The study of susceptibility and resistance of HIV integrases to integrase strand transfer inhibitors and the development of novel single domain antibody targeting HIV integrase

Xiaoju Ni

THESE DE DOCTORAT DE

L'ECOLE NORMALE SUPERIEURE DE CACHAN

Présentée par

Mademoiselle Xiaoju NI

Pour obtenir le grade de

DOCTEUR DE L'ECOLE NORMALE SUPERIEURE DE CACHAN

Sujet de la thèse:

The study of susceptibility and resistance of HIV integrases to integrase strand transfer inhibitors and the development of novel single domain antibody targeting HIV integrase

Thèse présentée et soutenue publiquement le 30 septembre 2011

Devant le jury composé de :

Dr. Xu-Guang XI	**Président du Jury**
Dr. Vincent PARISSI	**Rapporteur**
Dr. Karen MOREAU	**Rapporteur**
Dr. Olivier DELELIS	**Examinateur**
Dr. Jean-François MOUSCADET	**Co-Directeur de thèse**
PI. Kai-Cheng QIAN	**Co-Directeur de thèse**

Laboratoire de Biologie et Pharmacologie Appliquée
(CNRS UMR 8113)
61, avenue du Président Wilson, 94235 Cachan CEDEX (France)

REMERCIEMENTS

J'ai effectuée ce travail de thèse dans l'équipe Interactions fonctionnelles des retrovirus au Laboratoire de Biologie et Pharmacologie Appliquées, UMR 8113 du CNRS. Je voudrais d'abord remercier Christian Auclair et Jean-François Mouscadet, qui ont assurée successivement la direction du laboratoire au cours de ces années, pour m'avoir accueillie au sein de leur unité de recherche et m'avoir permis de travailler dans un environnement scientifique aussi exceptionnel.

Je suis très sensible à l'honneur que m'ont fait les membres du jury de thèse en acceptant de lire et d'évaluer ce travail. Je les remercie d'avoir apporté un intérêt à mon travail. Ce mémoire a beaucoup bénéficié des commentaires qu'ils ont tous effectués.

Je tiens également à exprimer toute ma gratitude à Jean-François Mouscadet pour son accueil chaleureux au sein de son équipe. Je tenais aussi à le remercier pour les conditions de travail idéales dans lesquelles s'est déroulée ma thèse.

Je tiens également à remercier tous les membres passés et présents de cette équipe qui m'ont soutenu tout au long de cette thèse. Merci à Olivier Delelis, Stéphanie Bury-Moné, Osamuede Osemwota, Gladys Mbemba, Gwenola Manic, Frédéric Subra, Sylvain Thierry, Aurélie Maurin-Marlin.

Je tiens, d'autre part à remercier le Professeur KaiCheng Qian de l'East China Normal University, directeur de Shanghai transfusion institute, qui m'a accueillie au sein de son labo en Chine. Je le remercie pour son aide lorsque j'ai rencontrée des problèmes administratifs en Chine et surtout pour toutes les discussions concernant mon avenir professionnel.

Je remercie également tous les membres du LBPA pour les temps inoubliables que j'ai passées en votre compagnie. En particulier, je souhaite remercier à Marie-Alix Poul pour tous les aides concernant mes travails d'anticorps. Merci également à mes collègues au LBPA : Luba Tchertanov, Hervé Leh, Françoise Simon, Eric Deprez, Françoise Chaminade, Safwat Abdel-Azeim, Janine, Anne-Marie, Martine, Marie-Christine, Yingying Chen, Xiaoqian Xu, Li Na, Huan Wang, Jiayao Li, Ahmad Khodr, Victoria Fairweather, Bianca Sclavi, Claude Nogues, Anne Olliver, Philippe Fossé,… Merci pour toutes les discussions scientifiques ainsi que pour les bons moments que nous avons pu

partager. Merci à tous pour m'avoir installée dans une ambiance amicale.

Je souhaite remercier particulièrement à Serge Benichou, Jerome Bouchet de Institut Cochin qui nous avez aidé sur les manipulations d'anticorps. Un grand merci également à Daniel Baty, Patrick Chames, Julie Matz à Marseille pour leurs conseils précieux et patiences.

Un grand merci également à Charlotte Charpentier, Diane Descamps d'Hôpital Bichat-Claude Bernard, et Anne-Geneviève Marcellin, Vincent Calvez d'Hôpital Pitié-Salpêtrière.

Je souhaite aussi remercier le programme de coopération entre l'GENS (Groupes des Ecoles Normales Supérieurs) et l'ECNU (East China Normal University), qui m'offre une formation complète et véritablement professionnelle. Merci à Madame Bogdana Neuville, Christine Rose et Brigitte Vitale à l'ENSC pour leur aide pour résoudre les problèmes administratifs et dans la vie quotidienne en France. Mes grands mercis également à Madame Yun-Hua Qian, Monsieur Hai-Sheng Li à l'ECNU pour leur aide concernant les affaires administratives en Chine.

Enfin, merci à tous mes amis en France qui me tiennent compagnie pendant certain périodes longues ou courts, Ting Wu, Hua Ren, Ya-Xin Pen, Zhong-wei Tang, Jian Chen, Hua Yi, Zhe Sun, Min Zhang, Yao Wang, Tong Bu, Sosuke Mizusaki, Jakub kallas..., pour leur humour, leur amitié, leurs conseils et soutiens de chaque instant lors des moments difficiles et les bon moments passés ensemble.

Mon grand et même le plus important merci est pour César Rodríguez qui est apparu dans ma dernière année de thèse, et fait cet année beaucoup plus jolie et plus facile à passer. Merci pour me rappeler de temps en temps que 'la vie est belle'.

Mes principaux remerciements s'adressent finalement à mes parents. Je vous remercie pour votre confiance que vous m'avez témoignée durant ce long parcours d'études. Merci de m'avoir donné la vie assez belle, de m'avoir aidée à croire en moi pendant toutes ces années d'études.

Merci à tous.

RESUME

Ce mémoire de thèse présente mes travaux sur la détermination de la susceptibilité et de la résistance des intégrases (INs) du virus de l'immunodéficience humaine (VIH) aux inhibiteurs de transfert de brins de l'IN (INSTIs) ainsi que le développement de fragments d'anticorps simple-chaîne (sdAbs) ciblant l'IN du VIH.

Tout d'abord, car les études antérieures ont suggéré que les variations significatives de l'IN de souche CRF02_AG pourrait avoir des effets consécutifs sur l'interaction entre l'inhibiteur et l'IN, la susceptibilité de l'IN de souche CRF02_AG du VIH-1 aux dernières INSTIs a été déterminée. Accord avec l'étude in silico, nous avons mis en évidence que l'activité de 3'-processing et de transfert de brin des INs de souche B et de souche CRF02_AG sont comparables. La susceptibilité des INs recombinantes de souche CRF02_AG aux INSTIs utilisés (Raltégravir-RAL, Elevitégravir-EVG et L-731, 988) est similaire à celle de l'IN de souche B, malgré les variations naturelles qui se produisent dans les INs de souche CRF02_AG. Le polymorphisme de l'IN de CRF02_AG n'a pas d'effet significatif sur la susceptibilité aux INSTIs.

Dans un second temps, la résistance de l'IN du VIH-2 au RAL, l'unique INSTI approuvé, a été confirmée in vitro avec des enzymes mutées portant des mutations de résistance. Les mutations aux positions 155 et 148 jouent un rôle similaire pour les VIH-1 et VIH-2, en rendant l'IN résistante au RAL. La mutation G140S confère peu de résistance, mais compense le défaut catalytique dû à la mutation Q148R. À l'inverse, Y143C seule ne confère pas de résistance au RAL excepté si la mutation E92Q est également présente. De plus, l'introduction de la mutation Y143C dans le mutant résistant N155H baisse le niveau de résistance de l'enzyme contenant la mutation N155H, ce qui pourrait expliquer l'absence de détection de ces deux mutations ensemble dans un seul génome.

Enfin, des anti-VIH sdAbs avec nombreuses propriétés intéressantes ont été sélectionnés pour développer des agents antirétroviraux. Après la sélection de sdAb ciblant l'IN du VIH, nous avons obtenu des qui sdAbs qui reconnaissent spécifiquement une vaste gamme d'INs in vitro, y compris le mutant G140S/Q148R résistant aux INSTIs. Néanmoins, l'activité inhibitrice des sdAbs n'a pas été observée. Les sdAbs ciblant l'IN du VIH peuvent être utilisés pour d'autres applications, telles que des réactifs ciblant des nanocapteurs. À l'avenir, en raison des avantages uniques des sdAbs, le développement de sdAbs anti-IN du VIH qui bloquent la réplication du VIH reste attractive pour l'obtenir des inhibiteurs efficaces de l'IN.

ABSTRACT

This thesis presents the determination of susceptibility and resistance of HIV integrases (INs) to IN strand transfer inhibitors (INSTIs) and the development of single domain antibody (sdAb) targeting HIV IN.

Firstly, the susceptibility of HIV-1 subtype CRF02_AG INs to the latest INSTIs was determined, since previous studies suggested that the significant variations of CRF02_AG IN may have consequential effects on the interaction between the inhibitor and IN. Consistent with in silico study, we found that 3'-processing and strand transfer activity of both HIV-1 subtype B IN and subtype CRF02_AG IN are comparable. The susceptibility of recombinant CRF02_AG INs to employed INSTIs (Raltegravir-RAL, Elevitegravir-EVG and L-731, 988) is similar to that of HIV-1 B IN. Hence, the polymorphism of CRF02_AG IN cannot significantly effect on the susceptibility to INSTIs.

Secondly, the resistance of HIV-2 IN to RAL, the unique approved INSTI, has been confirmed in vitro with mutated enzymes harboring resistance mutations. Mutations at positions 155 and 148 played a similar role in HIV-1 and HIV-2, rendering the IN resistant to RAL. The G140S mutation conferred little resistance, but compensated for the catalytic defect due to the Q148R mutation. Conversely, Y143C alone did not confer resistance to RAL unless E92Q is also present. Furthermore, the introduction of the Y143C mutation into the N155H resistant background decreased the resistance level of enzyme containing the N155H mutation, possibly accounting for the lack of detection of these two mutations together in a single genome.

Finally, anti-HIV IN sdAb that is endowed with many attractive properties was selected for developing antiretroviral agents. After the selections, we have obtained some sdAbs that specifically recognize a broad range of INs in vitro, including INSTI-resistance mutant G140S/Q148R. However, the inhibition activity of anti-HIV IN sdAbs has not been observed yet. Anti-HIV IN sdAbs can be applied for other application, such as targeting reagents for nanosensor. In future, development of anti-HIV IN sdAbs which are able to block HIV replication remains attractive for obtaining efficient inhibitor of IN.

TABLE OF MATERIALS

ABBREVIATIONS

AIDS: Acquired immunodeficiency syndrome
ART: Antiretroviral therapy
AZT: Zidovudine
BAF: Barrier-toautointegration factor
CA: Capsid
CCD: Catalytic core domain
CCR5: C-C Chemokine Receptor Type 5
CDK9: Cyclin-dependent kinase 9
CDR: Complementarity determining region
cPPT-CTS : central polypurine tract-central termination sequence
CRF: Circulating recombinant form
CTD: Carboxy-terminal domain
CXCR4: C-X-C chemokine receptor type 4
CYCT1: Cyclin T1
CypA: Peptidyl-prolyl isomerase cyclophilin A
DC: Dendritic cell
DKA: Diketoacid moiety
dNTP: deoxynucleoside triphosphate
DTT: Dithiothreitol
EED: Embryonic ectoderm development
ELISA: Enzyme-linked immunosorbent assay
EVG: Elvitegravir
FI: Fusion inhibitor
FV: Foamy virus
HAART: Highly active antiretroviral therapy
HCAb: Heavy-chain antibody
HIV: Human immunodeficiency virus
HMGA1: High mobility group protein A1
HR: Heptat repeat
HSP60: Heat-shock protein 60
HTLV-III: Human T-Cell leukemia virus III
IBD: Integrase binding domain
IN: Integrase
INI: Integrase inhibitor
INI1/hSNF5: Integrase interactor 1
Inr: Initiator
INSTI: Integrase strand transfer inhibitor
LEDGF/p75: Human lens epithelium-derived growth factor
LTR: Long Terminal Repeat
MA: Matrix

MHC: Major histocompatibility complex
MiRNA: micro-RNA
MOI: Multiplicity of infection
NC: Nucleocapsid
Nef: Negative effector
NF-B: Nuclear factor-B
NLS: Nuclear localisation signal
NNRTI: Non-nucleoside reverse transcriptase inhibitor
NPC : Nucleoprotein complex
NRTI: Nucleoside reverse transcriptase inhibitor
NTD: Amino-terminal domain
N-TEF: Negative transcription elongation factor
ODN: Oligonucleotides
PCR: Polymerase chain reaction
PI: Protease inhibitor
PIC: Pre-integration complex
PPT: Polypurine tract
3'-P: 3'-processing
P-TEFb: Positive transcription elongation factor b
RAL: Raltegravir
Rev: Regulator of virion gene expression
RNAPII: Polymerase II
RT: Reverse transcriptase
RTC: Reverse transcription complex
scFv: single-chain Fv
sdAb: Single-domain antibody
SIV: Simian immunodeficiency viruses
SPRi: Plasmon resonance imagery
ST: Strand transfer
TAM: Thymidine associated mutations
TAR: Trans-activation response element
Tat: Transcriptional transactivator
Th: T helper
TRIM5 : Ttripartite motif protein 5
TRN-SR2 : Transportin SR-2
TSG101: Tumour-susceptibility gene 101
URF: Unique recombinant form
VH: Variable domain
VHH: Variable domain of heavy chain of heavy chain antibody
Vif: Viral infectivity factor
Vpr: Viral protein r
Vpu: Viral protein u

INTRODUCTION

I. Human immunodeficiency virus (HIV)

Nearly 30 years after the first reported cases of the acquired immunodeficiency syndrome (AIDS) and the discovery of the etiologic agent, effective control of the AIDS pandemic which is arguably the most serious infectious disease remains elusive. It is estimated that 33.3 million people worldwide are living with HIV in 2009, and 2.6 million people newly infected with HIV in that year alone.

The origin of this challenge is the molecular pathogenesis of HIV, a virus that has evolved a number of mechanisms to escape from host immune control. Among the most notable strategies are the virus' direct targeting of the CD4+ molecule expressed by the key T lymphocyte in human immune system; the high rate of mutation which enable the virus to evade the host immune system (mutational escape); the highly glycosylation of the external glycoprotein, which protects neutralization epitopes; and integration into the human genome, which implies that cells which are not killed are permanently infected.

I.1. Identification of HIV

AIDS was first recognized in the United States in 1981, following an increase in the incidence of usually rare opportunistic infections such as the pneumonia caused by Pneumocystis carinii in homosexual men which were caused by a general immune deficiency. A dramatic reduction of CD4+ lymphocyte cells in peripheral blood was the common characteristic of all the patients.

In 1983 HIV was isolated by the team of L.Montagnier of Pasteur Institute at Paris (Barré-Sinoussi et al., 1983). The activity of reverse transcription was identified from a sample of patient with lymphoadenopathy. The presence of a retrovirus initially named as lymphoadenopathy associated virus (LAV) was confirmed by the visualization using electronic microscopy. Simultaneously, some American teams identified two virus considered as the pathogenic agent responsible of AIDS, human T-Cell leukemia virus III (HTLV-III) and AIDS associated-virus (Gallo et al., 1984; Levy et al., 1984). In 1985, the scientific community concluded that these 3 viruses are identical. This virus was nominated as HIV-1 in 1986 by the discovery of the second similar virus with slightly different genome structures, isolated from two West-African patients with AIDS and then named as HIV-2. By its genomic sequences and its proteins, this virus is different from the HIV-1 isolated from U.S.A., Europe and Central Africa. It differs also from simian immunodeficiency virus (SIV), but displays an antigenic relationship with the latter

virus, at the level of its external envelope protein (Clavel et al., 1986).

I.2. Taxonomy of the retrovirus family

The reference of classification is from the International committee of taxonomy of viruses (ICTV, http://www.ictvonline.org/virusTaxonomy.asp). The affiliation of one family is defined by common taxonomic denominators: genetic information as well as the replicative properties. Certain structure elements, as the presence of envelope and pathogenesis are equally counted in some cases.

Retroviruses are further divided into seven groups defined by evolutionary relatedness, each with the taxonomic rank of genus. Five of these groups represent retroviruses with oncogenic potential (formerly referred to as oncoviruses), and the other two groups are the lentiviruses and the spumaviruses **(Table 1)**.

Retroviruses can be broadly divided into two categories, simple and complex, distinguishable by the organization of their genomes (Murphy et al., 1994). All oncogenic members except the human T-cell leukemia virus-bovine leukemia virus (HTLV-BLV) genus are simple retroviruses. HTLV-BLV and the lentiviruses and spumaviruses are complex.

HIV-1 and HIV-2 are both characterized by extensive genetic diversity. HIV-1 is phylogenetically divided into three groups, major (M), outlier (O), and nonmajor and nonoutlier (N), with the M group further split into 9 subtypes and 15 circulating recombinant forms. Today, group M has a near global distribution, whereas groups N and O are restricted to individuals of West African origin. HIV-2 is also most common in individuals from West Africa and is composed of seven subtypes. Despite its initial association with homosexual men, it is clear that HIV-1 and HIV-2 are now primarily transmitted by heterosexual intercourse and from mother to infant.

