CHAPTER I – HOST FACTORS INVOLVED IN HIV-1 REPLICATION

Background. The study of retroviruses has had a broad impact on diverse areas of biology. The simple retroviral genome, the total dependency on the host cell and other useful properties have made them powerful tools for studying basic cell function. Retrovirology revolutionized our concept of genetic information flow in that it established that genetic information can be transferred from RNA to DNA. Retroviruses have helped make major advances in the field of cellular growth control and led to the discovery of oncogenes.

Much of contemporary retrovirology is devoted to HIV-1, the main causative agent of the Acquired Immune Deficiency Syndrome (AIDS). Many major breakthroughs in basic biology have come from study of HIV-1, including advances in understanding of transcriptional elongation (126), nuclear export (43), and membrane vesicle sorting (49). HIV-1/AIDS is now the leading cause of death in sub-Saharan Africa and is the fourth biggest killer worldwide (WHO, 2002). In 2005, 40.3 million people were living with HIV-1, of which 4.9 million were newly infected, and 3.1 million died in 2005 (WHO AIDS epidemic update, 2005). The effectiveness of existing anti-HIV-1 drugs is limited by toxicity and selection of drug resistant virus. Understanding HIV-1 replication mechanisms would provide knowledge helpful for developing new anti-HIV-1 drugs.

Retroviruses: origin and features. Unlike other RNA viruses, Retroviridae do not utilize their genomic RNA for immediate viral replication. Instead, it is reverse-transcribed into a DNA intermediate (6, 119), which is integrated into the genome of the host, thus making retroviral infection permanent – “a retrovirus is forever” (29). The

reverse transcribed and integrated viral cDNA, termed as “the provirus”, then serves as a template for viral genomic and messenger RNAs. This unique strategy of replication brought about the designation “retro”.

Members of the Retroviridae family are enveloped viruses that carry two identical copies of a single stranded RNA genome that is non-covalently linked at the 5’ termini. The outermost shell of the virus, referred to as the envelope, is derived from the membrane of the host cell (29). The virions are approximately 80-100 nm in diameter, with the length of the genomic RNA typically ranging from 7 kb to 12.3 kb (29). Retroviridae are further subdivided into 7 genera, five of which encompass retroviruses with oncogenic potential, and the remaining two designated lenti- and spumaviruses (26).

HIV-1 belongs to the genus of lentiviruses. The designation lenti (slow, *lat*) comes from their especially long clinical latency period before the outbreak of disease. All lentiviruses share common morphological characteristics (a cylindrical core) and genetic features (several genes with regulatory roles), and are further divided into groups according to their tropism (26). HIV-1 belongs to the group of primate lentiviruses along with HIV-2, which is a distinct virus that also causes AIDS in humans, and a number of SIV viruses that exist in other human and non-human primates. A further subdivision of primate lentiviruses is phylogenetic: it is based on parts of genomic sequences of the viruses. This subdivision usually reflects information about the species- and subspecies origin of the virus. Thus, all known SIV viruses fall within one phylogenetic lineage and then diverge according to the primate species from which the virus originated, which then in turn form their own monophyletic clades (105).

Unlike other primates, humans do not have their own monophyletic HIV-1 lineage: instead, they seem to have acquired two different lineages, named HIV-1 and HIV-2, both further subdivided into several groups (105). Both lineages are examples of a zoonotic transmission of a virus that existed before in a different host: all three HIV-1 groups (M, N and O) crossed over into the human species from chimpanzees on three different occasions during the early 20th century (69, 99), while HIV-2 seem to have arisen from different transmissions from sootey mangabeys to humans (Fig.1.1), (105). Group M (main group) represents the most common HIV-1 isolates, while Group O (outlier group) isolates come from a distinct geographic location in west Africa. Only a few subtypes within the N (not M not O) group of HIV-1 viruses have been isolated so far, all of them reported in Cameroon. These observations are consistent with the hypothesis of three separate introductions of HIV-1 into the human population (105).

