (MMTV), can import genetic material via a TNPO3-independent pathway in non-dividing cells and can integrate randomly when using a split genome due to low viral titer (Konstantoulas and Indik 2014). Since TNPO3 is the binding partner of IN, low viral production is expected in TNPO3 depleted cells.

Due to disagreement regarding possible interaction of importin α1 with IN in vitro (Woodward et al. 2009; Ao et al. 2010), the contributions of other members of the importin family was investigated in HeLa cells, CD4+ C8166 T cells, and primary macrophages by quantitative RT-PCR, as they have 50-80% sequence similarities. The depletion of importin α3 decreases 2-LTR levels (Ao et al. 2010). Both TNPO3 and importin α1 (Rch1), acting concurrently or independently, are required for successful nuclear import of IN (Levin et al. 2010a). Another homolog, importin 7, was thought to be dispensable in HIV-1 or SIV infected macrophages (Zielske and Stevenson 2005). However, in a cell based co-immunoprecipitation assay, a significant reduction in HIV-1 production was noted in importin 7 knockout cells, indicating that the IN bipartite NLS interacts with importin 7 and not importin 8 in vivo (Ao et al. 2007). Another NLS-containing HIV-1 protein, Rev, is rich in the basic amino acid arginine (instead of lysine) and prefers to interact with transportin and importins 5, 7, and 9 in vitro independent of importin α (Arnold et al. 2006; Hutten et al. 2009). This preference may be dependent on the concentration of each import receptor or competition between import cargo. Additionally, it may be temperature dependent since efficient nuclear import of GST-Rev was observed in digitonin...
permeabilized cells in the presence of cytosol and Ran wild type (WT) protein at 18°C but not in the absence of cytosol or in the presence of a Ran mutant Q69L, a predominantly GTP-bound form (Arnold et al. 2006). Although this mutant inhibits import via importin β, importin 5, and transportin, the insensitivity to importin 7 suggests a lower affinity to GTP-bound Ran protein (Arnold et al. 2006). Surprisingly, Rev can induce the separation of preassembled Ran-GTP-importin β complexes (Henderson and Percipalle 1997; Arnold et al. 2006), based on competitive binding of Rev and importin α to importin β. Experimentally, addition of importin β binding proteins, such as c-Fos or snurportin, to the complex causes GTP hydrolysis of Ran (Arnold et al. 2006).

Currently, there is debate regarding the stage of the retroviral life cycle in which TNPO3 plays a prominent role. Some research suggests involvement before nuclear entry in the uncoating step, based on in vitro work (Shah et al. 2013), while other studies suggest action during and after the integration step (Cribier et al. 2011; Ocwieja et al. 2011; Schaller et al. 2011; Valle-Casuso et al. 2012; De Iaco et al. 2013). Additionally, Krishnan et al. (2010) claimed that the capsid, and not IN, is the genetic determinant for utilization of TNPO3 during HIV-1 infection because CA mutants with altered capsid stability (e.g., E45A, N74D) are less reliant on TNPO3 for infection. Notably, cyclophilin A (CypA) and tripartite motif 5 (TRIM5) are speculated to have the opposite effect on regulation of CA core stability. CypA destabilizes the CA lattice, accelerating reverse transcription. However, the CA A92E mutant is not sensitive to CypA and causes blockage between the reverse transcription and integration steps (Ylinen et al. 2009; Li et al. 2009). TRIM5 recognizes and destabilizes the CA core, promoting uncoating. Moreover, TNPO3 depletion can alter the sensitivity to inhibition by PF-3450074 (PF74), a small-molecule HIV-1 capsid-targeting compound that mediates capsid lattice destabilization. In the presence of PF74,
WT HIV-1 virus has relatively lower provirus production than nuclear import, while the N74D or A105T CA mutants do not (De Iaco and Luban 2011). Therefore much is still unknown about the preferred conditions, timing, and significance of TNPO3-HIV-1 capsid, IN, and potentially the binding to other proteins.

Cleavage and polyadenylation specific factor 6 (CPSF6). In eukaryotes, most mature mRNAs are generated from immature pre-mRNA through endonucleolytic splicing of introns and the addition of a poly (A) tail at the 3’ end (Gruber et al. 2012). Previously, biochemical studies revealed that at least six factors are necessary for this modification to occur efficiently in vitro, including the cleavage and polyadenylation specificity factor (CPSF), the cleavage stimulation factor (CstF), and cleavage factors I_m and II_m (CFI_m and CFII_m, respectively). CPSF, poly (A) polymerase, and the nuclear poly (A) binding protein 1 (PABPN1) are involved in polyadenylation (Cardinale et al. 2007). The 3’ end processing factor CFI_m is a heterodimer composed of a small subunit (25 kDa) and a larger subunit of 59, 68, or 72 kDa. Surprisingly, the structure of CFI_m 68 kDa (CPSF6) is reminiscent of SR proteins, a family of pre-mRNA splicing factors involved in both constitutive and alternative mRNA splicing (Rüegsegger et al. 1998). CPSF6 is comprised of a proline-rich region flanked by an N-terminal RNA recognition motif (RRM) and an RS-like C-terminal region enriched with RS/D/E dipeptides. Full-length CPSF6 is found only in the nucleus of NIH3T3 cells, while the CPSF6 C-terminal truncated protein (CPSF6-358) is localized to the cytosol, suggesting that the RS-like domain is required for its nuclear transport (Lee et al. 2010).