Subfamily	Genus name	Type species name
00.061.1. Orthoretrovirinae	00.061.1.01. Betaretrovirus	00.061.1.01.001. Mouse mammary tumor virus
	00.061.1.02. Gammaretrovirus	00.061.1.02.001. Murine leukemia virus
	00.061.1.03. Alpharetrovirus	00.061.1.03.001. Avian leukosis virus
	00.061.1.05. Deltaretrovirus	00.061.1.05.001. Bovine leukemia virus
	00.061.1.06. Lentivirus	00.061.1.06.001. Human immunodeficiency virus 1
	00.061.1.08. Epsilonretrovirus	00.061.1.08.001. Walleye dermal sarcoma virus
00.061.2. Spumaretrovirinae	00.061.2.07. Spumavirus	00.061.2.07.001. Simian foamy virus

Table 1: Composition of retroviridae family. (ICTVdB - The Universal Virus Database, version 3. http://www.ncbi.nlm.nih.gov/ICTVdb/ICTVdB/)

I.3. Origins of HIV

The key to understanding the origin of HIV was the discovery that closely related viruses, the simian immunodeficiency viruses (SIVs), were present in a wide variety of African primates. Collectively, HIV and SIV comprise the primate lentiviruses, and SIVs have been isolated in more than 20 African primate species.

The evolutionary history of HIV-1 and HIV-2 has been reconstructed in great detail by inferring phylogenetic trees of the primate lentiviruses. It was discovered that the two human viruses are related to different SIVs and therefore have different evolutionary origins **(Figure 1)**. Specifically, HIV-1 is most closely related to SIVcpz, which is found in some sub-species of chimpanzee that inhabit parts of equatorial Western and Central Africa, respectively (Gao et al., 1999; Santiago et al., 2002). This HIV-1 progenitor probably was passed from chimpanzees to human hunters through blood borne transmission.

Phylogenetic analysis of HIV-1 and related viruses from non human primates suggests that three independent transmission events early in the 20th century spawned three HIV-1 groups: M, N and O. Although strains related to the M and N groups have been found in chimpanzees, recent evidence suggests that group O HIV-1 may have originated in gorillas, in which the closest relatives of this group have been identified (Taylor et al., 2008). It is speculated that the virus then spread among humans along the Congo River into Kinshasa, Zaire, where the earliest documented case of HIV-1 infection (with group M strain) in humans has been traced to a blood sample from 1959 (Zhu et al., 1998).

By contrast, HIV-2 is most closely related to SIVsm (Gao et al., 1992), which is found at

high prevalence in sooty mangabey monkeys. Sooty mangabeys are most frequent in the regions of West Africa where HIV-2 is likely to have emerged.

Molecular phylogenies also show that there have been many cross-species transmissions to humans, because there is a mixing of HIV and SIV lineages. Although the vagaries of sampling make it difficult to determine exactly how many cross-species transfers have occurred, three jumps from chimpanzees to humans are considered to explain the current diversity of HIV-1, such that groups M, N and O each have an independent origin (Gao et al., 1999).

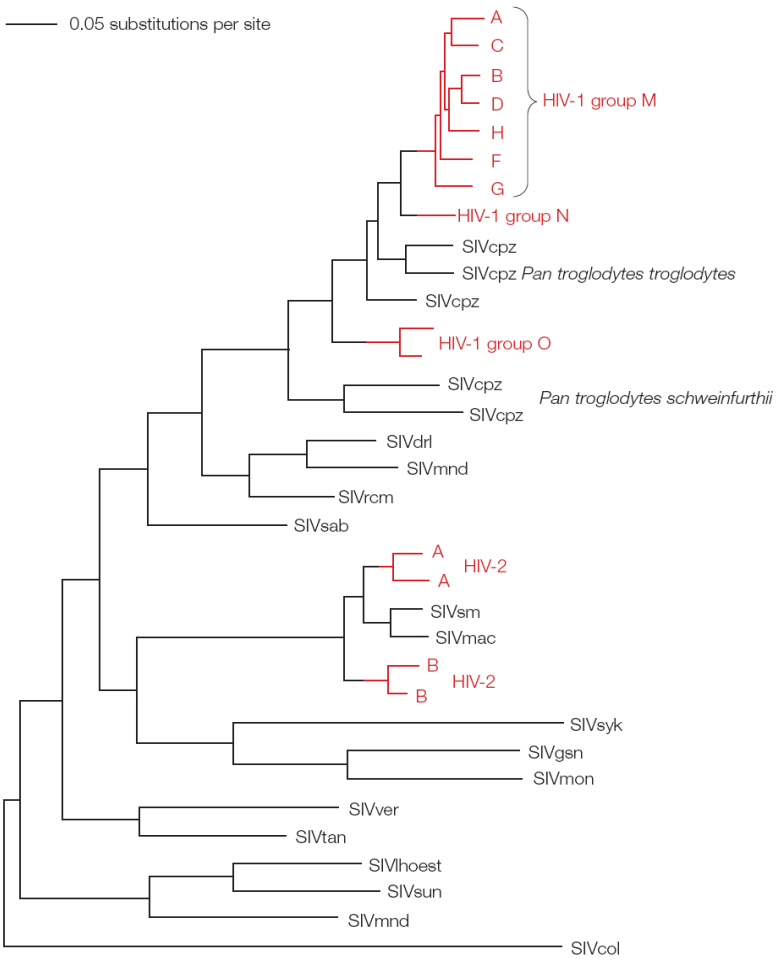


Figure 1: Evolutionary history of the HIV-1 and HIV-2.

Because both the HIV-1 and HIV-2 lineages (red branches) fall within the SIVs that are isolated from other primates, they represent independent cross-species transmission events. HIV-1 groups M, N and O represent separate transfers from chimpanzees (SIVcpz), again because there is a mixing of the HIV-1 and SIV lineages. Similarly, HIV-2 seems to have been transferred from sooty mangabey monkeys (SIVsm). Only some subtypes of HIV-1 and HIV-2 are shown for clarity (Rambaut et al., 2004).

I.4. Basics of HIV replication

HIV enters the body through the exchange of bodily fluids, and it infects mainly T helper (Th) cells, macrophages and, to some extent, microglial cells and dendritic cells (DCs). This tropism is determined at the level of viral entry by the use of CD4 as a primary receptor and the use of co-receptors which are strain and target specific **(Figure 2)**. HIV R5 strains use C-C chemokine receptor type 5 (CCR5) as their co-receptor and can, therefore, enter macrophages, DCs and T cells, whereas X4 strains of HIV use C-X-C chemokine receptor type 4 (CXCR4) as a co-receptor and can infect T cells only (Doms and Trono, 2000). Early in infection, only R5 viruses can be detected in infected individuals. With time, X4 viruses come to predominate, which hastens the demise of Th cells, the hallmark of AIDS.

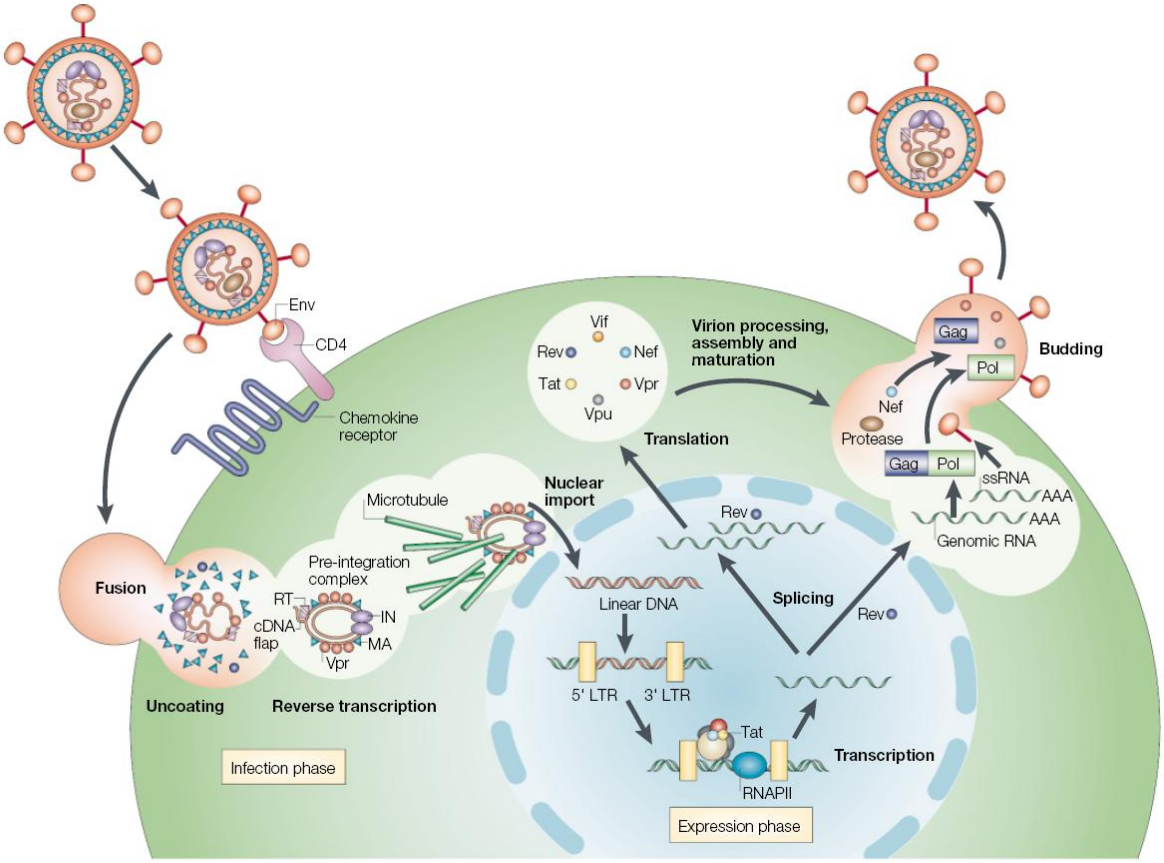


Figure 2: The replicative cycle of HIV (Peterlin and Trono, 2003).

Upon membrane fusion, the viral core, containing the viral capsid (CA) and nucleocapsid (NC) along with the viral genome, reverse transcriptase (RT), integrase (IN), protease (PR), the viral accessory proteins Vif, Nef and Vpr, and regulatory proteins is released into the cytoplasm. Collectively called reverse transcription complex (RTC), this assembly traffics along the microtube network from the cell periphery toward the centrosome. Within the RTC, RT converts the viral RNA genome into a double-stranded viral complementary DNA, usually occurring upon entry into the host cell.

Reverse transcription and viral core disassembly are concomitant, tightly coupled processes (Zhang et al., 2000). More recent results indicate that viral core disassembly might occur later near the centrosome (Lehmann-Che et al., 2005) or the nuclear envelope (Arhel et al., 2007), likely depending on the virus and/or the cell metabolism. Uncoating might be the rate-limiting step that determines nuclear import of the viral genome (Lewinski et al., 2006).

Reverse transcription process is quite complex and requires different catalytic activities. The process is initiated from the 3' end of a primer tRNA3Lys partially hybridized to the primer binding site (PBS) which is located near the 5' end of the viral genome. It results in the synthesis of a short stretch of minus, single-stranded DNA, which then relocates and hybridizes to the repeat sequence at the 3' end of the RNA genome. After this strand transfer, the synthesis of the first DNA strand can continue. The synthesis of the first DNA strand requires multiple enzymatic activities: RNA-primed RNA-dependent DNA polymerization (RDDP), DNA-primed RDDP, strand-transfer, and RNaseH activity. The RNaseH which is embedded in the RT enzyme hydrolyzes the copied RNA template prior to strand-transfer. The synthesis of the second DNA strand uses RNA fragments still hybridized to the first DNA strand, like the polypurine tract (PPT), as primers, because of the RNaseH resistant conformation of those RNA/DNA duplexes. The second DNA strand synthesis also requires a strand transfer reaction. While reverse transcription proceeds, the RTC undergoes a progressive remodeling to become a pre-integration complex (PIC) as the viral nucleoprotein complex competent for proviral integration.

By undergoing an active phase of nuclear import through the nuclear pore, HIV does not require disassembly of the nuclear membrane during cell division, which is unlike to the other retroviruses. This property has been ascribed to the presence of karyophilic elements present in viral nucleoprotein complexes (NPCs) (Bukrinsky et al., 1992; Suzuki and Craigie, 2007). However, the role of different compenents of NPC in nuclear import remains controversial. Human lens epithelium-derived growth factor (LEDGF/p75) which contains the basic and classical nucleusnuclear localization signal (NLS) consistent stretch

146RRGRKRKAEKQ156 is involved in tethering the viral PIC to chromatin (Maertens et al., 2004b; Maertens et al., 2003; Meehan and Poeschla, 2010; Vanegas et al., 2005). IN residues 161 to173 had been thought as a transferable karyophilic signal (Bouyac-Bertoia et al., 2001); however, further studies suggested that no NLS is present at this location (Dvorin et al., 2002; Limón et al., 2002) . Moreover, it was shown that a three-stranded central DNA flap which is a structure present in neo-synthesized viral DNA, specified by the central polypurine tract-central termination sequence (cPPT-CTS), acts as a cis-acting determinant of HIV-1 genome nuclear import (Arhel et al., 2007; Rivière et al., 2010; Zennou et al., 2000). These elements may mediate nuclear import directly or via the recruitment of the host's proteins. Indeed, several cellular proteins have been found to influence HIV-1 infection during nuclear import, such as the karyopherin α2 Rch1 (Gallay et al., 1996); importin 7, an import receptor for ribosomal proteins and histone H1 (Fassati et al., 2003); the transportin SR-2 (TRN-SR2) as a binding partner of HIV-1 IN (Brass et al., 2008; Christ et al., 2008); or the nucleoporins Nup98 (Ebina et al., 2004), Nup358/RANBP2, and Nup153 (Brass et al., 2008; König et al., 2008). By interacting directly with the NPC via the binding of PIC-associated IN to the C-terminal domain of Nup153, HIV-1 sub-viral particles gain access to the nucleus (Woodward et al., 2009).

In vivo, integration is a rare event. The integrated viral genome represents only 5 to 10% of the total viral DNA content of an infected cell (Brussel and Sonigo, 2004). Both the extra linear molecules and circular molecules, containing 1 or 2 long terminal repeats (LTR), represent the most abundant forms of the viral DNA genome. Only the linear form cDNA can integrate into the host genome, yielding the LTR flanked provirus. The linear double-stranded cDNAs have IN and chromatin-remodelling complexes at their termini (Miller et al., 1997b). Broadly sketched, lentiviruses favor integrating in active transcription units without favoring the promoter regions in particular. HIV integration is also strongly disfavored in centromeric heterochromatin, which is transcriptionally repressed and a poor substrate for viral transcription (Poeschla, 2008). The full-length linear cDNA decays rapidly if not integrated (estimated half life of 1 to 2 days) (Zhou et al., 2005), while 1- and 2-LTR circles accumulate in the nucleus at later times of infection. These circular forms, which are considered dead-end products of aborted integration events, can also lead to viral gene expression (Gelderblom et al., 2008).

Once integrated in the host genome, the provirus behaves like any human gene or so Pol II-transcribed genes. At this point, the 5’ LTR which contains enhancer and promoter sequences with binding sites for several transcription factors, behaves like any eukaryotic promoter; and the 3’ LTR acts as the polyadenylation and termination site (Peterlin and

Trono, 2003). Moving upstream from the transcription start site, the initiator (Inr), the TATA Box and three SP1-binding sites are found. These elements position RNA polymerase II (RNAPII) at the correct site for initiating transcription. Further upstream is the enhancer, which binds nuclear factor-B (NF-B) and nuclear factor of activated T cells (NFAT), as well as members of the E-twenty six (ETS) family of transcription factors (Jones and Peterlin, 1994). These activators ensure that the virus replicates at a high level in activated T cells and differentiated macrophages.

Additionally, the presence of a strong regulatory element located 3' to Inr is the most unusual feature of the LTR. This RNA structure, which is found at the 5' end of all viral transcripts, is known as the trans-activation response element (TAR) and it binds the transcriptional transactivator (Tat) (Taube et al., 1999). In the absence of Tat, HIV transcription begins, but elongation is inefficient**.** Tat and its cellular cofactor, positive transcription elongation factor b (P-TEFb), cooperate to bind TAR with high affinity, allowing RNAPII to produce full-length viral transcripts. P-TEFb contains two components: cyclin T1 (CYCT1) and cyclin-dependent kinase 9 (CDK9). Once recruited to the nascent HIV RNA, CDK9 phosphorylates the carboxy-terminal domain of RNAPII and negative transcription elongation factor (N-TEF), thereby allowing efficient elongation (Price, 2000). The ability of Tat to recruit P-TEFb through an RNA sequence is unique among transcriptional activators, and it renders HIV replication particularly sensitive to inhibition by compounds that target CDK9. In contrast, Tat is inefficient at recruiting RNAPII, and it requires a strong basal promoter for optimal effects (Jones and Peterlin, 1994). NF-B, which also recruits P-TEFb, can substitute partially for Tat (Barboric et al., 2001). Furthermore, Tat and P-TEFb affect the ability of HIV to establish latency and affect the transcription of MHC class II genes.

Viral proteins are synthesized from more than 30 mRNA species, which are all derived from the same primary transcript. These transcripts that encode the late viral proteins, which are mainly structural and enzymatic components of the virion and factors that fine-tune infectivity, are singly spliced or unspliced, while those encoding the most of early regulatory proteins are fully spliced.