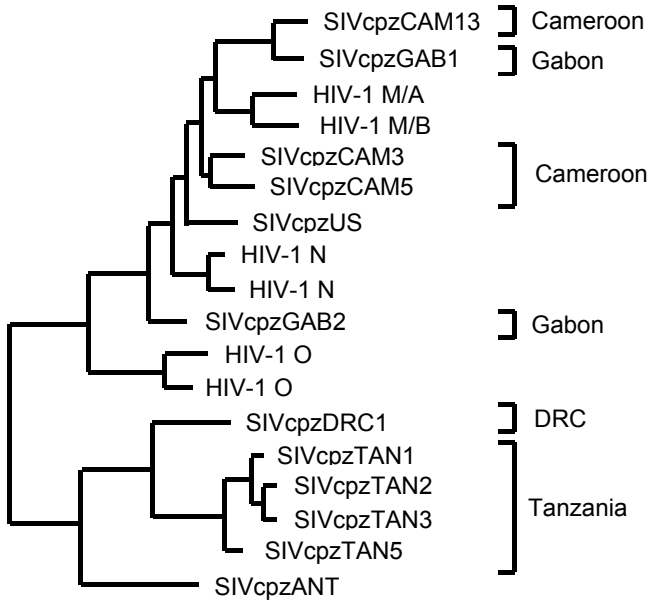


Figure 1.1. Phylogeography of SIVcpz (adapted from (105)). Brackets on the right indicate the country of origin of the various SIVcpz strains; SIVcpzUS and SIVcpzANT are of unknown geographic provenance.

Genomic organization. Retroviruses are powerful infectious entities, yet their genome is of fascinating simplicity. All retroviruses can be classified as either simple or complex depending on the organization of their genome (28). Simple retroviruses contain the three open reading frames (ORFs) common to all retroviruses: *gag, pol* and *env* (Fig. 1.2-A). The first open reading frame produces a set of structural proteins of the virus: it is transcribed and translated as the Gag precursor polyprotein $Pr55^{Gag}$. $Pr55^{Gag}$ is cleaved during maturation into four proteins: matrix (MA), which lines the viral lipid envelope; capsid (CA), which forms the viral core; nucleocapsid (NC), which resides inside the viral core and lines the viral genomic RNA, and the protein p6, which is implicated in late stages of the viral assembly. The *pol* ORF codes for three distinct proteins with enzymatic activity: protease (PR), which generates the cleaved products of *gag* and *pol* genes during virion maturation; reverse transcriptase (RT), an RNA dependent DNA polymerase which synthesizes the viral cDNA and possesses RNaseH activity; and integrase (IN), which mediates insertion of the viral cDNA into the host cell genome (Fig. 1.2-B). The *env* gene encodes surface glycoproteins that mediate viral entry through interaction with the cellular receptor CD4 and a chemokine receptor, usually CCR5 or CXCR4. In the case of HIV-1, a precursor polyprotein gp160 is synthesized and cleaved by a cellular protease into the surface unit gp120, also referred to as SU, and the transmembrane unit gp41 which is also called TM (Fig. 1.2-B). The two subunits remain loosely attached to each other via non-covalent interactions (29).

All lentiviruses are complex retroviruses. In addition to *gag*, *pol* and *env*, they encode a number of auxiliary proteins, which are involved in regulation of viral replication such as promotion of transcriptional elongation (Tat), nuclear import and cell

cycle arrest (Vpr), defense against cellular antiviral inhibitors (Vif), nuclear export of RNA (Rev) and modulation of signaling pathways (Nef) (Fig. 1.2-A). Some of the auxiliary genes overlap with ORFs of other proteins, and transcripts for these proteins are generated via differential splicing (29).

The viral genomic RNA resembles a cellular RNA in that it is capped at the 5' end and polyadenylated at the 3' end. After reverse transcription, the newly synthesized viral cDNA contains two identical direct repeats, termed Long Terminal Repeats (LTRs), which flank the viral genes. The LTRs contain regulatory elements that control initiation of transcription as well as transcriptional control elements anda polyadenylation signal. The newly synthesized viral RNA serves as mRNA for viral protein synthesis by cellular ribosomes, or as genomic viral RNA that is packaged into nascent virions. A complex structure at the 5' end of the RNA, referred to as the packaging signal ψ, promotes RNA incorporation into the virions (29).

