# Identification and characterization of new cellular interacting proteins of HIV-1 Integrase

By

Md. Kamal Uddin Parvez

A Thesis submitted to the Faculty of Graduate Studies of
The University of Manitoba
in partial fulfilment of the requirements of the degree of

**MASTER OF SCIENCE**

Department of Medical Microbiology & Infectious Disease

University of Manitoba

Winnipeg

# ABSTRACT

HIV-1 integrase (IN) enzyme employs several viral and cellular proteins for nuclear translocation and crucial integration of viral cDNA. Successful identification of new viral/cellular interactions may shed light for better understanding of HIV-1 replication. 293T cells were transiently transfected with pYEF-1-TAP-IN and cell lysate were subjected to Tandem Affinity Purification system to pull down putative IN-interacting cellular partners. A number of distinct bands from the Coomassie-stained gel were excised followed by in-gel digestion and mass spectrometry. Putative cellular partners of HIV-1 IN were heat shock protein 60 (HSP60), β-tubulin, γ-actin, ATP synthase alpha subunit and histone H1.2 were identified by mass spectrometry. Additionally, SF3A3 (splicing factor 3A3), another previously reported factor, was successfully co-immunoprecipitated with IN. The C-terminal portion of IN was found to be the region of interaction with SF3A3. Overall, this study has provided better understanding of IN dynamics enriching existing knowledge of HIV-1 IN biology.

## ACKNOWLEDGEMENT

For the successful completion of my thesis and master's program, I would like to cordially thank my supervisor Dr. Xiao-Jian Yao for his brilliant and dynamic mentorship. Besides, I am very much thankful to my graduate advisory committee members: Dr. Keith Fowke, Dr. Zhikang Peng and Dr. Keding Cheng for their timely and invaluable feedback and advice during the course of my research work.

I am also very grateful to my lab members; Dr. Zhujun Ao, Kallesh Dinappa Jayappa, Xiaoxia Wang, Binchen Wang and specially Yingfeng Zheng for their help, advice and support in my lab experiments. Besides, I would like to thank Derek Riley Stein from HIV research group for his generous help in performing different proteomics experiments. I do acknowledge the generous help of Garrett Wesmacott, PhD, Chief, Proteomics & Mass Spectrometry Core Facility, National Microbiology Laboratory, Public Health Agency of Canada Winnipeg, Manitoba and his lab members for performing the mass spectrometry experiments. Special thanks to Stuart McCorrister of same facility for his skillful mass spec data analysis and feedback from time to time. I also cordially thank Zinnatun Nabi for critical help in statistical analysis. Besides, I am grateful to my loving parents and family members for their mental support, to the faculties, students and supporting administrative and technical staffs of the Department of Medical Microbiology for their generous help during my course work and study.

Finally, I am extremely grateful to Manitoba Health research Council (MHRC) for funding my research through graduate studentship award.

*Dedicated*

*to*

*my beloved parents*

# Table of contents

# List of Figures

# List of Table(s)

# Abbreviations

| | |
|---|---|
| AIDS | : Acquired Immuno-Deficiency Syndrome |
| APOBEC3G | : Apolipoprotein B mRNA editing enzyme, catalytic polypeptide-like 3G |
| BAF | : Barrier –to-autointegration Factor |
| BTF3b | : Basic transcription factor 3 isoform b – component of RNAP II complex |
| HMGN2 | : High mobility group nucleosomal binding domain 2 |
| CCD | : Catalytic Core Domain |
| CCR5 | : Chemokine Co-Receptor 5 |
| cDNA | : Complementary DNA |
| CTD | : C Terminal Domain |
| CXCR4 | : Chemokine (C-X-C motif) Receptor 4 |
| DC | : Dendritic Cell |
| ESCRT | : Endosomal Sorting Complex Required For Transport |
| GFP | : Green Fluorescent Protein |
| HAART | : Highly Active Anti-Retroviral Therapy |
| HAT | : Histone Acetyl Transferase |
| hHSP60 | : Human Heat Shock Protein 60 |
| HIV | : Human Immuno-deficiency Virus |
| HMGI (Y) | : High Mobility Group type protein (Y) |
| Imp7 | : Importin 7 |
| IN | : Integrase |
| INI-1 | : Integrase Interactor 1 |
| LEDGF/p75 | : Lens Epithelium –Derived Growth Factor/p75 |
| MA | : Matrix |
| NC | : Nucleocapsid |
| Nef | : Negative factor |
| NTD | : N- Terminal Domain |
| PIC | : Pre-Integration Complex |
| PR | : Protease |
| PRC | : Pericardin |
| RNAi | : RNA interference |
| RRE | : Rev Response Element |
| RT | : Reverse Transcriptase |
| SF3a3 | : Splicing factor 3a3 |
| SF3b2 | : Splicing factor 3b2 |
| THRAP3 | : Thyroid hormone receptor protein 3 |
| TSG101 | : Human Tumor susceptibility Gene 101 |
| VHL | : von Hippel Lindau binding protein |
| Vpr | : Viral protein R |
| YFP | : Yellow Fluorescent protein |

CHAPTER I

# INTRODUCTION

"No event in documented world history, before or since, has had as dramatic an impact on the human population — until the arrival of HIV. Nobody knows exactly how many people perished in the Black Death, but experts believe that HIV has already surpassed the numbers of people sickened by the plague. And when the currently HIV-infected cohort of 40 million people have succumbed to the disease, Acquired Immunodeficiency Syndrome (AIDS) will rank as the worst plague of all human history" (1). It is Human Immuno-deficiency Virus (HIV) that causes this fatal disease i.e. AIDS (2). AIDS has been termed "modern-day plague" for its unparallel potential to impose severe and unprecedented morbidity and mortality upon human race (3).

HIV-1 is one of the two species (i.e. HIV-1 and -2) comprising the viral genus Lentivirus (Latin 'Lenti' = 'slow': viruses with characteristic slow growth and long incubation period). Lentiviruses are under the Retroviridae family: RNA viruses that replicate in host cells via viral reverse transcriptase enzyme that catalyze production of cDNA or complementary DNA from viral RNA genome. Again, according to the Baltimore classification, HIV-1 comprises the group VI viruses: viruses with single-stranded positive sense RNA, reverse transcriptase enzyme and with DNA intermediate in life-cycle (4, 5).

AIDS was first clinically reported in homosexual men in USA in 1981 (6). In 1983, HIV, a retrovirus, was confirmed as causative agent for AIDS (7, 8). *In vivo*, HIV can replicate

only in hematopoietic cells expressing CD4 and co-receptors CXCR4 or CCR5, and mainly infects T-helper lymphocytes, macrophages and dendritic cells(4). *In vitro*, CCR5-tropic (R5) viruses can infect primary cultures of lymphocytes and macrophages (9). *In vivo*, R5 strains infect mainly memory T cells but X4 isolates can predominantly infect naïve T-cell lines and at least some populations of macrophages (reviewed in 4). Presence of replication-competent virus has been reported in the monocytes of patients treated with HAART (highly active antiretroviral therapy) indicating that HIV can infect monocytes also (9).

The virus is not itself lethal, most infected but untreated individuals mainly die from subsequent opportunistic infections or malignancies which are characteristically associated with the progressive and systematic failure of immune system (10).Most infected, untreated individuals will progress to AIDS within, approximately, ten years of HIV infection. Anti-retroviral therapy can increase the life expectancy of patients but can not cure the disease (11). Without antiretroviral therapy, infected and untreated individuals typically with AIDS die within a year (12). The unique feature that distinguishes HIV-1 from other retroviruses is it's ability to infect both dividing and non-dividing host cells and, like other retroviruses, the viral intermediate genome is integrated into the host cell genome inducing a persistent latent infection. Latent HIV1 reservoirs constitute a principal barrier to eradication, despite the presence of HAART (13).

AIDS was first reported in 1981 (6) and since then, almost 60 million people have become infected with HIV and HIV-related infections have cost life of 25 million people world-wide, so far. According to the UNAIDS report on the global AIDS epidemic on

2010: in 2009 alone, some 33.3 million [i.e. between 31.4 million-35.3 million] people were living with HIV, 2.6 million [2.3 million-2.8 million] had new infections and 1.8 million [1.6 million-2.1 million] died by AIDS-related complicacies. (14).

## 1.1 Literature Review:

After the discovery of HIV/AIDS, so far lots of scientific data have been generated regarding biology, pathology and treatment of HIV/AIDS. The following mini-review is intended to provide an over-view of HIV-1 life-cycle, host cell factors and IN-inhibitors that have been reported so far.

## 1.2 HIV-1 virion structure:

HIV-1 is an enveloped virus and its mature virion has diameter of roughly 110 nm. Glycoprotein gp120 and gp41 protrude from viral lipid-envelope which has been derived from the host cell membrane. A thin layer of matrix protein provides structural support to the envelope. The virion harbors HIV-1 genome, i.e. two identical copies of single-stranded RNAs of 9 kb each. The genome is surrounded by conical capsid composed of roughly 2000 copies of viral protein p24. Various other viral proteins like reverse transcriptase (RT), integrase (IN), nucleocapsid (NC), and accessory viral proteins e.g. Vif, Vpr, p6 and Nef are present within the capsid (Fig 1) (15,16).

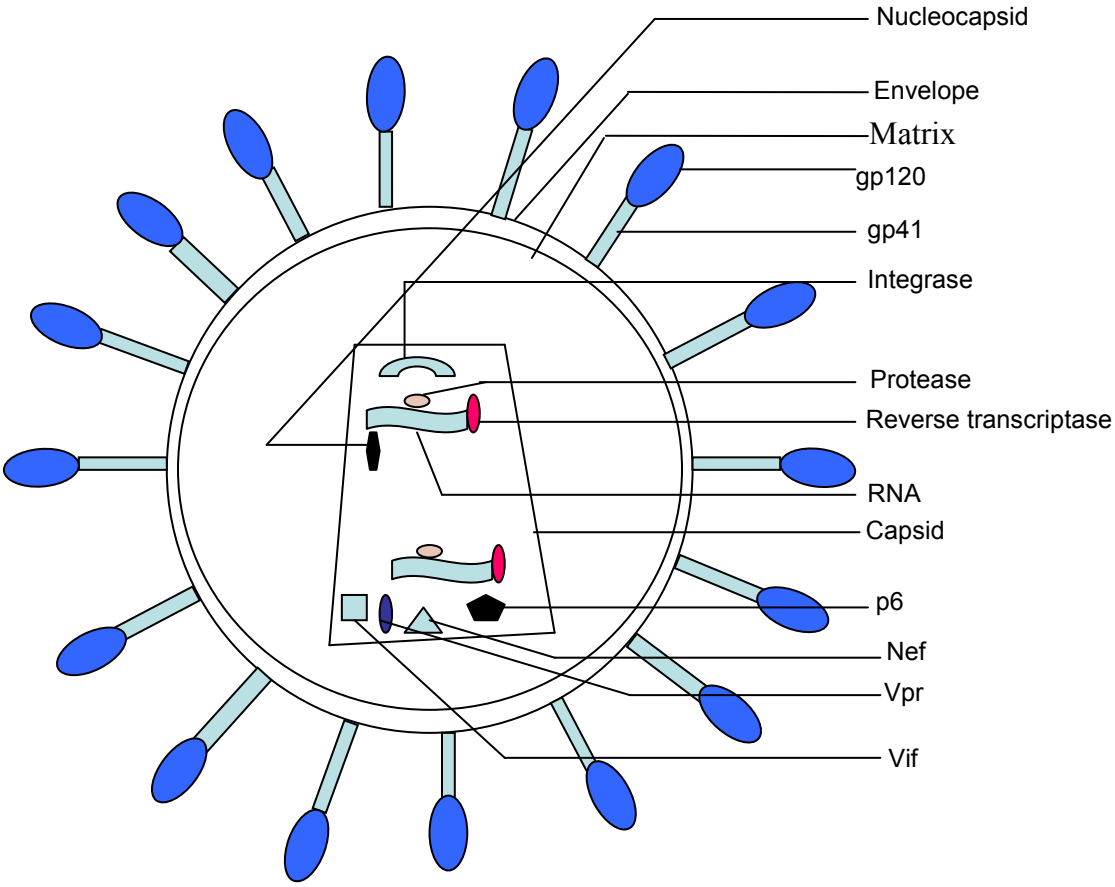

Nucleocapsid
Envelope
Matrix
gp120
gp41
Integrase
Protease
Reverse transcriptase
RNA
Capsid
p6
Nef
Vpr
Vif

**Fig 1: Schematic illustration of HIV-1 mature virion structure:** HIV-1 is an enveloped virus and surface protein gp41 protrudes from the envelope. gp120 is anchored by gp41 and interacts with CD4 receptor plus co-receptors. Beneath the envelope is the matrix and inner capsid proteins housing the viral RNAs (2 copies), and other viral proteins such as integrase, reverse transcriptase, protease,Vif, Vpr, Nef and p6 (15, 16).

## 1.3 Genomic organization of HIV-1:

HIV-1 viral genome has two identical copies of single-stranded RNA (each strand being 9 kb long). The HIV-1 genome encodes nine structural as well as six non-structural proteins. The three structural genes are *gag, pol* and *env*. Gag encodes for a polyprotein, $Pr55^{Gag}$. This polyprotein is further cleaved by viral protease enzyme into smaller proteins like p6, nucleocapsid (NCp7), matrix (MAp17) and capsid (CAp24) etc. The pol gene encodes viral enzymatic proteins which are, at first, synthesized as part of gag-pol polyprotein precursor $Pr160^{GagPol}$ (17). During translation of Pr55, a frame shift event, induced by the cis-acting RNA motif, yields Pr160 which encodes for viral enzymatic proteins reverse transcriptase (RT), protease (PR) and integrase (IN) which are also cleaved by viral protease (18). The env gene encodes for viral surface glycoprotein gp120 and viral transmembrane glycoprotein gp41 which are originally synthesized as a polyproteins precursor, gp160 which also undergo proteolytic processing by cellular protease (15).

## 1.4 Non-structural genes of HIV-1 and their products:

HIV-1 genome encodes two regulatory and four accessory proteins.

The Rev gene encodes a regulatory protein, Rev. The two separate exons of the rev gene partly overlap with tat exons. The 19 kDa protein facilitates export from nucleus to cytoplasm, splicing and stabilizes viral RNA transcript. Rev directly binds with RRE (Rev Response Element) located in env coding domain and premature RNA-splicing is prevented by this binding before nuclear export. A strong leucine-rich NLS or nuclear

export signal allows translocation of Rev between cytoplasm and nucleus. Up regulation of production of gag and env structural genes, and induction of transition from early to late stage of viral replication have also been reported by this protein (15, 16, 19).

Tat is another viral regulatory protein which is encoded by two exons, separated by env gene, that are located at 3′-end of viral genome. The first exon encodes a 14 kDa polypeptide composed of 72 amino acids containing functional domains capable of transactivating transcription. The second exon encodes a two kDa polypeptide of 14 amino acids and it has been proposed to help binding of Tat to target cell surface by integrin-mediated cell adhesion. Mainly, Tat protein binds to TAR or transactivating response RNA element located in LTR promoter then regulates viral gene expression in transcriptional level. It also impairs cellular processes by repressing host promoters, and allows expression of full-length viral transcripts (15, 16, 20).

### 1.5 Accessory proteins of HIV-1:

HIV-1 has four accessory proteins named Vif, Vpu, Vpr and Nef. They are called accessory because they are not required for viral replication in an *in vitro* infection system. However, numerous studies have clearly demonstrated that they all play important roles in HIV infection *in vivo* (21-23).

Vif: This 23 kDa protein is localized both in target cell cytoplasm and virion. Vif mainly counteracts the innate antiviral activity of APOBEC3G by forming complex with it, inducing it's ubiquitin-mediated proteasomal degradation so that APOBEC3G can not be

incorporated in progeny virion in order to mutate viral cDNA during the next round of infection (15, 16, 24).

Vpu: This is a 16 kDa integral trans-membrane protein and it associates with both host cell membrane and endoplasmic reticulum (ER) membrane. Functions of Vpu include degradation of CD4 allowing transport of gp41 and gp120 from ER to cell surface so that they can be incorporated into budding viral membrane. Besides, Vpu detaches viral particle from host cell surface and thus enhances their release (15, 16). Interestingly, Vpu has been reported to neutralize cell-surface antiviral factor tetherin/BST-2 by binding it and directing its beta-TrCP2-dependent degradation (25). More recently, Vpr has been reported to replace cell-surface tetherin and to re-localize residual pool of tetherin in a perinuclear compartment affecting its outward trafficking and/or recycling from the TGN (trans-golgi network). This sequestration of the anti-viral factor, i.e. tetherin, away from the plasma membrane so that the anti-viral factor can not exert its effect on the critical step of viral assembly (26).

Vpr: Is a 14 kDa protein. It interacts with classical Impα/β heterodimer pathway and thus assists in PIC nuclear import in non-dividing cells. Vpr can activate ATR cellular pathway and thus can induce cell cycle arrest in G2 phase (27).

Abnormal accumulation of CD4+ T-cells in G2/M phase can be observed in HIV-infected individuals indicating that Vpr-mediated G2 arrest likely plays an important role *in vivo* for viral replication or pathogenesis. Indeed, Vpr has been reported to upregulate the

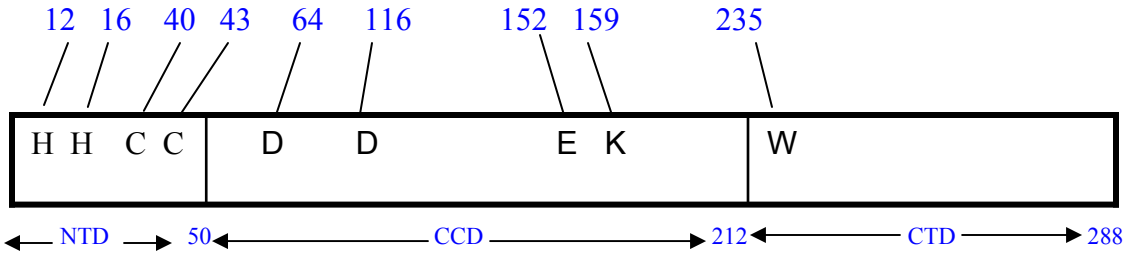
expression of ligands for the activating NKG2D receptor and thus promoting natural killer (NK) cell-mediated killing by a process that is dependent on ability of Vpr to induce a G2 arrest. This phenomenon suggests an immunomodulatory role for Vpr both in HIV-1-induced CD4+ T-lymphocyte depletion and in HIV-1-induced NK cell dysfunction (28).

**Nef:** This 27 kDa protein is a membrane-associated phosphoprotein. Nef redirects MHC1 and CD4 molecules from cell surface and Golgi apparatus to lysosome for degradation (15). It can bind to several members of Src-family proteins and can regulate their tyrosine-kinase activity and ultimately Nef can enhance viral infectivity and can prevent apoptosis through interaction with signal-regulating kinase Arabidopsis skp1-like1 (ASK1) protein (29).

### 1.6 Structure and function of HIV-1 Integrase:

HIV-1 Integrase (IN) is 32 kDa in size and has 288 amino acids. It is derived by proteolytic cleavage from the C-terminus of Gag-Pol polyproteins. It is involved in early stages of viral replication such as reverse transcription, nuclear import and, most importantly, integration of viral cDNA in host cell genome. IN functions in a multimeric form and has three structurally and functionally distinct domains (Fig 2). The N-terminal domain or NTD spans from residues 1-50 harboring a conserved zinc-finger binding site and its functions include multimerisation of IN that contribute in catalytic activity of the enzyme. The catalytic core domain or CCD spans from amino acid residues 51-212 which is relatively less conserved and catalyzes each 3′-OH end processing of viral

cDNA and also catalyzes DNA strand transfer, thus mediating crucial integration of viral cDNA into host genome. Additionally, the amino acid residues from 212-288 comprises the C-terminal domain or CTD which is involved in non-specific DNA-binding. Although, *in vivo*, IN works as a tetrameric complex, individual expression of each IN domain, *in vitro*, forms dimeric complexes that are consisted of two monomers (reviewed in 30).




Multimerisation
DNA binding
Zn binding

Multimerisation
DNA binding
Active site

Multimerisation
DNA binding

Fold: Helix-turn-helix : RnaseH : Src homology region

**Fig. 2: Domains of Integrase enzyme also showing conserved amino acid residues**. The enzyme is comprised of three domains: (i) the N-terminal domain (NTD, residues 1–49), (ii) catalytic core domain (CCD, residues 50–212) and (iii) C-terminal domain (CTD, residues 213–288). The DDE motif is required for catalytic function of the enzyme.

## 1.7 HIV-1 replication:

The replication of HIV-1 can be divided into two stages: early stage and late stage. Early stage of viral replication includes viral entry into the target cell, uncoating, reverse transcription of viral RNA into cDNA, nuclear import of pre-integration complex or PIC, and, finally, integration of cDNA into host genome (Fig. 3). Late stage replication comprises stages like viral gene expression, assembly of viral proteins, budding and maturation of immature virion etc (15, 31).

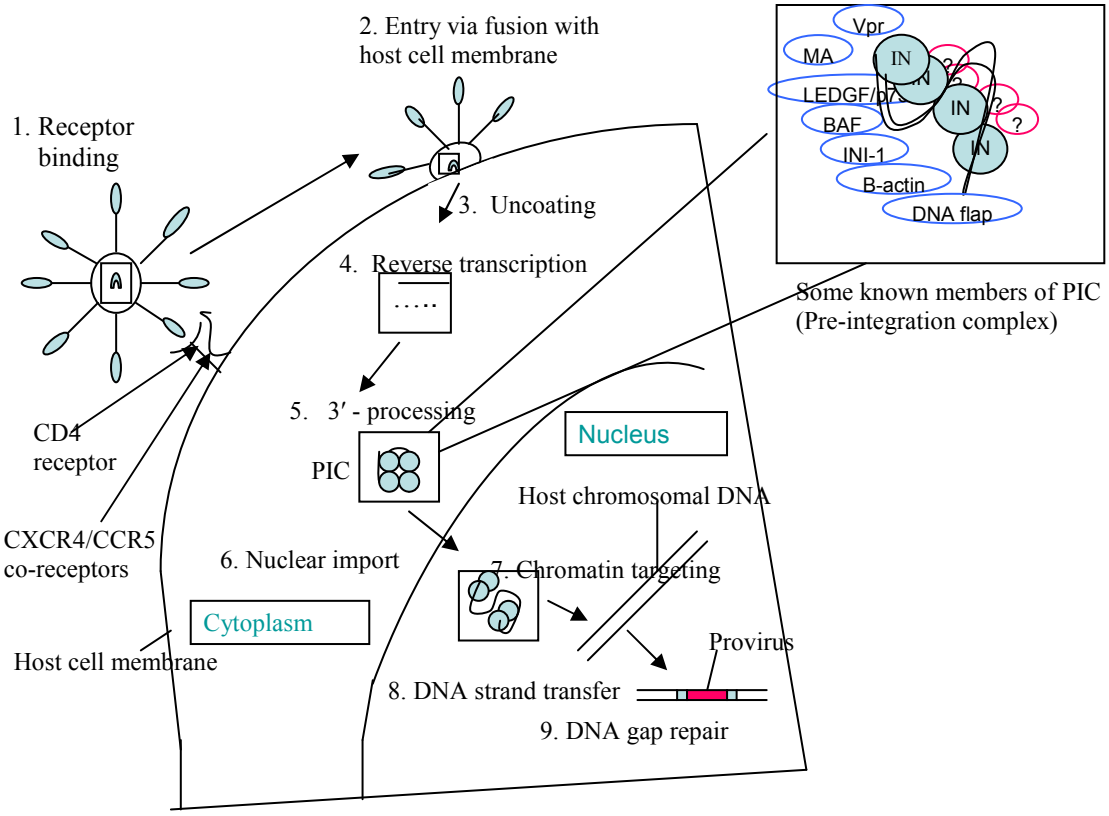

1. Receptor binding
2. Entry via fusion with host cell membrane
3. Uncoating
4. Reverse transcription
5. 3′ - processing
PIC
Nucleus
Host chromosomal DNA
6. Nuclear import
7. Chromatin targeting
8. DNA strand transfer
9. DNA gap repair
Provirus
CD4 receptor
CXCR4/CCR5 co-receptors
Host cell membrane
Cytoplasm
Vpr
MA
IN
LEDGF/p75
BAF
INI-1
B-actin
DNA flap
?
Some known members of PIC (Pre-integration complex)

**Fig. 3: Schematic illustration of steps in the early stages of HIV-1 life-cycle:** After binding of viral surface glycoprotein gp120 with CD4 receptors and co-receptors, the virion enters the cell cytoplasm via fusion of viral envelope with host cell membrane. After uncoating, viral RNA is reverse transcribed to cDNA by viral reverse transcriptase. Meantime pre-integration complex (PIC) is formed where viral IN cleaves two extreme terminal bases of each 3′-end of each viral cDNA strand. PIC is a multi protein complex and some of the known members are shown in the enlarged view. PIC mainly helps nuclear translocation and chromatin targeting of viral cDNA (31, 15, 16).