The timely production of viral gene products requires the regulator of virion gene expression (Rev), which transports slightly spliced and unspliced genomic transcripts from the nucleus to the cytoplasm. To function optimally, this protein needs to reach threshold levels (Pomerantz, 1992), the splicing of genomic transcripts must be slow and an active Crm1/RanGTP complex must be present in the cell. For HIV and other lentiviruses, the only really fundamental post-integration peculiarity in the nucleus is a necessary

mechanism to bypass the cells splicing checkpoint for the viral genome RNA and some coding mRNAs.

After their translation, viral structural and enzymatic proteins travel to the plasma membrane, where immature virions assemble in cholesterol-rich lipid rafts (Nguyen and Hildreth, 2000). The carboxy-terminus of Gag, p6, is ubiquitylated, and it recruits components of multivesicular bodies, such as tumour-susceptibility gene 101 (TSG101) and vacuolar protein sorting 4 (VPS4), which facilitate the release of progeny virions from the cell. Processing of Gag and Gag-Pol yields mature HIV particles. Some accessory viral proteins such as Vpr and Nef, as well as cellular components, e.g., MHC class I and II molecules, and some CD proteins, are incorporated into virions (Esser, 2001). The envelope glycoprotein is an essential viral component that allows the infection of new cells. The high cholesterol content of the virion, which is a consequence of its budding through rafts, is crucial for this process as well.

I.5. Genomic organization and structure of HIV

I.5.1. Genomic organization

The genome of HIV composes a double copy of positive-sense, single stranded RNA of about 9 kilo nucleotides and encodes 15 distinct proteins **(Figure 3)**. The two RNA molecules are linked by their 5' extremity and classic proprieties of eukaryotic mRNA: a 5' cap (m7GpppGm) and a poly (A) tail at 3' extremity. The RNA is flanked by repeated R sequence and U5 region at 5' and U3 region at 3'. Reverse transcription produces a duplication of region U5 and U3 when RNA is transformed into DNA double strands. The new extremity called long terminal repeat (LTR) contains some essential elements for viral replication (e.g., promoter region at 5' U3 and poly-adenylation site at 3' U3) and sites for integration (sequence ACGT at 5' and CAGT at 3').

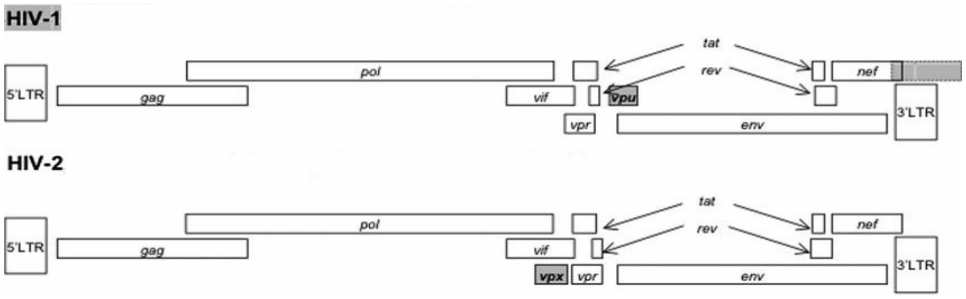
From the point of genetic view, HIV is a complex virus which codes for the characteristic retroviral genes (*gag, pol* and *env*) but also for regulatory proteins. The gene *gag* (group antigen) codes for polyprotein $Pr55^{gag}$. This precursor is digested by viral protease to release different virion elements. The viral protease, integrase, and reverse transcriptase are always expressed within the context of a Gag-Pol fusion protein. The Gag-Pol precursor (p160) is generated by a ribosomal frame shifting event, which is triggered by a specific cis-acting RNA motif (Dulude et al., 2002). When ribosomes encounter this motif, in approximate 5% of the cases, they shift to the pol reading frame (at

position -1 according to gag reading frame) without interrupting translation.

I.5.2. Structure of HIV

The viral particles exist as spheres of approximately 80 to120 nm in diameter with a spherical morphology **(Figure 4)**. The cone-shaped core, which contains 2 copies of genomic RNA, reverse transcriptase, integrase, protease, and etc, is surrounded by lipid matrix containing key surface antigens and glycoproteins. Those proteins that are derived from the *gag* (group-specific antigen), *pol* (polymerase) and *env* (envelope) genes are classical structural and enzymatic factors that are required by all retroviruses.

A



B

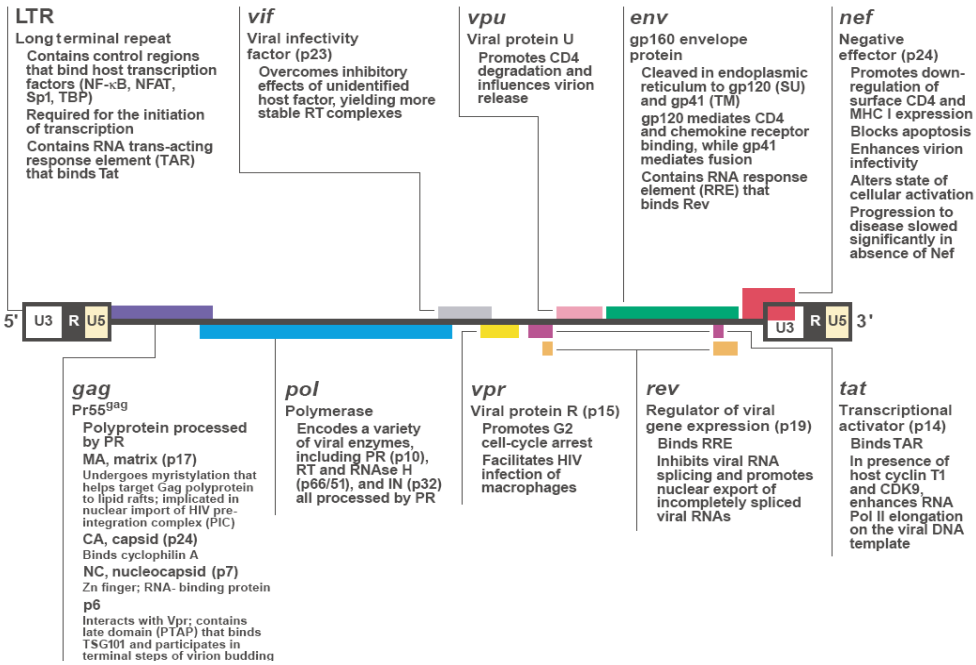


Figure 3: Genomic organization of HIV-1 and HIV-2 (A) and a summary of the functions of 9 genes encoding 15 proteins of HIV provirus (B) (Greene and Peterlin, 2002).

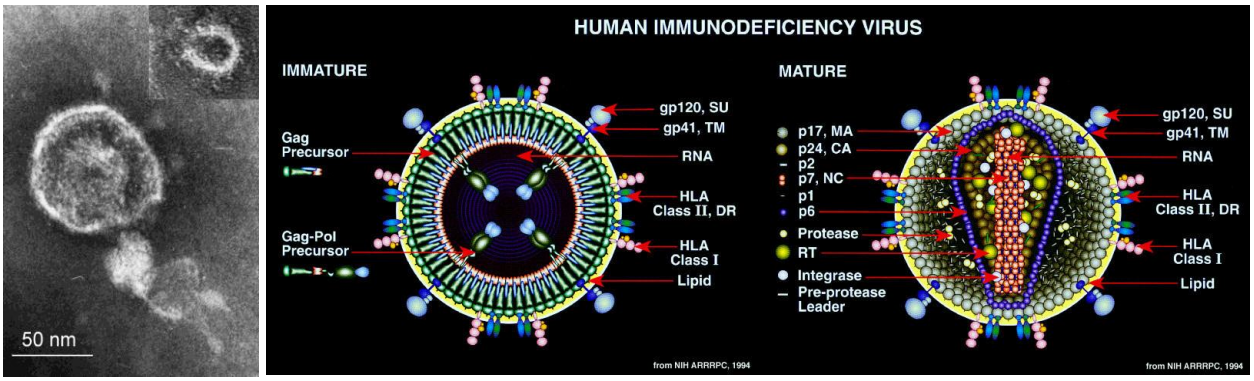


Figure 4: Photo and schematic structure of viral particle.

http://www.ictvdb.org/WIntkey/Images/dg_HIV.jpg

Viral envelope (Env)

The envelope (Env) spikes on HIV-1 define the viral tropism, mediate the fusion process and are the prime target of the humoral response. Three gp120 subunits comprise the 'head' of Env and three gp41 subunits comprise the 'stalk' and other membrane-associated elements.

The gp120 moiety has five hypervariable regions, designated V1 through V5 loop. V3 loop is not involved in CD4 binding, but is rather an important determinant of the preferential tropism of HIV-1 for either T lymphoid cell lines or primary macrophages (Hwang et al., 1991). Sequences within the V3 loop interact with the HIV chemokine co-receptors CXCR4 and CCR5, which partially determine the susceptibility of cell types to given viral strains (Deng et al., 1996). The V3 loop is also the principal target for neutralizing antibodies that block HIV-1 infectivity.

Structural proteins

Gag precursor protein, also called p55, is expressed from the unspliced viral mRNA. The membrane-associated Gag polyprotein recruits two copies of the viral genomic RNA along with other viral and cellular proteins that triggers the budding of the viral particle from the surface of an infected cell. After budding, p55 is cleaved by the virally encoded protease during the process of viral maturation into four smaller proteins designated matrix (MA, p17), capsid (CA, p24), nucleocapsid (NC, p7) and p6. MA is a 132 amino acid protein and

post-translationally myristylated at the N-terminus. MA and CA play critical roles in both late and early stages of the viral replication cycle. NC contains two successive zinc fingers, promotes viral RNA dimerization and encapsidation and activates annealing of the primer tRNA to the initiation site of reverse transcription (Morellet et al., 1992). The p6 protein regulates the final abscission step of nascent virions from the cell membrane by the action of two late assembly (L-) domains. The 52 amino acid peptide of p6 protein binds at least to two cellular budding factors (Tsg101 and ALIX), and mediates the incorporation of the HIV-1 accessory protein Vpr into viral particles (Votteler et al., 2011).

Nonstructural proteins

During viral maturation, the virally encoded protease cleaves Gag-Pol precursor, releases the Pol polypeptide away from Gag and further digests it to separate the protease, reverse transcriptase, and integrase.

- ***Protease (Pro)***

HIV protease activity is not required for virus production and release per se, but is essential for viral maturation leading to infectious viral particles. It generates mature infectious virus particles through cleavage of the viral Gag and Gag-Pol precursor proteins.

HIV protease is in the family of aspartic proteases (with an aspartic acid in the active site at position 25) and is a symmetrically assembled homodimer consisting of two identical subunits of 99 amino acids (Navia et al., 1989; Wlodawer et al., 1989). The centre of the enzyme is formed by the substrate-binding cleft, which interacts with the different substrate cleavage site sequences in the Gag and Gag-Pol proteins. Dimerization of Gag-Pol precursor proteins is required by the activation of viral protease which is embedded in the Gag-Pol protein. Despite its critical function in viral maturation and infectivity, HIV protease shows great plasticity and polymorphisms which are observed at one-third of the 99 amino acids (Rhee et al., 2003).

HIV-1 protease recognizes the asymmetric shape of the peptide substrates, rather than a particular amino acid sequence of the cleavage site (Prabu-Jeyabalan et al., 2002). The peptide substrates have a secondary structure, yielding a substrate ‘envelope’ which fits within the protease substrate-binding region. However, a few differences between the cleavage site substrates in which amino acid side chains protrude out of the ‘envelope’. These amino acid side chains protruding out of the ‘envelope’, which makes small differences between cleavage sites, contribute to the highly ordered and regulated process in which all individual cleavages occur at different rates. The ordered cleavage appears to

be regulated by the amino acid sequence near the actual protease cleavage site.

- ***Reverse transcriptase (RT)***

The reverse transcriptase is an asymmetric heterodimer comprising a p66 subunit (560 amino acids) and a p51 subunit (440 amino acids) (Kohlstaedt et al., 1992). Both subunits are encoded by the same sequence in the viral genome. The DNA- and RNA-dependent DNA polymerase activity is catalyzed by the p66 subunit, which is responsible for the critical step of conversion of the minus stranded RNA viral genome into double stranded DNA. RNase H consists of the last (carboxy terminal) 120 amino acids of the p66 subunit, which correspond to the p15 fragment cleaved from the p66 subunit by the viral protease to generate the p51 subunit. RNase H activity removes the original RNA template from the first DNA strand, allowing synthesis of the complementary strand of DNA.

Like other polynucleotide polymerases, the overall three-dimensional structure of the p66 subunit is often compared to a right hand, with a fingers (amino acids 1-85 and 118-155), a palm (amino acids 86-117 and 156-237) and a thumb (amino acids 238-318) domain (Steitz, 1999). The palm domain contains the polymerase active site with its three aspartic acids (110, 185 and 186) and the YMDD characteristic motif. The incoming dNTP binds between the palm and the finger subdomains and the ribose and base make important contacts with residues including L74, Y115, M184 and Q151. The catalytic pocket is formed by the fingers folding down into the palm domain (Huang et al., 1998).

Next to the catalytic domain, the p66 subunit also contains the RNaseH domain (amino acids 427-560), linked by the connection domain (amino acids 319-426). The connection domain is also involved in interactions with the nucleic acid and the p51 subunit.

A two-metal model has been proposed for the phosphoryl transfer reaction of polymerases including HIV RT (Steitz, 1999). In the RT active site, two Mg^{2+} ions are coordinated by the catalytic triad of D110, D185 and D186 that, along with residues including K65, interact with triphosphate and 3’-terminus of the primer (Huang et al., 1998).

Structural studies support an induced fit model where proper base pairing by the incoming dNTP results in formation of a closed polymerase, primer/template, and dNTP complex with the incoming dNTP appropriately aligned to be attacked by the 3’ OH at the terminus of the elongating primer strand (Doublie et al., 1998). Kinetic studies showed a corresponding rate limiting step that might reflect this conformation change. The dependence of the rate-limiting step on correct dNTP binding and base-pairing forms the basis for the fidelity of polymerization (Joyce and Benkovic, 2004).

Despite their sequence homology, the p66 subunit assumes a flexible and open structure,

whereas the p51 subunit is rather compact, and seems to play a structural role, devoid of catalytic activity, with the three aspartic acids buried inside.

- ***Integrase (IN)***

The IN protein mediates the insertion of the HIV proviral DNA into the genomic DNA of an infected cell. The structure and function of IN will be described in more detail in part II. HIV integrase and integration.

Regulatory proteins and accessory proteins

HIV encodes two regulatory proteins: the transcriptional transactivator (Tat) and the regulator of virion gene expression (Rev), and four so-called accessory proteins: the ill-named 'negative effector' (Nef), viral infectivity factor (Vif), and the viral proteins r (Vpr), u (Vpu) for HIV-1 and u (Vpx) for HIV-2.

Rev is a 13 kDa sequence-specific RNA binding protein. Produced from fully spliced mRNAs, Rev regulates the transition from the early to the late phases of viral gene expression by allowing the export of incompletely spliced and unspliced viral mRNA species from the nucleus to the cytoplasm (Kim et al., 1989).

Tat is capable to interact with a couple of cellular proteins. Tat binds to a short-stem loop structure, known as the transactivation response (TAR) element that is located at the 5' terminus of HIV RNAs. Tat binding occurs in conjunction with cellular proteins that contribute to the effects of Tat. The binding of Tat to TAR activates transcription from the HIV LTR at least 1000 fold (Feng and Holland, 1988; Kao et al., 1987).

Nef has been shown to have multiple activities, including the down regulation of the cell surface expression of CD4, the perturbation of T cell activation, and the stimulation of HIV infectivity. Tat and Nef are not only crucial for high levels of HIV replication, but also have important roles in promoting viral immune evasion.

Vpr is a multifunctional accessory protein. It plays a role in the ability of HIV to infect non dividing cells by facilitating the nuclear localization of the PIC (Heinzinger et al., 1994). Other critical functions include transcriptional coactivation of viral and host genes, induction of cell-cycle arrest, both direct and indirect contributions to T-cell dysfunction, and regulation of nuclear factor kappa B (NF-κB) activity (Kogan and Rappaport, 2011).

Vpu of 16 kDa is unique to HIV-1. There is no similar gene of Vpu in HIV-2 or SIV. HIV-1 Vpu is an oligomeric, type I integral membrane phosphoprotein. Two distinct biological activities, via distinct molecular mechanisms, have been attributed to Vpu: enhancement of viral particle release from the plasma membrane of infected cells and the

specific degradation of the HIV-1 receptor molecule CD4 in the endoplasmic reticulum (ER) (Subbramanian and Cohen, 1994; Willey et al., 1992). Vpu can efficiently antagonize the restriction of releasing enveloped viruses from infected cells which is imposed by Tetherin (BST-2/CD317), an interferon-inducible antiviral host protein (Neil et al., 2008; Van Damme et al., 2008). **Vpx** of 12 kDa was found only in HIV-2/SIVagm. Vpx function in relation to Vpr is not fully elucidated.

Vif is a 23 kDa polypeptide that is essential for the replication of HIV in peripheral blood lymphocytes, macrophages, and certain cell lines (Strebel et al., 1987). Vif inhibits the antiviral activity of the cellular protein APOBEC3G from entering the virion during budding from a host cell by targeting it for ubiquitination and proteasomal degradation (Lecossier et al., 2003; Sheehy et al., 2002; Yu et al., 2003).