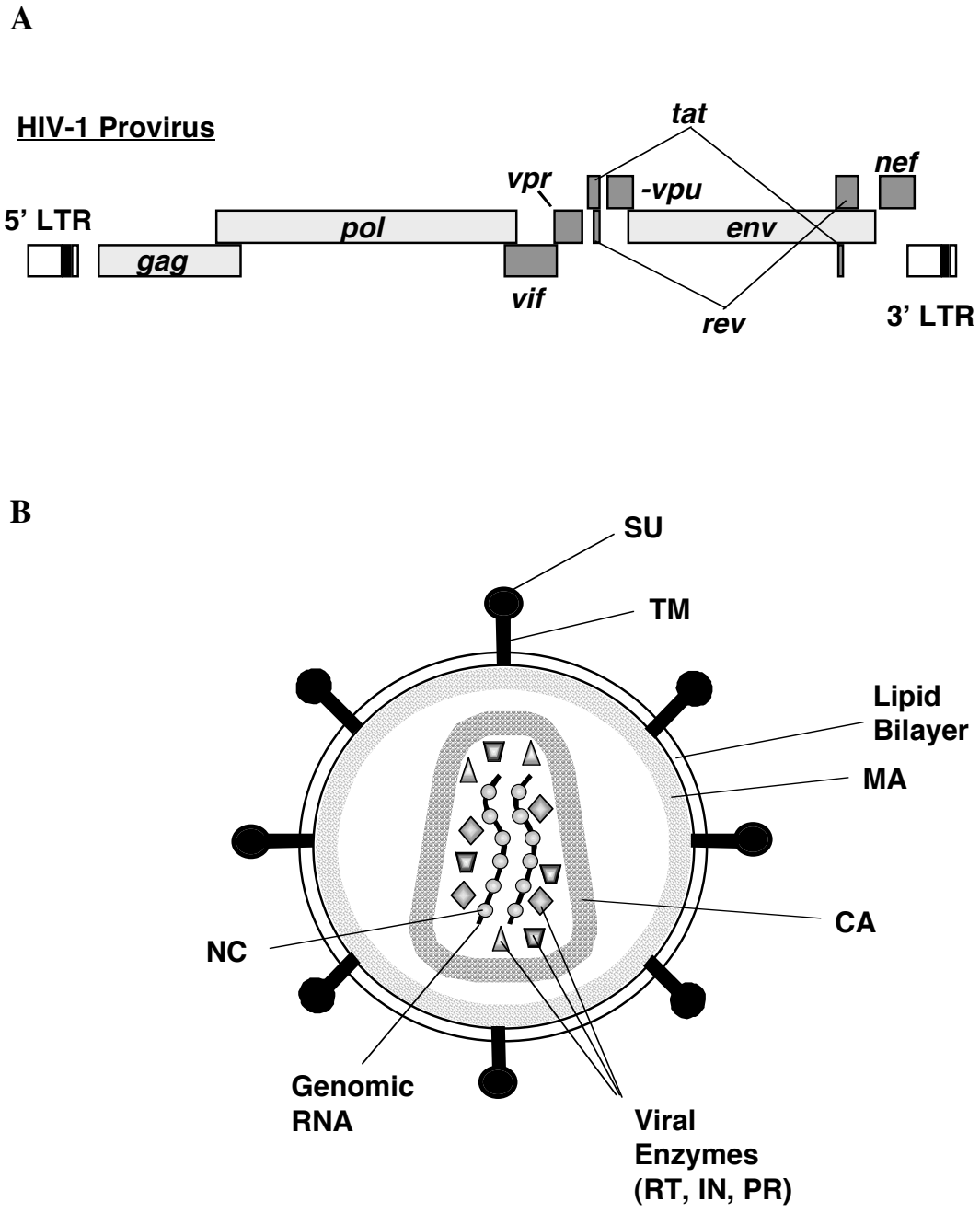


Figure 1.2. (A) Schematic of HIV-1 provirus genomic organization, (B) simplified mature virion. SU, TM - envelope glycoprotein surface and transmembrane components; MA - matrix; CA - capsid; NC - nucleocapsid; RT - reverse transcriptase; IN - integrase, PR - protease.