Price et al. (2014) suggested that CPSF6 binds to the N-terminal region of monomeric HIV-1 CA and accommodates NUP153 only when CA exists in its assembled hexameric form, resulting in both being transported into the nucleus by TNPO3 (Price et al. 2014). Thus, CPSF6 interacts
with intact cores, which are larger than the nuclear pore channel and impedes nuclear entry. However, CA N74D is insensitive to CPSF6-358 restriction by preventing interaction of CA with CPSF6-358 because N74D requires ligand-specific contact and binds to CPSF6 but not NUP153 (Price et al. 2012, 2014). The protein interface is conserved both structurally and functionally across diverse lentiviruses including HIV-2, SIVmac, and FIV. Price group proposes that CA-CPSF6 interaction occurs post cellular entry. According to previous data, the CPSF6 residues 301-327 are required for restriction in a Trim-fusion assay (fusion of CPSF6 residues 301 to 358 to rhesus TRIM5α), and conserved residues 313-327 are necessary for function (Lee et al. 2012). PF74 was originally assumed to be a potent inhibitor of CPSF6 due to structural similarities; however, now PF74 is known to interact with the N-terminal domain of CA, completely blocking the CA-CPSF6313-327 interaction (Price et al. 2012).

TNPO3 KD can cause mislocalization of CPSF6 in the cytoplasm, inhibiting HIV-1 replication (De Iaco et al. 2013). The role of cytosolic CPSF6 was investigated through different approaches, such as deletion of the CPSF6 NLS and fusion of an additional nuclear export signal. Targeting CPSF6 to the nucleus with a heterologous NLS in the deletion mutant CPSF6-358 can rescue viral infection, suggesting that cytosolic CPSF6 binds to a pocket in CA where N74 and A105 are located and causes delayed disassembly, resulting in stabilization of the capsid cores (De Iaco et al. 2013). The suitable physiological conditions for cytoplasmic CPSF6 accumulation remain unknown but may be specific to cell type or stimulus and thus CPSF6 may be required for disassembly only in certain conditions (De Iaco et al. 2013).

Although binding efficiency of the D750R/D751R TNPO3 mutant to IN and ASF/SF2 was similar to that seen with the WT, HIV-1 production and sensitivity to CPSF6 were impaired (Maertens et al. 2014). In other experiments, the mutant exhibited slightly reduced affinity for IN.
but strong impairment of binding to unphosphorylated ASF/SF2, indicating its lower cargo binding ability (De Houwer et al. 2014).

**Nucleoporins (NUPs).** Several genome-wide RNA interference screening experiments identified numerous NUPs as potential nuclear transport factors for the *Retroviridae* family, including NUP358/RAN BP2, NUP153, NUP98, and NUP214/CAN (NUP 50, 62, 85, 107, 133, 155,160, 210, ELYS, and TRP) (Ebina et al. 2004; Brass et al. 2008; König et al. 2008; Woodward et al. 2009; Lee et al. 2010; Zhang et al. 2010; Matreyek and Engelman 2011, 2013a; Di Nunzio et al. 2013; Matreyek et al. 2013b). Among them, NUP358 and NUP214 are located exclusively at the cytoplasmic face. NUP153 is found at the nuclear side of the NPC, whereas NUP98 is found on both sides (Walther et al. 2002; Liang and Hetzer 2011). The vertebrate NPC is a macromolecular structure (about 120 MDa) comprising 30 different NUPs. The transport receptor-cargo complex docks on NUPs, allowing passage across the NPC. Hepatitis B virus, herpes simplex virus, influenza virus, and adenovirus also use NUPs for nuclear entry (Trotman et al. 2001; König et al. 2008, 2010; Copeland et al. 2009; Schmitz et al. 2010), whereas poliovirus and cardiovirus induce modifications of NUPs (Gustin and Sarnow 2001; Porter and Palmenberg 2009). Previous LC-MS/MS studies revealed that KTN1, NUP43, NUP358, NUP45, and NUP54 expression is downregulated on HIV-1 infection while NUP358, NUP98, and TPR expression is upregulated (Chan et al. 2007; Monette et al. 2011).