Cellular proteins in viral structure

Certain cellular compositions are packed during virus assembly. This encapsidations are due to the interaction of viral components and cellular proteins. For example, the peptidyl-prolyl isomerase cyclophilin A (CypA), binding a proline-rich loop on the surface of HIV-1 CA, is involved in assembly. This interaction increases HIV-1 infectivity in humans but promotes an anti-HIV-1 restriction activity in non-human primates (Sokolskaja and Luban, 2006). CypA is a ubiquitous cytoplasmic protein that catalyzes the cis/transisomerization of peptidyl-prolyl bonds. Via isomerization of a peptide bond in CA conformation (Bosco et al., 2002), CypA modulates HIV-1 virion core detection by a class of innate pattern recognition molecule tripartite motif protein 5α (TRIM5α) (Stremlau et al., 2004). It is speculated that the rapid turnover of TRIM5α and presumably TRIM5α-virus complexes by proteasomes leads to an early block to infection, before the virus has had the opportunity to reverse transcribe. How exactly TRIM5α renders the virus uninfectious remains unclear (Towers, 2007).

During budding, viral particle is enveloped with the help of lipid bilayer of infected cell. It is also possible to find some cellular proteins on the surface of virus. Among them, the antigen of major histocompatibility complex (MHC) can be numerous and as important as the viral glycoproteins (Arthur et al., 1992).

I.6. HIV molecular epidemiology and mechanism of genomic diversity

I.6.1. HIV molecular epidemiology and circulating recombinant forms

HIV-1 subtypes, also called clades, are phylogenetically linked strains of HIV-1 that are approximately the same genetic distance from one another; in some cases, subtypes are also linked geographically or epidemiologically. Genetic variation within a subtype can be 15 to 20%, whereas variation between subtypes is usually 25 to 35%. Group M is the predominant circulating HIV-1 group which has a near global distribution. It has been divided into subtypes, denoted with letters, and sub-subtypes, denoted with numerals: subtypes and sub-subtypes A1, A2, A3, A4, B, C, D, F1, F2, G, H, J, and K.

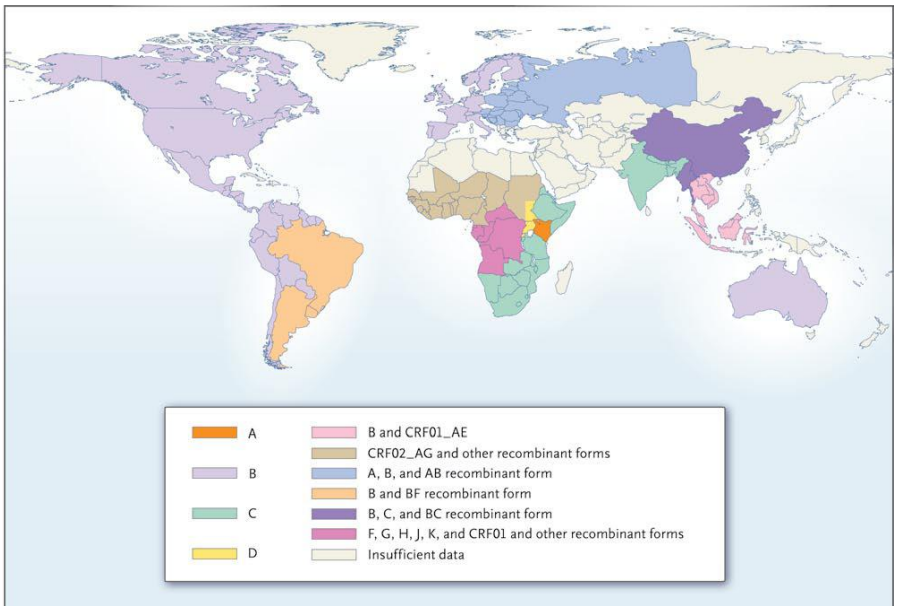
Over the past decade, advances in full-genome sequencing of HIV have led to the identification of circulating and unique recombinant forms (CRFs and URFs, respectively). These are the result of recombination between subtypes within a dually infected person, from whom the recombinant forms are then passed to other people. The recombinant progeny are classified as circulating recombinant forms if they are identified in three or more people with no direct epidemiologic linkage; otherwise they are described as unique recombinant forms (Taylor et al., 2008). Among the 40 identified CRFs, the most described are CRF01_AE playing an important role in epidemic in Southeast Asia. CRF02_AG is responsible for the epidemic in West Africa, and more and more frequently found in other countries (Machado et al., 2009; Malet et al., 2008c; Njai et al., 2006; Ye. J. et al., 2011).

The global distribution of subtypes and CRFs reflects the complexity of the molecular epidemiology of HIV-1 (**Figure 5**). In 2004-2007, subtype C accounted for nearly half (48%) of all global infections, followed by subtypes A (12%) and B (11%), CRF02_AG (8%), CRF01_AE (5%), subtype G (5%) and D (2%). Subtypes F, H, J and K together cause less than 1% of infections worldwide. Other CRFs and URFs are responsible for 4% of global infections respectively, rendering the combined total of worldwide CRFs to 16% and all recombinants (CRFs along with URFs) to 20% (Hemelaar et al., 2011). Analysis of the global distribution of HIV-1 demonstrated a broadly stable distribution of HIV-1 subtypes worldwide with a notable increase in the proportion of CRFs (Hemelaar et al., 2011). CRF02_AG isolates showed a slightly but not significantly higher replication advantage compared with subtype A or G isolates. This advantage in replicative fitness of CRF02_AG may contribute to its dominant spread over A and G subtypes in West and West Central Africa (Njai et al., 2006). Recombination that may contribute to rescue high-fitness HIV-1 variants is a common feature among retroviruses and, as well, between HIV-1 strains, particularly clade A (Spira et al., 2003).

From 1985, it was demonstrated that the genetic variability of HIV is very important. Some further studies showed that it does not exist two identical viral strains. And in one

individual the virus presents in the form of a polymorphic viral population with a multitude of different genomes called 'quasi species'. In a recently infected patient, the circulating viruses are genetically homogeneous. But the viral population evolutes over time with a global rate of change (estimated at 1% each year for gene env, 0.5% for gene *gag*). A complex mixture of variants progressively appears and evolutes in different and independent way in different tissues and cells over time. The initial view that the virus is classifiable into distinct subtypes or clades now needs to reflect the reality of a dynamic genetic evolutionary process, through which new HIV-1 strains are constantly emerging (Taylor et al., 2008).

A



B

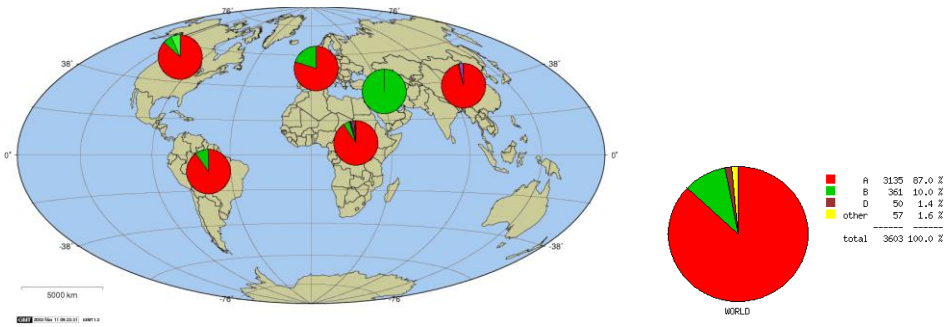


Figure 5: Global distribution of HIV.

(A) Global distribution of HIV-1 subtypes and recombinant forms (Taylor et al., 2008). (B) Distribution of all HIV-2 sequences. Note that this map only includes sequences for which the sampling country is known. Subtype distributions represent the frequency in the HIV Database and not the population. (http://www.hiv.lanl.gov/components/sequence/HIV/geo/geo.comp).

I.6.2. Mechanisms of HIV genomic diversity

Mathematical models that predicted the eradication of virus from a patient within 2 to 3 years failed to adequately take into account two key aspects of HIV biology: viral reservoirs and evolution (Perelson et al., 1997). HIV has several intrinsic mechanisms, e.g., frequent random mutation and genetic recombination, which ensure rapid viral evolution, in response to both host immune activity and antiretroviral therapy.

- ***Frequent random mutation***

The reverse transcriptase of HIV lacks exonucleolytic proofreading activity which confirms that the DNA transcript produced is an accurate copy of the RNA template, and confers a mutation rate of approximately 3.4×10^{-5} mutations per base pair per replication cycle. Since the HIV genome is an estimated 10^5 base pairs in length, and HIV infection is characterized by a very high replication rate, with the production of 1 to 10 billion new virus particles per day in an untreated infected individual, millions of viral variants are produced within any infected person in a single day, and a mutant at each nucleotide position in the viral genome is produced every day (Béthune, 2010; Perelson et al., 1996).

- ***Genetic recombination***

The recombination phenomenon is possible due to the capacity of reverse transcriptase to jump from one RNA molecule to another one, creating a recombinant DNA. Recombinogenic template switching is an even more common occurrence with an estimated three recombination events occurring per genome per replication cycle during HIV-1 DNA synthesis thereby exceeding the introduction of base substitution errors rate per replication. The discovery that most infected cells harbor two or more different proviruses, and the evidence for dual infection, set the stage for recombination to have a central role in generating HIV diversity. HIV-1 recombination can lead to further viral diversity and occurs when one person is coinfected with two separate strains of the virus that are multiplying in the same cell **(Figure 6)**. Recombination is an inherent property of retroviral DNA synthesis, and hence the vast majority of HIV-1 DNAs are biochemical recombinants (An and Telesnitsky, 2002).

Most mutations lead to the production of defective virions, certain of them confers a great power of adaptation permitting virus to escape from its host immune system. Only the viruses best adapted could be selected and multiply in organism. The antiretroviral therapy also confers a selection in viral population during viral replication, which favors the most resistant variants to that drug under treatment. Moreover, it has been suggested

that the genetic diversity of HIV may, at least under certain conditions, be promoted by the host DNA deaminase APOBEC3G that is known to induce G-to-A mutations in HIV genome (Mulder et al., 2008).

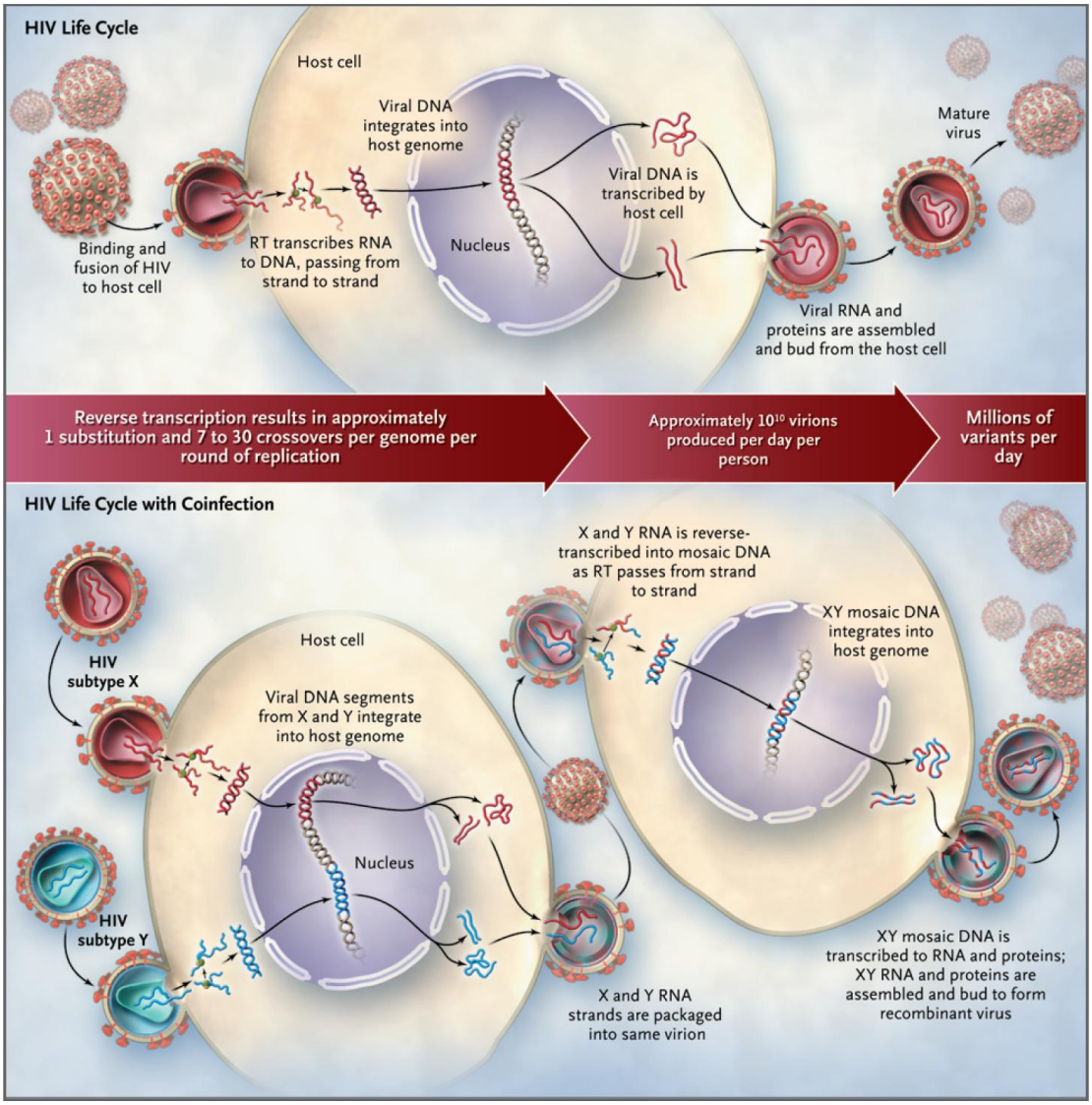


Figure 6: Evolution of diversity in HIV-1 during the viral life cycle and creation of unique recombinant forms in the context of coinfection with two subtypes (Taylor et al., 2008).

I.6.3. Virus evolution and selection of drug resistance

The capacity of evolving rapidly is a key biological property of HIV-1 which results in extensive genetic diversity, and facilitates the emergence of drug resistance. Within a chronically infected patient without effective treatment, continuous high level replication, the error-prone nature of HIV reverse transcriptase, and recombination events lead to the generation of massive numbers of genetically distinct viral variants, referred to as a viral quasispecies. Within the viral quasispecies, wild-type is the most fit and common individual sequence. However, the total number of HIV variants is orders of magnitude higher than the number of wild-type viruses present in a patient. The entire viral quasispecies are subject to selection and evolution. Response to different selection pressures is allowed by its genetic flexibility. For example, a mutant that is less fit may demonstrate an increased fitness under the selection pressure imposed by antiretroviral drug and would be considered as a drug-resistant variant. The presence of antiretroviral drugs, particularly at suboptimal concentrations in certain celluar compartments, exerts a positive selection pressure for the expansion of drug-resistant viral isolates (Kepler and Perelson, 1998). Moreover, recombination events can accelerate the emergence of multidrug-resistant HIV isolates **(Figure 7)**.

The evolutionary potential of HIV was severely underestimated by those proposing highly active antiretroviral therapy (HAART) as a cure for AIDS. This occurred for two reasons. First, though the evolution of drug resistance in single drug therapy had been documented, it was thought that the new triple-drug combination therapy would prove too much difficulty to evolve a solution for escape, because mutations were believed to occur singly and to accumulate in an ordered fashion (Molla et al., 1996). However, this ignored recombination, which allows the virus to accumulate and exchange drug resistant mutations in a nonlinear fashion, leading to rapid evolution of drug-resistance mutants, even between different reservoirs. Second, early studies from patients on HAART concluded that the virus was not undergoing evolution in those with ‘undetectable’ levels of viral load. In fact, the data from these studies clearly indicated that the virus was evolving (Rambaut et al., 2004).

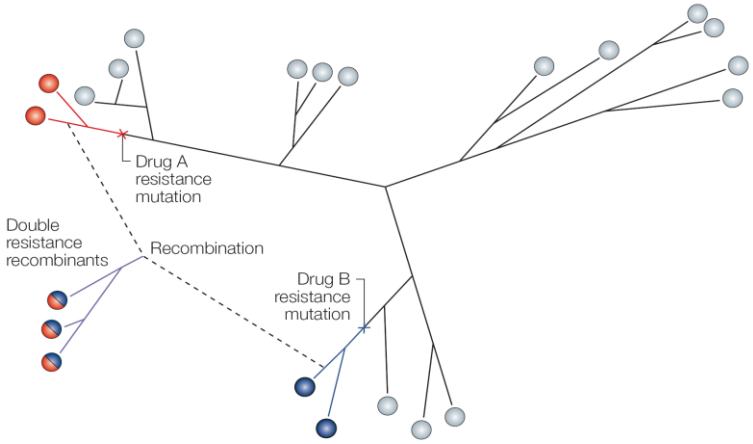


Figure 7: Multiple drug resistance induced by recombination.