The retroviral life cycle. The retroviral life cycle begins with the interaction of the viral envelope glycoproteins with a cell surface molecule that serves as a viral receptor for entry (Fig. 1.3). In the case of HIV-1, the membrane surface CD4 molecule, which is mainly present on the surface of $CD4^+$ T lymphocytes, is utilized as a primary viral receptor (32, 66). In addition, a presence of a co-receptor is required for entry. Various chemokine receptors can serve as co-receptors, with CCR5 (27, 35) being the principal co-receptor and CXCR4 ((38) the secondary co-receptor in vivo. The chemokine receptors are expressed on many different cells, but expression of CD4 is limited to mainly T-cells and cells of monocyte/macrophage lineage, making these cell populations susceptible to HIV-1 infection. Conformational changes in gp120 induce insertion of the fusion peptide by gp41 into the membrane of the target cells, causing the viral and the cellular membranes to come in close proximity. The viral entry occurs via direct fusion of both membranes (8).

Following entry, the viral core containing the genomic RNA and viral proteins is released into the cytoplasm of the cell. Viral cDNA synthesis is initiated inside the core by RT, which utilizes a cellular $tRNA^{Lys}$ derived from the virus producer cell as a primer (118). The core undergoes disassembly or "uncoating", a poorly understood process during which the CA proteins that form the core of the virus are thought to dissociate, releasing the newly synthesized viral cDNA and the viral proteins into the cytoplasm. The cDNA is transported and actively imported into the nucleus, where the viral integrase mediates formation of the provirus by inserting the viral cDNA into the host's genomic DNA (88). After that, the promoter in the viral LTR drives RNA synthesis by a cellular RNA polymerase II, aided by NF-κB transcription factors and other cellular proteins for

initiation of transcription (63). The synthesized RNA is exported into the cytoplasm, where it is translated by cellular ribosomes into viral proteins. The viral proteins and RNA assemble at the cell membrane to form new viral particles, which bud from the cell membrane. After budding, the virions undergo maturation, during which the viral protease present in the virions cleaves Gag and Gag-Pro-Pol precursor polyproteins into single proteins (for review, see (84)). The mature virions are ready to infect new cells.