Furthermore, TNPO3 and NUP358/RanBP2 are involved in site selection in chromatin during integration (Ocwieja et al. 2011; Schaller et al. 2011). The absence of either protein alters the preference for integration to regions with lower gene density. Recently, the observed decrease in HIV-1 infectivity and reduced 2-LTR and provirus formation without significant change in early reverse transcribed transcripts in NUP358-depleted cells suggests that Nup358 likely functions
prior to nuclear import but after reverse transcription (Schaller et al. 2011). Previously HIV-1 CA, but not SIVmac CA, was shown to compete with the CypA domain of NUP358 for binding to CypA. The interaction between NUP358 and the preassembled CA-NC complex in vitro serves as a docking station for the HIV-1 core (Di Nunzio et al. 2012). Several CA mutants, N57A and N74D, or CypA binding mutants, G89V and P90A, show no or less sensitivity to NUP358 depletion, suggesting that they can be transported via a NUP358/TNPO3-independent route. However, they showed integration site propensity according to the density of transcription units and relevant factors. N57A and N74D integrate into areas with a lower density of transcription units, while G89V and P90A mutants exhibit selectivity to regions with a higher density of transcription units (Schaller et al. 2011).

NUP153 is a flexible structure in which the N-terminal region is required for targeting the cargo to the NPC, and the C-terminal FxFG region is important for binding to importin α/β-bound proteins before release by Ran-GTP binding (Moroianu et al. 1996; Ball and Ullman 2005; Woodward et al. 2009). In addition, the nuclear basket of the NPC is mainly composed of Tpr, which binds to NUP153 at its N-terminal region. HIV-1 replication is significantly reduced in the absence of either Tpr or Nup153 in the cells (Lelek et al. 2015). As expected, NUP153 is required for HIV-1 nuclear import. Tpr is not involved in reverse transcription, nuclear import, or integration, but super-resolution microscopy images show that it is essential for interaction with active genes. Depletion or overexpression of Tpr causes remodeling of chromatin proximal to the NPC by stabilization of LEDGF and changes the integration site preference (Lelek et al. 2015). Nup153 binds directly with HIV-1 IN, CA, and Vpr in an importin α/β independent manner for import (Varadarajan et al. 2005; Woodward et al. 2009; Di Nunzio et al. 2013), but CA import may be Nup153/TNPO3 dependent (Paulillo et al. 2005; Cardarelli et al. 2012;
Matreyek et al. 2013b). PF74 and/or CPSF6 compete with NUP153 for binding to the HIV-1 CA pocket, thus higher concentrations of PF74 can block HIV-1 infection. However, the dependency of different viruses is not the same. For example, NUP153 KD significantly inhibits RSV (53.7%) and MPMV (38.7%) infection but has little to no effect on MLV, FIV, and HIV-1 (<4%). Compared with the 3-5 fold defect in infection observed in either TNPO3 or NUP153 KD, double knockdown decreases infection of VSV-G WT virus by 40-fold (Matreyek and Engelman 2011).

Interestingly, NUP62 also interacts with HIV-1 IN and binds to chromatin both in the nuclear and chromatin-bound extract of 293T cells and T cells (Ao et al. 2012). In pull-down assays, NUP62 binds only with IN and not with MA or Vpr. Although C-terminal residues 358 to 522 compete with NUP62 for binding to HIV-1 IN, reduce IN-chromatin complex formation, and decrease HIV-1 infection, the coiled-coil domain of the C-terminal domain of NUP62 assists NUP62-IN association (Ao et al. 2012). NUP62 also binds TRN-SR2, which might provide a waiting spot before transport for host protein-bound cargo complexes. In normal cells, NUP62 resides only in the nuclear envelope. During HIV-1 infection, viral Gag, RNA, and Rev resides with NUP62, which mediates shifting to the cytoplasm and plasma membrane, suggesting that NUP62 is likely rearranged as its binding partner, the viral RNA-Gag complex, is exported from the nucleus (Monette et al. 2011).

Depletion of NUP68 also impairs the infection capacity of HIV-1 and MLV, but inhibition of the latter may be due to dependence on cell cycle progression as it enters the nucleus during the metaphase step (Di Nunzio et al. 2012).
Despite the postulation that NUP98 is not involved in nuclear import and weakly in HIV-1 integration (König et al. 2008), the efficient import of HIV-1 cDNA is blocked in NUP98-depleted cells, when siRNA is used and when vesicular stomatitis virus matrix protein (VSV M) is used as an inhibitor protein (Ebina et al. 2004). The VSV M interacts with the phenylalanine-glycine (FG) repeat region of NUP98. Over-expression of NUP98 can rescue the inhibitory effects on infection and decrease 2-LTR formation. Concurrently, the import system is mediated by the importin α/β transport pathway (Ebina et al. 2004).