Hypothetical example shows how recombination will be an important mechanism to generate drug resistance in HIV. Two different HIV strains that are resistant to drug A (in red) and drug B (in blue) recombine to produce a new strain resistant to both drugs (Rambaut et al., 2004)

II. HIV integrase and integration

II.1. Structure of integrase

Mechanistically and structurally, integrase (IN) belongs to a diverse family of polynucleotidyl transferases (Dyda et al., 1994b), which notably includes RNaseH and the transposases from Tn5 and eukaryotic mobile element Mos1. In each newly formed virion, approximately 50 to 100 copies of the IN enzyme are packaged. IN is a 32 kDa protein that comprises three structurally independent domains: the amino-terminal domain (NTD), the catalytic core domain (CCD) and the carboxy-terminal domain (CTD) **(Figure 8)**.

- ***Amino-terminal domain (NTD)***

The amino-terminal domain encompasses residues 1 to 49. It has a non-conventional HHCC zinc finger motif consisting of 4 α-helices arranged as a three-helix bundle which is common to all retroviral integrases. The HHCC motif formed by His12, His16, Cys40 and Cys43, with the last 43 to 49 amino acids disordered. Binding of one Zn^{2+} to the HHCC motif promotes IN multimerization which is a key process in integration: more specifically, the N-terminal tail and the first half of helix α3 mediate dimer interface through hydrophobic amino acids F1, L2, I5, P29, L31, V32 and hydrophilic Q35 (Eijkelenboom et al., 1999; Eijkelenboom et al., 2000; Lee et al., 1997)

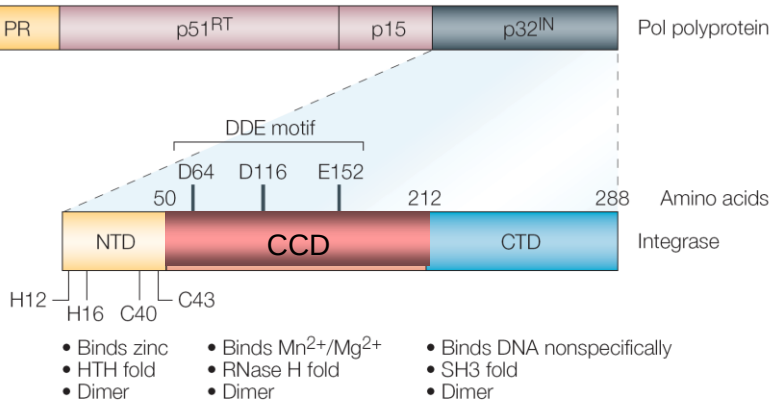


Figure 8: Structural domains of HIV-1 integrase (Pommier et al., 2005).

- ***Catalytic core domain (CCD)***

The catalytic core domain or central domain (CCD) (amino acids 50 to 212) is structurally similar to other retroviral integrases and to RNase H (Dyda et al., 1994a). This family of polynucleotide transferases contains a canonical three-amino acid D, D(35)E motif corresponding to D64, D116 and E152 in HIV IN and involved in DNA substrate recognition. Mutation of any of these residues abolishes IN enzymatic activities and viral replication. These residues D64, D116 and E152 coordinate presumably two divalent metal ions (Mg^{2+} or Mn^{2+}) in complex with the viral and host DNA (Chiu and Davies, 2004). The core domain has a mixed α/β structure, with five β-sheets and six α-helices (Dyda et al., 1994a). The active site residues D64, D116 and E152 are located in different structural elements: β-sheet (β1), coil and helix (α4), respectively.

Another structural feature of the catalytic core domain is the 10 amino acid disordered 'flexible loop' (residues G140 to G149) which is probably stabilized by DNA binding. Conformational changes in flexible loop are required for 3'-processing and strand-transfer reactions. Those two glycines potentially act as hinges for the overall movement of the loop that may serve as a clamp for the binding of the viral DNA ends to the catalytic site of IN. Binding of the viral DNA extremities mainly via the residues Q148, K156 and K159. Q148 has been shown to bind selectively to the penultimate cytosine at the 5' end of the viral DNA (Johnson et al., 2006). Moreover, Q148 is also a key residue for IN catalytic activity and resistance to raltegravir and elvitegravir (Delelis et al., 2009a; Sichtig et al., 2009a).

The CCD also contains other functional domains and residues such as an RNase H-fold (Rice and Baker, 2001); the $K_{186}R_{187}K_{188}$ multimerization motif at the dimmer/dimer interface (Wang et al., 2001); and several important residues involved in the chemical bond and hydrophobic contacts with LEDGF/p75 (Cherepanov et al., 2002; Poeschla, 2008).

- ***Carboxy terminal domain (CTD)***

The C-terminal domain (amino acids 213 to 288) is the least conserved among retroviral integrases but has an overall SH3 fold (Chiu and Davies, 2004). CTD binds DNA non-specifically and helps to stabilize the IN/DNA complex. Mutagenesis data support a role for the CTD in proper folding of the IN protein (Moreau et al., 2003).

Integrase functions in a multimeric form, at least a tetramer. Dimers associated with either end of the viral DNA molecule. Pairs of dimers bring together the two ends of the viral DNA and form a tetramer, stable synaptic complex for concerted integration. Viral DNA is processed within these complexes which go on to capture the target DNA and

integrate the viral DNA ends (Li and Craigie, 2005 ; Li et al., 2006).

- ***Overall structure of IN and model for target DNA binding***

Despite its acute importance for antiretroviral drug discovery and decades of rigorous research, the complete structure of HIV IN, either as a separate protein or in the context of the functional intasome defined as the minimal functional complex involving viral DNA and IN, is lacking, due to poor solubility and interdomain flexibility problems. However, several structures of isolated domains or of two consecutive domains have been reported (Mouscadet et al., 2010a). Recently a crystal structure of full-length IN from the prototype foamy virus (PFV), another lentivirus, in complex with its cognate DNA has recently been reported (**Figure 9**) (Hare et al., 2010b). In this complex, the retroviral intasome consists of an IN tetramer tightly associated with a pair of viral DNA ends. The inner subunits of the tetramer are responsible for all contacts involved in tetramerization and viral DNA binding. The CCDs of the outer subunits seem to provide supporting function. Within the PFV intasome, the CCD-CTD linker adopts an extended conformation for most of its length, tracking parallel to the NTD-CCD linker from the same subunit. Each CTD makes contact with the phosphodiester backbone of both viral DNA molecules, essentially crosslinking the structure. The overall conformation of the assembled intasome is well constrained.

The active sites of the inner IN subunits, engaged with the 3’ termini of the viral DNA, are located deep within the dimmer/dimer interface. Therefore, the only mode of host chromosomal DNA (target DNA) binding that would not require marked rearrangement of the intasome or severe DNA bending is along the cleft between IN dimers **(Figure 10)**. Note that the CTD, juxtaposed to the target DNA in this model, is known to possess sequence non-specific DNA binding activity (Hare et al., 2010b).

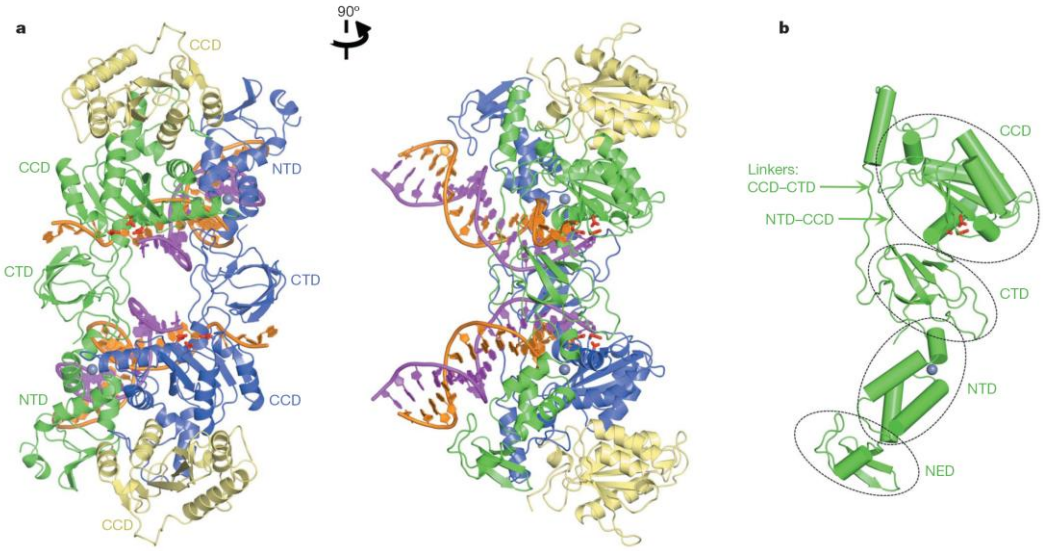


Figure 9: Overall structure of the foamy virus intasome with full-length integrase.

a. Views along (left) and perpendicular (right) to the crystallographic two-fold axis. The inner subunits of the IN tetramer, engaged with viral DNA, are blue and green; outer IN chains are yellow. The reactive and non-transferred DNA strands are magenta and orange, respectively. Side chains of Asp 128, Asp 185 and Glu 221 active-site residues are red sticks; Zn atoms are grey spheres. Locations of the canonical IN domains (NTD, CCD and CTD) are indicated. **b.** Inner (green) IN chain with domains and linkers indicated. The orientation is the same as in the right panel of a (Hare et al., 2010b).

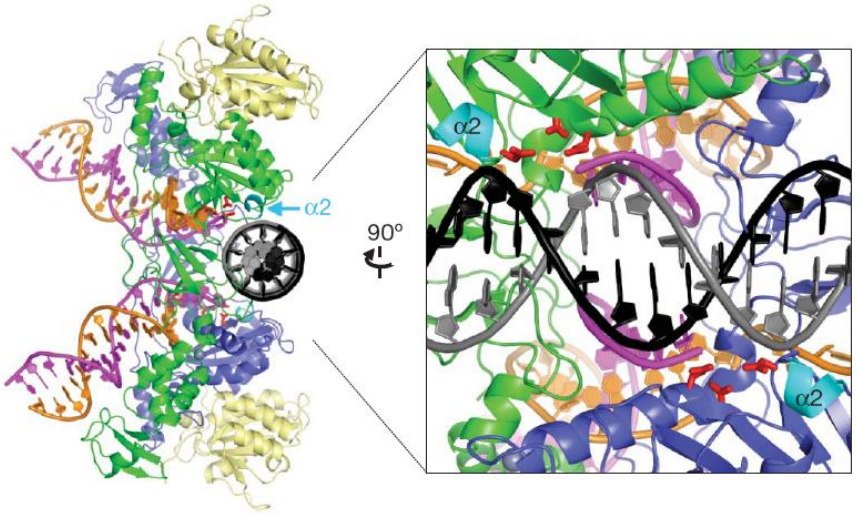


Figure 10: Predicted binding orientation of target DNA to integrase.

The DNA molecule modeled in black and grey shows the most likely orientation for target DNA binding. Cartoon representation of the intasome as in Figure 9 with active-site side chains shown as red sticks and α2 helices, known to contribute to target DNA binding, in cyan (Hare et al., 2010b).

II.2. Functions of integrase

IN has a variety of effects on viral replication. Accordingly, IN mutants are generally classified into two types. Class II mutants can have pleiotropic effects on the viral life cycle, including effects on particle budding, infectivity, and reverse transcription. In contrast, nonpleiotropic (class I) mutations affect only at the integration step in the viral life cycle. In HIV-1, mutations of any of three residues in the catalytic triad (D64, D116, and E152), which is broadly conserved in retroelement INs, induce properties of class I (Engelman, 1999).

II.2.1. Catalytic activities

- ***3'-processing***

3'-processing takes place in the cytoplasm following reverse transcription **(Figure 11, a to b)**. Integration requires correct sequences at both ends of the proviral DNA (att sites) **(Figure 12)**. The shaded A/T sequences, which start approximately ten base pairs from the viral DNA ends, are also conserved across HIV strains. IN catalyzes a endonucleolytic cleavage at the 3'-site of the conserved CA using water as the nucleophile in the presence of Mg^{2+} (the presumably physiological metal) or Mn^{2+}. Generally, two dinucleotides 5' GT 3' are released (one from each viral 3' end), unless the 3'-end has been made longer during reverse transcription (Yoder and Bushman, 2000). This reaction generates recessed reactive CA 3'-hydroxyl ends at both ends of the viral DNA that provide the nucleophile groups required for the second step, strand transfer. At least a dimer of IN is needed for 3'-processing (Faure et al., 2005). Integrase multimers remain bound to the ends of the viral DNA as the pre-integration complexes (PICs) translocate to the nucleus.

- ***Strand transfer***

Following PIC migration into the nucleus, the second reaction **(Figure 11, c to d)** catalysed by integrase is strand transfer (3' end joining) which inserts both viral DNA ends into a host-cell chromosome simultaneously. Genomic integration is random but tends to occurs in transcribed genes. Strand transfer is coordinated in such a way that each of the two 3'-hydroxyl viral DNA ends acting as nucleophiles attacks a DNA phosphodiester bond on each strand of the host DNA acceptor with a five-base-pair stagger across the DNA major groove. Strand transfer leaves a five-base, single-stranded gap at each junction between the integrated viral DNA and the host acceptor DNA, and a two-base overhang at

the 5’ ends of the viral DNA. Gap filling and release of the unpaired 5’ ends of the viral DNA are carried out in coordination with cellular DNA repair enzymes, although RT has been proposed to be involved in this reaction (Marchand et al., 2006).

- ***Disintegration***

Disintegration may be considered to be the reverse of the strand transfer reaction. Unlike the 3’-processing and strand transfer reactions which require the full-length protein, disintegration is the only reaction catalyzed by the isolated CCD (Chow et al., 1992). Currently no experimental evidence is available to suggest that this reaction occurs in vivo. However, disintegration activity was widely used for testing the competitive mechanisms of certain inhibitors.

- ***Internal cleavage activity***

A specific and internal cleavage catalysed by the full-length HIV IN has been observed in vitro More recently (Delelis et al., 2007). Oligonucleotides mimicking the palindromic sequence found at the LTR-LTR junction of the 2-LTR circles (found in infected cells) were efficiently cleaved at internal positions by HIV-1 IN, with cleavage kinetics comparable to the 3’-processing reaction. This reaction requires a symmetrical organisation of the DNA substrate with a strong cleavage at the CA dinucleotide as well as a tetrameric organisation of the protein, whereas the 3’-processing reaction is efficiently catalysed by a dimer. From a structural point of view, this reaction is related to the endonucleolytic reaction of a restriction enzyme. This activity seems to be generalized to other retroviral IN as reported earlier for PFV-1 IN.

II.2.2. Noncatalytic activities

Non catalytic functions of IN are also crucial for the viral replication cycle. For instance, a functional interaction between IN and RT has been observed, suggesting that IN is involved, at least indirectly, in controlling the synthesis of viral DNA (Zhu et al., 2004).

Another non catalytic activity of IN is related to the PIC trafficks towards the nucleus, suggested by the identification of components of the microtubule network associated with IN (Ao et al., 2007; Depienne et al., 2001).

IN could be indirectly involved in the regulation of transcription of integrated provirus. After the integration, IN could be tightly bound to the integrated DNA and then, the degradation of IN by the proteasome-ubiquitin pathway was proposed to regulate the transcription of viral genes (Mousnier et al., 2007).

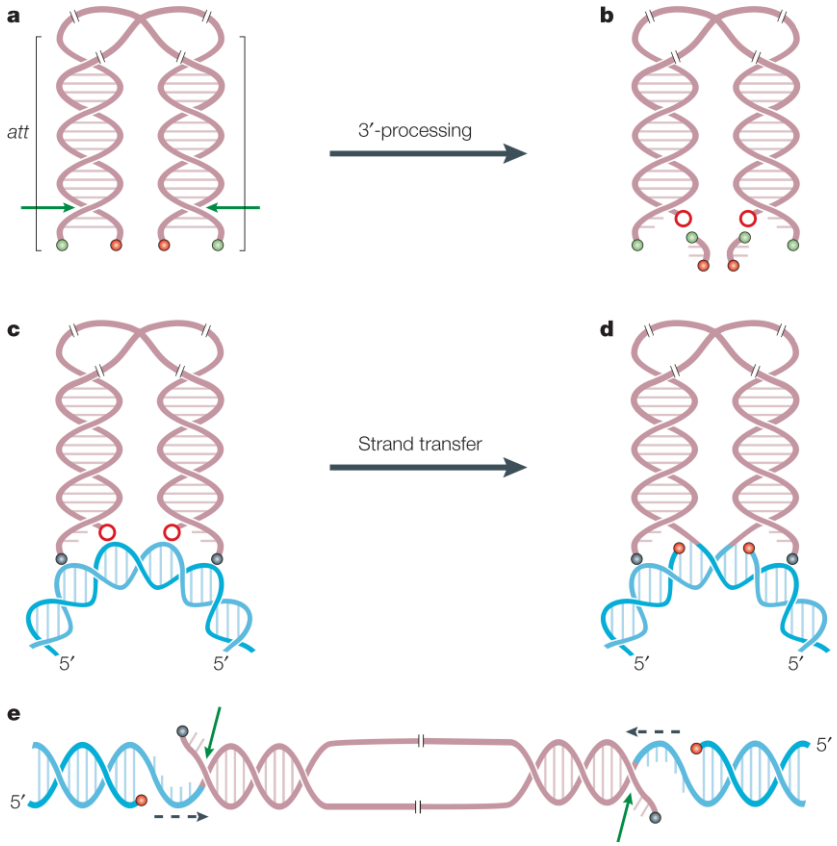


Figure 11: 3'- processing and strand transfer occur at viral DNA recombination (att) sites.