### 1.7.1 Early stages of HIV-1 life cycle:

Viral entry:

Entry of the infectious HIV-1 viral particle occurs through the binding of viral surface glycoprotein gp120 with target cell surface receptor CD4 and co-receptor which leads to membrane fusion and penetration of virion core into the host. Highly specific binding between gp120 and CD4 induces signals for conformational changes in gp120 altering it's specificity and increasing it's affinity for co-receptors like CCR5 and CXCR4 (32, 33). Actually, gp120 has a V3 loop that determines the viral tropism to either CCR5 or CXCR4. CCR5, which is a β-chemokine receptor, is typically found on surface of macrophages and primary lymphocytes but not on T cell lines. On the other hand, CXCR4, an α-chemokine receptor, is expressed on surface of T cell lines and primary lymphocytes, but not on macrophages. Additional binding to co-receptor broadens HIV-1 host range of immunologic cells which express variable receptors (34, 35).

After binding of gp120 with CD4 and CXCR4/CCR5; gp41 undergoes a conformational change leading to insertion of N-terminus of gp41 into host cell membrane (15, 30, 34, 36). Then, after formation of a stable six-helix bundle by C and N helices of tightly folded in gp41, this bundle positions host and viral membranes in place to facilitate HIV-1 core to undergo fusion and delivery inside host cytoplasm (30).

Uncoating:

After entry of the viral capsid into the host cytoplasm, viral uncoating step begins. This poorly understood process is characterized by partial and gradual disassembly of viral core resulting in the release of viral RNA genome as well as RTC (reverse transcription complex) into the cytoplasm (37). But there is conflicting report on uncoating on it's being an early event or not. It has been reported that reverse transcription occurs inside intact capsid shell while it's on the way to nucleus. And, once the PIC reaches the nucleus, viral cDNA flap facilitates uncoating. Interestingly, the linear viral DNA remains trapped within capsid, in absence of cDNA flap (38).

Reverse transcription:

Like other retroviruses, HIV-1 can convert its genomic RNA into viral cDNA after entry and uncoating into the host cell. The process of reverse transcription is mediated by the viral enzyme reverse transcriptase (RT), a 66 kDa/51 kDa complex cleaved from by the viral protease from $Pr160^{GagPol}$ polyprotein (15, 39). Different cis-acting host and viral elements participate in series of steps to form a reverse transcription complex (RTC) which carries out the reverse transcription reaction yielding linear, double-stranded, full length DNA molecule (40, 41). During the initial stage of reverse transcription; host tRNA anneal to the 5′ UTR (untranslation region) region of viral LTR. It functions as a primer for the synthesis of negative strand of viral DNA. Methylated viral protein NC (Nucleocapsid) has been proposed to facilitate RNA rearrangement which is required for this event (40). Viral capsid protein CA might play some roles in this process as HIV-1 particles with mutation in the N-terminus of CA show impaired reverse transcription (40,

42, 43). Among the viral accessory proteins, Vif shows activity during reverse transcription as degradation of host antiviral factor APOBEC3G by Vif resists APOBEC3G-mediated impairment of reverse transcription (39).

Cytoplasmic trafficking:

After reverse transcription, viral cDNA is translocated into nucleus via PIC or pre-integration complex. The exact underlying mechanism of cytoplasmic trafficking is not well-understood but already compiled data suggest that both viral and host factors participate in the formation of PIC and its cytoplasmic trafficking. The viral memebers of the PIC are IN, RT, MA, Vpr, viral cDNA, CA, NC while the host cell members are barrier-to-autointegration factor (BAF), high mobility group protein HMGI(Y), human lens epithelium growth factor/transcription co-activator p75 (LEDGF/p75), and lamina-associated polypeptide 2α (LAP2α) etc (44 -47) . Besides, host cell proteins β-actin and β-tubulin have been implicated in cytoplasmic trafficking of PIC (48).

Nuclear import of PIC (pre-integration complex):

HIV-1 can infect both dividing and non-dividing cells and the hallmark is the ability of the virus, in non-dividing cells, to translocate viral cDNA complexed within the PIC from cytoplasm to nucleus through nuclear pore complex (NPC) (13, 15). Although the exact mechanism of PIC nuclear import is largely unknown, it has been reported that viral members of the PIC e.g. Vpr, MA, IN, and central DNA flap etc have nuclear localization signals and so have karyophilic properties that enable them exploit cellular nuclear import machinery such as Impα, Impβ, Imp7and RanGDP/GTP gradient resulting in

nuclear import of PIC using both non-classical and classical pathways (reviewed in 49). This unique ability of hijacking host cellular factors/pathways makes HIV-1 different from other oncoretroviruses as HIV-1 can not only infect dividing cells but also non-dividing cells like quiescent T cells, mucosal dendritic cells, tissue macrophages etc (50-52).

Tethering to host chromosome and integration:

Once the PIC is inside the nucleus, viral cDNA has been reported to exist in three forms: linear, 1-LTR, 2-LTR circles (53). Before intergration LEDGF/p75 helps the viral cDNA to tether to the target sequence of host DNA (54, 55). Integration occurs in three sequential steps: (i) 3′- end processing, (ii) DNA strand transfer, and (iii) gap repair between the host DNA and inserted DNA fragment. In the 1$^{st}$ step, IN cleaves last two bases of each 3′-end of each strand of viral cDNA (from a 5′- CAGT-3′ motif) to yield a CA-3′-OH end which is active intermediate of subsequent DNA strand transfer reaction.

The 2$^{nd}$ step of integration process is characterized by a staggered cleavage of target DNA (of transcriptionally active genes) mediated by IN leading to annealing of sticky CA-3′-OH ends of viral cDNA to the cleaved 5′ -end of target DNA sequence. Finally, during the gap repair step, cellular enzymes fill the gap between 5′-viral DNA end and target DNA end. The inserted viral DNA is called provirus (reviewed in 30).

### 1.7.2 Late stages of HIV-1 life-cycle :

Viral gene expression:

Once the viral cDNA is integrated, viral and host factors start to interact with cis-acting elements present in the proviral 5′-LTR promoter inducing the viral gene transcription in two stages: i) early Tat-independent and ii) late Tat-dependent (15, 16, 56-58). Host cell factors that participate in proviral DNA transcription are AP-1, nuclear phosphatase-1, IRF, Sp1, NF-κB etc (59-62). Binding of NF-κB to 3′-LTR region activates the viral promoter located in LTR to produce small and incomplete RNA transcript as the RNA polymerase II can not function efficiently under strong influence of local host cell transcription environment (63). So, in this Tat-independent stage, not full mRNA but only basal amount of viral transcriptional activators such as tat, rev and nef are transcribed only. Tat can recruit a cellular factor called positive transcriptional elongation factor-b (P-TEF-b) to the TAR element, and phosphorylate RNAPII C-terminal domain - completing full RNA transcription resulting in three classes of RNA transcripts (64-68):

a) Fully spliced RNA: this group comprises small 1.7 -2.0 kb fragments of RNA with HIV-1 introns removed and they can express Rev, Tat or Nef.
b) Partially-spliced RNA: This group characteristically possesses 1 intron, has a size of 5 kb, and can express Vif, Env, Vpr, Vpu.
c) Unspliced RNA: This group of primary transcripts has size of 9 kb. Gag and GagPol proteins are generated from them.

HIV-1 viral assembly requires accumulation of unspliced and partially spliced mRNA in cytoplasm and viral protein Rev facilitates the transport of these RNAs by binding to rev response element (RRE) present within unspliced RNAs (64,19,69).

Viral assembly:

Assembly of retrovirus including HIV-1 virion assembly has been reported to be driven by viral structural protein Gag. Experimental evidences suggest that viral assembly process comprises multiple steps including: (a) targeting of Gag to the site of assembly i.e. typically plasmamembrane, (b) interaction between gag and lipid bilayer membrane, (c) multimerisation of Gag, (d) encapsidation of genomic RNA, (e) incorporation of Env into virus particles, and (f) budding and release of nascent particles (reviewed in 70).

In HIV-1, Gag is synthesized as precursor polyprotein Pr55Gag, which contains four major structural domains CA, MA, NC, and p6, and two spacer peptides, SP1 and SP2 (71). These structural domains which are present within Pr55Gag have been found to be similar to mature viral-protease-processed Gag proteins. Regions determining Gag targeting, membrane binding, and Env incorporation are all present in MA, but both CA and NC regions determine efficient Gag multimerization. CA has been reported to promote Gag–Gag interaction through its dimerization, whereas RNA bound to NC has been postulated to serve as scaffolding. NC zinc finger domains mediate viral genomic RNA encapsidation. p6, due to presence of specific peptide sequences, can recruit cellular protein complexes that are essential for efficient virion particle release. However, the order of steps in virus assembly and interdependence between them is a matter of

further research. It is noteworthy that packaging of two copies of full - length dimeric viral genomic RNA is principally mediated by interactions between the NC domain of Pr55Gag and a 50 segment of the viral genome variously termed the packaging signal, encapsidation element (E), or C site (71,72).

Viral budding and release:

After successful assembly, budding is the next critical step of HIV-1 life-cycle. Budding and release of mature virion require the cellular ESCRT (endosomal sorting complex required for transport) machinery, normally implicated in cytokinesis and endosomal sorting. Budding may occur from platforms giving rise to multiple consecutive budding events, as has been reported for Rous sarcoma virus, or may occur mostly from unique sites, as suggested for Moloney murine leukemia virus. (reviewed in 73).

To promote virion particle release from plasma membrane, the C-terminal p6 domain of HIV-1 Gag has been reported to directly bind to two components of infected cell's budding machinery, i.e. ESCRT, through L (late-assembly) domains. The primary HIV-1 L domain is consisted of a PTAP (Pro-Thr-Ala-Pro) motif that binds ESCRT-I component Tsg101. This PTAP-Tsg101 interaction is critical for both correct proteolytic processing of Gag and infectious virion release, the latter being downregulated nearly 500-fold in case of a disrupted PTAP motif (reviewed in 74).

CD317 (BST-2/Tetherin/HM1.24) is a cellular antiviral restriction factor (α-interferon-inducible). In case of HIV-1, the restriction mechanism involves retention (tethering) of

mature HIV-1 particles at the producer cell surface, then, subsequent internalization, thus preventing virion release. Viral accessory protein Vpu can overcome this restriction, possibly by proteasomal and/or endolysosomal degradation of CD317 or reducing its cell surface exposure by sequestration of CD317 to the *trans*-Golgi network. Which mechanism will be used is determined by expression level of CD317 and Vpu, and/or availability of specific degradation pathways in a given target cell (reviewed in 75).

Viral maturation:

After successful assembly and release the HIV-1 virion must undergo a maturation step to produce infectious virion particle which are capable of infecting new target cells. This step takes place during or immediately after viral budding and is characterized by cleavage of $Pr160^{Gag - Pol}$ and $Pr55^{Gag}$ into individual viral proteins by the viral - encoded proteolytic enzyme protease (PR). The maturation of immature virion to an infectious particle is coincided by contemporary phenomena like continuous association of MA with viral membrane, condensation of NC with dimeric RNA genome, and reassembly of CA forming a closed conical capsid shell which surrounds the NC–RNA complex (reviewed in 76).

Proteolytic cleavage of Pr55Gag follows a controlled and sequential series of events. Initial cleavage separates the NC - SP2 - p6 fragment and MA - CA - SP1 fragment and subsequent cleavage is carried out at the MA - CA and SP1 - p6 sites. Finally cleavage happens at the CA - SP1 and NC - SP2 sites. Accurate maturation by proteolytic

cleavage of Gag is important to produce competent and infectious virions, as mutations in cleavage sites result in compromised particles with reduced infectivity (reviewed in 71).

A second assembly event involves release of mature CA resulting in formation of a, typically cone-shaped, core structure by a proportion of CA. Disruption of core stability or formation blocks reverse transcription on infection of a new cell. Although CA–CA interactions play a central role, the exact mechanism of core formation is yet to be fully defined. Morphologically the most significant change during maturation is CA release from Pr55Gag; however, the generation of mature NC and MA also has important impact on virus maturation. Role of mature NC has been implicated in condensation of ribonucleoprotein complex (RNP) and stabilization of viral dimeric RNA, and both are important for efficient HIV-1 replication (reviewed in 70).

### 1.8 Genetic variability in HIV-1:

A remarkable feature of HIV-1 (as well as HIV-2) is the exhibition of extraordinary genetic diversity. HIV-1 strains are classified into three distinct groups: M, N and O. Approximately, 98% of the global HIV infections have been reported to be caused by HIV-1 group M, whereas group N and O are, comparatively rare, and mainly confined to Cameroon and surrounding countries (139). The group M comprises nine subtypes (A to I) and many circulating recombinant forms (140). Actually, HIV-1 exists as “snowflake-like quasispecies in which nearly every virus in a population differs from every other one” (141). The genetic diversity in HIV-1 populations is caused by recombination and

point mutations. Point mutations are introduced mainly by error-prone DNA synthesis or by functions of host antiviral factors such as APOBEC3's cytidine deaminase activity (reviewed in 141).

The fast replicative cycle of the virus in the host cell generates $10^9$ to $10^{10}$ virions per day. As the reverse transcriptase enzyme of HIV has no proof-reading activity the error during reverse transcription may reach 1 per $10^5$ nucleotides whereas for cellular DNA polymerase it is 1 per $10^9$. Therefore, in the course of one day, in a single infected individual, many variants of HIV can be generated. Moreover, infection of a single cell by more than one strains of HIV further compound the situation (75-77). In fact, Preston et al. have reported approximately "five to ten errors per HIV-1 genome per round of replication in vivo" (138). As a retrovirus, HIV-1 has a very high frequency of recombination, a property attributed to recombinogenic template switching during reverse transcription and recombination-prone nature of HIV-1 replication machinery. It has been stated that "the best current estimate of HIV-1 recombination rates is about four to five crossovers per genome, a truly astounding rate that suggests that recombination is a normal part of the generation of each viral DNA" (141).

The impact of genetic diversity in HIV-1 population is that it enables the virus to counter-balance immune clearance and to develop resistance to avert actions/effect of antiviral drugs. Although, highly active antiretroviral therapy (HAART), introduced in mid-1990s, has been significantly successful to reduce HIV-related morbidities and prolong the lifespan of people living with HIV, the growing emergence of drug-resistance to the

existing antiretroviral drugs is causing treatment failure undermining the global fight against this lethal virus. HAART can not completely eradicate the virus and due to high frequency of mutation, no vaccine approach, also, had been successful against HIV, so far (77,118).

Recently reported mutations in reverse transcriptase gene conferring HIV-1 resistance to USFDA-approved drugs include M41L, D67N and K70R to zidovudine. Multi-drug resistant mutation such as M46I in the viral protease gene yield resistance to Indinavir, Atazanavir and Lopinavir. Mutations such as E92Q, Y143R, Q148H and N155H in the IN gene have been associated with resistance against only FDA-approved IN inhibitor Raltegravir (142).

### 1.9 Role of Integrase in HIV-1 life-cycle:

HIV-1 replication is carried out by viral machinery consisted of three virus-encoded enzymes: the reverse transcriptase (RT), the integrase (IN), and the protease (PR). The viral genomic RNA is used as a template, by viral reverse transcriptase, to produce a copy of viral complementary DNA by reverse transcription. The protease cleaves Gag and Gag-pol polyproteins into mature viral proteins. Among the HIV-1 proteins, IN is of particular importance because of its role in viral DNA integration (i.e. covalent insertion) into the host genome (13).This integrated form of the viral DNA is called the provirus. Once integrated, the provirus can persists in the host, beyond the reach of host immune system, and can serves as a template for the future transcription of viral genes and simultaneous replication of the viral genome. This leads to the production of new

progeny viruses upon release which infect new immune cells and keep the infection active (13).

Integrase has two major catalytic activities: (i) 3'-processing i.e. endonucleolytic cleavage at each 3'-OH extremities of the viral cDNA, and (ii) a strand transfer reaction leading to the insertion of the processed viral cDNA into the target site in the host DNA by a trans-esterification mechanism (29). Recently, IN has been reported to catalyze a specific internal cleavage of retroviral palindromic DNA sequence (78). Besides, IN has been proposed to catalyze a "disintegration' reaction (inverse reaction of strand transfer), but its occurrence *in vivo* is not clear (reviewed in 30).

*In vivo*, integrase oligomer, in association with viral cDNA, reverse transcriptase (RT), matrix protein (MA), Vpr, nucleocapsid protein (NC), and cDNA flap as well as known and unknown cellular partners forms a PIC or pre-integration complex (79-83). IN stabilizes the PIC and help translocation of it into the nucleus. Besides, functional interaction between HIV IN and viral RT has been reported which indicates an indirect role, at least, of IN in controlling viral cDNA synthesis (84-86).

### 1.10 Role of cellular proteins on IN activity:

The following discussion should clarify that different host cell factors interact with HIV-1 IN in different stages of its function and the interaction may be direct or indirect. Ultimately, HIV-1 exploits these interactions for successful replication and infection of the host.

### 1.10.1 In establishing IN solubility and favorable conformation:

Role of several cellular proteins have been reported to enhance IN solubility and establish favorable conformation. For example, Parissi et al. (2001) showed that hHSP60 stimulated the *in vitro* processing and joining activities of IN. They suggested that hHSP60-IN interaction might help IN to prevent aggregation and adopt a more competent conformation (87). LEDGF/p75, Lens Epithelial Derived Growth Factor, has been reported to interact with IN and stimulate the whole integration process. Indeed, it has been reported that LEDGF-IN complex actually shows more favorable solubility profiles than that of free IN (88). Besides, LEDGF/p75 has been reported to increase stabilization of the tetrameric state of IN which is required for the concerted integration (89).

### 1.10.2 Role of cellular proteins in nuclear import of PIC:

Prior to integration, the PIC has to cross the nuclear membrane, especially in the non-dividing host cells (e.g. macrophages) to translocate viral cDNA into the host chromosome (30). HIV-I has been reported to interact with actin for short distance trafficking and microtubule network for long distance trafficking (90). Specifically, β-actin and β-tubulin have been proposed to interact with IN and help cytoplasmic trafficking of PIC (reviewed in 31). The role of the members of PIC in HIV-1 life-cycle is very important. Because viral components of this PIC e.g. IN, Vpr, MA and cis-acting determinant cDNA flap have karyophilic properties which enable them to facilitate nuclear import of the PIC by interacting with host karyopherins. Host karyopherins like Importin α1, Importin 7 (Imp7) and transportin-SR2 (TRN-SR2) interact with HIV-1 IN and are involved in viral nuclear import (reviewed in 49).

Co-immunoprecipitation assay done by Ao et al (2007), in our lab, clearly indicated that IN interacts specifically with host importin 7 (Imp7) *in vivo* to carry out nuclear import, but does not interact with importin 8 (Imp8) or importin α (Rch1). To further elucidate the contribution of the IN/Imp7 interaction to HIV-1 replication, they also showed that an Imp7-binding defective IN mutant virus lost infectivity and displayed defects during both reverse transcription and nuclear import, and HIV-1 produced from Imp7-depleted cells, exhibited a 2.5–3.5-fold reduced infection in Imp7-knockdown susceptible cells (91). Additionally, Woodward et al. (2009) have reported that interaction of IN with nucleoporin NUP153 to facilitate nuclear import of HIV-1 PIC (92). BAF (Barrier-to-autointegration factor), a cellular component of the HIV-1 PIC, has been reported to stimulate the integration reaction in the PIC complex (93). Li et al (2000) have suggested that HMG I(Y), another host partner of the HIV-1 PICs, binds on multiple sites on viral cDNA, compacts it and thus facilitates formation of an active integrase-cDNA complexe (94).

One group has reported binding of INI-1 (IN-interactor-1) to IN and enhancement of increases the strand transfer activity of IN (95). However, Boese et al. (2004) found no effects on viral integration in INI-1-depleted cells (96). Again, another group has proposed requirement of INI-1 for Tat-mediated transcription activation (97). So, further research is necessary to work out the controversial role of such IN interactors and to discover new cellular partners of IN that will help us to understand the functional dynamics of the PIC.

Very recently Ao et al. (2010) from our lab, investigated role of importin alpha family members on HIV-1 replication. Using short hairpin RNA (shRNA)-mediated Impalpha3 knockdown (KD) they found significant impairment of HIV infection in different cell lines. They also showed that integrase (IN) interacted with Importin alpha3 both in a 293T cell expression system and in HIV-infected CD4-positive C8166 T cells, through a region 250 to 270 in the C-terminal domain of IN. Overall, this study reports another cellular interacting factor of HIV integrase, Importin alpha 3, which is required for successful nuclear import and replication of HIV-1 in both non-dividing and dividing cells (49).

### 1.10.3 Cellular factors in intra-nuclear trafficking of PIC and integration process:

Some cellular factors have been proposed to help intra-nuclear trafficking of PIC. For example, interaction between BAF and emerin (an internal-inner-nuclear-envelope protein) might facilitate integration by favoring access of PIC to chromatin (98). Additionally, interactions of IN with cellular proteins play important role in targeting of viral cDNA integration. A systematic study of the sites of HIV DNA integration into the host DNA has shown that integration is not entirely random. *In vivo* analysis of integration sites indicate that HIV cDNA tends to integrate into transcriptionally active sites (99).