NUP98, NUP358, and other NUPs such as NUP214/CAN and NUP88 are involved in the export of RNA and the leucine-rich nuclear export signal (NES)-containing viral protein Rev from the nucleus, which involves the Crm1 export protein (Zolotukhin and Felber 1999; Hutten and Kehlenbach 2006). Interestingly, NUP214 assists the import of adenovirus DNA through the NPC (Trotman et al. 2001), but no such evidence exists for HIV-1. Therefore, more extensive studies are required to elucidate the interaction of NUPs and their exact role in the import of HIV-1 proteins.

**Heat-shock protein 70 (Hsp70).** Hsp70 is a member of the cellular chaperone family of proteins present in both the cytoplasm and nucleus. Heat shock proteins are expressed in response to a variety of physiological and environmental insults (Daugaard et al. 2007; Multhoff 2007). There is evidence that Hsp70 promotes the import of HIV-1 PIC into the nucleus (Agostini et al. 2000) in a manner similar to that for Vpr, at least in macrophages (Kogan and Rappaport 2011). Hsp70 is required for the transport of karyophilic proteins into the nucleus (Okuno et al. 1993). As expected, their absence from the cytosolic extract retards protein transport, while induction of Hsp70 restores protein transport, indicating the relatedness of Hsp70 with the karyopherin protein family and nuclear import. Previously, in vitro binding assays revealed that the C-
terminal of karyopherin α is required for transcription factor STAT1 and basic type NLS-containing protein binding, while the N-terminal domain is a “hotspot” for Hsp60 and Vpr binding (Agostini et al. 2000). Furthermore, as Hsp70 cannot bind the HIV-1 MA, which has affinity to karyopherin α, the interaction of Hsp70 with karyopherin α likely mediates the interaction between karyopherin α and HIV-1 MA, resulting in the formation of a heterotrimeric complex (Agostini et al. 2000).

**Lens epithelium-derived growth factor (LEDGF).** LEDGF/p75, a ubiquitous nuclear protein, binds ectopic HIV-1 IN in 293T cells (Cherepanov et al. 2003; Liano et al. 2004a, 2004b). A PWWP domain in the N-terminus binds chromatin, acting as a tethering agent of IN to chromatin (Hombrouck et al. 2007). LEDGF/p75 contains a functional SV40-like NLS in its N-terminal domain (148GRKRKAEEKQ156) and is transported to the nucleus in an importin α/β- and temperature-dependent manner (Maertens et al. 2004). In a transient knockout via siRNA, LEDGF was found to be essential and sufficient for the nuclear transport of HIV-1 IN. Unlike LEDGF, its alternative splice variant, p52, neither binds with nor imports IN. Moreover, an NLS-defective mutation in at least one amino acid (Lys150) position causes abrogation of HIV-1 IN nuclear import (Maertens et al. 2014). Consistently, stable knockout by shRNA perturbs the interaction between HIV-1 IN and LEDGF, thereby preventing nuclear entry, compatible with its roles in tethering IN and chromatin. Comparable levels of 2-LTR observed in both non-dividing and wild type cells indicate that LEDGF can bind tightly to multiple lentiviral INs and tether them to chromatin in an NLS-independent manner. However, LEDGF might be involved in nuclear import in either an NLS-dependent or independent manner. More intense studies are needed to confirm its genuine role in the lentiviral life cycle (Llano et al. 2004a; Levin et al. 2010a, 2011).
To clarify retroviral nuclear shuttling, further understanding is needed regarding the role of myxovirus resistance 2 (MX2, also known as MXB), barrier-to-autointegration factor (BAF), and other proteins, especially transcription related proteins and RNA binding proteins, in the host cell. BAF is a 10 kDa protein ubiquitously found in eukaryotic cells (Segura-Totten and Wilson 2004). BAF binds directly to dsDNA, lamin A (essential in nuclear structure and chromatin arrangement), transcriptional activators, and retroviral PICs. BAF also binds retroviral MA to become a component of the HIV-1 PIC. BAF dimers have two binding sites for dsDNA and likely stop retroviral DNA from autointegration via IN (Lee and Craigie 1998). As a component of PICs, BAF can bind to DNA and inhibit autointegration after reverse transcription but before and during the integration step (Segura-Totten and Wilson 2004).