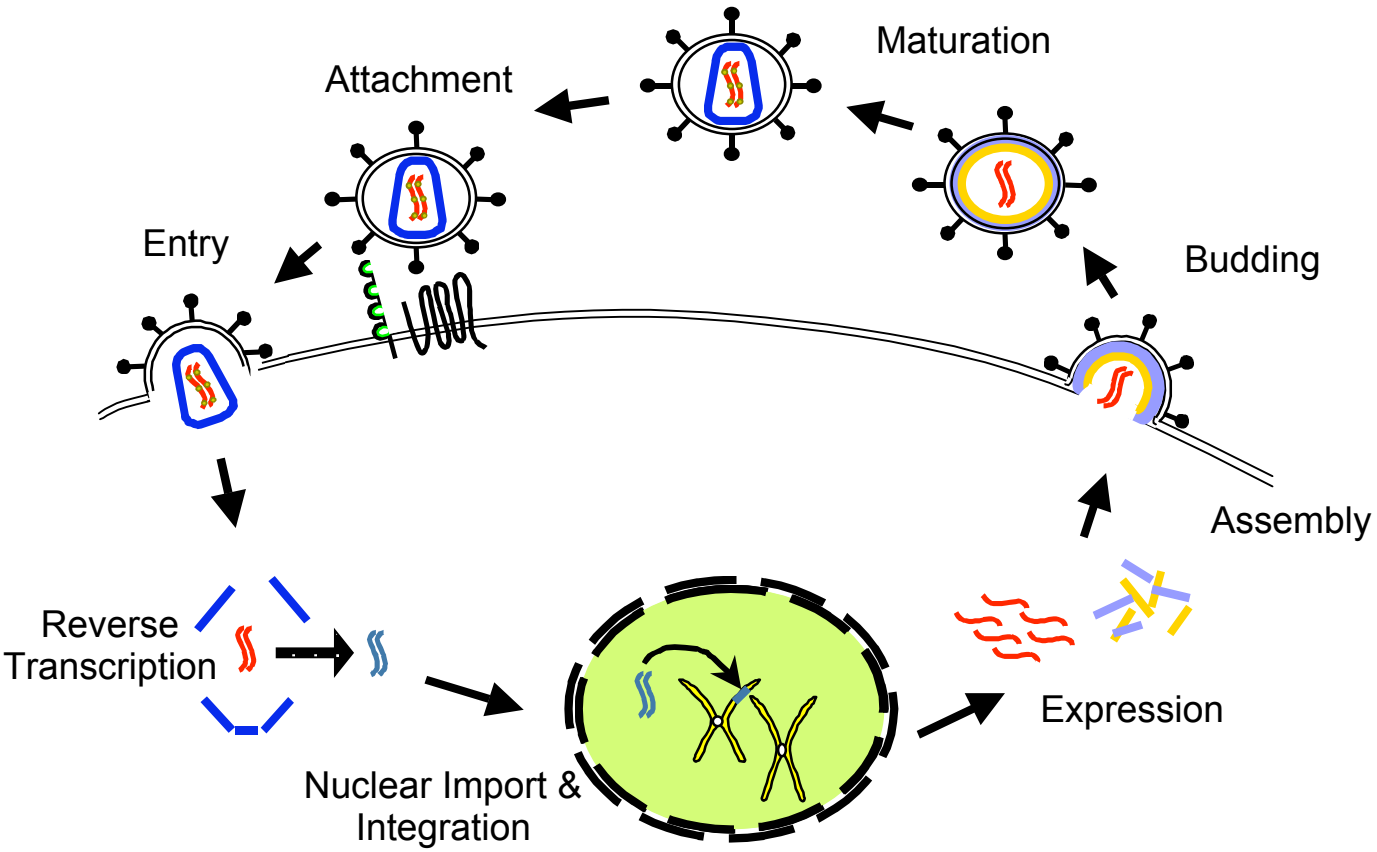


Figure 1.3. Retroviral life cycle. The major steps in the life cycle of a retrovirus are shown.

Cyclophilin A: a host factor modulating HIV-1 infectivity. In search for cellular factors interacting with the Gag precursor polyprotein $Pr55^{gag}$, a yeast-two-hybrid screen against a human cDNA library identified Cyclophilin A (CypA) as a protein that specifically binds to $Pr55^{gag}$ (77) (Fig. 1.4-A). The interaction was then shown to be direct by using recombinant protein in vitro, and confirmed in vivo, as purified HIV-1 virions were shown to specifically incorporate CypA (46, 120). The CypA binding site mapped to a proline-rich loop located in the center of CA (Fig. 1.4-B), where a single proline residue P90, or the immediately preceding glycine residue G89 were found to be crucial for this interaction, since mutation of each of these residues resulted in disrupted binding of CypA to $Pr55^{gag}$ (21, 46, 120), (Fig. 1.4-B). The proline-rich loop in CA is highly conserved in HIV-1 lineage, but does not occur in other retroviruses, including HIV-2, SIV and MLV (77, 79, 120). As expected, these viruses do not bind CypA (77, 120). In addition, CsA and its derivatives, the non-immunosuppressive CsA-analog $MeIle^{4}$-CsA (SDZ NIM811) and the structurally unrelated CypA-binding drug Sanglifehrin all act as competitive inhibitors of the CypA-CA interaction (36, 39, 46, 97, 120).

CypA is a small cytoplasmic protein that belongs to a large family of peptidyl-prolyl-isomerases. Members of this family catalyze *cis-trans* isomerisation of the peptidyl-prolyl bonds, which is a rate-limiting step in protein folding (reviewed in (42, 60)). It has been originally discovered as a cellular protein that binds and forms a tight complex with cyclosporine (CsA), an immunosuppressive drug that is used to prevent organ rejection in allograft recipients (52). The CypA-CsA complex binds and inhibits a cellular phosphatase calcineurin, blocking the T-cell receptor mediated signaling cascade (75). Although CypA is expressed ubiquitously in all eukaryotic organisms, its exact

function in vivo is not known.