Several cellular cofactors like LEDGF/p75 (100), INI-1 (101), Ku (102) and BAF (103), have been reported to make interaction with PIC in nucleus and the first two make direct physical interaction (104,105). LEDGF/p75 is mainly responsible for targeting

integration and it has been called the chromatin tethering factor for IN (54, 55). As LEDGF/p75 is essential for HIV-1 replication and LEDGF/p75 interacts directly with IN, the domain of interaction between these two proteins is, therefore, a promising target for the development of integrase ligands with antiviral activity (106) and such compounds/peptides have been already developed (107).

### 1.10.4 Interaction of cellular factors with acetylated IN:

In relationship with chromatin, it has been reported that p300 histone acetyl transferase (HAT) binds directly to HIV-1 integrase *in vitro* and acetylates three Lys residues in the CTD of integrase, at positions 264, 266 and 273. Acetylation increased integrase DNA-strand-transfer activity approximately two fold *in vitro*, indicating a potential role for p300-dependent acetylation of integrase during virus infection. Consistent with this notion, mutant virus harboring three Lys-to-Ala substitutions in integrase supported normal levels of reverse transcription and PIC nuclear import, but only ~4–7% of wild-type integration (108).

Interestingly, using a yeast-two-hybrid system, Allouch & Cereseto (2009) have shown that some cellular factors like BTF3b (basic transcription factor 3 isoform b – component of RNAP II complex), HMGN2 (High mobility group nucleosomal binding domain 2 – chromatin remodeling factor) and THRAP3 (Thyroid hormone receptor protein 3 – subunit of large transcription mediator TRAP complex) bind efficiently to unmodified IN and do not require IN acetylation while IN acetylation enhances binding affinity of Exp2 (Exportin 2 --- mediator of Importin α binding protein, mediator of Impα nuclear export

after NLS cargo release into the nucleus ), eIF3h (eukaryotic translation initiation factor 3 subunit H) and KAP1 (Kruppel associated factor 1 – transcription co-repressor and DNA damage response factor) to IN (109). Very recently, Terreni et al (2010) have showed that GCN5-mediated acetylation of HIV-1 integrase can enhance viral integration (110).

### 1.10.5 New cellular factors reported to be involved in interaction with IN:

Studamire et al. (2008), using yeast-two-hybrid system, have reported strong interactions, *in vitro*, between HIV-1 IN and yeast proteins like Fen-1, Ku-70, PRC, Zn finger p38, SLU7, SF3b2, SF3a3, SMN, Ddx p18, Kif3A, Radixin and Ran bp 10 (111). Finally, von Hippel Lindau binding protein 1 has been reported to interact and degrade HIV-1 IN by the prefoldin-VHL-proteasome pathway and this degradation of IN induces an integration-transcription transition in viral replication cycle (112).

Interestingly, interactions of IN with some cellular factors may be dispensable or indispensable. For example, down- regulation of INI-1 by RNAi or silencing the gene expression by gene knockout have shown that INI-1 is indispensable for HIV-1 replication (reviewed in 31). On the other hand , in our lab previous work by Ao et al (2007) showed that IN interacts specifically with host importin 7 (Imp7) *in vivo* to carry out nuclear import, but, HIV-1 produced from Imp7-depleted cells, exhibited a 2.5–3.5-fold reduced infection in Imp7-knockdown susceptible cells, not 100% reduction of infection (91). More interestingly, recent work by Zheng et al (2010) in our lab have found that although LEDGF/p75 functions as a chromatin tethering factor for HIV-1 PIC, IN is capable of binding to chromatin in a LEDGF/p75-independent manner sustaining a low level of viral replication (113).

All these developments imply that there might be some other unidentified cellular partners of HIV-1 IN whose interactive functionality with IN could compensate knock-down of other factors (91).

### 1.10.6 IN as a target for anti-HIV approach:

Although, HIV/AIDS has inflicted unprecedented mortality and morbidity on the human race, still there is no effective vaccine against HIV/AIDS. The only method of prevention has been described as ABC (Abstinence, Be faithful, and use a Condom) (reviewed in 114). The most effective, so far, antiviral therapy is called HAART or "Highly Active Anti-Retroviral Therapy" which consists of a drug cocktail of at least three different groups of anti-HIV drugs including nucleoside analogue reverse transcriptase inhibitors, non- nucleoside analogue reverse transcriptase inhibitors and protease inhibitors. But, HAART can only suppress HIV-1 replication in the infected individual, and can not cure the patient totally from the disease. Other drawbacks of the HAART includes the drugs have to be taken for a life long period, toxicity of the drugs, metabolic disorders induced by the drugs and emergence of drug-resistant viral strains etc (reviewed in 115). In this ordeal, new anti-HIV drugs with less or no toxicity, more effectiveness and more potential drugs are required in the arsenal of anti-viral therapy.

All these observations have been the object of pharmacological research and paved the way for the development of inhibitors targeting the interactions between IN and cellular cofactors (106). These wealth of data suggests that new host cell interactors of HIV-I IN are being discovered and that still some other interactions may be left to be revealed.

Growing resistance against already existing anti-IN drugs and ani-retroviral therapy issues a strong urge to identify more interaction of HIV with host cell. Identification of newer interactions between IN and host cell factors may help designing new targets for future and further development of integrase inhibitors.

These catalytic functions of the integrase are crucial for overall integration process and have thus been the target of intensive pharmacological research oriented to develop anti-HIV drug and so far several IN-inhibitors have been developed (30). It is noteworthy that two IN-inhibitory compounds — MK-0518 (raltegravir) and GS9137 (elvitegravir) — have clinically shown great promise and raltegravir has been approved by FDA for patients who fail to HAART (116). However, resistance has been reported against raltegravir already (117).

The massive global mortality and morbidity caused by HIV has been the basis for experimental research oriented to better understanding of HIV biology and, ultimately, to design effective anti-HIV therapeutic approach against this deadly virus (118). The outcome of these research data has established the notion that HIV-1 exploits a lots of protein-protein interactions between host cell factors and viral proteins for successful replication of the virus (13, 119). Importantly, many of these identified interactions have been the target of anti-HIV drugs developed so far (14). Growing evidences suggest that there might be more other cellular factors to be elucidated that play major impact in HIV-1 life-cycle (112, 119). This thesis will address, investigate and review the role of host cellular factors in different stages of HIV-1 life-cycle, and explore whether these new

information can be exploited to enrich our knowledge on HIV-1 dynamics *in vivo*. The primary focus will be to identify new interactions between HIV-1 enzyme integrase (IN) and new IN-interacting partners in host cell with the intention that new data will shed light for better understanding of early stages of HIV-1 life-cycle in host cell.

Previous research in this particular field has shown that HIV-1 IN interacts with a number of host cell factors as well as other viral proteins to form a pre-integration complex (PIC) for successful translocation of viral cDNA, nuclear import and, finally, integration in host genome (30). However, different groups have reported different results in terms of relevance/requirement of some these cofactors for IN function. As for example, cofactors (such as INI-1 and Importin-7) that initially were reported to stimulate IN activity, later it was found that HIV-1 infection was significantly unaffected by potent downregulation of these proteins (reviewed in 31, 91). Even there are conflicting reports on the importance of LEDGF/p75 which has been termed as chromatin tethering factor of viral cDNA (54, 55, reviewed in 31). Therefore, different aspects of interactions between HIV-1 IN or PIC with host cell factors is still poorly understood. Additionally, Studamire et al. have reported 27 yeast proteins that interacted with HIV-1 IN fusion proteins (111). However, no further research has been reported whether human homologues of these yeast factors do interact with IN or not. Additionally, as resistance against existing anti-IN drugs continues to be reported, further research is a must to develop 2$^{nd}$ generation anti-IN drugs and, more importantly, to develop alternative therapeutic approaches such as peptide-based inhibitors of IN. Better understanding of PIC composition, specific interactions of the members of PIC with IN are matter of further research.

So, in order to fill the existing knowledge gap in HIV-1 IN functionality, in this study attempt was made for identification of new HIV-1 IN interactors using Tandem affinity Purification (TAP) system, a relatively new approach to study protein-protein interaction.

# HYPOTHESIS

Given the fact that HIV-1 IN interacting partners, identified so far, do not account for all of the activities of IN, we hypothesize that some more cellular factors may be required or involved as part of the pre-integration complex (PIC) for successful cytoplasmic trafficking of viral cDNA within the PIC through the cytoplasm, its nuclear import and overall integration process.

# OBJECTIVES

The aim of this project was to test the hypothesis that there might be some other unidentified cellular partners of HIV-1 IN and to examine the efficacy of the newly developed tandem affinity purification (TAP) system to study protein-protein interaction(s) between HIV-1 IN and host cell co-factors. The particular objectives of this research work were:

(i) To set up optimum condition for Tandem Affinity Purification (TAP) system to identify new cellular binding partners of IN.

(ii) To confirm the interaction by *in vivo* co-immunoprecipitation assay.

(iii) Functional analysis of identified co-factors during HIV-1 replication

CHAPTER II

# MATERIALS AND METHODS

## 2.1 Construction of different plasmid vectors:

In the research work, one of our objectives was to pull-down and detect new host cell factors that might be interacting with HIV-1 IN. To do that, three TAP-IN sequence containing plasmids were constructed, previously: (i) pYEF-1-TAP-IN (ii) pNTAP-B-IN and (iii) pMSCVhyg-TAP-IN. The TAP (Tandem Affinity Purification) tag allows *in vivo* binding and pull-down of host cell factors with the protein of interest. The TAP tag was comprised of two sequences: (a) Streptavidin binding peptide or SBP (amino acid sequence: MDEKTTGWRGGHVVEGLAGELEQL RARLEHHPQGQREPSGGCKLG) and (b) Calmodulin binding peptide or CBP (KRRWKKNFIAVSAANRFKKISSSGAL). The full length cDNA of HIV-1 IN was generated from full length HIV-1 clone, HIV-1 HxBru.

For ectopic expression of TAP-tagged Integrase (i.e. TAP-IN); plasmid vector pYEF-1-TAP-IN was constructed with the backbone from pEF-1-PCS-Puro (7.5 kb). The IN cDNA was cut out from CMV-YFP-IN (described in 91) with *BglII* and *BamH1*. Then it was subcloned into pNTAP-B (4.5 kb) (Stratagene, Catalog # 240103) with *Nhe1* and *BamH1* to yield $P^{NTAP\text{-}B\text{-}IN}$ (transgene expression is driven by CMV promoter). Then the TAP-tagged IN sequence was cut out from pNTAP-B-IN with *Nhe1* and *BamH1* and was subcloned into pEF-1-PCS-PURO between *Nhe1* and *BamH1* sites and the newly

constructed plasmid was named pYEF-1-TAP-IN. To construct pMSCVhyg-TAP-IN, BamH1/BamH1 fragment of TAP-IN was cut from pYEF-1-TAP-IN and was subcloned into pMSCVhyg-MCS-Puro.

**Construction of codon-optimized pYEF-1-TAP-IN$_{opt}$ and pAcGFP1-C1-IN$_{opt}$ :**

Codon-optimized IN cDNA from pUC57-IN was cut out with *BamH*1 and *Xho*1 and subcloned in pNTAP to construct pNTAP-B-IN$_{opt}$. From pNTAP-B-IN$_{opt}$ vector TAP-IN sequence was cut out with *Nhe*1 and *Xho*1 and was subcloned in pEF-1-MCS-Puro ultimately to construct pYEF-1-TAP-IN$_{opt}$. To construct a codon-optimized pAcGFP1-C1-IN$_{opt}$: optimized IN cDNA was cut out from pNTAP-B-IN with *Bam*H1 and *Xho*1 and was inserted into pAcGFP1-C1 between *Bam*H1/*Bam*H1 sites.

**Construction of SVCMV-YFP-IN fusion expressor:**

The SVCMV-YFP-IN fusion expressor was constructed from the amplified HxBru IN cDNA, which was pre-amplified from HIV-1 HxBru provirus. The primers used for PCR were 5′-*Bgl*II (5′-GCCAGATCTTTCTTAGATGGAATAGATAAG-3′) and 3′-*Bam*H1 primers (5′-CTAAACGGATCCATGTTCTAA-3′), respectively. After the PCR, the amplified fragments were digested with restriction enzymes *Bgl*II and *Bam*H1. The digested fragment was then cloned in frame to the 3′ end of the EYFP cDNA of a pEYFP vector (BD Biosciences Clontech).

To generate C-terminal deletion mutant construct of IN i.e. SVCMV-YFP-$IN_{1-212}$: cDNA fragment encoding amino acids from 1 to 212 of IN was produced by PCR with primers 5′-*Bgl*II (5′- CAATTCCCGGGTTTGTATGTCTGTTTGC-3′) and 3′-*Xma*I IN212 (5′-CAATTCCCGGGTTTGTATGTCTGTTTGC-3′), respectively. This fragment was then digested with and inserted into the pEYFP-C1 vector at *Bgl*II and *Xma*I sites.

**SVCMVin-IN-YFP substitution mutation construct:**

The substitution mutants YFP-$IN_{K159P}$, YFP-$IN_{V165A}$, YFP-$IN_{A179P}$, YFP-$IN_{KR186,7AA}$ were generated by a two-step PCR-based method (120) using primers 5′- IN-HindIII-ATG and 3′-IN-Asp718, and complementary primers containing required mutations. The amplified, mutant cDNAs were cloned into pEYFP-C1 vector. Then all the constructs were subcloned into SVCMV vector so ensure better expression efficiency of each fusion protein under the influence of strong cytomegalovirus immediate early gene promoter. Presence of corresponding mutations was confirmed by sequencing.

**Construction of T7-SF3A3 plasmid:**

To generate a T7-SF3A3 fusion expressor, the full-length wild-type SF3A3 (Origene) cDNA was amplified by PCR using a 5′-BamH1 primer (5′-CTAAGGGGGATCCGAGACAATACT-3′) and a 3′-Not1Cla1 primer (5′-GCCCAAAAGCTGAGCATCGATGCGGCCGCCTAGAGC-3′). The amplified SF3A3 fragment was cloned in-frame at the 3′ end of the T7 cDNA in a CMV-T7 vector.

## 2.2 Virus production:

The 4.3-BruΔBgl/Luc (Luc+/env-) single cycle replication-competent provirus was constructed where nef gene was replaced by a firefly luciferase gene. RT/IN/Env-defective HIV-1 provirus was used as a backbone to generate a proviral clone with genotype 5′-LTR (Long Terminal Repeat) + gag+ pol+ vif+ vpr+ tat+ rev+ vpu+ env-nef-. From the backbone provirus; the sequence from ApaI site through to SalI (i.e. 1556 to 5329, NL4.3 transcription initiation begins at +1 of 1556) was removed and replaced with a corresponding sequence from HIV-1 provirus HxBru. The 4.3 GFP+/nef- virus was a generous gift from Dr. Eric Cohen from the University of Montreal, and possess a GFP gene in place of nef.

## 2.3 Cell lines:

293T human embryonic kidney cells and HeLa cells were maintained and cultured in Dulbecco's modified Eagle's medium (DMEM) supplemented with 10% fetal bovine serum (10% FBS) and 1% penicillin/streptomycin (1% P/S), at 37 ˚C in 5% $CO_2$ incubator.

## 2.4 Transfection method:

Transfection (or co-transfection during co-immunoprecipitation) of relevant vector DNA in 293T cells was carried out after a confluency of 50% to 60% was achieved using standard calcium phosphate DNA precipitation method. To optimize protein expression, 0.5 µl of proteasomal inhibitor MG-132 (Calbiochem) was added to the plates approximately 16 hours before cell-lysis.

### 2.4.1 Direct lysis method:

To check the ability and efficiency of the expression vectors (i.e. pYEF-1-TAP-IN, pNTAP and pMSCV-hyg) to express TAP-IN, and also, to check the expression of GFP/YFP (green fluorescent protein/yellow fluorescent protein), 293T cells were transfected with relevant plasmid. 48 hours after transfection cells were lysed with 0.5% NP-40 (in M199 medium) for 30 minutes on ice with manual shaking in every 5 minutes. Then, after centrifugation in 14,000 rpm for 20 minutes, the supernatant was collected, mixed with 1X Laemmli buffer, and run on 12.5% SDS-PAGE gel. Then the gel was subjected to western blot analysis to check the expression of desired protein.

## 2.5 Modified Tandem Affinity Purification (TAP) protocol :

The Tandem Affinity Purification (TAP) system is comparatively a new tool to study protein-protein interactions. This method is characterized by combination of two purification steps of a desired protein or protein complex of interest by using two different affinity purification tags in a tandem manner with subsequent mass spectrometric identification of unknown components of the protein complex. In this method two consecutive purification steps are performed using two different affinity purification tags (streptavidin binding peptide tag and calmodulin binding peptide tag) Treatment of cell lysate with streptavidin and calmodulin beads/resin and subsequent mild washing and elution conditions allow for isolation of desired interacting proteins without destabilizing the targeted complex.

We decided to employ this system to study and identify cellular *in vivo* interactions between HIV-1 IN and other cellular factors. For each plasmid vector, 293T cells were grown in six big culture dishes (in order to get, finally, $10^7$ cells) and were transfected with pYEF-1-TAP-$IN_{opt}$ or with pYEF-1-TAP (as control). Then the cells were lysed and cell-lysate was subjected to Tandem Affinity purification (TAP) protocol as per supplier's instruction.

**Collection of 293T cells:**

Transfected 293T cells were washed three times with PBS and then, using 5 ml ice-cold PBS, cells were collected by centrifuging for 10 minutes at 1500 × g. The PBS was removed before purifying the protein complexes.

**Preparation of the Protein Extracts:**

All steps were performed at 4˚C to prevent the interacting proteins from dissociating. Cells were resuspended in 6 ml of TAP lysis buffer (Stratagene, cat. no. 240107 ) in a polypropylene centrifuge tube and then were subjected to three successive rounds of freeze-thawing by incubating the cells on dry ice for 10 minutes to freeze the cells, followed by incubating the cells for 10 minutes in cold water to thaw the cells. After centrifuging at 16,000 × g for 10 minutes to pellet the cell debris the supernatant was collected. To each milliliter of cell lysate, 0.5 M EDTA and 14.4 M β-mercaptoethanol (provided) were added.

**Preparing the Streptavidin Resin:**

300 ul 50% streptavidin resin (taken out by sterile cut-tips) slurry was centrifuged at 1500 × g for 5 minutes. The resin was resuspended in 1 ml of streptavidin binding buffer (SBB) and was re-centrifuged at 1500 × g for 5 minutes to collect the resin. The resin was resuspended in 150 ul streptavidin binding buffer (SBB).

**Purifying the Protein Complexes Using Streptavidin Resin:**

The washed streptavidin resin (50% slurry), 150 ul, was added to the cell lysate. The tubes containing cell lysate was rotated at 4°C for over-night to allow the tagged proteins to bind to the resin. The resin was collected by centrifugation at 1500 × g for 5 minutes. The resin was resuspended in 1 ml of SBB by rotating the tube at 4°C for 5 minutes and this wash step was repeated. The resin was collected by centrifugation at 1500 × g for 5 minutes. 600 ul streptavidin elution buffer (SEB) was added to the resin and was incubated for one and half hour at 4˚C to elute the protein complexes. The resin was collected by centrifugation and the supernatant was carefully transferred to a fresh tube (the supernatant contained the eluted proteins). The supernatant (600 ul) was concentrated down to 20 ul by a DNA SpeedVac$^{TM}$ SpeedVac Model DNA110 (from Savant) for approximately 30 minutes. On the other hand, the beads were washed three times with (1 ml X 3 times) with streptavidin binding buffer (SBB).

## 2.6 SDS-PAGE analysis:

**Preparation of samples:**

20 ul of sample was mixed with 20 ul of NuPAGE® LDS sample buffer (4X) (Invitrogen, Cat. # NP0007), 8 ul of NuPAGE® Reducing agent (10X) ((Invitrogen, Cat. # NP0004)) and 32 ul of sterile deionized water. Again, in case of 150 ul of Streptavidin beads: 37.5 ul of NuPAGE® LDS sample buffer (4X), 15 ul of NuPAGE® Reducing agent (10X) and 97.5 ul of sterile deionized water mixed. Samples were boiled in 100˚C and then centrifuged in 7000 rpm for 3-4 min.

**Preparation of running buffer:**

1X SDS running buffer was prepared by adding 50 ml 20X NuPAGE® MOPS SDS running buffer (Invitrogen, Cat # NP0001) to 950 ml of deionized water.

**Loading of running buffer:**

The XCell Surelock™ Mini-Cell (Invitrogen, Cat # EI10001), i.e. the gel tank, was set up for SDS-PAGE run and the upper buffer chamber of the tank was filled with 200 ml 1X NuPAGE® MOPS SDS running buffer. 500 ul NuPAGE® antioxidant (Cat. # NP0005) was added in the buffer of the upper buffer chamber. The lower buffer chamber was filled with 600 ml 1X NuPAGE® MOPS SDS running buffer.

**Loading and run of the sample:**

25 ul of diluted protein sample was loaded in the wells of NuPAGE® Novex 4-12% Bis-Tris gel (8X8 cm) (Invitrogen, Cat# NP0321Box) set up in the gel tank. The gel was then run in 80 V for approximately 45 min, and then in 116 V for approximately three hours.

## 2.7 Silver Staining Protocol:

For mass spectrometry-compatible silver staining; SilverQuestTM Silver Staining Kit from Invitrogen (Cat. # LC6070) was used (all incubation steps were done on a rotary shaker rotating at a speed of one revolution/sec at room temperature. To avoid keratin contamination; precautions were taken as much as possible using hand gloves, face mask, head covers, lab coat and covering the gel container with aluminium foil etc). Briefly, after electrophoresis, the gel was rinsed shortly with ultrapure water. Then the gel was fixed in 100 ml of fixative solution (ethanol 40 ml, acetic acid 10 ml and water to make 100 ml) for 20 minutes. Then the fixative solution was decanted and the gel was washed in 30% ethanol for 10 minutes. After decanting the ethanol the gel was sensitized in 100 ml of sensitizing solution (ethanol 30 ml, sensitizer 10 ml and water to 100 ml) for 10 minutes. After decanting the sensitizing solution the gel was again washed in 100 ml of 30% ethanol for 10 minutes with subsequent wash in ultrapure water for 10 minutes. Then the gel was incubated with the staining solution (stainer 1 ml, water to 100 ml) for 15 minutes and after decanting the staining solution the gel was washed with ultrapure water for 20-30 seconds. Then the gel was incubated in 100 ml of developing solution (developer 10 ml, developer enhancer 1 drop, water to 100 ml) for 4-8 minutes until bands started to reappear and the desired band intensity was reached. As soon as the

desired staining intensity was achieved, 10 ml of stopper solution was added onto the gel still immersed in the developing solution, rotated for further 10 minutes. Color change from pink to colorless indicated stop of the development. Finally, the stopper solution was decanted and the gel was washed with 100 ml of ultrapure water for 10 minutes.

### 2.8 Destaining For Mass Spectrometry:

After silver staining the silver ions were removed from the protein bands, before the bands were subjected to in-gel trypsin digestion, by the following protocol provided with the SilverQuest$^{TM}$ Silver Staining Kit from Invitrogen (Cat. # LC6070):

After completion of silver staining, the gel was washed thoroughly with ultrapure water. Then the gel was placed on a box lamp and distinctive bands of interest were excised using separate clean scalpels and the gel piece was placed into separate 1.5 ml microcentrifuge tube. 50 ul destainer A and 50 ul destainer B were added to each microcentrifuge tube and the contents of the tubes were mixed thoroughly by vortexing at room temperature for 15 minutes. After a quick centrifuge, the supernatant was removed using sterile pipette tip. Then, to wash the gel pieces, 200 ul of ultrapure water was added to the tube and vortexed for 10 minutes at room temperature. The wash step was repeated twice, then samples were treated for in gel digestion.