Like BAF, another nuclear inner membrane protein, emerin (EMD), is necessary for HIV-1 infection in non-dividing cells (Jacque and Stevenson 2006). Silencing of EMD in macrophages produces abortive integration into the host chromatin, although nuclear entry of PIC and infection are not interrupted. However, the opposite result is not seen in dividing cells, suggesting that at the end of mitosis, the nuclear envelope is regenerated and EMD may induce chromatin engagement because the binding of HIV-1 RT copies to chromatin is a prerequisite for entry into the nucleus. Consistent with these data, restriction of MLV infection relies on nuclear protein lamina-associated polypeptide 2 (LAP2) but not EMD. Based on the dependency of HIV-1 on both EMD and lamina-associated polypeptide α (LAPα) for its infection, the functions of these proteins are likely redundant because the malfunction of either can be rescued by the other. Intriguingly, lentiviruses and oncoretroviruses may use LAPα throughout their life cycle, but EMD is required for lentiviruses (Jacque and Stevenson 2006). Moreover, upon viral infection, EMD is phosphorylated, which is dependent on both mitogen-activated protein kinase (MAPK)
and MEK1 kinase (Bukong et al. 2010). Particularly, MAPK binds with the virion and induces Gag-MA phosphorylation (Cartier et al. 1997; Bukong et al. 2010). Inhibition of these two proteins in macrophages causes a reduction in EMD phosphorylation and viral integration, affecting the viral infection (Bukong et al. 2010). Therefore, these proteins are likely directly or indirectly related to each other and required for HIV-1 PIC import, at least in a cell-specific manner.

Import-mediating viral components

**Gag.** In lentiviruses, Gag is one of the essential polyproteins. Gag is cleaved to generate four proteins: MA, CA, NC, and p6 (Coffin et al. 1997). There is debate whether CA is present in the viral PIC or not. Most recently, compartmentalization of CA in the nucleus of infected cells was determined by direct interaction with the FG-repeat enriched NUP153 (Matreyek et al. 2013b). CA binds with several cellular proteins, including TNPO3, NUP358, NUP153, NUP158, TRIM5α, CPSF6, and Cyp A, all of which are involved in nuclear transport machineries (Bieniasz 2012; Matreyek et al. 2013b). Because CA does not possess its own NLS, it may provide anchoring sites or pockets for PIC proteins, allowing transport of PIC into the nucleus. In addition, study of the chimeric HIV-1/MLV viruses and some CA mutants highlight the role of CA in nuclear import. Replacement of HIV-1 CA and MA with the same proteins from MLV decreases HIV-1 infectivity through accumulation of 2-LTR in non-dividing cells (Yamashita and Emerman 2004). Several HIV-1 CA mutants and MLV CA generate larger PICs than WT HIV-1, thereby slowing disassembly and disabling transport of the PIC through the NPC (Dismuke and Aiken 2006; Fassati 2006; Matreyek and Engelman 2013a). Therefore, CA may
change the accessibility of PIC to the host or viral proteins that determine nuclear entry, ultimately determining the fate of nuclear import in non-dividing cells.

Recently yeast two-hybrid screening identified putative CA binding proteins (MAP1A, MAP1S, CKAP1, and WIRE), which are known cytoskeletal components. Depletion of MAP1A/MAP1S leads to lower viral infectivity because of defective nuclear entry (Fernandez et al. 2015). HIV-1 CA interacts with human MAP1A and MAP1S, which are microtubule-associated proteins that presumably mediate tethering of HIV-1 CA to cytoskeletons. Otherwise, a large amount of capsids accumulate throughout the cytoplasm (Fernandez et al. 2015).

Most recently, the crystal structure of the novel HIV-1 inhibitor BI-2 revealed that, similar to the small molecule inhibitor PF74, BI-2 also binds in the site 2 pocket of the N-terminus of CA. Although BI-2 is required at higher concentrations (about 10-fold) than PF74 to inhibit HIV-1 infection, it is not toxic to host cells (Fricke et al. 2014). PF74 prevents the binding of CPSF6 to the same pocket in CA, thus destabilizing the HIV-1 core (Fricke et al. 2013, 2014). Additionally, a novel assay for measuring HIV-1 CA-NC complex stability showed that the cytosolic extracts, especially CPSF6, can stabilize the CA-NC complex like the drugs PF74 and BI-2, while the peptide inhibitor CAI, an inhibitor of HIV-1 assembly in vitro (Sticht et al. 2005), is similar to CypA in its ability to destabilize the complex (Fig. 3). Interestingly, PF74 decreases both 1-LTR and 2-LRT formation, whereas BI-2 interferes with 2-LTR formation but has no effect on late reverse transcription. These data suggest that PF74 functions during the reverse transcription step, while BI-2 works after reverse transcription but prior to nuclear entry (Fricke et al. 2014) (Fig. 3).
There are two NLSs in HIV-1 MA: NLS-1 (25GKKKYKLLKH33) and NLS-2 (110KSKKK) (Fouchier et al. 1997; Reil et al. 1998). MA mediates nuclear import through interaction with karyopherin α. Truncated or full deletion mutants, and even those with point mutations at different positions (9, 26, 27, 67, 72, or 77), display moderate interference with viral infection in both dividing and non-dividing cells, indicating that MA is dispensable in nuclear entry (Bukrinsky 2004; Kaushik and Ratner 2004; Fassati 2006). Interestingly, phosphorylation of Tyr-132 and Ser-111 in MA by the ERK family member ERK2 is critical for HIV-1 infection (Suzuki and Craigie 2007) and nuclear localization of the PIC, which was confirmed through the use of Ser-Thr phosphatase inhibitors (Kaushik and Ratner 2004, Giroud et al. 2011). Rather, MA is involved in Crm1-dependent nuclear export due to the presence of an NES (Baluyot et al. 2012). The coexistence of an NLS and NES in MA is puzzling, raising questions about its major function. Comparison to other cellular proteins with the same properties, such as STAT1 and ERK5, provided the theoretical insight that the putative NLSs of MA may undergo unknown biophysical or biochemical modification upon phosphorylation and display a conditional NES, which uses CRM1 as a carrier protein. MA phosphorylation likely unmask or activates the NLS, stimulating the directional movement of the PIC to the nucleus (Nardozzi et al. 2010). This is likely only one of several factors relevant to this import process.