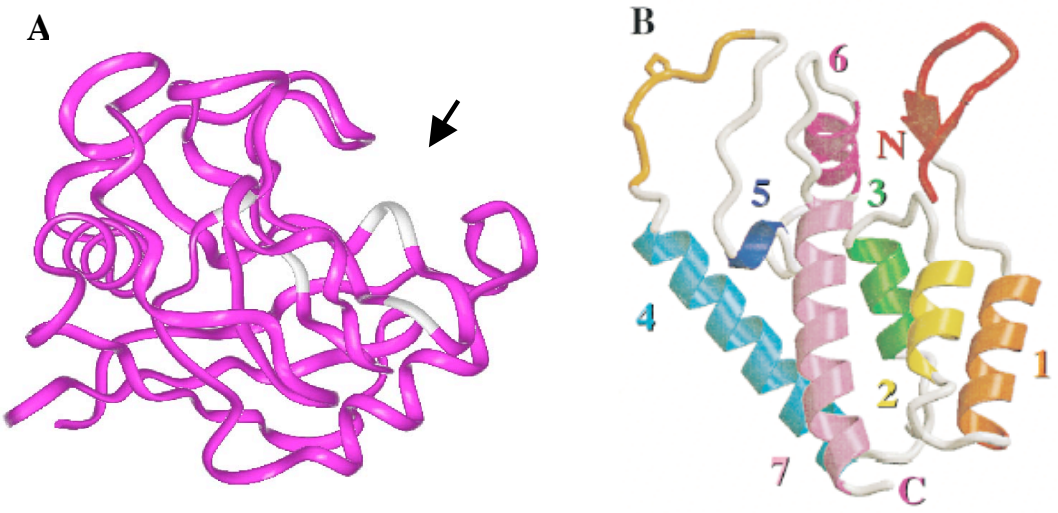


Figure 1.4. (A) Ribbon representation of the CypA structure. The arrow indicates the hydrophobic pocket that serves as a CA binding site. (B) Ribbon representation of the HIV-1 CA structure as determined by NMR-spectroscopy (48). α-Helices 1-7 are colored, the N-terminal β-hairpin is shown in red. The CypA binding loop connecting helices 4 and 5 and the position of the Pro90 within the loop are shown in yellow.

The importance of CypA for efficient spread of HIV-1 in tissue culture was assessed by disrupting CypA incorporation into virions by using several approaches. Replication assays with virus bearing CA mutants disruptive of CypA binding demonstrated that such mutations have a detrimental effect on HIV-1 replication (21, 46, 120). In addition, competitive inhibitors of CypA-CA interaction significantly inhibited HIV-1 spread in human T-cells in vitro (19, 21, 45, 97, 120). Finally, replication in human T-cell lines homozygous for deletion of the CypA gene demonstrated that HIV-1 replication in these cells is attenuated if compared to wild-type T-cells (23).

Examination of CypA-deficient virions failed to detect significant abnormalities in virion protein content, viral precursor protein processing kinetics, viral genomic RNA

packaging, endogenous reverse transcriptase activity, or virion ultrastructure (21, 23, 46, 67, 120, 128). Nonetheless, disruption of CypA incorporation into virions by any of the approaches mentioned above decreased virion infectivity. Experiments quantifying viral cDNA synthesis of HIV-1 virions with abolished CypA binding detected a block occurring early after infection, following membrane fusion at the stage of reverse transcription (21, 23, 120).

Producer cell versus target cell CypA. The results described above were consistent with the hypothesis that $Pr55^{gag}$-associated, intravirion CypA renders HIV-1 virions fully infectious. The presence of CypA in the virion at the time of infection was thought to promote an early step in the life cycle of the virus, such as disassembly of the viral core, after viral entry into the target cell. Thus, it was hypothesized that CypA weakens the core structure by interfering with contacts between CA monomers. In the absence of CypA, the viral core forms a tight structure that does not disassemble by itself; CypA is stoichiometrically incorporated into virion particles at a ratio of 1:10 (CypA:$Pr55^{gag}$), thus promoting destabilization necessary for uncoating, hereby allowing the newly synthesized cDNA to enter the cytoplasm (21, 133).