### 2.9 Coomassie Blue Gel Staining Protocol:

Mass spectrometry compatible Coomassie blue staining was also done for a gel using the following protocol available on line: (http://sr.burnham.org/sr/homepage/proteomics/

coomassie.html) (To avoid keratin contamination precautions were taken as much as possible using hand gloves, face mask, head covers, lab coat and covering the gel container with aluminium foil etc). The gel was incubated in 100 ml of Coomassie stain solution (0.2 % Brilliant Blue R-250 from FisherBiotech, Cat. # BP 101-25, 20% methanol, 0.5% acetic acid dissolved in ultrapure water) for 60 minutes. Then the gel was rinsed briefly in ultrapure water. To destain, the gel was rotated in 100 ml of 30% methanol with replacement of destaining solution in every 30 minutes over a destaining period of 2 hours. After adequate destaining, the gel was stored in 0.5% acetic acid. Then appropriate gel bands were excised with clear, separate scalpels, put in sterile eppendorf tubes and were cut into smaller pieces inside the tubes so that trypsin can reach to the proteins within the gel well. Later, the gel bands were subjected to in-gel digestion.

### 2.10 In-gel Digestion:

The gel bands in eppendwarf tubes were initially washed with 600 µl of 100 mM ammonium bicarbonate (AB) and 400 µl of acetonitrile (ACN) by vortexing for 10 min. After a quick centrifuge supernatant was removed. Then 200 µl of ACN was added to the pellet, mixed for 2 minutes and after a quick centrifuge the supernatant was removed. Again, 600 µl of AB was added and mixed for 5 minutes. Then 400 µl of ACN was added and mixed for 10 min. After a quick centrifuge supernatant was removed. 200 µl of ACN was added and vortexed for 2 min. After a short spin, the supernatant was removed. Whole wash steps were repeated until no stain was visible in gel pieces.

**Reduction & Alkylation:**

100 µl freshly prepared dithiothreitol (DTT) was added to the gel pieces and were incubated for 45 min at 56 °C (or 1 hr at 37 °C). After a short spin, the supernatant was removed and cooled for 5 min. Then 50 µl Iodoacetamide (IA) was added and was incubated for 30 minutes in dark at room temperature. Then after short spin, supernatant was removed. Then, the gel pieces were washed with AB and ACN as described earlier.

**Digestion:**

The contents in the tubes were vacuum dried (until gels were not over-dried) by a DNA SpeedVac$^{TM}$ SpeedVac Model DNA110 (Savant). Then 50 µl trypsin solution (5 ng/µl) was added to the tubes. It was ensured that gel-pieces were swollen and submerged in trypsin solution. Then the tubes were incubated in ice for 2 hrs (this step was supposed to allow enough time to enable trypsin to reach proteins inside the gel). After a short centrifuge the tubes were incubated for overnight at 37 °C.

**Collection of Peptide:**

80 µl 0.1% formic acid and 80 µl acetonitrile were added to the tubes and mixed for 20 min. To the supernatant, 200 µl Acetonitrile was added and mixed for 2 min to shrink the gel. Again, the supernatant was collected and whole wash step was repeated as above with formic acid and acetonitrile, until gel is not completely shrunk. Then the collected extract was vacuum dried. Finally, samples were analyzed by mass spectrometry.

## 2.11 LC-MS/MS Configuration:

Nanoflow LC (Liquid Chromatography) of tryptic peptide samples was preformed with an Agilent 1100 nanoflow LC system equipped with a $C_{18}$ pre-column (Zorbax 300SB-C18, 5μm, 5 mm × 0.3 mm, Agilent) and a $C_{18}$ analytical column (Zorbax 300SB-C18, 3.5 μm, 15 cm × 75 μm, Agilent). The aqueous mobile phase (solution A) contained 5% acetonitrile and 0.1% formic acid, and the organic mobile phase (solution B) contained 95% acetonitrile and 0.1% formic acid. Samples (5-μl injected) were loaded and washed on the pre-column for 5 minutes with solution A at 50 ul/min. Peptides were then eluted off the pre-column and through the analytical column with a 125-minute gradient from 1 to 40% solution B, 5-minute gradient from 40% to 95% solution B, and a 5-minute rinse with 95% solution B at a flow rate of 250 nl/min. The columns were immediately re-equilibrated for 10-minutes at initial conditions (100% solution A for the pre-column and 1% solution B for the analytical column). Eluting peptides were injected via nanospray source into a QStar XL Qq-TOF (Applied Biosystems). The ion source was equipped with a 50-μm inner-diameter, fuse-silica needle with a 15-μm tip (PicoTip Emitter, New Objective). Data dependant acquisition was used with a 10 second cycle: 1-second interval for acquiring intact peptide signal (MS), and three 3-second intervals for collision induced dissociation of the 3 most intense peptides signals in the initial 1-second interval (MS/MS). The MS *m/z* range was 350 to 1500, and the MS/MS *m/z* range was 70 to 2000. Collision energy was automatically determined by the data acquisition software (Analyst QS 1.1). MS/MS data was acquired for the entire LC run.

## 2.12 Data analysis:

The Mascot search engine (Matrix Science, London, UK; version 2.1.03) was used to search the NCBInr (20070922) database with the MS/MS data. The search parameters were as follows: taxonomy was restricted to Homo sapiens, protein molecular weight was unrestricted, fixed modification was carbamidomethyl (C), variable modification was Oxidation (M), peptide and fragment mass tolerance was ± 0.4 Da, and up to one missed cleavage was allowed.

## 2.13 Criteria for Protein Identification:

Scaffold (version Scaffold-01_06_19, Proteome Software Inc., Portland, OR) was used to validate MS/MS based peptide and protein identifications. Individual Peptide identifications were accepted if they could be established at greater than 95.0% probability as specified by the Peptide Prophet algorithm (119), while protein identifications were accepted if they could be established at greater than 80% probability and had at least 2 identified peptides. Protein probabilities were assigned by the Protein Prophet algorithm (121). Proteins that contained similar peptides and could not be differentiated based on MS/MS analysis alone were grouped.

## 2.14 Co-immunoprecipitation:

293T cells were co-transfected with specific or relevant expression plasmids and 18 hours before cell lysis 50mM MG-132 (proteasome inhibitor from Calbiochem) was added to each dish. After 48 hours of transfection, cells were lysed by 0.25% NP-40 (0.25% NP-40 in M-199 medium plus protease inhibitor cocktail from Roche) on ice for 30 minutes

(with gentle mixing in every 5 minutes). After centrifugation in 14,000 rpm for 5 minutes, supernatants were treated with glass-beads to ensure mechanical breakage of cell-lysate. After centrifugation in 14,000 rpm at 4˚C, supernatant was pre-cleared by 100 ul sephadex beads by rotating the tubes containing the supernatant at 4˚C for minimum 2 hours. After centrifugation, 0.8 ul anti-GFP antibody (rabbit) was added to each supernatant and was again rotated at 4˚C overnight followed by rotation of supernatant with 100 ul protein A sepharose for 2 hours at 4˚C. Then, the beads were collected by centrifugation at 3000 rpm for 2 minutes at 4˚C. Finally, the beads were washed with 800 ul of lysis buffer for 3 times. For western analysis: co-immunoprecipitated proteins were re-suspended in 30-40 ul of Laemmli buffer (4X) followed by boiling of the tubes for 5 minutes at 100˚C and centrifuge at 7000 rpm for 3-5 minutes. Then the supernatant was run on 12.5 % SDS-PAGE gel.

### 2.15 Western blot:

Western blot was done to detect the TAP-purified proteins or the co-immunoprecipitated proteins. Having resolved on an SDS-PAGE gel, the proteins were transferred for over-night onto a nitrocellulose membrane using transfer buffer (20% methanol, tris and glycine). After blocking in 5% skim milk for 1 hour at room temperature and subsequent wash in 1X PBS, the membrane was rotated with relevant primary antibody [e.g. anti-integrase 1: 4000 (rabbit) (Cat. no. 757, obtained through the AIDS Research Reference Reagent Program, Division of AIDS, NIAID, National Institutes of Health.), anti-T7 1:5000 (mouse, Novagen, Cat. # 69522-3), anti-SF3A3 1:2000 (mouse, Abcam, Cat. # ab56823), anti-GFP 1:4000 (rabbit, Invitrogen, A6455), anti-γ-actin 1:1000 (rabbit,

Abcam, ab84479), anti-β-tubulin 1:10,000 (rabbit, Abcam, Cat. # 15568), anti-histone H1 1:250 (mouse, Abcam, Cat. # ab4269), anti-histone H4 1:1000 (mouse, Abcam, ab17036) etc] for overnight. After washing four times with PBS for 15 minutes each, the membrane was rotated with relevant HRP-conjugated secondary antibody for 1 hour at concentrations specified by manufacturers. Finally, the membrane was washed again four times followed by incubation with Western Lightning$^{TM}$ Chemiluminescence Reagent Plus (PerkinElmer, Cat. No. NEL103) according to the manufacturers protocol. The film was exposed for 10 seconds to 20 minutes (depending on the protein of interest) and was developed in a developer machine (Bio-Rad, Gel Doc 2000) to check the expression and/or binding of desired protein bands.

### 2.16 Establishment of shRNA-mediated knockout cell line:

**Transfection:**

1.3 X $10^6$ 293T cells were cultured in small dishes and after 50-60 % of confluency cells were co-transfected with plasmids expressing VSV-G (vesicular stomatitis virus G) glycoprotein, Δ8.2 (Gag-Pol), respectively and also with lentiviral vector pLKO.1 expressing shRNA against SF3A3 (from Open Biosystems; 1$^{st}$ clone ID TRCN0000050533, sequence:5′-CCGGGCCAACACG TAGACAGACTATCTCGAGA TAGTCTGTCTACGTGTTGGCTTTTTG-3′, 2$^{nd}$ clone ID TRCN0000050534,sequence: CCGGCCAAGGTTCCATTAAGTGAATCTCGAGATTCACTTAATGGAACCTTGGT TTTTG, 3$^{rd}$ clone ID TRCN0000050535, sequence: CCGGCGTGCATACCTATGGCTT GATCTCGAGATCAAGCCATAGGTATGCACGTTTTTG). Supernatant, containing viral-like particles (VLPs), was collected 48 hours and 72 hours post-transfection. After

clearing the supernatant by centrifuging for 15 minutes in 3000 rpm at 4˚C, the VLPs were pelleted from supernatant by ultra centrifuge in 35,000 rpm for 90 minutes at 4˚C. Then the viral-like pellets were re-suspended in 1 ml of DMEM (Dulbecco's Modified Eagles Medium from Thermo Scientific) and stored in -80˚C.

**Transduction:**

0.5 X $10^6$ HeLa cells were cultured in 6-well plate and after 50-60% of confluency, 1.5 ml media was removed from the wells and cells were transduced with 100 ul of VLP (100 pmol).

**Selection with puromycin and check for efficient knock down:**

After adequate incubation (usually more than 24 hours) puromycin was added in all wells (initially 2 μg/ml of media, then 1 ug/ml). If cells were dead in the control well, some cells, from treatment wells, were lysed and establishment of SF3A3-knockout cell line (i.e. absence of SF3A3) was checked by western blot. Rest of the cells in treatment wells were continued to culture for luciferase assay.

## 2.17 Luciferase assay:

In 293T cells, a single-cycle replicating virus, with VSV-G pseudotyped, was produced as described previously (reviewed in 113). Briefly, 293T cells were co-transfected with an RT/IN/Env-deleted HIV-1 provirus NLlucΔBglΔRI, each CMV-Vpr-RT-IN (wt/mutant) expression plasmid and a VSV-G expression plasmid. After 48 h of transfection, viruses were collected and concentrated from the supernatants by

ultracentrifugation at 35,000 RPM for 2 h. Virus titers were quantified using HIV-1 p24 Antigen Capture Assay Kit (purchased from the NCIFrederick AIDS Vaccine Program). Equal amounts of viruses (adjusted by virion-associated p24 levels) were used to infect SF3A3-knockout HeLa cells overnight at 37°C. At 48 h post-infection, $1 \times 10^6$ cells from each sample were collected and lysed with 50 μL of luciferase lysis buffer (Fisher Scientific Inc). A 10 μL aliquot of cell lysate was subjected to the luciferase assay by using a POLARstar OPTIMA (BMG LABTECH, Germany), and the luciferase activity was valued as relative light units (RLU).

### 2.17.1 Statistical Analysis:

A two-tailed, two-sample Student t test was used to calculate p values for differences in means between two groups. The data are expressed as means ± SEM. For statistical inference, a p-value of 0.05 was considered significant.

CHAPTER III

# RESULTS

## 3.1 Generation and efficiency testing of TAP-IN expressing vectors :

Three expression vectors were constructed by fusing HIV-1 IN cDNA to the 3′ end of TAP tag in each respective plasmid: (a) pYEF-1-TAP-IN, (b) pNTAP-B-IN and (c) pMSCV-hyg-IN. To check the ability and efficiency of these expression vectors to express TAP-IN, 293T cells were transfected with these expression vectors separately. 48 hours after transfection cells were lysed with 0.5% NP-40 and was subjected to direct lysis method as described in "Materials and Methods" section. Another CMV-YFP (expressing yellow fluorescent protein) vector was used as positive control to check the efficiency of transfection. Western blot analysis with anti-Integrase antibody (rabbit, 1:2000) showed that only pYEF-1-TAP-IN could express the Tap-tagged Integrase having a molecular weight of 40 kDa (Figure 4). The observed band of TAP-IN was too strong, may be, due to too higher concentration of anti-integrase antibody and inadequate washing during western blot. The other expression vectors, i.e. pNTAP-B-IN and pMSCV-hyg-TAP, failed to express TAP-IN as was evident from western blot analysis (Fig.4). Later, a codon-optimized pYEF-1-TAP-IN was constructed and used in subsequent pull-down experiment using the tandem affinity purification system.

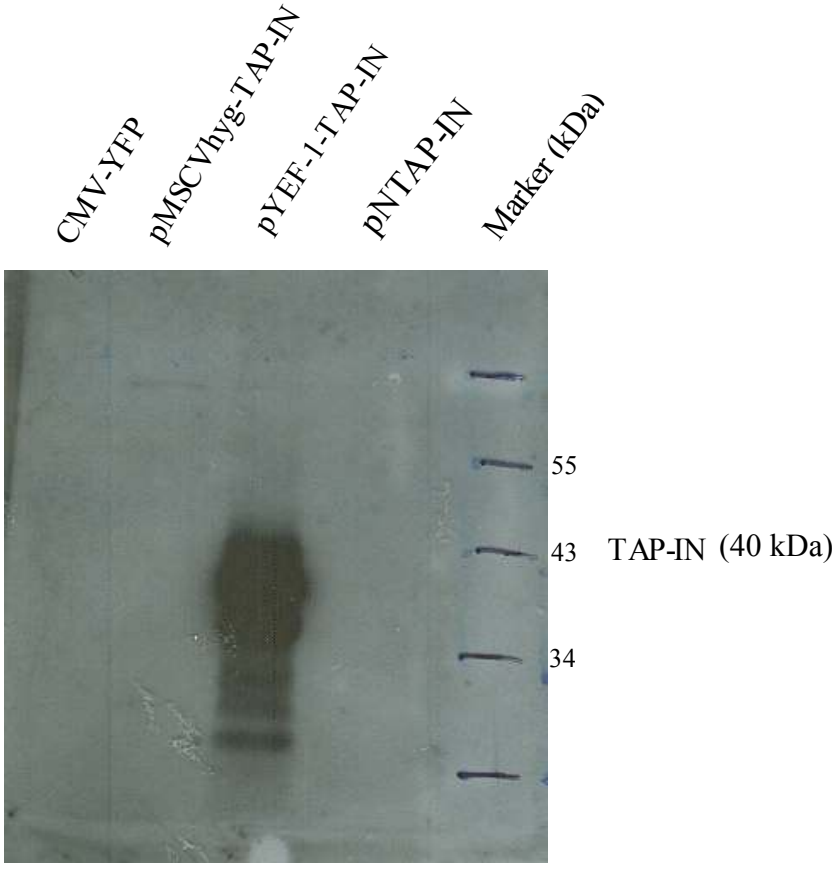
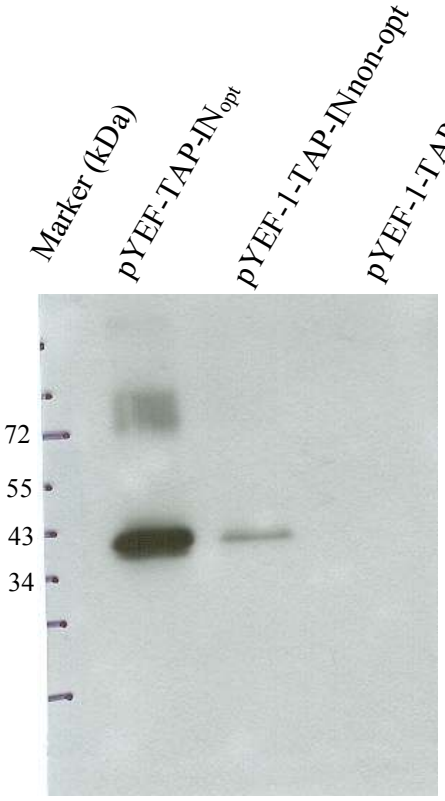

CMV-YFP
pMSCVhyg-TAP-IN
pYEF-1-TAP-IN
pNTAP-IN
Marker (kDa)
55
43
34
TAP-IN (40 kDa)

**Fig. 4: Western blot analysis to determine suitable expression vector expressing TAP-IN:** 293T cells were transfected separately with 5 ug of each vector and 48 hours post-transfection cells were lysed with 0.5% NP-40. During western blot the membrane was probed with anti-IN antibody. It is clearly evident that only pYEF-1-TAP-IN showed expression of TAP-IN, while other vectors failed to express. CMV-YFP was used as positive control to check efficiency of expression.

### 3.2 Codon-optimized HIV-1 IN cDNA can confer robust expression of IN:

After observing expression only in pYEF-1-TAP-IN, we presumed that codon-optimized HIV-1 IN (Integrase) cDNA would have better expression of Integrase protein and so better interactions would occur between HIV-1 IN (Integrase) and host cell factors. We used a codon optimized (i.e. humanized) HIV-1 IN (integrase) (codon optimization was done by Genscript) supplied in pUC57-IN. IN cDNA from this plasmid was cut out with *BamH*1 and *Xho*1 and subcloned in pNTAP to construct pNTAP-$IN_{opt}$. From pNTAP-B-$IN_{opt}$ vector TAP-IN sequence was cut out with *Nhe*1 and *Xho*1 and was subcloned in pEF-1-MCS-Puro ultimately to construct pYEF-1-TAP-$IN_{opt}$.

Then 293T cells were transfected with pYEF-1-TAP-$IN_{opt}$ (IN was optimized), pYEF-1-TAP-$IN_{non\text{-}opt}$ (IN was non-optimized) and pYEF-1-TAP separately. 48 hours after transfection, cells were lysed with TAP lysis buffer using a lysis protocol outlined in Materials and Methods section. After electrophoresis on a 12.5% SDS-PAGE gel, proteins were probed by western blot using anti-IN (rabbit, 1:5000) antibody. Optimized IN showed strong expression of IN, while non-optimized IN showed very weak expression (Fig. 5). This experiment validates use of codon-optimized Integrase that should allow more efficient and broader range of interactions between integrase and host cell factors.


Marker (kDa)
pYEF-TAP-INopt
pYEF-1-TAP-INnon-opt
pYEF-1-TAP
72
55
43
34

W.B: Anti-IN

**Fig. 5: Western blot analysis showing robust expression of codon-optimized IN:** 293T cells were transfected with corresponding plasmid vectors (as shown in the picture) and 48 hours after transfection cells were lysed by TAP lysis buffer. During western blot, codon-optimized TAP-IN showed robust expression, while non-optimized TAP-IN showed poorer expression. pYEF-1-TAP was used as a negative control.

### 3.3 To set up optimum conditions for Tandem Affinity Purification (TAP) system to identify new cellular binding partners of IN.

The next objective was to optimize empirical conditions for the tandem affinity purification (TAP) system to study protein-protein interactions between HIV-1 IN and host cell factors. After transfecting 293T cells with pYEF-1-TAP-$IN_{opt}$ (that showed strong expression of IN), collection and lysis of cells , extraction and purification of protein complexes (all steps were done according to the TAP protocol mentioned in the Material and Methods section); we ran samples from different stages of the protocol e.g. samples (both streptavidin beads and supernatant) after elution with streptavidin elution buffer, samples (both calmodulin beads and final supernatant) after elution with calmodulin elution buffer. But for repeated times, after silver/Coomassie staining of SDS-PAGE gel, we got no bands from samples from calmodulin beads and or the final supernatant and got protein bands only from streptavidin beads and relevant supernatant (Fig. 6). So, it was determined that treatment of the cell lysate with streptavidin beads was enough for such study and although, many protein bands were commonly present in both lanes, only those protein bands were considered, to be sent for mass spectrometry, which were present only in pYEF-1-TAP-IN lane but not in pYEF-1-TAP (i.e. control lane). Subsequent in-gel trypsin digestion and mass spectrometry revealed presence of previously reported as well as new but putative cellular interacting partners of HIV-1 integrase (Table 1).