**Integrase.** HIV-1 IN, a 32 kDa protein processed from Pol polyprotein, is composed of three functional domains: the N-terminal domain (NTD)/Zn-binding motif (residues 1-50), the central core domain (CCD) (residues 51-212), and the C-terminal domain (CTD)/DNA-binding motif (residues 213-288). IN integrates viral cDNA into the host genome with the help of diverse cellular proteins. Cellular factors such as TNPO3 (Larue et al. 2012), barrier-to-autointegration factor (Lin and Engelman 2003), LEDGF/p75 (Cherepanov et al. 2003), high-mobility group
protein (HMG I(Y)) (Llano et al. 2004b), IN interactor protein-1 (Yung et al. 2001), DNA-dependent protein kinase (Daniel et al. 1999), and DNA repair protein hRad18 (Mulder et al. 2002) all bind IN (Llano et al. 2004b). Lentiviral HIV-1 and FIV INs depend on LEDGF for nuclear localization and protection from ubiquitin-proteasome, while the IN of the oncoretrovirus MoMLV does not, although all of them can be co-immunoprecipitated with LEDGF (Llano et al. 2004a). Although HIV-1 IN contains a basic bipartite-type NLS at residues 186-188 and 211-219 and an atypical NLS in the CCD at 161-173 (Tsurutani et al. 2000; Bouyac-Bertoia et al. 2001; Ao et al. 2004), its dispensable role in nuclear trafficking might be attributed (Petit et al. 2000). Similarly, NLS-properties were also found in the arginine-and lysine-rich C-terminus of foamy virus IN (Hossain et al. 2014). Nuclear localization of over-expressed full-length IN and disability in 2-LTR production by IN-CTD deletion mutants provide strong evidence of its karyophilic role. Recent studies show that the IN CTD interacts with both importin α3 and importin 7, as well as with TNPO3, for passage through the NPC (Levin et al. 2010a). The IN CTD-importin 7 interaction mediated import is specific for digitonin-permeabilized cell systems. A reduction in HIV-1 infectivity was observed in HeLa and glioma cells (Fassati et al. 2003; Bukrinsky 2004; Ao et al. 2004). While slipping through the NPC, IN can also bind with Nup153 (Woodward et al. 2009) and Nup62 (Ao et al. 2012). Although KD of Nup153 interferes with nuclear translocation of viral cDNA, Nup62 KD impairs only cDNA integration and not reverse transcription or the nuclear import step (Ao et al. 2012). The molecular mechanism of such biased behavior is not yet understood.

Rev. Regulator of expression of the virion (Rev), one of the HIV-1 regulatory proteins, harbors both an NLS and NES and contributes to the export of Rev responsive element (RRE)-containing spliced or unspliced mRNA (Cullen 2003; Strebel 2003; Zhang et al. 2010). Interestingly, Rev
can bind with IN and, as proposed in a recent model, might balance the movement of IN between
the cytoplasm and nucleus (Levin et al. 2010b). Over-expression of Rev or Rev-IN complexes
inhibits nuclear translocation of IN, while deletion mutants of Rev hamper nuclear import of IN
(Levin et al. 2010b), thereby suggesting the direct involvement of Rev in IN movement.
Additionally, Rev relies on NUP BP2 for its own nuclear entry (Zhang et al. 2010).