Studies focusing on disruption of CypA-CA interaction utilizing CypA-binding drugs solely during virion assembly have made similar observations with respect to the CypA requirement for HIV-1 infectivity (4). These reports consolidated the hypothesis that CypA is required to be packaged into the virus particle for HIV-1 to gain its full infectivity.

However, although CypA-CA interaction was originally discovered as CypA

binding to $Pr55^{gag}$, CypA also binds the mature CA protein, which is cleaved from the $Pr55^{gag}$ by the viral protease inside the virion upon maturation, after virion release from the producer cell (77). Therefore, it could not be excluded that the stimulatory effect of CypA on HIV-1 infectivity is mediated by target cell CypA binding to CA after virion entry into the cell.

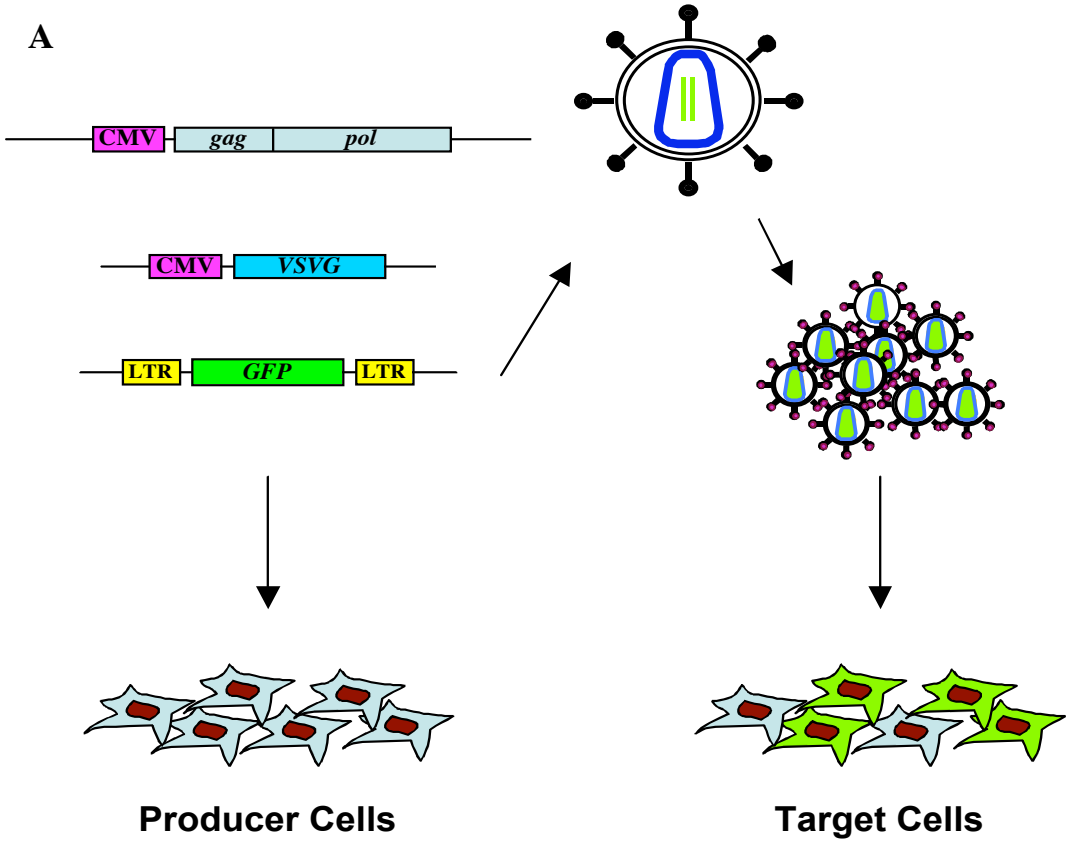
Methods to study viral infectivity. Experiments described above provided important information about CypA requirement for HIV-1 infectivity. However, they were not designed to differentiate between effects of CypA in the producer and in the target cells for several technical reasons. Thus, spreading infection assays are initiated in a pool of susceptible cells, which is inoculated with a replication competent virus. The newly produced virus is released into the cell culture medium and spreads from infected cells to new cells as they proliferate; the virus replication is measured by monitoring reverse transcriptase activity in the cell culture supernatant over a prolonged period of time. These assays offer insight into the "total" replication fitness of the virus; however, they do not provide information about distinct stages of the viral life cycle, and are therefore not suitable to determine the stage of infection at which the inhibition of replication occurs. In addition**,** spreading infection assays utilizing CsA to disrupt CypA-CA interaction require constant presence of the drug in the culture medium, thus disrupting both producer cell CypA incorporation into the virions as well as the target cell CypA binding to mature CA after virion entry. CA mutants that disrupt the CypA-CA interaction also prevent both producer and target cell CypA from interacting with CA. Finally, Jurkat cells engineered to be CypA deficient by gene targeting were helpful tools