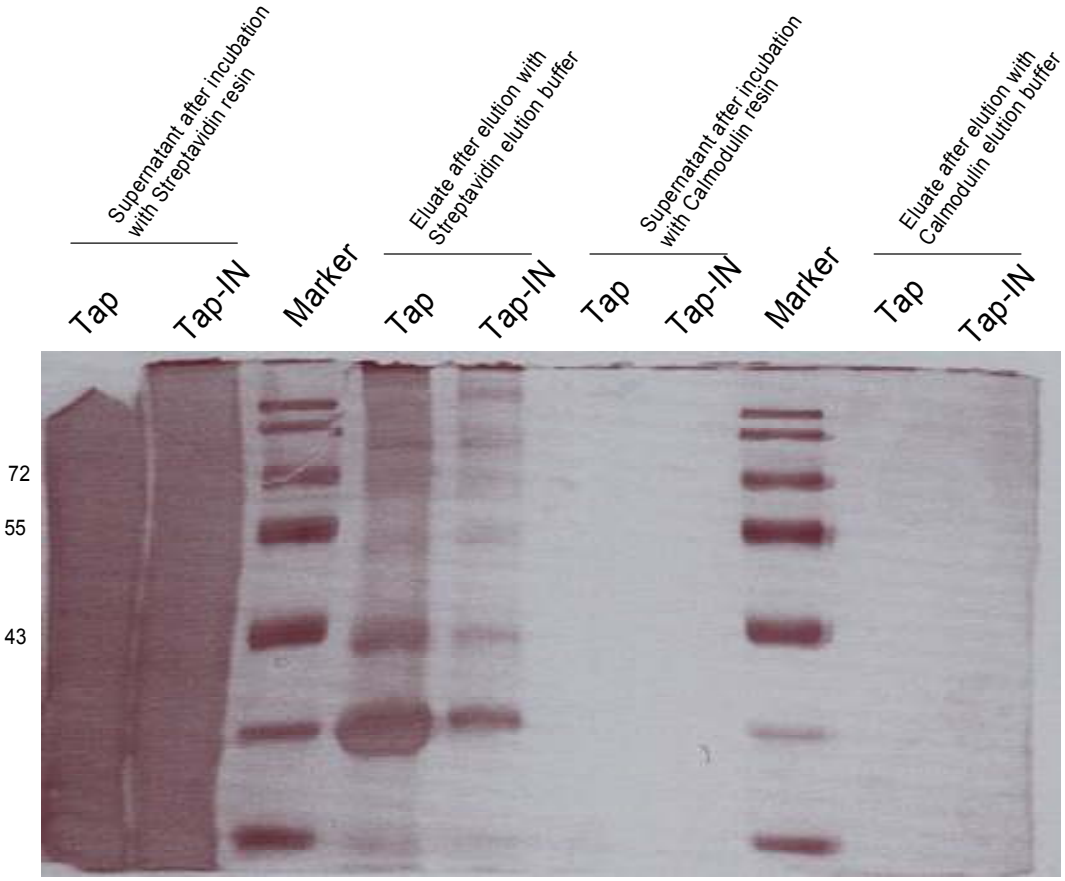

Supernatant after incubation with Streptavidin resin
Eluate after elution with Streptavidin elution buffer
Supernatant after incubation with Calmodulin resin
Eluate after elution with Calmodulin elution buffer
Tap
Tap-IN
Marker
Tap
Tap-IN
Tap
Tap-IN
Marker
Tap
Tap-IN
72
55
43

**Fig. 6: Silver-stained gel shows presence/absence of TAP-purified protein samples in different fractions of the TAP system:** 293T cells were separately transfected with pYEF-1-TAP-$IN_{opt}$ and pYEF-1-TAP, lysed by TAP lysis buffer and interacting proteins were extracted and purified after step-wise incubation with streptavidin beads, washing, elution, again incubation with calmodulin beads, washing and elution – as per manufacturer's protocol. Finally, extracted samples were run on a 12.5 % SDS-PAGE gel and the gel was stained by silver staining protocol. Only eluate after elution with streptavidin elution buffer showed some visible, distinctive protein bands. (TAP = pYEF-1-TAP, TAP-IN = pYEF-1-TAP-$IN_{opt}$).

### 3.4 Silver staining and Coomassie staining of TAP-purified protein samples:

Having treated the cell-lysate by modified Tandem Affinity Purification (TAP) system, the final sample supernatant was run on NuPAGE® Novex 4-12% Bis-Tris gel. After silver staining (Fig.7) or Coomassie staining (Fig. 8) of the gel as described in the materials and methods section, a good no. of prospective bands were found in pYEF-1-TAP-IN lane but not in pYEF-1-TAP lane.

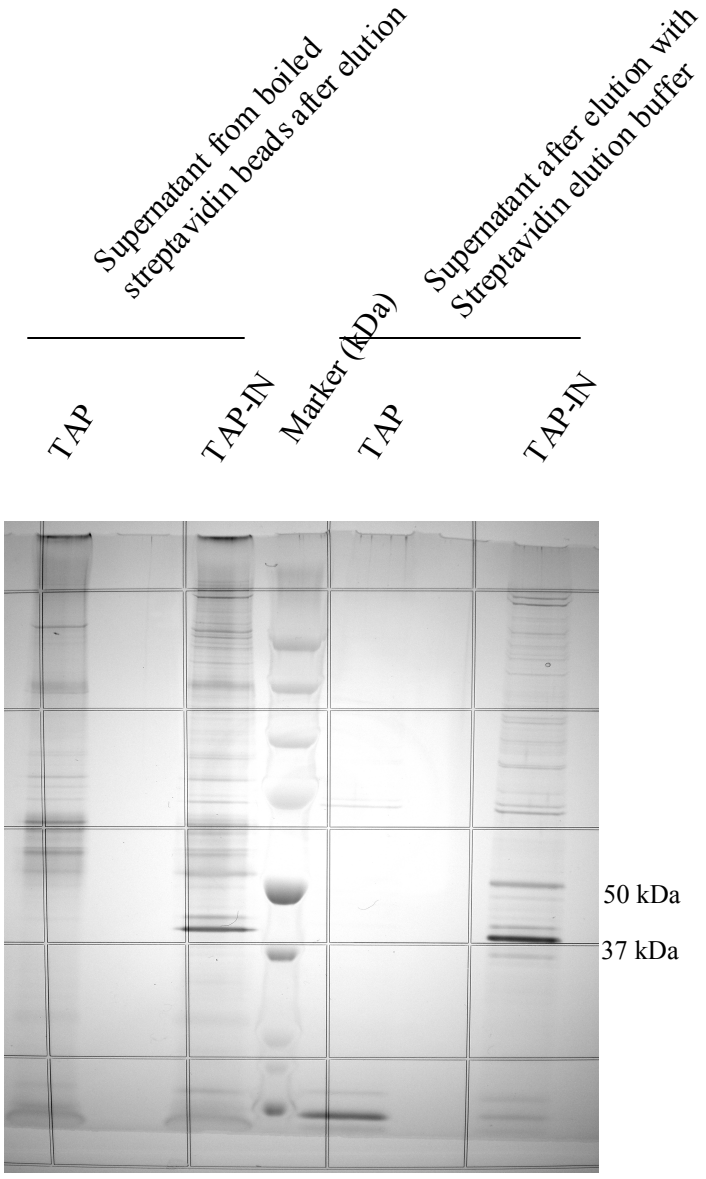

Supernatant from boiled streptavidin beads after elution
Supernatant after elution with Streptavidin elution buffer
TAP
TAP-IN
Marker (kDa)
TAP
TAP-IN
50 kDa
37 kDa

**Fig. 7: Silver stained gel shows better resolution of the TAP-purified protein complexes.** 293T cells were separately transfected with pYEF-1-TAP-$IN_{opt}$ and pYEF-1-TAP, lysed by TAP lysis buffer and interacting proteins were extracted and purified by incubating only with streptavidin beads, after washing and then eluting with streptavidin elution buffer. After elution: streptavidin beads and final supernatant (i.e. eluate) were run on a NuPAGE® Novex 4-12% Bis-Tris gel. After silver staining of the gel as described in the materials and methods section, a good no. of prospective bands were seen in TAP-IN lane but not in TAP lane. (TAP = pYEF-1-TAP, TAP-IN = pYEF-1-TAP-$IN_{opt}$)

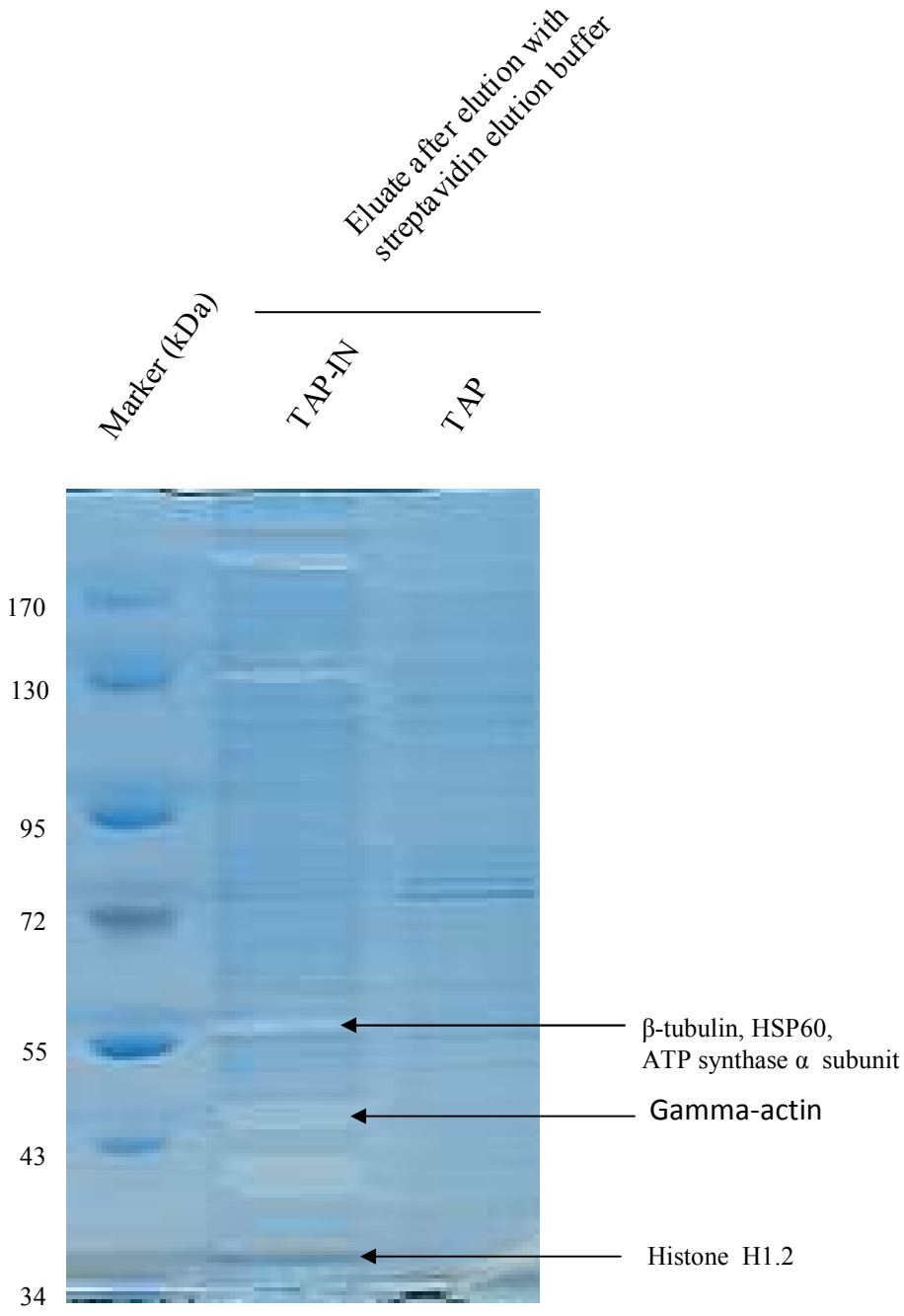

Eluate after elution with streptavidin elution buffer
Marker (kDa)
TAP-IN
TAP
170
130
95
72
55
43
34
β-tubulin, HSP60,
ATP synthase α subunit
Gamma-actin
Histone H1.2

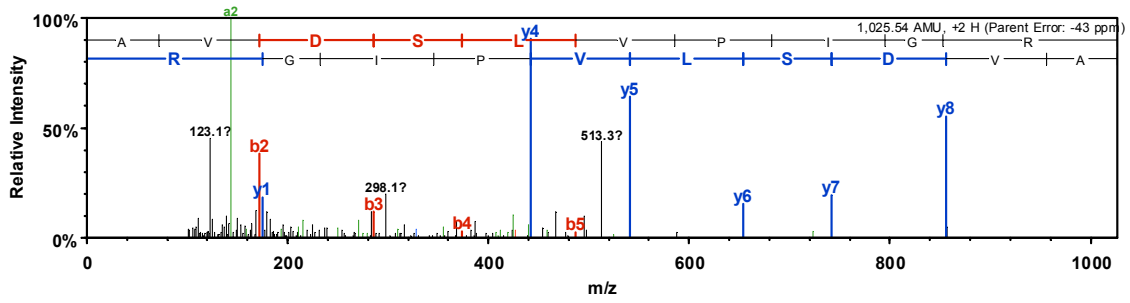
**Fig. 8: Coomassie-stained NuPage™(4-12%) Bis-Tris gel shows pulled-down putative HIV-1 IN interactors.** 293T cells were separately transfected with pYEF-1-TAP-$IN_{opt}$ and pYEF-1-TAP, lysed by TAP lysis buffer and interacting proteins were extracted and purified by incubating only with streptavidin beads, after washing and then after eluting with streptavidin elution buffer. After elution: streptavidin beads and final supernatant (i.e. eluate) were run on a NuPAGE® Novex 4-12% Bis-Tris gel. After Coomassie staining of the gel as described in the materials and methods section, a good no. of prospective bands were seen in TAP-IN lane but not in TAP lane. Putative bands were cut, in-gel digested and subjected to mass spectrometry. (TAP = pYEF-1-TAP, TAP-IN = pYEF-1-TAP-$IN_{opt}$). Of note: accidentally, prospective gel bands were cut first, then photograph was taken.

### 3.5 Identification of new putative cellular partners of HIV-1 Integrase (IN) by mass spectrometry:

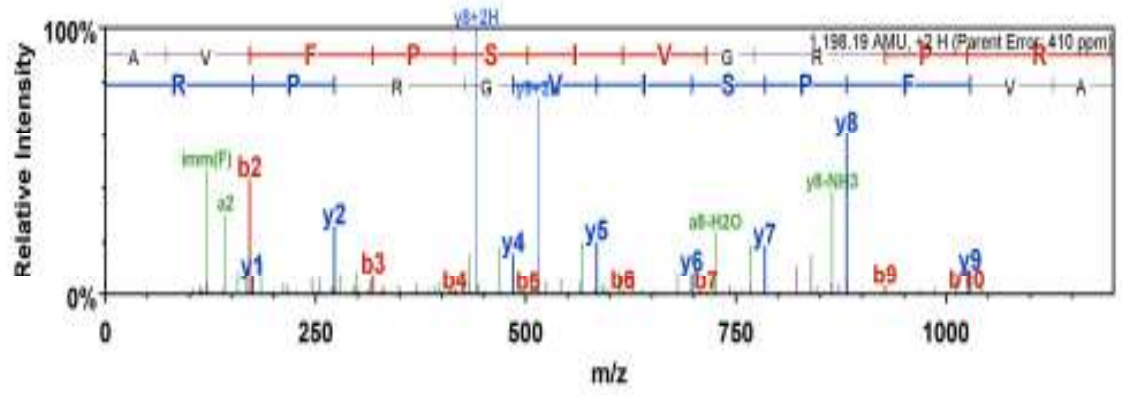
After Coomassie staining of the gel, putative bands were cut out and were exposed to in-gel trypsin digestion. Finally, the digested proteins from gel-bands were subjected to mass spectrometry [LC-MS/MS, QStar XL Qq-TOF (Applied Biosystems)]. From the mass spectrometric data we got two previously-identified cellular interacting partners of HIV-1 IN: (i) Chaperonin (Heat Shock Protein 60 or HSP60) (*Homo sapiens*) (ii) β-tubulin (*Homo sapiens*). Identification of previously reported cellular co-factors of IN validates the overall modified TAP system.

However, the most important observation from the mass spectrometric data was identification of three new but putative cellular co-factors of HIV-1 Integrase (IN) : (i) ATP synthase, H+ transporting, mitochondrial F1 complex, alpha subunit 1, cardiacmuscle [*Homo sapiens*] (ii) Histone H1.2 (*Homo sapiens*) and (iii) Gamma-actin [*Homo sapiens*]. The proteins identified by mass spectrometry are summarized in the table 1 as well as the mass spectra are given below :

A


1,025.54 AMU, +2 H (Parent Error: -43 ppm)
Relative Intensity
m/z
123.1?
298.1?
513.3?

B


Relative Intensity
m/z

C

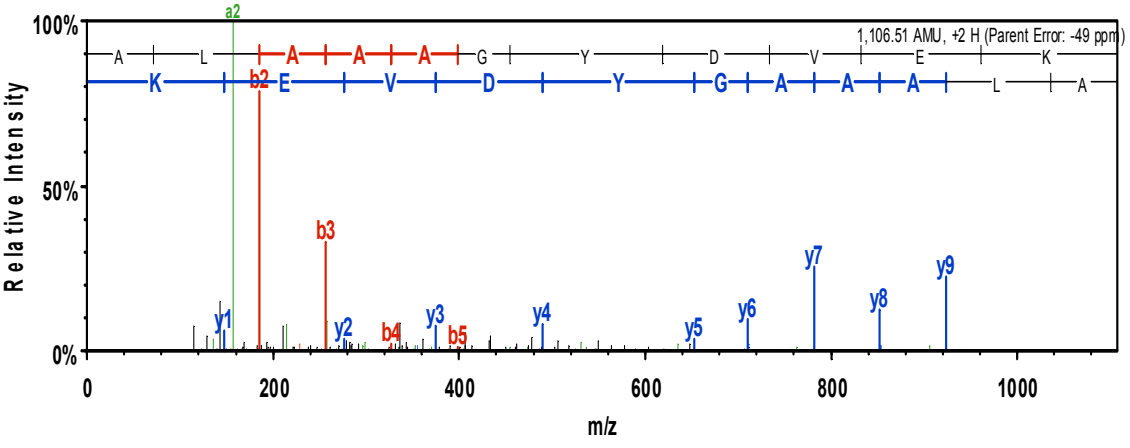

1,106.51 AMU, +2 H (Parent Error: -49 ppm)
Relative Intensity
m/z

**Fig. 9: Representative LC/MS/MS data for host interacting partners of HIV-1 IN.** (A) ATP synthase, H+ transporting, mitochondrial F1 complex, alpha subunit 1 peptide AVDSLVPGR (B) Gamma actin peptide AVFPSVGRPR (C) Histone H1.2 peptide ALAAAGYDVEK. The 'b" and "y" ion series provide amino acid sequence information. The b-ion series (shown in red) is read from the N-terminus to C-terminus, while the y-ion series (shown in blue) is read from the C-terminus to N-terminus, providing thus complementary sequence information. Other minor fragments resulted from peptide fragmentations at other sites are shown in green.

| # | Data File Name & Link to Mascot Results | Top Listed Hit | Description (*Homo sapiens*) | Mass | Score | Total protein hits | Total unique peptides | Previously reported ? |
|---|---|---|---|---|---|---|---|---|
| 1 | 20091222-004-0043-00412+ gel y1 | gi\|306890 | Chaperonin (HSP60) | 61157 | 642 | 23 | 42 | Yes (Parissi et al. 2001) |
| 2 | 20091222-005-0043-00412+ gel y2 | gi\|127798841 | ATP synthase, H+ transporting, mitochondrial F1 complex, alpha subunit 1, cardiac-muscle | 59785 | 398 | 7 | 14 | No |
| 3 | 20091222-006-0043-00412+ gel y3 | gi\|196167663 | HIV-1 IN | 31598 | 481 | 64 | 71 | -- |
| 4 | 20091222-007-0043-00412+ gel d1 | gi\|4885375 | Histone H1.2 | 10359 | 422 | 10 | 23 | No |
| 5 | 20091222-011-0043-00412+ gel d5 | gi\|4501887 | Gamma-actin | 42151 | 745 | 106 | 118 | No |
| 6 | 20091222-012-0043-00412+ gel d6 | gi\|29788785 | Beta-tubulin | 50095 | 1040 | 160 | 151 | Yes (Turlure et al. 2004) |

**Table 1: Putative interacting partners of HIV-1 Integrase (IN) identified by mass spectrometry:** Mass spec data analysis was done using the Mascot search engine (Matrix Science, London, UK; version 2.1.03). Note that chaperonin (HSP60) and β-tubulin were reported earlier, while the others were not, making them new, putative candidate as cellular cofactor for HIV-1 IN.

### 3.6 Identification of HIV-1 Integrase (IN) itself in TAP-purified protein sample:

To check the expression of HIV-1 Integrase (IN) itself in TAP-purified protein samples, TAP-purified protein samples were run on an SDS-PAGE gel and probed with anti-Integrase antibody (rabbit, 1: 4000) by western blot analysis. Western blot analysis revealed the presence of TAP-tagged HIV-1 Integrase (IN) in the TAP-purified protein complex in a position of 40 kDa (32 kDa Integrase plus 8 kDa TAP tag) (Fig. 10). This identification clearly proved that the cellular factors identified by mass spectrometry were cellular co-factors of HIV-1 Integrase (IN).

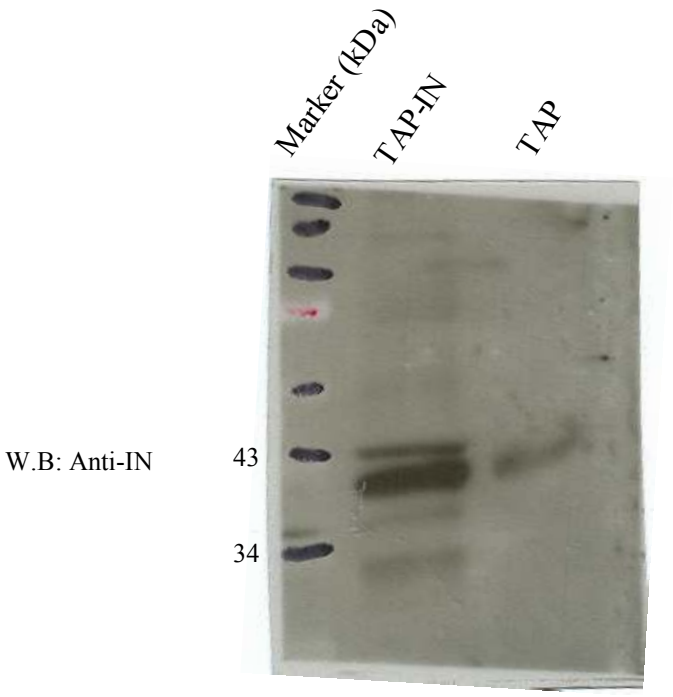
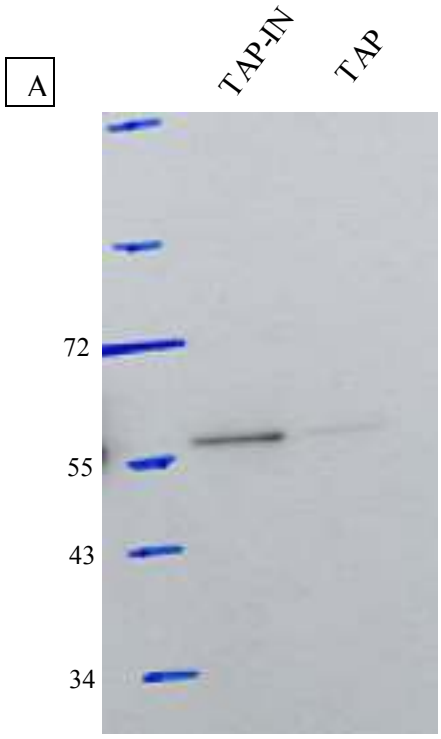

Marker (kDa)
TAP-IN
TAP
W.B: Anti-IN
43
34

**Fig. 10: Western blot analysis can detect HIV-1 IN itself in TAP-purified protein complex:** 293T cells were transfected with pYEF-1-TAP-$IN_{opt}$, lysed by TAP lysis buffer and interacting proteins were extracted and purified by incubating only with streptavidin beads, after washing and then after eluting with streptavidin elution buffer. After elution: streptavidin beads and final supernatant (i.e. eluate) were run on a NuPAGE® Novex 4-12% Bis-Tris gel. In western blot analysis, the membrane was probed with anti-IN (rabbit) antibody which showed IN band in TAP-IN lane but not in TAP lane. (TAP = pYEF-1-TAP, TAP-IN = pYEF-1-TAP-$IN_{opt}$).