Vpr. Viral protein R, or Vpr, has many functions: it regulates apoptosis, arrests the cell cycle at
the G2 stage by binding to DDB1 and Cul4-associated factor 1 (DCAF1), and RNA splicing
(Nitahara-Kasahara et al. 2007; Kogan and Rappaport 2011). Despite the absence of a functional
NLS, it exhibits karyophilic properties during PIC nuclear transport. Primarily, it interacts with
three forms of importin α (α1, α3, α5) and with nuclear pore proteins in non-dividing cells in an
importin β-independent manner. Vpr is insensitive to high concentrations of ATP or lower NTP
levels, suggesting its energy-independent movement. NMR studies reveal that it is comprised of
three helices, α-H1 (13-33), α-H2 (38-50), and α-H3 (55-77), which have different functions.
However, another range of residues, 17-34 and 46-74, shows independent nuclear localization,
which can be blocked by any mutation among these regions (Kamata and Aida 2000). Vpr
oligomerization determines the sub-cellular localization of Vpr but not its pro-apoptotic activity
(Fritz et al. 2008). Vpr requires interaction with the C-terminus of p55 Gag polyprotein precursor
(Pr55gag) to anchor with assembling virion particles (Bachand et al. 1999). However, the PIC is
ubiquitously found to be associated with proteins of the microtubular network, such as actin and
dynein, during cytoplasmic migration. A Vpr-GFP fusion construct was used to track the role of
Vpr in PIC movement in living cells, revealing movement throughout the cytoplasm along
cytoskeletal filaments and accumulation in the perinuclear region (McDonald et al. 2002).
Although Vpr is an integral component of the PIC, it is unclear whether Vpr assists in PIC
movement. Vpr can associate with Nup153 (Woodward et al. 2009), Pom121 (Matreyek and Engelman 2013a), and human nucleoporin CG1 (Le Rouzic et al. 2002; Matreyek and Engelman 2013a) both in vitro and in transfected cells. Using a photobleaching experiment in living cells, the nucleocytoplasmic shuttling of Vpr, which prominently occurs in the nuclear envelope, was confirmed (Le Rouzic et al. 2002). Therefore, this protein may influence cleavage of the nuclear envelope, facilitating the travel of PICs through the NPC, which might explain how Vpr induces apoptosis.

Central DNA flap. During reverse transcription, binding of tRNA to the primer binding site (PBS) initiates formation of the minus-strand strong-stop DNA and subsequent elongation. The central DNA flap is generated as a result of (+) strand initiation at the central polypurine tract (cPPT) and termination after about 100 bp strand displacement at the central termination sequence (CTS) (Iglesias et al. 2011). In lentiviruses, positive-strand strong-stop DNA initiation at two distinct sites leads to displacement of the downstream strand containing about 99 nucleotides of overlap until the CTS is reached. This generates a single-stranded overlapping DNA fragment called the central DNA flap, which acts as a cis-determinant of HIV-1 DNA nuclear import (Sirven et al. 2000). Although the genome of the reverse transcription complex (RTC) can consist of either RNA or RNA-DNA intermediates, PIC does not contain RNA species (Suzuki and Craigie 2007). Instead, along with viral karyophilic proteins and several cellular proteins, the overlapped portion of viral cDNA causes the nuclear import of PIC and integration into cellular DNA. When a sequence of the DNA flap is introduced in an HIV-derived vector in human CD34+ hematopoietic stem cells, there is increase in gene transduction ability (Sirven et al. 2000). Mutations in either the HIV-1 cPPT or C-terminal sequence can impair viral replication in MT4 cells, CEM cells, and PBMCs (Zennou et al. 2000; De Rijck et
al. 2005). Insufficient ability of flap negative HIV-1 to infect HeLa cells could be due to delayed 2-LTR formation. However, the role of the flap in nuclear import is contested (Limón et al. 2002; Dvorin et al. 2002). The observation that HIV-1 replication in a single-cycle replication system is defective but increases 5-7 fold in dividing and non-dividing cell lines indicates its specificity for strain and host cell types (Follenzi et al. 2000). Fluorescence in situ hybridization (FISH) analysis of the sub-cellular distribution of viral cDNA between wild type and DNA flap-defective mutant demonstrated that wild-type DNA including the DNA Flap accumulates predominantly inside the nucleus of transduced cells 48 hr post-transduction, and more specifically within open regions of the chromatin. However DNA flap defective viral genomes accumulate on the cytoplasmic side of the nuclear membrane (Arhel et al. 2006). Flap may have a protective role in viral cDNA stability, based on the observation that PIC collected at 10 hours, but not that from earlier samples, is resistant to DNase I digestion (Ao et al. 2004). However, both cPPT and CTS are essential for a functional central DNA flap, as viral infectivity and gene transfer are impaired in mutants. Most disrupted DNA flap mutants can maintain residual nuclear import, which accounts for 5-15% of the WT HIV-1 (Iglesias et al. 2011). Although the entry of residual amounts of PIC in the absence of the central DNA flap occurs with very low efficiency, reinsertion of cPPT or CTS sequences recovers the viral infectivity and gene transfer efficacy by increasing nuclear import (Follenzi et al. 2000; Zennou et al. 2000; Iglesias et al. 2011).

**Conclusions and future perspectives**

More than 30 years after the discovery of HIV-1, researchers have achieved significant successes in the control and treatment of HIV-1 viremia but are still a long way from finding a suitable treatment leading to cure. AIDS drug design and selection are limited due to side effects occurring because of utilization of cellular proteins by viral proteins. Interestingly, viral particle
and host protein interactions have been targeted for new therapeutic targets over the last two decades. This review summarizes the progress in the understanding of nuclear import from both cellular and viral aspects.