(a) Green arrow in a shows an endonucleolytic cleavage, 3′-hydroxyl ends; 5' phosphate ends are shown as red and green dots, respectively. (b) Red circles show reactive CA 3' hydroxyl ends of the viral DNA. (c) Acceptor DNA in blue. (e) Arrows show unpaired 5' ends of the viral DNA (Pommier et al., 2005).

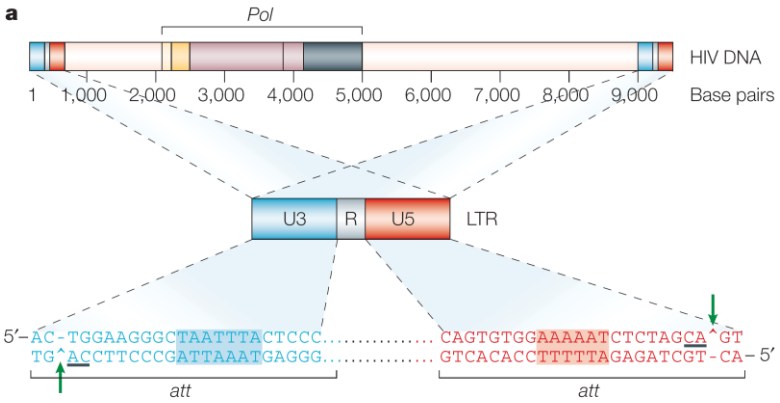


Figure 12: Long terminal repeat sequences and 3'-processing (Pommier et al., 2005).

II.2.3. Kinetics of integrase

IN is characterized by overall slow cleavage efficiency. The formation of stable complexes with both DNA substrate and DNA product limits multiple turnover of IN. The polynucleotidyl-tranferases such as transposases seem to share a peculiar enzymatic property: they have evolved to catalyse multi-sequential steps in a single active site. This multi-sequential reaction requires a strong binding of the enzyme to DNA product after each catalytic step to optimize the whole process, while consequently reduces the overall enzymatic efficacy in term of turnover. In vivo, the tight binding of IN to the processed viral DNA likely allows the complex to remain associated after the 3'-processing reaction long enough for subsequent integration. This weak catalytic activity is not detrimental in the cellular context, because a single event of integration or transposition is sufficient for the overall function (Delelis et al., 2008).

II.3. Preintegration complex

Although purified recombinant IN is sufficient to carry out 3'-processing and strand transfer in vitro, a variety of viral and cellular proteins have been considered as important partners in establishing the integrated provirus in the infected cell. In the preintegration complex, IN tetramer is associated with viral and cellular co-factors. The numerous proteins that have been suspected to influence integration can be grouped into three main categories: those that (i) interact with IN directly; (ii) bind to the viral DNA to influence the process; or (iii) participate in the final gap repair step (Poeschla, 2008).

LEDGF/p75 is involved in nuclear import and chromosome tethering of the PIC, increasing the efficiency of concerted integration. Moreover, LEDGF/p75 is able to stabilize the tetrameric form of IN, protect IN from degradation (a pair of IN tetramers and two subunits of LEDGF/p75 comprise a symmetrical complex), and increase IN solubility by binding to IN. It is assumed that LDEGF/p75 strongly influences the genome-wide pattern of HIV integration **(Figure 13)** (Poeschla, 2008). In the absence of p75, promoter regions and CpG islands, the latter genomic feature is enriched in the unusual CpG dinucleotide and is associated with transcription start sites and regulatory regions, are relatively favored by HIV, suggesting that removing p75 uncovers a gamma-retroviral-like targeting proclivity (Marshall et al., 2007).

LEDGF/p75 is a member of the hepatoma-derived growth factor family. The amino-terminal domain ensemble participating in chromatin binding is primarily the

PWWP domain and the AT hook domain, with CR1-CR3 being relatively charged regions that influence chromatin binding to a lesser extent **(Figure 14)**. The PWWP domain functions as a protein-protein interaction domain and/or DNA-binding domain. The protein is predominantly localized in the nucleus, where it is intimately associated with the chromosomes. Nuclear localization of LEDGF/p75 is the outcome of two mechanisms, NLS-mediated nuclear import and chromatin trapping (Maertens et al., 2004a). Chromatin linkage of IN is mediated by specific domains located at opposite ends of p75. A C-terminal IN binding domain (IBD) binds to IN. The IBD, which is highly conserved in vertebrates, forms a right-handed α helix bundle that resembles so-called HEAT domains from other proteins. A less-structured inter-helical loop of the IBD extends to bind in a pocket at the IN CCD dimer interface. Ile365, Asp366 and Phe406 were identified as amino acids in LEDGF/p75 that are important for the interaction with HIV-1 IN (Maele et al., 2006).

The gene encoding LEDGF/p75 also encodes a smaller splice variant, p52, which shares a region of 325 residues with LEDGF/p75 at the N terminus but contains eight additional amino acids. In contrast to LEDGF/p75, p52 has a stronger and more general transcriptional co-activator activity but fails to interact with HIV-1 IN in vitro or in living cells (Maertensm et al., 2003).

The **barrier-to autointegration factor (BAF)** is a host cellular protein probably involved in chromatin organization, which prevents autointegration and stimulates chromosomal integration (Lin and Engelman, 2003). BAF bridges and compacts viral cDNA inside the PIC. Although BAF is directly involved in stimulating the intermolecular integration in the in vitro PIC assay, the in vivo role would be indirect (Maele et al., 2006).

Emerin is an integral inner-nuclear-envelope protein. It bridges the interface between the inner nuclear envelope and chromatin, and may be necessary for chromatin engagement by viral cDNA before integration. BAF, a binding partner of emerin, was required for the association of viral cDNA with emerin, and for the ability of emerin to support virus infection (Jacque and Stevenson, 2006). Nevertheless, further study showed that emerin is not a universally important regulator of HIV-1 infectivity (Shun et al., 2007).

The **high mobility group protein A1 (HMGA1)**, a host DNA binding protein involved in the chromosomal architecture, is another component of the PIC that facilitates IN binding by unwinding the LTR termini, and stimulates concerted integration by bridging and compacting the viral cDNA (Hindmarsh et al., 1999). Conversely, cells knocked-out for the HMGA protein exhibited wild-type levels of integration, indicating that either HMGA is not important in vivo or it is redundant with other factors (Beitzel and Bushman,

2003).

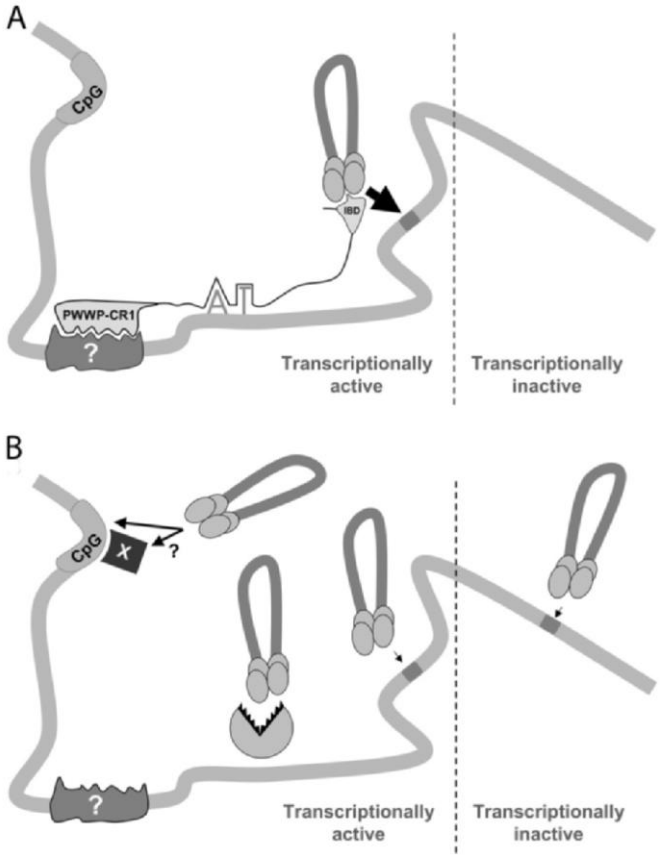


Figure 13: Schematic models for LEDGF/p75 function in HIV integration.
(A)The chromatin fiber is depicted without nucleosome or solenoid detail. p75 engages chromatin via the PWWP-CR1 elements and AT hook domains primarily. CR2 and CR3 (not shown) contribute to chromatin interaction. Unknown p75 ligand(s) may include elements of the general transcription machinery. The relative contribution of direct DNA binding versus protein binding is unknown. In the presence of p75, integration is favored into active transcription units (large arrow). (B) In the absence of p75, integration is reduced (arrows are smaller) and active transcription units are not favored. The IBD interaction may also protect IN from degradation. The speculative black box (X) indicates this could be mediated by a tethering factor or factors analogous to p75. More complicated scenarios, such as interaction with p75 prior to chromatin engagement, are not depicted, nor are p75-independent palindrome-like local sequence preferences illustrated (Poeschla, 2008).

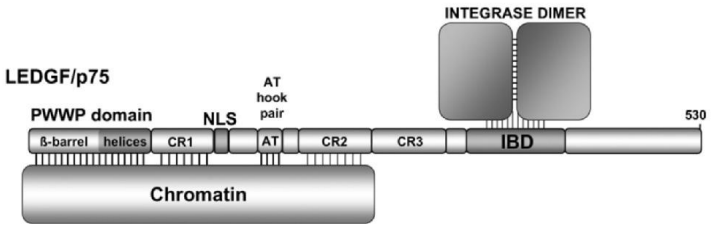


Figure 14: The modular basis for IN-to-chromatin tethering: molecular domain structure of LEDGF/p75 (Poeschla, 2008)**.**

The **integrase interactor 1 protein (INI1/hSNF5)**, a subunit of the SWI/SNF chromatin remodeling complex, was the first cofactor identified to interact with IN (Kalpana et al., 1994). INI1/hSNF5 also interferes with early steps of HIV-1 replication (Maroun et al., 2006; Masson et al., 2007). Although INI1 originally was found to stimulate IN activity in vitro, no strong evidence has been provided for a possible role in the HIV-1 integration in vivo. INI1 probably interacts with IN within the context of Gag-Pol (Maele et al., 2006).

Ku interacts with retroviral PICs and is thought to interfere with the retroviral replication cycle, particularly the formation of 2- LTR viral DNA circles, viral DNA integration, and transcription (Li et al., 2001). Ku may be involved in targeting retroviral elements to chromatin domains prone to gene silencing, though not strictly required for retroviral integration (Masson et al., 2007).

Several other proteins have been reported as potential cellular co-factors of retroviral integration and could be part of the PIC. The polycomb group embryonic ectoderm development (EED) protein, HRP2, the heat-shock protein 60 (HSP60) and the p300 acetyltransferase all interact with IN (Cereseto et al., 2005b; Parissi et al., 2001; Violot et al., 2003). The p300 acetyltransferase, known to acetylate histones and regulate chromatin conformation, has been found to acetylate three specific lysines in the C terminus of IN and to modify its activity (Cereseto et al., 2005a).

Some viral proteins are also part of the PIC. **MA** interacts with IN and BAF, and is implicated in the nuclear import of the PIC. The **Vpr** exhibits karyophilic properties, induces apoptosis after cell-cycle arrest and reduces viral mutations through an interaction with the uracil DNA glycosylase. **NC** is a nucleic acid chaperone and a viral co-factor of RT that stimulates integration. And **RT** could regulate the activity of IN by direct interactions (Carteau et al., 1999; Gao et al., 2003; Mansharamani et al., 2003; Miller et al., 1997a).

III. Anti HIV therapy

III.1. Overview of current anti-HIV therapy

AIDS is one of the greatest challenges to humankind. AIDS and HIV infection represent global health hazards, complex scientific puzzles and obvious targets for drug discovery and vaccination. It took years to identify the pathogenic virus, HIV; develop sensitive tests to detect infected people during the latency period and to introduce the first rationally designed effective therapy, zidovudine (AZT). The life expectancy for AIDS patients was less than 1 year before AZT was introduced in 1987 (Mitsuya et al., 1985). The prognosis of AIDS patients who have full access to current therapies has completely changed since the first cases of AIDS were reported. This dramatic change is due to the development of effective therapies, to early detection of HIV-positive individuals, and to a sustained effort to analyze and understand viral-resistance mechanisms, which can be overcome by rational drug development and combination therapy.

The primary goal of antiretroviral therapy is to achieve a durable supression of viral replication below the level of detection for currently approved diagnostic tests (i.e. <50 RNA copies/ml) together with immune reconstitution measured by the increase of CD4+ cells while minimizing adverse effects (Estéa and Cihlarb, 2010). The maximal and durable suppression of plasma viremia can minimize the selection of drug resistance mutations, preserve the CD4+ T cell count and confer overall clinical benefits to patients.

Six major factors have led to effective anti HIV therapies: first, early recognition of the severe population health problem posed by AIDS; second, adequate government prioritization and funding for basic research; third, elucidation of the genetics and life cycle of HIV; fourth, identification of viral-specific drug targets and the development of screening assays for drug discovery; fifth, the development of clinical tests for measuring viral load and therefore evaluating therapeutic efficacy; and sixth, drug combinations to overcome innate and acquired drug resistance resulting from the high-mutator phenotype of retroviruses (Pommier et al., 2005).

Current therapy of AIDS involves the use of a combination of at least three antiviral drugs (commonly referred as highly active antiretroviral therapy, HAART) to inhibit HIV-1 and 2 (in short HIV) replication. Targeting distinct areas of the HIV lifecycle, HAART consists of three or more HIV drugs, most commonly two nucleoside reverse transcriptase inhibitors in combination with a non-nucleoside reverse transcriptase inhibitor, protease inhibitor or, most recently, IN inhibitor. Compounded with the routine

use of HIV RNA viral load and CD4+ T cell counts as surrogate markers of drug efficacy and disease progression (Mellors et al., 1996), HAART has brought a dramatic effect on the morbidity and mortality of HIV-infected individuals. Combination therapy also reduces the emergence of drug resistant viruses, as the multiple mutations that are needed to overcome drug resistance decrease viral fitness (Pommier et al., 2005).

III.2. Antiretroviral drugs

III.2.1. Overview of present antiretroviral drugs

25 years after the discovery of the antiviral effect of AZT (Mitsuya et al., 1985), there are 25 approved single antiretroviral drugs in 6 mechanistic classes that can be used to design multiple combination regimens **(Table 2)**. These six classes include the nucleoside/nucleotide reverse transcriptase inhibitors (NRTIs), non-nucleoside reverse transcriptase inhibitors (NNRTIs), protease inhibitors (PIs), entry/fusion inhibitors (FIs), and integrase inhibitors (INI) **(Figure 15)**. Some of these drugs are minimally used, because of their side effect profiles, high pill burden and/or inconvenient administration schedules, or difficult to manage drug-drug interactions.

However, due to the emergence of drug-resistant virus isolates to antiretroviral medications, the need for lifelong antiretroviral therapy (ART) for HIV-infected individuals, and the goal of minimizing ART-related adverse effects and toxicity, the need for new antiretrovirals in existing classes and development of therapies aimed at novel targets are paramount (Dickinson et al., 2009).

Encouragingly, a number of agents have recently gained approval in existing and new drug classes while others are still at the developmental stage. For instance, entry inhibitor maraviroc is a CCR5 co-receptor antagonist that prevents viral interaction with the CCR5 coreceptor. It is approved for use in ART-experienced patients with CCR5-tropic HIV. Vicriviroc is another CCR5 co-receptor antagonist that is in late clinical trials. Fusion of the viral membrane with the target cell membrane can be blocked by the peptidic inhibitor enfuvirtide, which prevents a conformational change in the viral Env protein needed to bring the two membranes into close proximity. Raltegravir is the first US FDA approved HIV IN inhibitor. Elvitegravir is another HIV-1 integrase inhibitor in clinical development. Other new antiretroviral agents in clinical development include PRO140, a monoclonal antibody against CCR5, and bevirimat, a maturation inhibitor that prevents late-stage gag polyprotein processing. A number of other drug targets, such as CCR5 co-receptor agonists,

CXCR4 co-receptor antagonists, novel fusion inhibitors, and alternative antiretroviral strategies, such as immune stimulation and gene therapy, are under investigation (Dau and Holodniy, 2009).