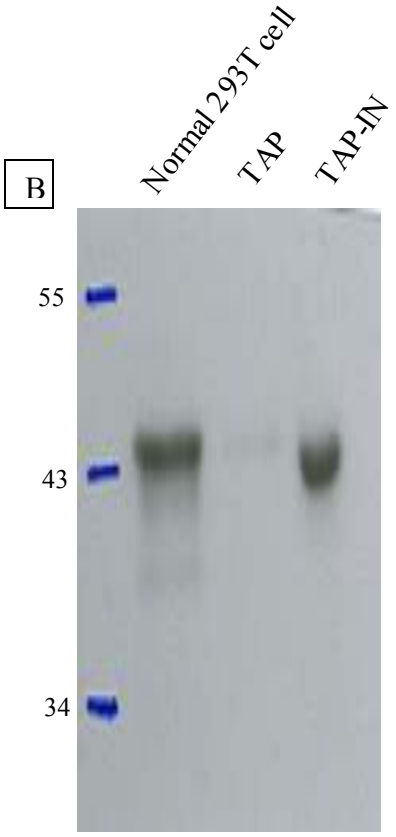
### 3.7 Identification of gamma-actin and β-tubulin in TAP-purified protein sample:

From tandem affinity purification experiment coupled with mass spectrometry, it has been shown that the new approach could identify totally five host cell factors. Among them two are novel: (i) Histone H1.2 and (ii) ATP synthase alpha subunit 1. The remaining three, previously reported by different groups, were: (i) Chaperonin (HSP60) (ii) Gamma-actin and (iii) β-tubulin.

The next step was to verify the interaction of these endogenous factors, with HIV-1 IN, in TAP-purified protein samples. The samples were run on an SDS-PAGE gel and by western blot they were probed with relevant antibodies i.e. anti-HSP60 (Santa Cruz Biotechnology, Inc., conc. 1:2000), anti-gamma-actin and anti-beta-tubulin. Although, chaperonin (HSP60) was not identified in the TAP-purified samples by western blot (data not shown), beta-tubulin (Fig. 11A) and gamma-actin (Fig. 11B) were identified. These results confirm previous reports that cellular beta-tubulin and gamma-actin interact with HIV-1 IN.




W.B: Anti-ß-tubulin




W.B: Anti- γ-actin

**Fig. 11: Western blot can detect (A) β-tubulin and (B) γ-actin in TAP-purified protein complex:** 293T cells were transfected with pYEF-1-TAP-$IN_{opt}$ and pYEF-1-TAP, lysed by TAP lysis buffer and interacting proteins were extracted and purified by incubating only with streptavidin beads, after washing and then eluting with streptavidin elution buffer. After elution: streptavidin beads and final supernatant (i.e. eluate) were run on a NuPAGE® Novex 4-12% Bis-Tris gel. During western blot, membrane was probed with relevant antibodies revealing the fact that TAP-IN can pull down β-tubulin (in A) and γ-actin (in B) but only TAP can not. (TAP = pYEF-1-TAP, TAP-IN = pYEF-1-TAP-$IN_{opt}$).

### 3.8 Identification of SF3A3 in TAP-purified protein complex :

Using yeast two-hybrid system and *in vitro* binding systems, Studamire et al (2008) has showed that a yeast splicing factor SF3A3 is an interacting partner of HIV-1 Integrase. It suggests that the human homologue of this yeast protein might also interact with HIV-1 Integrase. To test this possibility, a western blot analysis of the TAP-purified protein complex with human anti-SF3A3 antibody was performed and the results revealed that a human SF3A3 protein was co-pulled down with TAP-IN fusion protein, but not with TAP alone (Fig. 12). This observation indicated that not only yeast SF3A3 but also human SF3A3 is a host cell partner of HIV-1 Integrase.

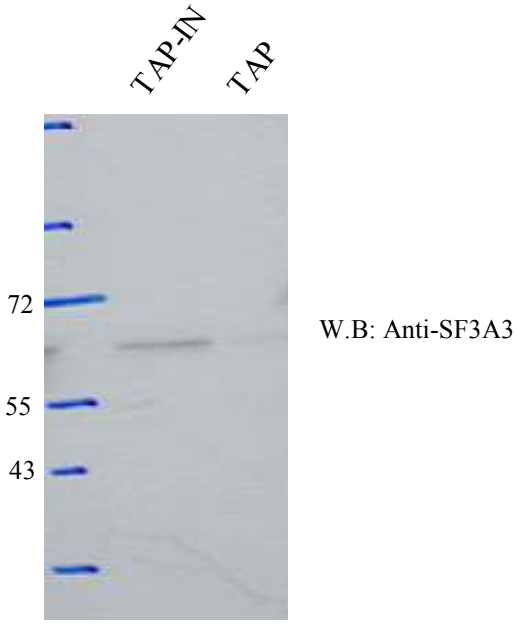
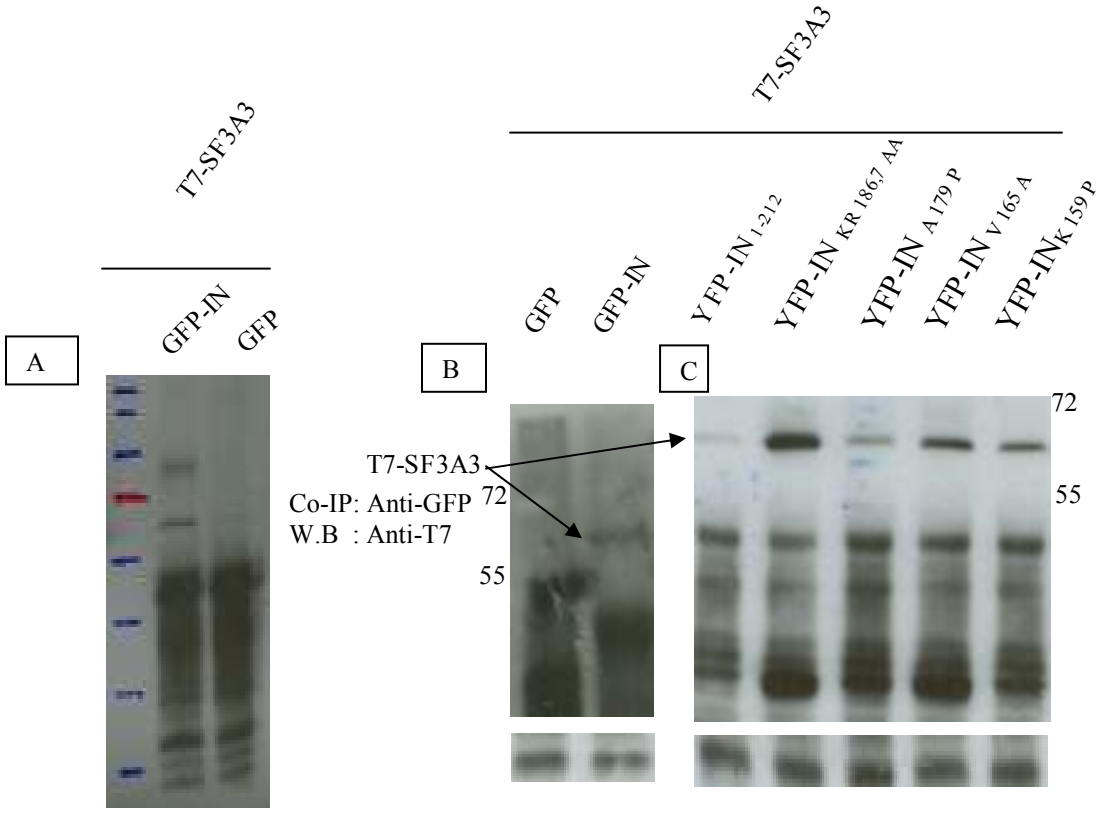

TAP-IN
TAP
72
55
43
W.B: Anti-SF3A3

**Fig. 12: Western blot analysis can detect human splicing factor SF3A3 in TAP-purified protein complex:** 293T cells were transfected with pYEF-1-TAP-$IN_{opt}$ and pYEF-1-TAP, lysed by TAP lysis buffer and interacting proteins were extracted and purified by incubating only with streptavidin beads, after washing and then eluting with streptavidin elution buffer. After elution: streptavidin beads and final supernatant (i.e. eluate) were run on a NuPAGE® Novex 4-12% Bis-Tris gel. During western blot, membrane was probed with anti-SF3A3 antibody showing the presence of SF3A3 in TAP-purified protein complex. (TAP = pYEF-1-TAP, TAP-IN = pYEF-1-TAP-$IN_{opt}$).

### 3.9 Proving the interaction between HIV-1 IN and SF3A3 by *in vivo* co-immunoprecipitation assay:

An *in vivo* co-immunoprecipitation assay was used to confirm the possible interaction between HIV-1 Integrase and human splicing factor SF3A3. 293T cells were co-transfected with GFP-IN (a reporter gene sequence of green fluorescent protein, GFP, was fused with HIV-1 Integrase sequence in a plasmid) and T7-SF3A3 (a plasmid that encodes a T7 tag fused with SF3A3 on the N-terminal). 48 hours after co-transfection, cells were lysed with 0.25 % NP-40 and cell lysate was subjected to an *in vivo* co-immunoprecipitation assay as described in Materials and Methods section. In the western blot analysis, it was found that GFP-IN could pull down T7-SF3A3 while only GFP could not (Fig. 13A).




GFP-IN can co-immu-noprecipitate T7-SF3A3

**Fig. 13: *In vivo* co-immunoprecipitation assay shows interaction between human SF3A3 and HIV-1 IN:** (A) GFP-IN can interact and pull-down T7-SF3A3. 293T cells were cotransfected with T7SF3A3 and GFP-IN or GFP. 48 hours after transfection, cells were lysed by 0.1% NP-40. After treatment with glass-beads, supernatant was pre-cleared with protein A sephadex beads with subsequent incubation with anti-GFP antibody. Finally, the mixture was incubated with protein A sepharose. After adequate washing, the supernatant was run on 12.5 % SDS-PAGE gel and in western blot the membrane was probed with anti-T7 antibody showing the presence of T7-SF3A3 in the range of app. 60 kDa in GFP-IN lane. (B) and (C) Co-IP experiment defines the binding region: Same Co-IP protocol was applied as mentioned in (A). Upper panels show that wild-type IN can but C-terminal deletion of IN (YFP-$IN_{1-212}$) can not pull down SF3A3 revealing that the interaction region is C-terminal domain of IN. (C) also shows that mutations that disrupt interaction with LEDGF/p75 have no effect on interaction with SF3A3. Lower panels show expression of YFP in each lane (YFP and GFP are same protein with a few amino acids change in GFP and so its fluorescence is comparatively stronger). (GFP = pAc-GFP-C, GFP-IN = pAc-GFP-C-$IN_{wt}$).

### 3.10 Defining the binding region of HIV-1 Integrase and role of different mutants of HIV-1 IN required for interaction with SF3A3:

The next step was to delineate the essential regions or amino acids in HIV-1 IN required for it's interaction with SF3A3. Different HIV-1 Integrase mutant expressors were constructed by fusing each of different IN mutant to C-terminal of YFP (yellow fluorescent protein) (e.g. K159P, V165A, A179P, KR186, 7AA) or fusing IN mutant YFP-IN1-212 to a C-terminus of YFP. Mutant constructs were co-transfected with T7-SF3A3 in 293T cells. 48 hours after co-transfection, cells were lysed and cell lysate were subjected to *in vivo* co-immunoprecipitation assay as described in Materials and Methods section. In the western blot, it was found that wild type YFP-IN could pull down T7-SF3A3 (Fig. 13B) while C-terminal deletion of HIV-I IN (i.e. YFP-$IN_{1-212}$) could not (Fig. 13C). Besides, point mutations in HIV-1 Integrase had roughly no negative effects on interaction between SF3A3 and HIV-1 IN and moreover, a mutant i.e. YFP-$IN_{KR186,7AA}$ showed more interaction with SF3A3 as was evident by a brighter band in the gel (Fig. 13C).

### 3.11 Determining the effect of SF3A3-knockdown on HIV-1 replication:

To check the effect of SF3A3-knockdown on HIV-1 replication by shRNA-mediated knockdown, a HeLa cell line that did not express SF3A3 was established. Then these HeLa cells were transduced with HIV-1 virus-like particles that harbor a luciferase reporter gene (as described in Materials and Methods). Then transduced cells were lysed and subjected to luciferase assay. The luciferase assay showed that while HeLa cells treated with scrambled shRNA had no effect on HIV-1 replication but cells treated with SF3A3-specific shRNA (e.g. SF3A3-knokedout) showed more than twenty times inhibition in viral replication (Fig. 14B). This data suggested that SF3A3 is necessary for HIV-1 replication.

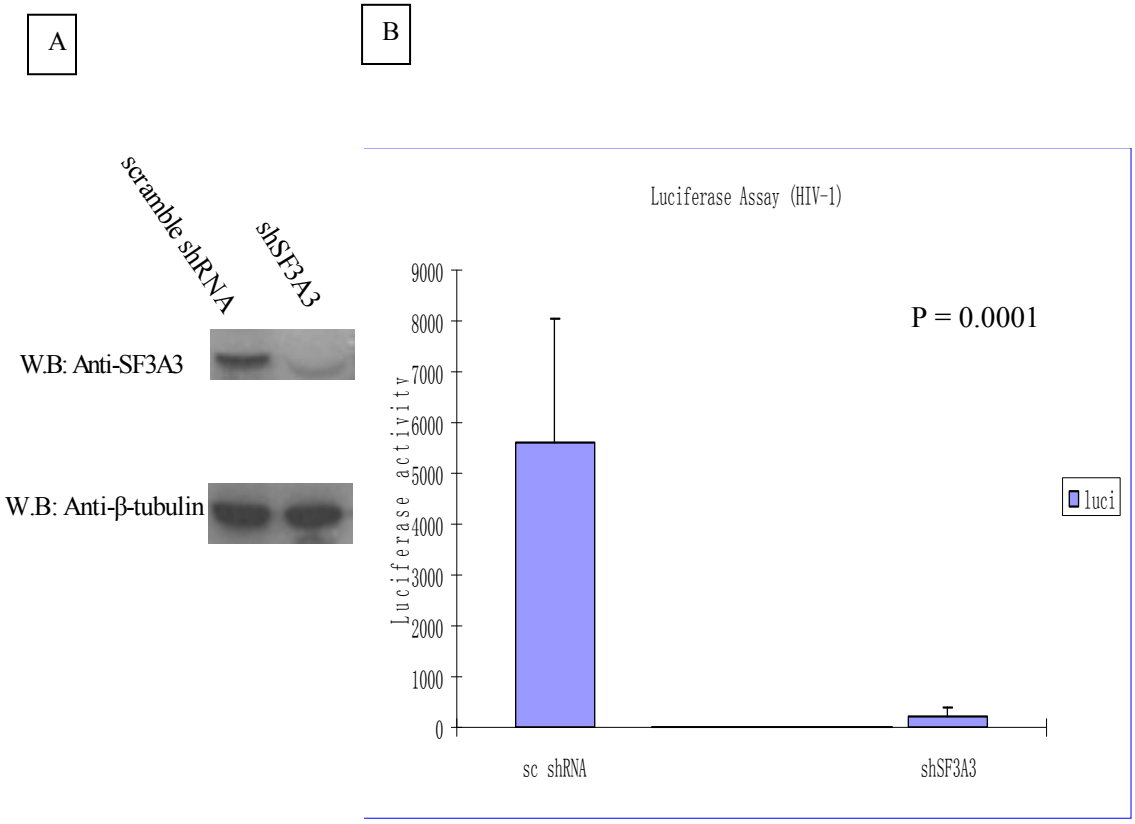

A
B
scramble shRNA
shSF3A3
W.B: Anti-SF3A3
W.B: Anti-β-tubulin
Luciferase Assay (HIV-1)
P = 0.0001
Luciferase activity
9000
8000
7000
6000
5000
4000
3000
2000
1000
0
sc shRNA
shSF3A3
luci

**Fig. 14: (A) ShRNA mediated knock-down of SAF3A3 in HeLa cells:** VLP (virus-like particles) was generated by co-transfecting 293T cells with VSV-G, Δ8.2 (Gag-Pol) and SF3A3-specific shRNA. VLP was transduced in HeLa cells, selected in puromycin, then live cells were cultured and SF3A3-knockdown was checked by western blot using anti-SF3A3 antibody (upper panel). Lower panel shows β-tubulin was not disrupted by shRNA-mediated silencing of SF3A3. (B) Luciferase assay shows much lower luciferase activity in SF3A3-knockout HeLa cells: SF3A3-knockout HeLa cells were transfected with HIV-1 provirus NLlucΔBglΔRI. At 48 h post-infection, $1 \times 10^6$ cells from each sample were collected and lysed with 50 μL of luciferase lysis buffer (Fisher Scientific Inc). A 10 μL aliquot of cell lysate was subjected to the luciferase assay by using a POLARstar OPTIMA (BMG LABTECH, Germany), and the luciferase activity was valued as relative light units (RLU). [Bars indicate the mean ± SEM of triplicate assays. The level of luciferase activity was found significantly different in treated cells compared to control cells at $p < 0.0001$].

CHAPTER IV

# DISCUSSION

HIV-1 IN has to interact/exploit cellular factors for establishment of successful infection of the virus. Although a lot of research data have been generated, recent studies give the indication that additional cellular factors may be involved in interaction with IN ultimately aiding the replication and infection of the virus in the host cell (28, 90,110). More importantly, these interactions are subject of pharmacological research in developing anti-IN i.e. anti-HIV therapeutic approach. Besides, Raltegravir (only FDA approved drug against IN action), other compounds have been reported which can disrupt these interactions (28,105). This ordeal prompted us to take an approach to identify new cellular interacting partners of HIV-1 IN. We hypothesized that further identification of new host cell partners of HIV-IN and possible inhibition of these new interactions along with other inhibitors may enrich our arsenal of anti-HIV therapeutic approach.

To identify new host cell partners of HIV-1 IN, a newly developed tandem affinity purification or TAP (Interplay$^{TM}$ Mammalian TAP system) system from Strategene was used. This approach has been reported to be an efficient tool to identify and study new protein-protein interactions including identification of a cellular partner (i.e. DDB1) of another HIV-1 protein Vpr (122-127).

Both wild-type and codon-optimized HIV-1 Integrase sequence were experimented and it was found that codon-optimized IN had far better expression in mammalian 293T cells (Fig. 2). So in most of the subsequent experiments we used codon-optimized IN assuming that robust expression of IN would increase possibility of range and intensity of interactions between HIV-1 IN and host cell factors.

Initially, all the steps of tandem affinity purification (TAP) protocol as recommended by the manufacturer (Stratagene) was applied. The TAP system has two incubation steps with streptavidin resin and calmodulin resin with subsequent elution steps for each. But from repeated experiments it was found that sample fractions eluted after incubation with streptavidin resin retained most of the proteins (perhaps all!) so that no distinctive protein bands (in silver or coomassie staining) were seen from samples eluted after incubation with calmodulin resin. Different parameters such as use of increased amount of calmodulin beads, longer (i.e. over night) incubation with calmodulin beads instead of 2 hours , increased time for elution (2 hours instead of 30 minutes) to optimize and bring a favorable condition were changed so that bands from calmodulin-treated samples could be found also. But no protein bands from samples eluted after incubation with calmodulin resin (Fig. 3 right lanes) were seen. So for final affinity purification experiment only the streptavidin resin was used and then eluted interacting proteins with elution buffer. After subsequent coomassie staining, in-gel digestion and mass spectrometry - two known and three new, putative host cell interactors of HIV-1 IN were identified. Identification of two already reported intercators of HIV-1 IN validates our curtailed approach.

Protein identification, extraction and purification are tough work in the field of proteomics. Functionality and efficiency of our modified but shorter protocol (to successfully identify proteins) has been proved empirically and this method is time-saving and less laborious.

Using tandem affinity purification system coupled with mass spectrometry, we identified totally five proteins. Among them two (i.e. Chaperonin or HSP60 and β-tubulin) were previously reported to be cellular cofactors of HIV-1 IN (reviewed in 31, 32). By western blot analysis, we could identify β-tubulin in sample extracted and purified by modified TAP protocol. Unfortunately, we did not see the presence of chaperonin (HSP60) in our purified protein sample. It can be assumed that the reasons behind this may be inadequate expression of HSP60 in cells in such an amount which can not be detected by TAP-system, or may be due to the lower sensitivity of specific antibody.

Identification of Histone H1.2 and ATP synthase alpha subunit 1 as putative cellular interactor of HIV-1 IN by mass spectrometry is an important development and worthy of further consideration. Although interaction of these two proteins with HIV-1 IN requires to be confirmed by *in vivo* co-immunoprecipitation assay, it is highly logical to hypothesize a possible interaction between Histone H1.2 (an isoform of H1) and HIV-I IN. Because histone H1 plays very crucial role in maintaining compact and higher order structure of chromatin by clumping the histone octamer and thus inhibiting their unnecessary and untimely relaxation of nucleosome (which are compact structure made of histone and DNA) beads. If a retroviral DNA sequence is targeted to be integrated into

a particular region of host genome, the higher order chromatin structure of that particular region must be broken, the nucleosomes must unwind the DNA of that particular region of genome. The whole complex process is mediated by nuclear histone acetyl tranferase (HAT) factors that acetylate histones inducing them to relax/unwind the target DNA (reviewed in 128).

Taganov et al (2004) showed that *in vitro* efficiency of integration catalyzed by HIV-1 integrase decreased after compaction of a pre-defined target DNA sequence, having 13 nucleosomes, with histone H1 (129). Thus it may be assumed that interaction of IN with histone H1, or more specifically H1.2, may affect target site selection for viral cDNA integration *in vivo* that results in crucial bias in favor of viral cDNA integration into actively transcribed regions of host DNA.

Apart from this, it can be presumed that HIV-1 IN interaction with histone H1.2 may help the infected DNA-damaged cell to avert apoptosis. Actually, it has been reported that in macrophages with DNA-damage, H1.2 is translocated to mitochondria and associates with a pro-apoptotic factor BAK (Bcl-2 homologous antagonist killer) (130). So, it can be hypothesized that HIV-1 IN might interact with histone H1.2 to sequester it so that it can not translocate to mitochondria to induce apoptosis after integration of viral cDNA into host genome.

Mass spectrometric identification of ATP synthase, H+ transporting, mitochondrial F1 complex, alpha 1 as cellular partner of HIV-I IN is really challenging to explain. ATP

synthase is a large multi-protein complex with many subunits and is a mitochondrial protein engaged in ATP synthesis (121). This protein was repeatedly identified in mass spectrometry but so far no scientific report has been published indicating any kind of interaction between this protein and HIV-1 Integrase. But the possibility of interaction may not be an impossible phenomenon as in this study we have found another mitochondrial protein HSP60 to interact with HIV-1 IN and, importantly, this interaction between HSP60 and HIV-1 IN has been reported previously, also (reviewed in 31, 30).