The nucleus is a selective permeable compartment and prevents free diffusion of random proteins. Retroviruses have evolved numerous mechanisms to evade the tight cellular regulatory systems, even circumventing several host proteins, to reach the nucleus.

Nuclear import is a critical step in the retroviral life cycle and consists of flexible utilization of cellular cofactors to access the nucleus (Lee et al. 2010). The viral capsid determines the requirement of TNPO3 during HIV-1 infection. Full length CPSF6 binds to the HIV-1 capsid by interacting with a distinct conserved interface, which dictates the dependence on TNPO3 and Nup358. Moreover, when CPSF6 is confined to the cytoplasm, the stability of the incoming viral core increases, similar to CypA (Li et al., 2009; De Iaco et al., 2013; Fricke et al., 2013; Shah et al., 2013). Current research indicates an orchestrated protein-protein network is required for retroviral nuclear transport. Since different cell lines have distinctive cell signaling pathways and cellular status, these parameters may influence the nuclear import of retroviral PICs (Levin et al. 2011).

The current research on HIV-1-related nucleocytoplasmic transport has been confined to nuclear localization signal (NLS)-containing proteins and their carriers (Stewart 2007). The search for NLS-independent pathways and more careful study of the unveiled NLS-exhibiting sequences in viral proteins are important areas for future research. Additionally, the simultaneous and flexible use of cellular cofactors by HIV-1 proteins is another important research topic. Research regarding debated results and models needs to be carried out as it is important that we do not
neglect controversial results. Some of these contested issues may be due to use of different cell lines (e.g., immature/adult, native/specialized cell lines) under different conditions (temperature, reagents, incubation time), or simply due to handling. Although there are some debated topics, the viral proteins for egress are known and agreed upon, such as importin β (Henderson and Percipalle 1997), NUP98 (Zolotukhin and Felber 1999), NUP214 (Zolotukhin and Felber 1999; Hutten and Kehlenbach 2006), and Hsp70 (Multhoff 2007). In this review, we compiled the data available in current literature on host-pathogen interaction in HIV-1 nuclear transport. This review will be helpful in determining new therapeutic targets.

Acknowledgements

This study was supported by a grant from the National Research Foundation of Korea (NRF), funded by the Korean government (2011-0011256).

Competing interests

The authors declare that they have no competing interests.

References


De Iaco, A., and Luban, J. 2011. Inhibition of HIV-1 infection by TNPO3 depletion is determined by capsid and detectable after viral cDNA enters the nucleus. Retrovirology, **8**:98. doi:10.1186/1742-4690-8-98. PMID:22145813.


Figure legends

Fig. 1. Model for early events of the retroviral life cycle. Early events of HIV-1 replication include fusion to the CD4 receptor on the cell membrane and entry. Subsequently, capsid removal or uncoating occurs, followed by reverse transcription of the viral RNA to cDNA or reverse transcripts, and production of the reverse transcription complex (RTC) with other viral proteins. These reverse transcripts either undergo recombination to form circular DNA, or form a pre-integration complex (PIC) through association with different cellular proteins. The complex moves along the cytoskeleton through the cytoplasm and enters the nucleus through the nuclear pore complex (NPC), where viral protein integrase (IN) incorporates the viral cDNA to form integrated DNA or proviruses.

Fig. 2. Cell-cycle dependent retroviral nuclear entry. (A) Entry of HIV-1 and simian immunodeficiency virus (SIV) through NPC into the nuclei of non-dividing cells and integration into the host genome. (B) Entry of murine leukemia virus (MLV) and foamy viruses such as prototype foamy virus (PFV) and feline foamy virus (FFV) into the nuclei of dividing cells. a. dissociation of nuclear membrane and access of PIC to cellular genomic DNA, followed by integration. b. Initiation of nuclear division and separation of the provirus-containing host genome. (C) Formation of two daughter nuclei containing integrated viral DNA.

Fig. 3. Competition of different components interacting with the HIV-1 capsid and their role in CA stabilization and nuclear entry. (A) Cyclophilin A (cyp A) and the drug CAI can bind
competitively with the capsid pocket to destabilize CA, resulting in premature release of viral RNA and inhibition of nuclear entry. (B) Cellular protein cleavage and polyadenylation specific factor 6 (CPSF6) and small molecule inhibitors such as PF-3450074 (PF74) and BI-2 compete for binding to CA, stabilizing the CA lattice to delay release of the viral genetic material, thereby resulting in impaired nuclear entry.
Fig 1
215x254mm (96 x 96 DPI)
Fig 2
214x87mm (300 x 300 DPI)
Fig 3
252x252mm (300 x 300 DPI)