FDA Approval	Brand Name	Generic Name	Manufacturer
Nucleosides Reverse Transcriptase Inhibitors (NRTIs)			
1987	Retrovir	zidovudine, azidothymidine, AZT, ZDV	GlaxoSmithKline
1991	Videx	didanosine, dideoxyinosine, ddI	Bristol-Myers Squibb
1992	Hivid	zalcitabine, dideoxycytidine, ddC	Roche Pharmaceuticals
1994	Zerit	stavudine, d4T	Bristol-Myers Squibb
1995	Epivir	lamivudine, 3TC	GlaxoSmithKline
1997	Combivir	lamivudine + zidovudine	GlaxoSmithKline
1998	Ziagen	abacavir sulfate, ABC	GlaxoSmithKline
2000	Trizivir	abacavir+ lamivudine+ zidovudine	GlaxoSmithKline
2000	Videx EC	enteric coated didanosine, ddI EC	Bristol-Myers Squibb
2001	Viread	tenofovir disoproxil fumarate, TDF	Gilead Sciences
2003	Emtriva	emtricitabine, FTC	Gilead Sciences
2004	Epzicom	abacavir+ lamivudine	GlaxoSmithKline
2004	Truvada	emtricitabine + tenofovir disoproxil fumarate	Gilead Sciences
Non-Nucleosides Reverse Transcriptase Inhibitors (NNRTIs)			
1996	Viramune	nevirapine, NVP	Boehringer Ingelheim
1997	Rescriptor	delavirdine, DLV	Pfizer
1998	Sustiva	efavirenz, EFV	Bristol-Myers Squibb
2008	Intelence	etravirine	Tibotec Therapeutics
Protease Inhibitors (PIs)			
1995	Invirase	saquinavir mesylate, SQV	Roche Pharmaceuticals
1996	Norvir	ritonavir, RTV	Abbott Laboratories
1996	Crixivan	indinavir, IDV	Merck
1997	Viracept	nelfinavir mesylate, NFV	Pfizer
1997	Fortovase	saquinavir (no longer marketed)	Roche Pharmaceuticals
1999	Agenerase	amprenavir, APV	GlaxoSmithKline
2000	Kaletra	lopinavir+ ritonavir, LPV/RTV	Abbott Laboratories
2003	Reyataz	atazanavir sulfate, ATV	Bristol-Myers Squibb
2003	Lexiva	fosamprenavir calcium, FOS-APV	GlaxoSmithKline
2005	Aptivus	tripranavir, TPV	Boehringer Ingelheim
2006	Prezista	darunavir	Tibotec Therapeutics
Fusion Inhibitors			
2003	Fuzeon	Enfuvirtide, T-20	Roche Pharmaceuticals & Trimeris
Entry Inhibitors			
2007	Selzentry	maraviroc	Pfizer
IN Inhibitors			
2007	Isentress	raltegravir	Merck

Table 2: Approved anti-AIDS therapies (Marchand et al., 2009)

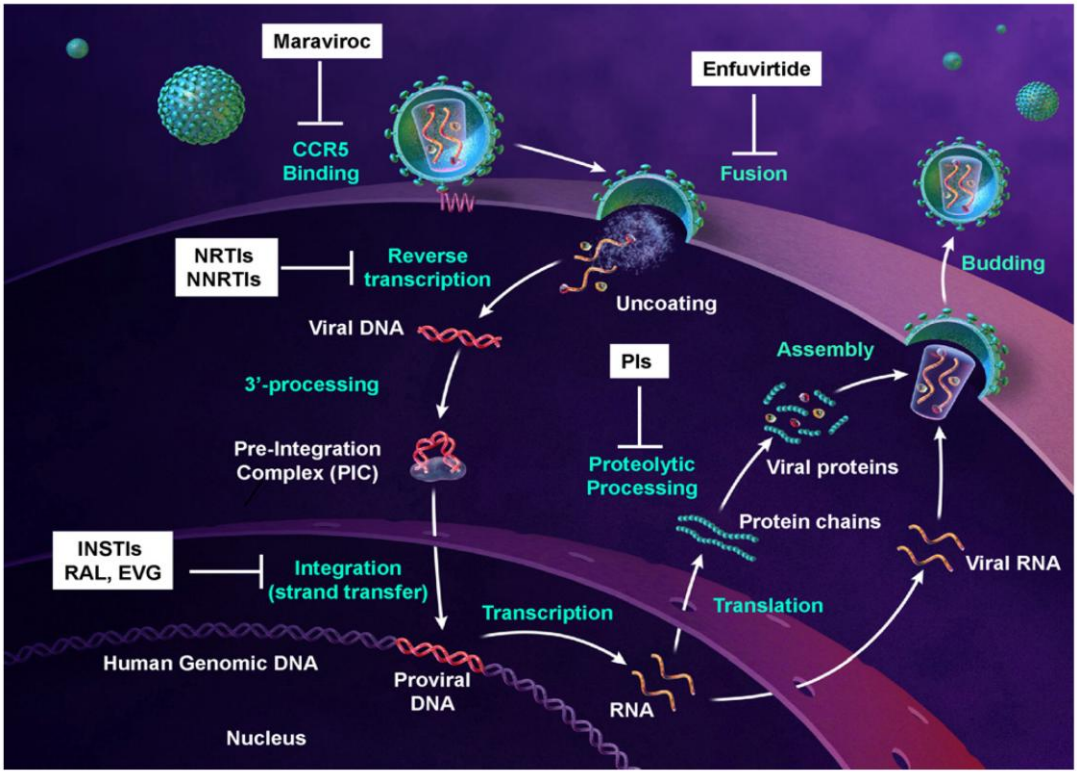


Figure 15: Overview of present treatment options (McColl and Chen, 2010).

III.2.2. Nucleoside and nucleotide reverse transcriptase inhibitors (NRTIs)

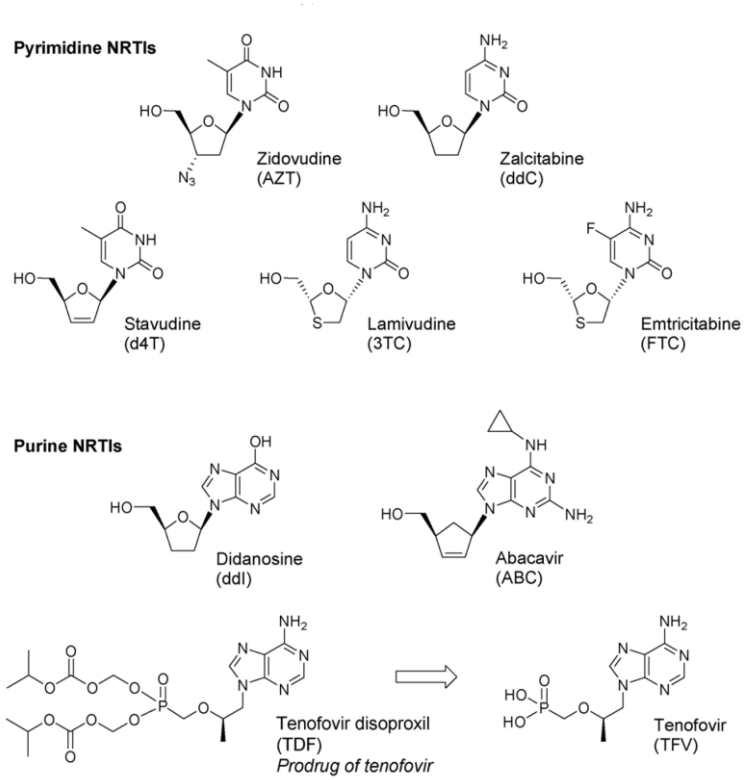
In 1985, two years after the identification of HIV and one year after the initial evidence about its etiological link to AIDS was reported, 3'-azidothymidine (AZT, zidovudine) was identified as the first NRTI acting through its triphosphate metabolite. The discovery of the anti-HIV activity of AZT provide the first proof of concept that the replication of HIV could be controlled by chemotherapy and thereby establishing the foundation of antiretroviral drug discovery research. Currently, a total of thirteen NRTI drug products are now approved by the FDA for the treatment of HIV: seven nucleosides and one nucleotide individual NRTIs, four fixed-dose combinations of two or three NRTIs, and one complete fixed-dose regimen containing two NRTIs and one non-nucleoside RT inhibitor **(Table 2)**.

- ***Mechanism of inhibition by NRTIs and the metabolism***

NRTIs are the backbone of current combination antiretroviral. NRTIs are analogs of endogenous 2'-deoxy-nucleosides and nucleotides. There is structural diversity in the eight

NRTIs that have been approved by regulatory agencies for HIV treatment **(Figure 16A)** and they are metabolized to analogs of all four natural dNTPs used during DNA synthesis. All currently approved NRTIs lack a 3'-hydroxyl and are obligate chain-terminators of DNA elongation. In contrast to the diverse chemical modifications of the ribose ring, very few base modifications are present in currently approved NRTIs.

A



B

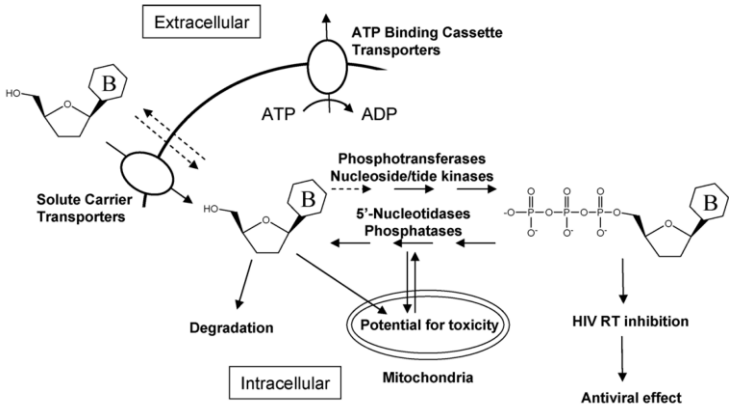


Figure 16: Structures of approved NRTIs (A) and general scheme for the reversible activation of NRTIs (B) (Cihlar and Ray, 2010).

NRTIs are inactive in their parent forms and require successive phosphorylation steps by host cell kinases and phosphotransferases to form deoxynucleoside triphosphate (dNTP) analogs capable of viral inhibition. NRTIs can thus be considered as pro-drugs. In their respective triphosphate forms, NRTIs compete with their corresponding endogenous dNTPs for incorporation by HIV RT. Once incorporated, they serve as chain-terminators of viral reverse transcripts, thus acting early in the viral replication cycle by inhibiting a critical step of proviral DNA synthesis prior to integration into the host cell genome. NRTIs are active against HIV-1 RT, and also against the RTs of other retroviruses, e.g., HIV-2, SIV, murine leukemia virus, visna virus, etc. Because NRTIs are analogues of normal nucleotides, they can also be incorporated into the DNA of the host, thereby resulting in toxicity.

The absolute dependence on host cell enzymatic processes for activation is a unique element in the pharmacology of NRTIs. Activation of NRTIs is first dependent on cellular entry by passive diffusion or carrier-mediated transport **(Figure 16B)**. Transport plays an important role in the disposition of NRTIs due to their hydrophilicity and somewhat limited membrane permeability. Once inside of cells, the distribution of phosphorylated metabolites of NRTIs is dictated by the balance between phosphorylating and dephosphorylating enzymes. The first, and often rate limiting, phosphorylation step for nucleoside analogs are most commonly catalyzed by the four deoxynucleoside kinases involved in salvage of endogenous nucleosides, cytosolic deoxycytidine kinase (dCK) and thymidine kinase 1 (TK1), and mitochondrial deoxyguanosine kinase and thymidine kinase 2. Addition of the second phosphate group to nucleoside monophosphate analogs is completed by the NMPkinases thymidylate kinase, uridylate-cytidylate kinase, adenylate kinases 1-5 and the guanylate kinase (Van Rompay et al., 2000). A variety of enzymes are able to catalyze the final phosphorylation step for NRTIs, including NDP kinase, phosphoglycerate kinase, pyruvate kinase and creatine kinase, resulting in formation of respective antivirally active triphosphate analogs. In addition to dephosphorylation, some NRTIs are subject to catabolism.

Following phosphorylation to their respective nucleoside triphosphate forms, NRTIs compete with endogenous dNTPs for incorporation by HIV RT. Despite the structural diversity present in NRTIs used clinically, their active triphosphates are able to effectively mimic the structural contacts of natural dNTPs in the HIV RT active site, allowing for efficient incorporation. Once incorporated, the lack of a 3’-hydroxyl in NRTIs results in obligate chain-termination of viral reverse transcripts.

The success of NRTIs in HIV therapy is due, at least in part, to the unique pharmacology.

One advantageous attribute of at least some NRTIs is the intracellular persistence. The intracellular retention of the active triphosphoraleted NRTI metabolites allows for more constant viral inhibition.

- ***Resistance to NRTIs***

While the unique pharmacology of NRTIs has helped them become the cornerstone of successful HAART, the effectiveness of NRTIs can be limited by drug-drug interactions, adverse events and emergence of drug resistance.

Two general modes of resistance to NRTIs have been elucidated. The first mode of resistance affects the binding and rate of incorporation of the incoming nucleotide analog and primarily involves residues of RT in direct contact with the incoming NRTI triphosphate (including K65R, L74V, Y115F, M184V/I and Q151M). The other set of mutations are in proximity to the triphosphate binding site and cause an increased rate of excision of an incorporated chain-terminating NRTI (including M41L, D67N, L210W, T215Y/F and K219Q/E/N/R) (Soriano and de Mendoza, 2002). Although the mutations enhancing the excision of NRTIs from terminated DNA were initially dubbed thymidine associated mutations (TAMs) due to their selection by combinations including AZT or d4T, they are truly multi-drug resistance mutations and their various combinations result in resistance to all approved NRTIs.

III.2.3. Non-nucleoside reverse transcriptase inhibitors (NNRTIs)

The first generation NNRTIs were approved by the FDA in 1996 (nevirapine), 1997 (delavirdine), and 1998 (efavirenz) **(Figure 17)**. Nevirapine and efavirenz are cornerstones of first line HAART, whereas delavirdine is barely used nowadays. They are characterized by a low genetic barrier to the development of resistance. (The genetic barrier to resistance development is usually defined by the number of resistance-associated mutations necessary to confer virological failure.) They need to be combined with at least two other fully active non-NNRTI antiretroviral drugs, and resistance to one of them precludes subsequent use of the other first generation NNRTIs. Etravirine was the first of next generation NNRTI to be approved by the FDA, in combination with other antiretroviral agents for the treatment of HIV-1 infection in treatment-experienced adult patients, who have evidence of viral replication and HIV-1 strains resistant to a NNRTI and other antiretroviral agents. Unlike the NRTIs, NNRTIs are active as such, and do not need intracellular metabolisation. NNRTIs are highly specific inhibitors of HIV and not active against other retroviruses.

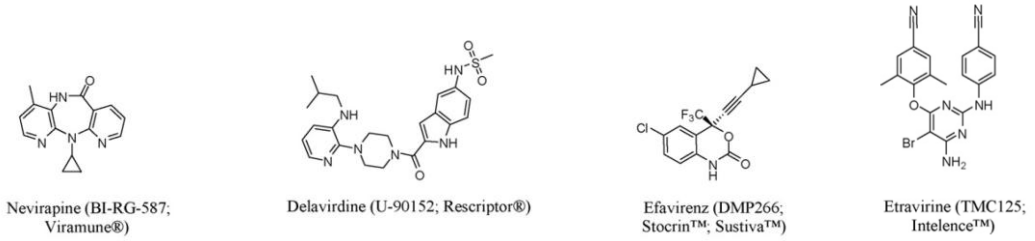


Figure 17: Chemical structures of first generation NNRTIs (nevirapine, delavirdine, and efavirenz) and next generation NNRTIs (etravirine) (Béthune, 2010)

- ***Mechanism of inhibition by NNRTIs***

Despite their chemical heterogeneity, NNRTIs inhibit RT by binding to the enzyme in a hydrophobic pocket. This pocket is located in the palm domain of the p66 subunit of the heterodimeric RT, approximately 10 Å from the catalytic site of the enzyme. In the crystal structures of unliganded RT, the NNRTI binding pocket is not observed, but it is created when an inhibitor binds to the enzyme (Kohlstaedt et al., 1992). Of note, the overall shape of the pocket does not vary significantly, even if the ligands are chemically very different. These conformation changes, induced by the binding of the NNRTIs to the RT are thought to be at the basis of their inhibitory action against the enzyme (Sluis-Cremer et al., 2004).

The conformation of the catalytic site of RT is impacted by NNRTI binding: the position of the YMDD motif is affected, especially of the D185 and D186 residues that coordinate the Mg^{2+} ions in the active site, as well as that of the structural elements that constitute the 'primer grip' of RT right hand model. Another major change observed upon NNRTI binding is the position of the thumb region in the p66 subunit. This change could result in either a decreased thumb mobility (the 'arthritic thumb' model), thus slowing or even preventing the translocation of the primer/template, and the elongation of the nascent DNA strand (Kohlstaedt et al., 1992), or the direction of movement of the thumb and other RT domains involved in primer/template translocation (Temiz and Bahar, 2002).

NNRTIs are non-competitive RT inhibitors with regard to the substrate, by binding to an allosteric site, different from the catalytic site without changing the apparent binding affinity of the enzyme for the substrate, and uncompetitive relative to the primer/template, the inhibitors bind only to the enzyme-substrate complex. In general, NNRTI binding does not inhibit and even seems to improve the binding of dNTP and primer/template to the RT. This could be a consequence of the repositioning of the thumb upon NNRTI binding. NNRTIs do not directly inhibit the phosphodiester bond formation but rather prevent the

correct positioning of the dNTP relative to the primer/template terminus (Xia et al., 2007). The increase in binding affinity of dNTPs, and hence NRTI-triphosphates, for RT in the presence of bound NNRTIs could account for the strong synergy observed between the two classes of RT inhibitors, and explain the potent efficacy of treatment regimens combining two NRTIs and an NNRTI, typically used as first line therapy.