Another important observation from the study is that the interaction between HIV-1 IN and human splicing factor SF3A3 has been proved. Yeast Splicing factor SF3A3 (also termed SF3a60) is an important member of SF3A heterotrimer complex, consisting of SF3a60, SF3a66, and SF3a120, whereas SF3A itself is a crucial member of functionally mature spliceosomal complex 17S U2 snRNP (131). Studamire et al (2008) showed using *in vitro* yeast-two-hybrid system assay that yeast splicing factor SF3A3 is an interacting partner of HIV-1 IN (112). In our study, not only human homologue of this protein in TAP-purified samples (Fig. 8) was identified, but also using *in vivo* co-immunoprecipitation assay it was shown that HIV-1 IN can interact with this cellular factor (Fig. 9A). These observations from both laboratories provide evidence that SF3A3 is an integrase-interacting cofactor.

Furthermore, using C-terminal deletion mutant and other point mutations in IN sequence (K159P, V165A, A179P, K186A) it was categorically showed that all these point mutants have no negative effect on their interaction with SF3A3 and, at least, the C-terminal

domain of HIV-1 IN is necessary for interaction with human SF3A3. (Fig. 9C). $IN_{K159P}$ is an integration-defective mutant (132). Other mutants of IN i.e. $IN_{V165A}$, $IN_{A179P}$, $IN_{K186A}$ etc. are defective in integration and interaction with host chromatin as well as has been reported to be unable to bind with a prime IN-interactor LEDGF/p75 (tethers viral cDNA complexed in PIC to the host chromosome) (133). We assume that as nuclear transport of these mutants was not affected by the point mutations and as LEDGF/p75 could not sequester them. It suggests that the binding region within HIV-1 IN to SF3A3 is different from that one required for IN nuclear translocation and from the binding region to LEDGF/p75.

Interestingly, from this data one can notice the strongest interaction between $IN_{KR186,7AA}$ and SF3A3 (evident by a larger band in the gel, Fig. 9C). This can be explained by the notion that as this basic amino acid residue (i.e. $IN_{K186}$) is important for tetramerization of IN to yield the active form of the enzyme (134), so after the point mutation no oligomerization of mutant IN molecules could happen and as a result more mutant $IN_{KR186,7AA}$ monomers were available for interaction with SF3A3.

Using a luciferase assay for the first time it has been shown that SF3A3 is necessary for viral replication as SF3A3 knock-out HeLa cells showed little replication efficiency (Fig. 10). It has been already mentioned that SF3A3 is a component of SF3A protein which itself is a crucial component of the cellular splicing machinery 17S U2 snRNP (small nuclear ribonuclear protein). We can hypothesize that interaction between SF3A3 and HIV-1 IN may be necessary to preferentially or selectively splice viral pre-mRNAs.

Although there is no report on whether SF3A3 can inhibit or augment any viral replication or not, recently SF3A3 has been reported to specifically inhibit transcriptional activity of a human protein called constitutive androstane receptor (CAR) (135). Further functional studies may reveal actual role of SF3A3 in HIV-1 replication.

Briefly, in this study we have successfully shown that modified tandem affinity purification system is an interesting approach to study protein-protein interaction. Using modified tandem affinity purification system coupled with mass spectrometry this study is the first to identify two new, putative cellular co-factor of HIV-1 IN i.e. nucleosomal protein histone H1.2 and mitochondrial protein ATP synthase (H+ transporting, mitochondrial F1 complex) alpha subunit 1. It also proved and confirmed interaction between human homologue of yeast splicing factor SF3A3 and IN by *in vivo* approach. Further research oriented to confirm the putative interactions would enrich our understanding of IN biology and existing knowledge in designing new IN-inhibitors for better ARV (antiretroviral) therapeutic approach.

The experimental data from this study can be culminated into a model (Fig. 15) where it may be hypothesized that gamma-actin, beta-tubulin and histone H1.2 become associated with the pre-integration complex. After formation of the PIC, interaction between integrase and gamma-actin helps the viral cDNA complexed in the PIC to migrate through the microfilament portion of the cytoskeleton. Then with the help of beta-tubulin, the PIC is trafficked through the microtubules and targeted to the nuclear pore. Having reached in the nucleus and subsequent targeting and tethering to the host chromosomal

DNA, integrase in the PIC interacts with histone H1.2 resulting in remodeling of nucleosomes that wind and compact the target gene. Ultimately, this interaction breaks the higher order structure of the chromatin and decondenses it to make the target gene accessible to viral and host factors of PIC for viral cDNA integration process. Additionally, interaction between histone H1.2 and IN, in an infected cell with DNA-damage due to integration process, may sequester H1.2 so that it can not translocate to mitochondria to interact with pro-apoptotic factor BAK (Bcl-2 homologous antagonist killer) that might induce apoptosis after integration of viral cDNA into host genome.

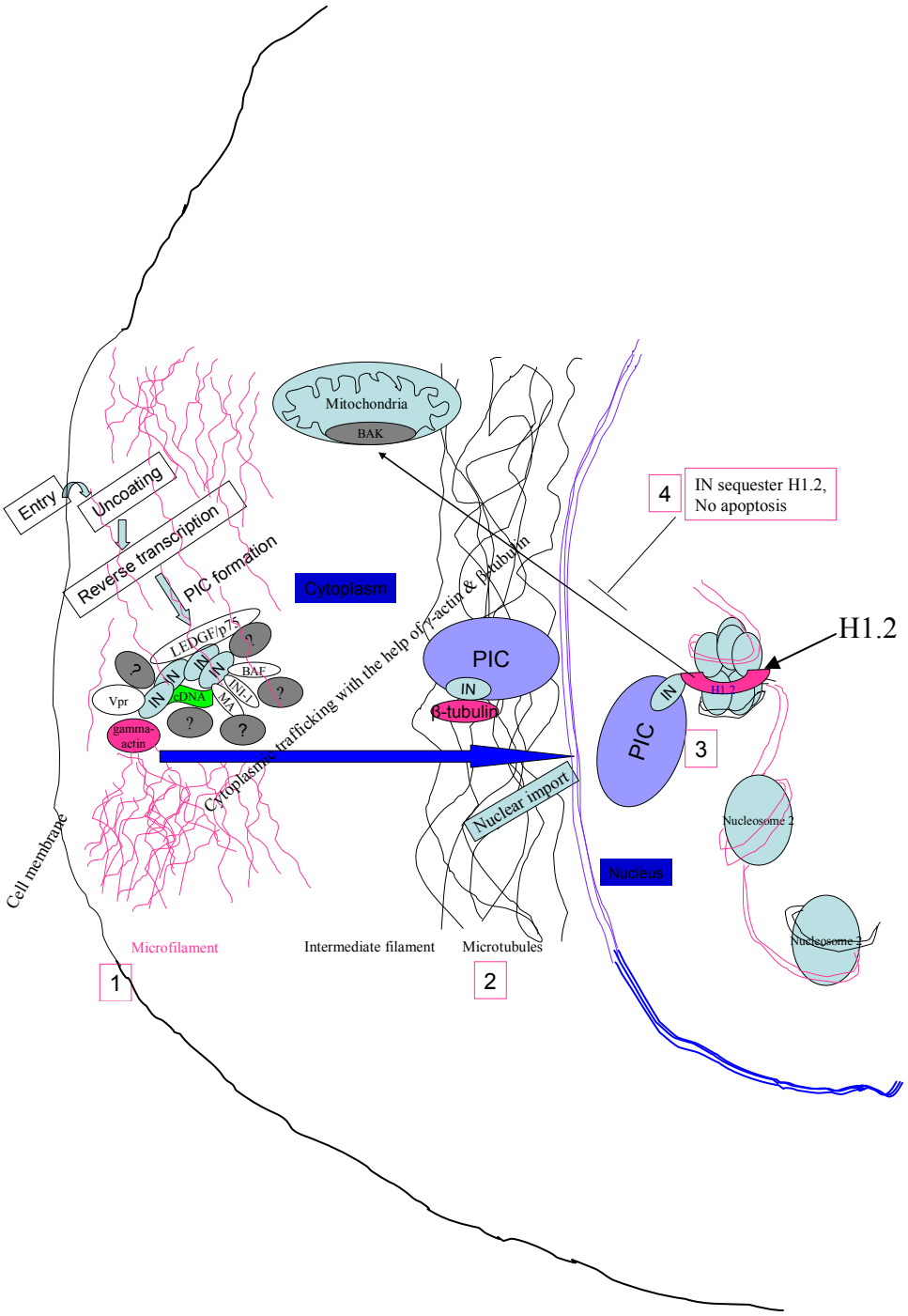

Entry
Uncoating
Reverse transcription
PIC formation
Cytoplasm
LEDGF/p75
Vpr
IN
cDNA
BAF
INI1
MA
gamma-actin
?
Cytoplasmic trafficking with the help of γ-actin & β-tubulin
Mitochondria
BAK
PIC
IN
β-tubulin
Nuclear import
4
IN sequester H1.2,
No apoptosis
H1.2
PIC
3
Nucleosome 2
Nucleosome 3
Nucleus
Cell membrane
Microfilament
Intermediate filament
Microtubules
1
2

**Fig 15: Proposed model describing interaction of HIV-1 IN with newly identified host cell factors**: 1.After entry of HIV-1 virion into the host cell, subsequent uncoating and reverse transcription, pre-integration complex is formed. Gamma-actin associates with the IN in the PIC and this interaction helps PIC migrate through the microfilament of cytoskeleton. 2. When PIC reaches microtubules portion of cytoskeleton, IN interacts with beta-tubulin and this interaction helps PIC's trafficking through the microtubules to the nuclear pore for nuclear import. 3. After nuclear import and tethering with target chromatin, IN interacts with histone H1.2. This interaction may induce remodeling of nucleosomes and decondensation of chromatin ultimately unwinding the target host DNA sequence. Altogether, interaction between IN and H1.2 renders the target DNA accessible to PIC and other host cell factors for successful integration of viral cDNA. 4. IN also sequester H1.2 so that, in a DNA-damaged cell, H1.2 can not translocate to mitochondria and can not interact with pro-apoptotic factor BAK (Bcl-2 antagonist killer) protein that might induce apoptosis. [Note that interaction between IN and beta-tubulin was reported previously (31)].

CHAPTER V

# FUTURE DIRECTIONS

In this study, several host proteins were identified as new HIV-1 IN-interacting cofactors. However, the mechanism how HIV-1 IN takes over these cofactors and what is their biological relevance during HIV replication remain to be elucidated. *In vivo* co-localization and co-immunoprecipitation studies can be done to examine the interaction of these newly identified proteins with HIV-1 IN. Additionally, efforts can be made to establish particular (i.e. each newly identified) protein-deficient cell line by siRNA-mediated knock-down and, once established, these cells can be transfected with HIV-1 provirus and subsequent luciferase assay can be used to check the impact of these proteins on viral replication.

Theoretically, for successful integration of HIV-1 provirus, the hostones/nucleosomes of the target site, in the host genome, have to be removed/remodeled ultimately leading to the latent stage of HIV infection. On the other hand, to initiate viral transcription from a latent stage, again, the histones/nucleosomes have to be removed. Actually, interactions between HIV-1 IN and histones have been well-addressed. It has been reported that HIV-1 transcription is restricted by the presence of a strictly positioned nucleosome (Nuc-1) positioned in close proximity to the viral RNA start site (+10 to +155). Experiments using models of chronic HIV infection has shown that increased accessibility of chromatin about the LTR has been associated with transcriptional activation. More recently, it has been reported that reactivation of HIV transcription requires remodeling

of the critical Nuc-1 by SWI/SNF (reviewed in 136). So, possibility of interaction between HIV-1 IN and histone H1.2 is not an impossible idea and further research to confirm the interaction and mechanism of interaction will help better understanding of epigenetic regulation of HIV-1 latency and transcriptional activation.

To confirm the putative protein-protein interaction between HIV-1 IN and gamma actin, Histone H1.2 and ATP synthase alpha subunit 1, respectively, as reveled by the mass spectrometry, future experiments can be designed to pull down Histone H1.2 and ATP synthase alpha subunit 1 from 293T cells by *in vivo* co-immunoprecipitation assay. Given that the interactions are confirmed, binding regions as well as critical amino acid residues for these interactions can be mapped in IN and relevant co-factor proteins by co-immunoprecipitation with mutant IN and relevant co-factors. Effect of siRNA-mediated deletion of gamma-actin, histone H1.2 and ATP synthase alpha subunit 1 in CD8166T cell on HIV replication, luciferase assay can be performed.

If the binding regions can be determined, synthetic peptides can be designed and tested to check if they can bind with these host factor binding sites of IN and whether can block the interaction between HIV-1 IN and these factors which may open a new avenue of therapeutic approach.

Besides, P-TEFb (positive transcription elongation factor b) has been proposed to be a histone H1 kinase as it preferentially phosphorylates Ser-183 in histone H1.1 *in vivo*. P-TEFb-mediated phosphorylation of H1 is necessary for both H1 mobility and HIV-1 transcription. Because P-TEFb phosphorylation of H1 destabilizes the interaction of H1

with chromatin, genes to be transcribed may be more readily remodeled (reviewed in 137).

Another experiment could be designed to see the effect of Lys-183 mutation in H1.2 in HIV-1 IN and H1.2 interaction. If such mutation(s) disrupts the interaction between HIV-1 IN and H1.2, it may be an important finding in explaining HIV-1 latency through epigenetic mechanisms and blocking the site of phosphorylation could lead a way for therapeutic approach.

Successful confirmation of the interaction between HIV-1 IN and histone H1.2 will open a new avenue of HIV research in the context of epigenetics. If viral replication can be inhibited or made defective by blocking these interactions with the help of synthetic peptides, this would be of hopeful benefit in our battle against this deadly virus.

# REFERENCES


1. Garret, L. (2005) *HIV/AIDS: The Modern-Day Plague ?* The Globalist (http://www.theglobalist.com/StoryId.aspx?StoryId=4711)

2. Weiss, R.A., *How does HIV cause AIDS?* Science, 1993. **260** (5112): p. 1273–9.

3. Morrow, C. D., J. Park, and J. K. Wakefield, *Viral gene products and replication of the human immunodeficiency type 1 virus*. Am. J. Physiol, 1994. **266**:C1135–C1156.

4. Clapham P.R., McKnight A., *Cell surface receptors, virus entry and tropism of primate lentiviruses.* J Gen Virol, 2002. **83**: 1809–1829.

5. Gilbert, P.B., et al., *Comparison of HIV-1 and HIV-2 infectivity from a prospective cohort study in Senegal*. Statistics in Medicine, 2003. **22** (4): 573–593.

6. Gottlieb M.S., Schroff R., Schanker H.M., Weisman J.D., Fan P.T., Wolf R.A., Saxon A., *Pneumocystis carinii* pneumonia and mucosal candidiasis in previously healthy homosexual men: evidence of a new acquired cellular immunodeficiency**.** N Engl J Med, 1981. **305:**1425-1431.

7. Barre-Sinoussi F., Chermann J.C., Rey F., Nugeyre M.T., Chamaret S., Gruest J., Dauguet C., Axler-Blin C., Vezinet-Brun F., Rouzioux C., Rozenbaum W., Montagnier L., *Isolation of a T-lymphotropic retrovirus from a patient at risk for acquired immune deficiency syndrome (AIDS).* Science, 1983. **220:**868-871.

8. Popovic M., Sarngadharan M.G., Read E., Gallo R.C., *Detection, isolation, and continuous production of cytopathic retroviruses (HTLV-III) from patients with AIDS and pre-AIDS.* Science, 1984. **224**:497-500.

9. Alkhatib G., Combadiere C., Broder C. C., Feng Y., Kennedy P. E., Murphy P. M. & Berger E. A., *CCCKR5: a RANTES, MIP-1a, MIP-1b receptor as a fusion cofactor for macrophage-tropic HIV-1.* Science*,* 1996. **272** :1955±1958.

10. Lawn S. D., *AIDS in Africa: the impact of coinfections on the pathogenesis of HIV-1 infection*. J. Infect. Dis., 2004. **48** (1): 1–12

11. Hogg R., Lima V., Sterne J.A., Grabar S., Battegay M., Bonarek M., D'Arminio Monforte A., Esteve A., Gill MJ., Harris R., Justice A., Hayden A., Lampe F., Mocroft A., Mugavero M. J, Staszewski S., Wasmuth J.C., van Sighem A., Kitahata M., Guest J., Egger M, May M., *Life expectancy of individuals on combination antiretroviral*

*therapy in high-income countries: a collaborative analysis of 14 cohort studies*. Lancet, 2008. **372**: 293–299.

12. Morgan D., Mahe C., Mayanja B., Okongo J. M., Lubega R, Whitworth J. A., *HIV-1 infection in rural Africa: is there a difference in median time to AIDS and survival compared with that in industrialized countries?* AIDS, 2002. **16** (4): 597–632.

13. Coiras M., Lo´pez-Huertas M. R., Pe´rez-Olmeda M., Alcamı´ J., *Understanding HIV-1 latency provides clues for the eradication of long-term reservoirs.* Nat Rev Microbiol, 2009. **7**: 798–812.

14. http://www.unaids.org/globalreport/Epi_slides.htm

15. Freed E. O., *HIV-1 replication.* Somat Cell Mol Genet, 2001. **26**(1-6):13-33.

16. Frankel A. D., Young J. A., *HIV-1: fifteen proteins and an RNA*. Annu Rev Biochem., 1998. **67**:1-25.

17. Delelis O., Carayon K., Saib A., Deprez E., Mouscadet J. F., *Integrase and integration: biochemical activities of HIV-1 integrase.* Retrovirology, 2008. **5**: 114.

18. Roldan, A., R. S. Russell, B. Marchand, M. Gotte, C. Liang, and M. A. Wainberg, *In vitro identification and characterization of an early complex linking HIV-1 genomic RNA recognition and Pr55Gag multimerization.* J. Biol. Chem., 2004. **279:**39886–39894.

19. Pollard V. W., Malim M. H., *The HIV-1 Rev protein*. Annu Rev Microbiol., 1998. **52**:491-532.

20. Kao S. Y., Calman A. F., Luciw P. A., Peterlin B. M., *Anti-termination of transcription within the long terminal repeat of HIV-1 by tat gene product.* Nature, 1987.**330**(6147):489-93.

21. Le Rouzic E., Benichou S., *The Vpr protein from HIV-1: distinct roles along the viral life cycle*. Retrovirology, 2005. **22**;2:11.

22. Favaro J. P., Maldarelli F., Arrigo S. J., Schmidt M. G., *Effect of rev on the cytoplasmic localization of intron-containing human immunodeficiency virus type 1 RNA*. Virology, 1999. **255**(2):237-49.

23. Emerman M., Malim M. H., *HIV-1 regulatory/accessory genes: keys to unraveling viral and host cell biology.* Science, 1998. **280**(5371):1880-4.

24. Wang J., Shackelford J. M., Casella C. R., Shivers D. K., Rapaport E. L., Liu B, Yu X. F., Finkel T. H., *The Vif accessory protein alters the cell cycle of human immunodeficiency virus type 1 infected cells.* Virology, 2007. **359**(2):243-52.

25. Mangeat B, Gers-Huber G, Lehmann M, Zufferey M, Luban J, et al. (2009) HIV-1 Vpu Neutralizes the Antiviral Factor Tetherin/BST-2 by Binding It and Directing Its Beta-TrCP2-Dependent Degradation. PLoS Pathog 5(9): e1000574. doi:10.1371/journal.ppat.1000574)

26. Dube´ M, Bhusan Roy B, Guiot-Guillain P, Binette J, Mercier J, et al. (2010) Antagonism of Tetherin Restriction of HIV-1 Release by Vpu Involves Binding and Sequestration of the Restriction Factor in a Perinuclear Compartment. PLoS Pathog 6(4): e1000856. doi:10.1371/journal.ppat.1000856)

27. Lai M., Chen J., *The role of Vpr in HIV-1 disease progression is independent of its G2 arrest induction function.* Cell Cycle, 2006. **5**(19):2275-80.

28. Belzile J-P, Abrahamyan LG, Ge´rard FCA, Rougeau N, Cohen E´A (2010) Formation of Mobile Chromatin-Associated Nuclear Foci Containing HIV-1 Vpr and VPRBP Is Critical for the Induction of G2 Cell Cycle Arrest. PLoS Pathog 6(9): e1001080. doi:10.1371/journal.ppat.1001080

29. Bukovsky A. A., Dorfman T., Weimann A., Göttlinger H. G., *Nef association with human immunodeficiency virus type 1 virions and cleavage by the viral protease.* J Virol. 1997. **71**(2):1013-8.

30. Warrilow D., Harrich D., *HIV-1 replication from after cell entry to the nuclear periphery*. Curr HIV Res., 2007. **5**(3):293-9.

31. Vandegraaff N., Engelman A., *Molecular mechanisms of HIV integration and therapeutic intervention.* Expert Rev Mol Med, 2007. **9**:1–19.

32. Ribeiro R. M., *Dynamics of CD4+ T cells in HIV-1 infection.* Immunol Cell Biol., 2007. **85**(4):287-94.

33. Vidricaire G., Gauthier S., Tremblay M. J., *HIV-1 infection of trophoblasts is independent of gp120/CD4 Interactions but relies on heparan sulfate proteoglycans*. J Infect Dis., 2007. **195**(10):1461-71.

34. Price P, Keane N, Gray L, Lee S, Gorry P. R., French M. A., CXCR4 or CCR5 tropism of human immunodeficiency virus type 1 isolates does not determine the immunological milieu in patients responding to antiretroviral therapy. Viral Immunol. 2006.**19**(4):734-40.

35. Yi L, Fang J, Isik N, Chim J, Jin T., *HIV gp120-induced interaction between CD4 and CCR5 requires cholesterol-rich microenvironments revealed by live cell fluorescence resonance energy transfer imaging.* J Biol Chem., 2006. **281**(46):35446-53.

36. Root M. J., Steger H. K., *HIV-1 gp41 as a target for viral entry inhibition*. Curr Pharm Des., 2004.**10**(15):1805-25

37. Dvorin J. D., Malim M. H., *Intracellular trafficking of HIV-1 cores: journey to the center of the cell.* Curr Top Microbiol Immunol., 2003. **281**:179-208.

38. Arhel N. J., Souquere-Besse S, Munier S, Souque P, Guadagnini S, Rutherford S, Prévost M. C., Allen T. D., Charneau P. *HIV-1 DNA Flap formation promotes uncoating of the pre-integration complex at the nuclear pore*. EMBO J., 2007. **26**(12):3025-37.

39. Yang Y., Guo F., Cen S., Kleiman L., *Inhibition of initiation of reverse transcription in HIV-1 by human APOBEC3F.* Virology, 2007. **365**(1):92-100.

40. Barraud P., Gaudin C., Dardel F., Tisné C., *New insights into the formation of HIV-1 reverse transcription initiation complex.* Biochimie, 2007. **89**(10):1204-10.

41. Castro H. C., Loureiro N. I., Pujol-Luz M., Souza A. M., Albuquerque M. G., Santos D. O., Cabral L. M., Frugulhetti I. C., Rodrigues C. R., *HIV-1 reverse transcriptase: a therapeutical target in the spotlight.* Curr Med Chem., 2006. **13**(3):313-24.