to study the role of CypA in spreading infection, but they were difficult to transfect or otherwise manipulate and were therefore not a suitable cell line for production of viral stocks.

Over the last years, a large number of experimental tools for studying viral infectivity became available, making possible to dissect the viral life cycle into distinct steps that can be studied separately. Thus, single-cycle viral infectivity assays utilizing replication-deficient viral particles provide a possibility to study early events in the life cycle of the virus (for review, see (125). These viral particles (vectors) are engineered to carry a reporter gene in their genome, typically a fluorescent protein or an antibiotic marker (34), and possess all necessary information in order to enter a cell and integrate the reverse transcribed cDNA of the reporter gene into the host genome, but are at the same time devoid of ability to form new infectious virions after integration; (125, 136) (Fig. 1.5-A). However, the integrated reporter gene is expressed, allowing identification of infected cells (Fig. 1.5-B). Thus, potentially confounding results of a secondary infection are eliminated, and the readout of the reporter gene reflects the efficiency of the post-entry events. In addition, reporter vectors can be pseudotyped with the vesicular stomatitis virus G envelope glycoprotein (VSV-G), a pantropic envelope that allows bypassing of the requirement for cell-specific viral receptors, enabling the vectors to efficiently infect any mammalian cell (37, 131, 132), (Fig. 1.5-A). When studying early post-entry effects of a replication-competent virus, it is possible to preclude virus spread to new target cells by chemicals that interfere with gp160-mediated fusion of viral and cell membranes (13). Finally, the recent development of RNAi in mammalian cells in combination with retroviral vector based gene delivery permits stable downregulation of

gene expression in transfectable cell lines, allowing to more readily assess the role of the gene of interest in viral infectivity (25).

In this study, we utilize the techniques described above to investigate the relative importance of the producer cell vs. target cell CypA for HIV-1 infectivity. In addition, we use previously described CA mutants and variants with altered CypA dependency to assess CypA role for HIV-1 replication (1, 21, 22, 44, 133).



B

Low Multiplicity of Infection **High Multiplicity of Infection**

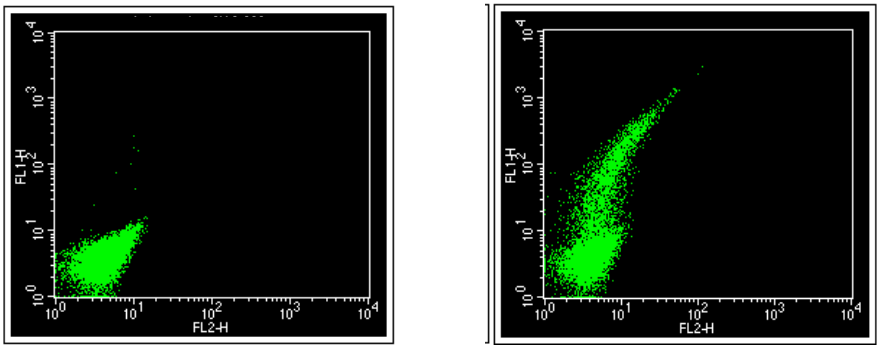


Figure 1.5. Single cycle viral infectivity assay. (A) Virions are produced by co-transfection of three constructs encoding viral packaging genes, the envelope glycoprotein and the GFP cDNAs into 293T producer cells. Virus is harvested from the supernatant, purified, and used to infect target cells. (B) Typical flow cytometry (FACS) readout measuring the percentage of infected (GFP-positive) cells. GFP fluorescence is detected on the FL1 channel.