42. Goldschmidt V, Paillart JC, Rigourd M, Ehresmann B, Aubertin AM, Ehresmann C, Marquet R., *Structural variability of the initiation complex of HIV-1 reverse transcription*.J Biol Chem., 2004. **279**(34):35923-31.

43. Tang S., Murakami T., Agresta B. E., Campbell S., Freed E. O., Levin J. G., *Human immunodeficiency virus type 1 N-terminal capsid mutants that exhibit aberrant core morphology and are blocked in initiation of reverse transcription in infected cells.* J Virol., 2001.**75**(19):9357-66.

44. Miller M. D., Farnet C. M., Bushman F. D., *Human immunodeficiency virus type 1 preintegration complexes: studies of organization and composition.* J Virol., 1997. **71**(7):5382-90.

45. Farnet C. M., Bushman F. D., *HIV-1 cDNA integration: requirement of HMG I(Y) protein for function of preintegration complexes in vitro*. Cell, 1997. **21**:88(4):483-92.

46. Llano M., Delgado S., Vanegas M., Poeschla E. M., *Lens epithelium-derived growth factor/p75 prevents proteasomal degradation of HIV-1 integrase*. J Biol Chem., 2004. **279**(53):55570-7.

47. Suzuki Y., Craigie R., *The road to chromatin - nuclear entry of retroviruses.* Nat Rev Microbiol, 2007. **5**(3):187-96.

48. Turlure, F., et al., (2004) Human cell proteins and human immunodeficiency virus DNA integration. Front Biosci, 2004. **9:** 3187-3208

49. Ao Z., Jayappa K. D., Wang B., Zheng Y., Kung S., Rassart E., Depping R., Kohler M., Cohen E. A, Yao X., *Importin {alpha}3 interacts with HIV-1 integrase and contributes to HIV-1 nuclear import and replication.* J Virol, 2010.**84**(17):8650-63.

50. McDonald D., Vodicka M. A., Lucero G., Svitkina T. M., Borisy G. G., Emerman M., Hope T. J., *Visualization of the intracellular behavior of HIV in living cells.* J Cell Biol., 2002. **159**(3):441-52.

51. Perno C. F., Svicher V., Schols D., Pollicita M., Balzarini J., Aquaro S., *Therapeutic strategies towards HIV-1 infection in macrophages.* Antiviral Res.,2006. **71**(2-3):293-300.

52. Kedzierska K., Crowe S. M., *The role of monocytes and macrophages in the pathogenesis of HIV-1 infection.* Curr Med Chem., 2002. **9**(21):1893-903.

53. Wu Y., *HIV-1 gene expression: lessons from provirus and non-integrated DNA*. Retrovirology, 2004.**1**:13.

54. Llano M., Saenz D. T., Meehan A., Wongthida P., Peretz M., Walker W. H., Teo W., Poeschla E. M., *An Essential Role for LEDGF/p75 in HIV Integration.* Science, 2006. **314**(5798):461-4.

55. Hombrouck A., De Rijck J., Hendrix J., Vandekerckhove L., Voet A., De Maeyer M., Witvrouw M., Engelborghs Y., Christ F., Gijsbers R., et al., *Virus evolution reveals an exclusive role for LEDGF/p75 in chromosomal tethering of HIV*. PLoS Pato, 2007. **3**:e47.

56. Al-Harthi L., Roebuck K. A., *Human immunodeficiency virus type-1 transcription: role of the 5'-untranslated leader region.* Int J Mol Med., 1998.**1**(5):875-81.

57. Dingwall C., Ernberg I., Gait M. J., Green S. M., Heaphy S., Karn J., Lowe A. D., Singh M., Skinner M. A., *HIV-1 tat protein stimulates transcription by binding to a U-rich bulge in the stem of the TAR RNA structure.* EMBO J, 1990. **9**(12):4145-53.

58. Marciniak R. A., Calnan B. J., Frankel A. D., Sharp P. A., *HIV-1 Tat protein trans-activates transcription in vitro.* Cell, 1990. **63**(4):791-802.

59. Copeland K. F., *Modulation of HIV-1 transcription by cytokines and chemokines.* Mini Rev Med Chem., 2005. **5**(12):1093-101.

60. Kim Y. S., Kim J. M., Jung D. L., Kang J. E., Lee S., Kim J. S., Seol W., Shin H. C., Kwon H. S., Van Lint C., Hernandez N., Hur M. W., *Artificial zinc finger fusions targeting Sp1-binding sites and the trans-activator-responsive element potently repress transcription and replication of HIV-1.* J Biol Chem., 2005. **280**(22):21545-52.

61. Ammosova T., Jerebtsova M., Beullens M., Voloshin Y., Ray P. E., Kumar A., Bollen M., Nekhai S., *Nuclear protein phosphatase-1 regulates HIV-1 transcription.* J Biol Chem., 2003. **278**(34):32189-94.

62. Romanchikova N., Ivanova V., Scheller C., Jankevics E., Jassoy C., Serfling E., *FAT transcription factors control HIV-1 expression through a binding site downstream of TAR region.* Immunobiology, 2003. **208**(4):361-5.

63. Jordan A., Defechereux P., Verdin E., *The site of HIV-1 integration in the human genome determines basal transcriptional activity and response to Tat transactivation.* EMBO J., 2001. **20** (7):1726-38.

64. Hayouka Z., Levin A., Maes M., Hadas E., Shalev D. E., Volsky D J., Loyter A., Friedler A,. *Mechanism of action of the HIV-1 integrase inhibitory peptide LEDGF 361-370*. Biochem Biophys Res Commun. 2010. **394**(2):260-5.

65. Harrich D., Ulich C., Gaynor R. B., A critical role for the TAR element in promoting efficient human immunodeficiency virus type 1 reverse transcription. J Virol., 1996.**70**(6):4017-27.

66. Nakatani Y., HIV-1 transcription: activation mediated by acetylation of Tat. Structure, 2002.**10**(4):443-4.

67. Fisher A. G., Feinberg M. B., Josephs S. F., Harper M. E., Marselle L. M., Reyes G., Gonda M. A., Aldovini A., Debouk C., Gallo R. C., et al., *The trans-activator gene of HTLV-III is essential for virus replication.* Nature, 1986. **320**(6060):367-71.

68. Anderson E.C., Lever A. M., *Human immunodeficiency virus type 1 Gag polyprotein modulates its own translation.* J Virol, 2006.**80**(21):10478-86.

69. Phuphuakrat A., Paris R. M., Nittayaphan S., Louisirirotchanakul S., Auewarakul P., *Functional variation of HIV-1 Rev Response Element in a longitudinally studied cohort*. J Med Virol., 2005. **75**(3):367-73.

70. Ono A., *HIV-1 assembly at the plasma membrane.* Vaccine. 2010. **28**, 2:B55-9.

71. Adamson C. S., Freed E. O., *Human immunodeficiency virus type 1 assembly, release, and maturation.* Adv Pharmacol, 2007.**55**:347–87.

72. Ivanchenko S., Godinez W. J., Lampe M., Kra¨usslich H-G., Eils R., *Dynamics of HIV-1 Assembly and Release*. PLoS Pathog. 2009. **5**(11)

73. Yoshiko Usami, Sergei Popov, Elena Popova, Michio Inoue, Winfried Weissenhorn and Heinrich G. Go¨ttlinger, *The ECRT pathway and HIV-1 budding.* Biochem Soc Trans., 2009.**37**(Pt 1):181-4.

74. Christine Goffinet, Stefanie Homann, Ina Ambiel, Nadine Tibroni, Daniel Rupp, Oliver T. Keppler, and Oliver T. Fackler, *Antagonism of CD317 restriction of human immunodeficiency virus type 1 (HIV-1) particle release and depletion of CD317 are separable activities of HIV-1 Vpu*. J Virol., 2010. **84**(8):4089-94

75. Robertson D. L., Hahn B. H., Sharp P. M., *Recombination in AIDS viruses*. J Mol Evol., 1995. **40** (3): 249–59.
76. Rambaut, A., et al., *The causes and consequences of HIV evolution*. Nat Rev Genet., 2004. **5**(1):52-61.

77. Luc Montagnier, *25 years after HIV discovery: Prospects for cure and vaccine*. Virology, 2010. **397**(2):248-54.

78. Delelis O., Parissi V., Leh H., Mbemba G., Petit C., Sonigo P., Deprez E., Mouscadet J. F., *Efficient and specific internal cleavage of a retroviral palindromic DNA sequence by tetrameric HIV-1 integrase. PLoS ONE,* 2007. **2**:e608.

79. Miller M. D., Farnet C. M., Bushman F. D., *Human immunodeficiency virus type 1preintegr ation complexes: studies of organization and composition. J Virol,* 1997. **71**:5382-5390.

80. Bukrinsky M. I., Sharova N., McDonald T. L., Pushkarskaya T., Tarpley W. G., Stevenson M., *Association of integrase, matrix, and reverse transcriptase antigens of human immunodeficiency virus type 1 with viral nucleic acids following acute infection. Proc Natl Acad Sci U S A,* 1993. **90**:6125-6129.

81. Nermut M. V., Fassati A., *Structural analyses of purified human immunodeficiency virus type 1 intracellular reverse transcription complexes.* J Virol, 2003. **77**:8196-8206.

82. Gallay P., Swingler S., Song J., Bushman F., Trono D., *HIV nuclear import is governed by the phosphotyrosine-mediated binding of matrix to the core domain of integrase.* Cell, 1995. **83**:569-576.

83. Dardalhon V., Herpers B., Noraz N., Pflumio F., Guetard D., Leveau C., et al., *Lentivirus-mediated gene transfer in primary T cells is enhanced by a central DNA flap*. Gene Therapy, 2001. **8**: 190–198.

84. Wu X., Liu H., Xiao H., Conway J. A., Hehl E., Kalpana G. V., Prasad V., Kappes J. C., Human immunodeficiency virus type 1 integrase protein promotes reverse transcription through specific interactions with the nucleoprotein reverse transcription complex. J Virol, 1999. **73**:2126-2135.

85. Zhu K., Dobard C., Chow S. A., *Requirement for integrase during reverse transcr iption of human immunodeficiency virus type 1 and the effect of cysteine mutations of integrase on its interactions with reverse transcriptase.* J Virol, 2004. **78**:5045-5055.

86. Dobard C. W., Briones M. S., Chow S. A., *Molecular mechanisms by which human immunodeficiency virus type 1 integrase stimulates the early steps of reverse transcription.* J Virol, 2007. **81**:10037-10046.

87. Parissi V., Calmels C., de Soultrait V. R., Caumont A., Fournier M., Chaignepain S., Litvak S., *Functional inter actions of human immunodeficiency virus type 1 integrase with human and yeast HSP60.* J Virol, 2001. **75**:11344-11353.

88. Busschots K., Vercammen J., Emiliani S., Benarous R., Engelborghs Y., Christ F., Debyser Z., *The inter action of LEDGF/p75 with integr ase is lentivirus-specific and promotes DNA binding*. J Biol Chem, 2005. **280**:17841-17847.

89. McKee C. J., Kessl J. J., Shkriabai N., Dar M. J., Engelman A., Kvaratskhelia M., *Dynamic Modulation of HIV-1 Integrase Structure and Function by Cellular Lens Epithelium-derived Growth Factor (LEDGF) Protein.* J Biol Chem, 2008. **283**:31802-31812.

90. Naghavi M. H., and Goff S. P., *Retroviral proteins that interact with the host cell cytoskeleton*. Curr Opin Immunol, 2007. **19**: 402-407.

91. Ao, Z., G. Huang, H. Yao, Z. Xu, M. Labine, A. W. Cochrane, and X. Yao, Interaction of human immunodeficiency virus type 1 integrase with cellular nuclear import receptor importin 7 and its impact on viral replication. J Biol Chem, 2007. **282:**13456-67

92. Woodward C. L., Prakobwanakit S., Mosessian S., Chow S. A., *Integrase interacts with nucleoporin NUP153 to mediate the nuclear import of human immunodeficiency virus type 1*. J Virol, 2009.**83**(13):6522-33

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

94. Li L., Yoder K., Hansen M. S., Olvera J., Miller M. D., Bushman F. D., *Retrovir al cDNA integration: stimulation by HMG I family proteins.* J Virol, 2000. **74**:10965-10974.

95. Kalpana G. V., Marmon S., Wang W., Crabtree G. R., Goff S. P., *Binding and stimulation of HIV-1 integr ase by a human homolog of yeast transcription factor SNF5.* Science, 1994. **266**:2002-2006.

96. Boese A., Sommer P., Gaussin A., Reimann A., Nehrbass U., *Ini1/hSNF5 is dispensable for retrovirus-induced cytoplasmic accumulation of PML and does not interfere with integr ation*. FEBS Lett, 2004. **578**:291-296.

97. Ariumi Y., Serhan F., Turelli P., Telenti A., Trono D., *The integrase interactor 1 (INI1) proteins facilitate Tat-mediated human immunodeficiency virus type 1 transcription.* Retrovirology, 2006. **3**:47.

98. Jacque J. M., Stevenson M., *The inner -nuclear -envelope protein emerin regulates HIV-1 infectivity*. Nature, 2006. **441**:641-645.

99. Marshall H. M., Ronen K., Berry C., Llano M., Sutherland H., Saenz D., Bickmore W., Poeschla E., Bushman F. D., *Role of PSIP1/LEDGF/p75 in lentiviral infectivity and integration targeting.* PLoS ONE, 2007. **2**:e1340.

100. Llano M., Vanegas M., Fregoso O., Saenz D., Chung S., Peretz M., Poeschla E. M., *LEDGF/p75 determines cellular trafficking of diverse lentiviral but not murine oncoretroviral integrase proteins and is a component of functional lentiviral preintegration complexes*. J Virol, 2004. **78**:9524-9537.

101. Yung E., Sorin M., Wang E. J., Perumal S., Ott D., Kalpana G. V., *Specificity of interaction of INI1/hSNF5 with retroviral integrases and its functional significance.* J Virol, 2004. **78**:2222-2231.

102. Lewinski M. K., Bushman F. D., *Retroviral DNA integration--mechanism and consequences*. Adv Genet, 2005. **55**:147-181.

103. Lin C. W., Engelman A., *The barrier-to-autointegration factor is a component of functional human immunodeficiency virus type 1 p eintegration complexes.* J Virol, 2003. **77:**5030-5036.

104. Kalpana G. V., Marmon S., Wang W., Crabtree G. R., Goff S. P. *Binding and stimulation of HIV-1 integrase by a human homolog of yeast transcription factor SNF5*. Science, 1994.**266**:2002-2006.

105. Cherepanov P., Ambrosio A. L., Rahman S., Ellenberger T., Engelman A., *Structural basis for the recognition between HIV-1 integrase and transcriptional coactivator p75*. Proc Natl Acad Sci U S A, 2005. **102**:17308-17313.

106. Mawsawi L., Q., Neamati N. *Blocking interactions between HIV-1 integrase and cellular cofactors: an emerging anti-retroviral strategy.* Trends in Pharmacological Sciences, 2007. **28** (10): p. 526-535.

107. Du L., Zhao Y., Chen J., Yang L., Zheng Y., Tang Y., Shen X., Jiang H., *D77, one benzoic acid derivative, functions as a novel anti-HIV-1 inhibitor targeting the interaction between integrase and cellular LEDGF/p75.Biochemical and Biophysical Research Communications*, 2008. **375**(1): p. 139-144.

108. Cereseto, A., et al., *Acetylation of HIV-1 integrase by p300 regulates viral integration*.EMBO J, 2005. **24** :3070-3081

109. Allouch A, Cereseto A., *Identification of cellular factors binding to acetylated HIV-1 integrase.* Amino Acids, 2009.

110. Terreni M., Valentini P., Liverani V., Gutierrez M. I., Di Primio C., Di Fenza A., Tozzini V., Allouch A., Albanese A., Giacca M., Cereseto A., *GCN5-dependent acetylation of HIV-1 integrase enhances viral integration.* Retrovirology, 2010. **7**:18.

111. Studamire B., Goff S. P., *Host proteins interacting with the moloney murine leukemia virus integrase: Multiple transcriptional regulators and chromatin binding factors.* Retrovirology, 2008. **5**: 48.

112. Mousnier A., Kubat N., Massias-Simon A., Ségéral E., Rain J. C., Benarous R., Emiliani S., Dargemont C., *von Hippel Lindau binding protein 1-mediated degradation of integrase affects HIV-1 gene expression at a postintegration step.* Proc Natl Acad Sci U S A, 2007.**104**(34):13615-20.

113. Zheng Y., Ao Z., Jayappa K.D., Yao X., *Characterization of the HIV-1 integrase chromatin- and LEDGF/p75-binding abilities by mutagenic analysis within the catalytic core domain of integrase.* Virol J. 2010.**7**:68.

114. Kis A. D., *ABC for AIDS prevention in Guinea: migrant gold mining communities address their risks*. AIDS Care, 2010. **22**(4):520-5.

115. De Clercq E., *Antiretroviral drugs.* Curr Opin Pharmacol, 2010.

116. Marchand C., Maddali K., Metifiot M., Pommier Y., *HIV-1 IN inhibitors: 2010 update and perspectives*. Curr Top Med Chem, 2009. **9**(11): 1016–1037.

117. Malet I., Delelis O., Valantin M. A., Montes B., Soulie C., Wirden M., Tchertanov L., Peytavin G., Reynes J., Mouscadet J. F., et al., *Mutations associated with failure of raltegravir treatment affect integrase sensitivity to the inhibitor in vitro*. *Antimicrob Agents Chemother,* 2008. **52**:1351-1358.

118. Wainberg M. A., Jeang K. T., *25 years of HIV-1 research – progress and perspectives***.** BMC Med, 2008. **6:**31.

119. Coiras M, López-Huertas M. R., Sánchez del Cojo M., Mateos E., Alcamí J, *Dual role of host cell factors in HIV-1 replication: restriction and enhancement of the viral cycle***.** AIDS Rev, 2010. **12**(2):103-12.

120. Ao Z., Fowke K. R., Cohen E. A., and Yao X., *Contribution of the C-terminal tri-lysine regions of human immunodeficiency virus type 1 integrase for efficient reverse transcription and viral DNA nuclear import.* Retrovirology, 2005. **2:** 62.

121. Ackerman S. H., Tzagoloff A., *Function, structure, and biogenesis of mitochondrial ATP synthase.* Prog Nucleic Acid Res Mol Biol., 2005. **80**:95-133.

122. D, Chen L., Johnson B., Bettano K., Jackson D., Taraszka J., Wang Y. K., Jones M. D., Morrissey M., Deeds J., Mosher R., Fordjour P., Lengauer C., Benson J. D., *Contribution of polycomb homologues Bmi-1 and Mel-18 to medulloblastoma pathogenesis*. Mol Cell Biol., 2007. **27**(13):4968-79

123. Wagenaar T. R., Moss B., *Association of vaccinia virus fusion regulatory proteins with the multicomponent entry/fusion complex.* J Virol., 2007. **81**(12):6286-93.

124. Schrofelbauer B., Hakata Y., Landau N. R., *HIV-1 Vpr function is mediated by interaction with the damage-specific DNA-binding protein DDB1.* Proc Natl Acad Sci U S A. 2007. **104**(10):4130-5.

125. Gallois-Montbrun S., Kramer B., Swanson C. M., Byers H., Lynham S., Ward M., Malim M. H., *Antiviral protein APOBEC3G localizes to ribonucleoprotein complexes found in P bodies and stress granules.* J Virol., 2007.**81**(5):2165-78.

126. Griffin M. J., Wong R. H., Pandya N., Sul H. S., *Direct interaction between USF and SREBP-1c mediates synergistic activation of the fatty-acid synthase promoter*. J Biol Chem., 2007. **282**(8):5453-67.

127. Chiu Y. L., Witkowska H. E., Hall S. C., Santiago M., Soros V. B., Esnault C., Heidmann T., Greene W. C., *High-molecular-mass APOBEC3G complexes restrict Alu retrotransposition.* Proc Natl Acad Sci U S A. 2006. **103** (42): 15588-93.

128. McBryant S. J., Lu X., Hansen J. C., *Multifunctionality of the linker histones: an emerging role for protein-protein interactions*. Cell Res., 2010. **20**(5):519-28.

129. Taganov K. D., Cuesta I., Daniel R., Cirillo L. A., Katz R. A., Zaret K. S., Skalka A. M., *Integrase-specific enhancement and suppression of retroviral DNA integration by compacted chromatin structure in vitro.* J Virol, 2004. **78**:5848-5855.

130. Okamura H., Yoshida K., Amorim B. R., Haneji T., *Histone H1.2 is translocated to mitochondria and associates with Bak in bleomycin-induced apoptotic cells.* J Cell Biochem., 2008. **103**(5):1488-96.

131. Nesic D, Krämer A., *Domains in human splicing factors SF3a60 and SF3a66 required for binding to SF3a120, assembly of the 17S U2 snRNP, and prespliceosome formation.* Mol Cell Biol., 2001. **21**(19):6406-17.

132. Cannon P. M., Wilson W., Byles E., Kingsman S. M., Kingsman A. J., *Human immunodeficiency virus type 1 integrase: effect on viral replication of mutations at highly conserved residues.* J Virol, 1994. **68:**4768-4775

133. Xu Z., Zheng Y., Ao Z., Clement M., Mouland A. J., Kalpana G. V., Belhumeur P., Cohen E. A., Yao X., *Contribution of the C-terminal region within the catalytic core domain of HIV-1 integrase to yeast lethality, chromatin binding and viral replication.* Retrovirology, 2008. **5**:102

134. Berthoux L., Sebastian S., Muesing M. A., Luban J., *The role of lysine 186 in HIV-1 integrase multimerization.* Virology, 2007. **364**:227-236

135. Yun H. J., Kwon J., Seol W., *Specific inhibition of transcriptional activity of the constitutive androstane receptor (CAR) by the splicing factor SF3a3.* Biol Chem., 2008. **389**(10):1313-8.

136. Margolis DM., *Mechanisms of HIV Latency: an Emerging Picture of Complexity*. Curr HIV/AIDS Rep. 2010;**7**(1):37-43.

137. O'Brien SK, Cao H, Nathans R, Ali A, and Tariq M. Rana TM.,*P-TEFb Kinase Complex Phosphorylates Histone H1 to Regulate Expression of Cellular and HIV-1 Genes.* J Biol Chem. 2010; **285**(39):29713-20.

138. Preston BD, Poiesz BJ, Loeb LA., *Fidelity of HIV-1 reverse transcriptase.* Science. 1988; **242**:1168-1171

139. Sharp PM, Hahn BH., *The evolution of HIV-1 and the origin of AIDS*. Philos Trans R Soc Lond B Biol Sci. 2010; **365**(1552):2487-94.

140. Roquebert B, Damond F, Brun-Vezinet F, et al. *Diversité génétique des VIH et ses conséquences*. Pathol Biol 2009,**57**:142-148. [HIV genetic diversity and its consequences]

141. Onafuwa-Nuga A, Telesnitsky A., *The remarkable frequency of human immunodeficiency virus type 1 genetic recombination.* Microbiol Mol Biol Rev. 2009; **73**: 451–480.

142. Johnson VA, Brun-Vézinet F, Clotet B, Günthard HF, Kuritzkes DR, Pillay D, Schapiro JM, Richman DD., *Update of the drug resistance mutations in HIV-1: December 2010.* Top HIV Med., 2010;**18** (5):156-63.