- ***Resistance to NNRTIs***

Over 40 amino acid substitutions have been identified to be associated with NNRTI resistance, in vitro and in vivo. Originally, NNRTI resistance mutations were all located in the NNRTI binding pocket. Those mutations are mainly observed in domains spanning amino acids 98 to 108, 178 to 190 and 225 to 238 of the p66 subunit (Soriano and de Mendoza, 2002). The most prevalent substitutions observed in viruses from patients who have been on a failing NNRTI containing regimen are K103N and Y181C. Nevirapine and delavirdine rely heavily on ring stacking interactions with Y181 and Y188 for their binding to the enzyme. It is therefore not surprising that substitutions by non-aromatic amino acids at those positions drastically reduce the binding affinity of these inhibitors for RT. This is not the case for efavirenz, which does not show extensive contact regions with Y181. Another mutation, Y318F, outside the originally identified domains has also been reported. Additionally, mutations at the E138 position cause resistance to NNRTIs, even when they are present in the p51 subunit only. More recently, other mutations, outside the NNRTI binding pocket have been reported in NRTI-treated patients: they are located in the connection domain of the p66 subunit, between the thumb and the RNaseH domains (Waters et al., 2009).

NNRTI resistance mutations impact the binding of the molecules in the NNRTI binding pocket, but some of those mutations have also been described to influence the other functionalities of RT other than DNA polymerization, i.e. the V106A and P236L mutations cause a slowing in the RNA 5’ end and DNA 3’ end directed RNaseH cleavage activities (Archer et al., 2000).

III.2.4. Protease inhibitors (PIs)

PIs have been designed to bind the viral protease with high affinity but tend to occupy more space than the natural substrates. Currently, there are nine PIs approved for clinical use: saquinavir, ritonavir, indinavir, nelfinavir, amprenavir, lopinavir, atazanavir, tipranavir and darunavir **(Figure 18)**. Most of them are prescribed with a concomitant low dose of ritonavir as boosting agent. All of them, with the exception of tipranavir, are

competitive peptidomimetic inhibitors, mimicking the natural substrate of the viral protease. The peptidomimetic inhibitors contain a hydroxyethylene core, which prohibits cleavage of the protease inhibitor by the HIV-1 protease. Instead of a peptidomimetic hydroxyethylene core, tipranavir contains a dihydropyrone ring as a central scaffold (Wensing et al., 2009).

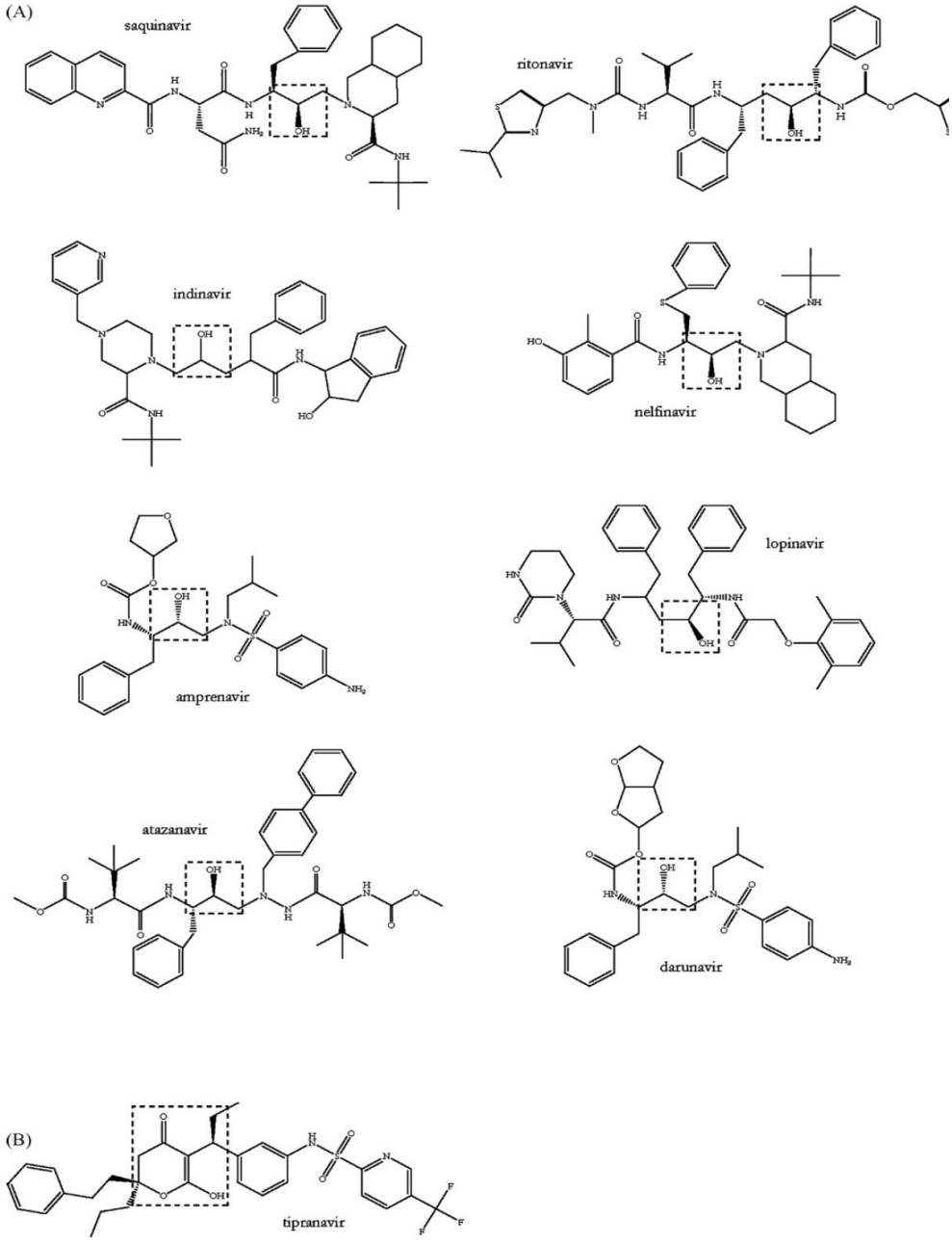


Figure 18: Chemical structures of the nine HIV-1 protease inhibitors approved for clinical use.

(A) Peptidomimetic protease inhibitors, characterized by a hydroxyethylene core, indicated with dashed-line boxes. (B) Non-peptidomimetic protease inhibitor characterized by a dihydropyrone rine, as indicated with a dashed-line box (Wensing et al., 2009).

- ***Resistance to PIs***

The development of protease inhibitor resistance is a stepwise process in which a substitution in the substrate-binding cleft of the viral protease is usually observed firstly. These resistance mutations in the viral protease result in an overall enlargement of the catalytic site of the enzyme. This leads to decreased binding to the inhibitor (causing a decrease in drug sensitivity) and, in parallel, to some decrease in binding to the natural substrate and thus to decreased viral replication (Mammano et al., 2000; Nijhuis et al., 1999). These mutations, which are selected first and individually reduce the susceptibility to a protease inhibitor, are called primary or 'major' resistance mutations (Johnson et al., 2008b). Secondary or 'minor' mutations generally emerge later and by themselves do not have a substantial effect on the resistance phenotype but improve replication of viruses containing major mutations. Some minor mutations are present as common polymorphic changes in HIV-1 non subtype B clades.

Protease inhibitor associated mutations have not only been observed in the viral protease itself, but also in the substrate of the viral protease, the Gag protein. They are commonly found in, or closely to, the protease cleavage sites and are thought to adapt the virus to the altered substrate-binding cleft of the mutant drug resistant viral protease.

Occasionally, amino acid insertions are selected during PI-based antiretroviral therapy. Insertions ranging from 1 to 6 amino acids have been detected at various sites in the viral protease sequence. Presence of these insertions is positively correlated with protease mutations associated with resistance to PIs. While the insertions lead to a decrease in PI susceptibility and modestly improve viral replication (Kim et al., 2001; Kozísek et al., 2008), they seem to contribute to PI resistance only in combination with other mutations either in the protease or in Gag.

III.2.5. Entry/Fusion inhibitors

Entry of HIV-1 into the host cell is essential for the virus to establish a productive infection and, therefore, represents a major target for preventing HIV-1 infection and transmission. Successful development of the entry step as a therapeutic target has been driven by basic research on the molecular mechanisms involved in virus entry. The central event in the entry process is the fusion between viral and cellular membranes, which delivers the viral genome into the target cell. The viral component that performs the fusion process is the envelope (Env) glycoprotein complex, the active form of which is composed of a heterotrimer containing three molecules each of the surface glycoprotein gp120 and the

transmembrane glycoprotein gp41.

New therapeutics targeting HIV entry are among the very recent additions to the current anti-HIV-1 agents. Besides numerous novel therapeutic candidates in preclinical and clinical development, two drugs have been approved for clinical use: the fusion inhibitor T20 (enfuvirtide, fuzeon or DP178) and the CCR5 antagonist maraviroc (selzentry) approved in 2007 for restricted clinical use in US **(Table 3)**.

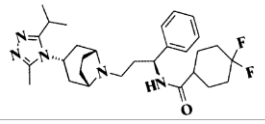
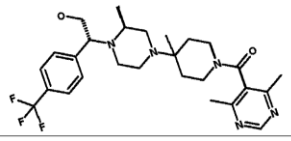
Inhibitor	Target	Status of clinical development	Structure
Enfuvirtide (T-20, fuzeon)	Fusion/entry (gp41)	FDA-approved	Ac-YTSLIHSLIEESQNQQEKNEQELLFLDKWASLWNWF-NH2
Maraviroc	Fusion/entry (CCR5)	FDA-approved	
Vicriviroc	Fusion/entry (CCR5)	Phase III	

Table 3: Entry/Fusion inhibitors (Adamson and Freed, 2008).

- ***Fusion inhibitor T20 (enfuvirtide, fuzeon)***

T20 was developed as the first HIV fusion/entry inhibitors. T20 have been licensed by the US FDA since 2003 for the treatment of HIV patients who failed to respond to the current antiretroviral therapeutics. It has been shown to potently inhibit virus replication both in vitro and in vivo. T20 is effective against a broad spectrum of HIV-1 strains, including those resistant to RTIs and PIs (Greenberg et al., 2004).

T20 is a 36-amino acid synthetic peptide, the sequence of which is derived from the gp41 heptad repeat 2 (HR2) (Wild et al., 1994). Targeting a transient structural intermediate in the fusion process, T20 blocks the fusion process by binding to HR 1, competitively inhibiting the HR1/HR2 interaction and blocking six-helix bundle formation.

- ***Resistance to T20***

Overall, T20 is considered to be a drug with a low genetic barrier for resistance development, despite the conserved nature of the gp41 HR1 region. Resistance to T20 usually maps to the peptide binding site in HR1, in a contiguous stretch of 10 amino acids between residues 33 to 43 (gp41 numbering) on the inner N-helices. Mutations at sites L33Q, V38E, Q40K, and Q41R are particularly detrimental (McGillick et al., 2010). Other limitations associated with T20 treatment include low potency and short half-life in vivo,

the high cost of the peptide manufacturing process and its mode of delivery (Pan et al., 2010).

- ***CCR5 inhibitors***

At present one CCR5 inhibitor, maraviroc, has been approved for clinical use **(Table 3)**. Several others are currently in preclinical and clinical development, including vicriviroc, which recently entered phase III clinical trials. Being developed to be administered in combination with a ritonavir-boosted protease inhibitor, early clinical trials have established the safety of vicriviroc in both treatment-naïve and -experienced R5-tropic HIV-1 infected individuals (Lenz and Rockstroh, 2010). Both maraviroc and vicriviroc are small molecules that bind CCR5, preventing its interaction with gp120.

By specifically targeting CCR5, they do not inhibit X4- or R5X4-tropic viral isolates (Strizki et al., 2005). Hence, FDA approval of maraviroc stipulates that it is used as therapy for treatment-experienced adult patients in which only R5-tropic HIV-1 is detectable. This baseline coreceptor screening is required because there is considerable concern about the clinical consequences of selection and/or expansion of X4-tropic viruses (Regoes and Bonhoeffer, 2005).

- ***Resistance to CCR5 inhibitors***

Resistance to CCR5 antagonists emerges by two mechanisms. Firstly, viruses that use X4 or are dual tropic emerge during treatment, and these are presumed to arise from a reservoir. This occurs in about two-thirds of failures and has not been associated with a rapid fall in CD4 count. The second mechanism is via substitutions of amino acid in V3 loop of the viral envelope protein, in which strongest genotypic determinants for tropism is located. Then the virus can bind to the CCR5 with the inhibitor bound in place. A possible mechanism which has not been observed is the evolution of the V3 loop such that it becomes CXCR4 tropic, a parallel situation to that observed during the normal course of disease (Cane, 2009).

III.2.6. Integrase inhibitors (INIs)

Chromosomal archiving of drug-resistant HIV genomes via integration in long lived cells can permanently exclude the utility of antiretroviral drugs or entire drug classes, which often prevents achieving the durable viral suppression necessary for stable health. The absence of known human homologue to IN suggests that IN inhibitors do not interfere with normal cellular processes, therefore probably accounts for the low toxicity and high

selectivity of the IN inhibitors. Of note, some diketo acids (DKAs) as well as blocking IN also block the activity of recombinases and RNase H, despite at higher drug concentrations (Melek M et al., 2001; Shaw-Reid et al., 2003). This cross-reactivity probably results from the mechanistic and structural similarities between recombinases, RNases and INs, although the structure-activity relationship for IN and other phosphotransferases are clearly distinct (Shaw-Reid et al., 2003).

Blocking integration was in the past postulated as potentially problematic. 50 to 100 copies of IN are brought into the cell by each incoming virion and the target process is a single pair of DNA cleaving and joining reactions per viral replication round. Unlike reverse transcription, where several thousand enzymatic turnovers by RT are needed to complete synthesis of the viral DNA, providing numerous 'shots at goal' for NRTIs and NNRTIs to block reverse transcription, the relatively small number of events needed to complete integration may potentially have made it difficult to inhibit. A timely new class of antiretroviral drugs, at this point best exemplified by the HIV integrase strand transfer inhibitor (INSTI) raltegravir (RAL, MK-0518, Isentress®) has proved that integration is a robust therapeutic target.

In terms of pharmacological development, two strategies have been considered for the design of IN inhibitors. The first strategy targets the free, unbound IN (3'-processing inhibitors), whereas the second targets the complex resulting from the association of IN with viral DNA (strand transfer inhibitors) (Hazuda et al., 2000). Both 3'-processing inhibitors and specific INSTIs have been identified in vitro, but only IN strand-transfer inhibitors (INSTIs) have turned out to be effective antiviral drugs in vivo (Mouscadet et al., 2010a).

III.2.6.1. Integrase strand-transfer inhibitors (INSTIs)

Several criteria define an INSTI or any other inhibitor of IN (Pommier et al., 2005). At least four criteria need to be met to conclude that IN is the cellular target of an antiviral inhibitor found to be active against recombinant IN. Firstly, the candidate INSTI must be active at the appropriate point in the viral lifecycle, after reverse transcription and before maturation (between 4 and 16 hours following infection), as defined by time-of-drug-addition experiments. Secondly, treatment of infected cells with a candidate INSTI should lead to an accumulation of 2-LTR circles. 2-LTR circles result from the accumulation of viral cDNA and its circularization by cellular enzymes. Successful integration and production of the proviral DNA (a process that can be measured by

Alu-PCR) should concomitantly be decreased. Finally, treatment of HIV-1 with a putative INSTI should lead to the selection of mutations in the IN gene in the selected viruses; and these viruses should show reduced susceptibility to the selecting compound. Transfer of these mutations to recombinant IN should also show reduced susceptibility of the resulting mutant enzyme to the inhibitor in strand transfer assays, in vitro.

A main impediment to discovery of INSTIs was in fact the relative complexity of the integration reaction itself. Once in vitro assays that mimicked functional integration became available, discovery in this field advanced relatively rapidly. The first INSTIs were based on a β-ketoacid fragment that efficiently chelated Mg^{2+}. Then the replacement of the α, γ, -diketoacid moiety with 8-hydroxy quinoline led to the generation of compounds with potent antiviral activity, providing the proof-of concept for INSTI activity in vivo. Further developments stemming from the replacement of the DKA fragment with a drug-compliant motif led to the discovery of RAL by Merk Research Laboratories (**Figure 19)**. RAL was approved by FDA in October 2007, and also approved for use in first-line antiretroviral therapy in June 2009. Now RAL is undergoing further phase 3 studies in antiretroviral treatment-naïve subjects including investigation of once daily dosing. The first approved INSTI appears to be as potent as any previously developed antiretroviral drug.

Further optimization of the dihydroquinoline motif derived from quinolone antibiotics gave rise to elvitegravir (EVG, JTK-303/GS-9137) **(Figure 19)**, a second INSTI in phase 3 clinical development in antiretroviral treatment-experienced subjects and is also undergoing phase 2 development in antiretroviral naïve subjects as part of a fixed dose combination regimen. EVG showed potent anti-HIV activity in vitro against HIV-1 of multiple subtypes (EC_{50} ranging from 0.1 to 1.26 nM) as well as against HIV-2 strains (EC_{50} 1.4-2.8 nM) (Shimura et al., 2008), SIV and murine leukemia viruses. Like RAL, EVG has shown potent antiviral activity in vivo against HIV-1 carrying resistance mutations to multiple antiretroviral drug classes (DeJesus et al., 2006b). EVG can be metabolized by Cyp3A4 and the pharmacokinetic profile of EVG could be boosted by ritonavir, increasing the EVG plasma half-life from three to nine hours.

A series of two-metal binding INSTIs based on a naphthyridinone scaffold was described more recently. One of these, S/GSK1349572 has recently shown impressive activity in a phase 2A study. Resistance data suggested that this compound may have an improved resistance profile on RAL and EVG selected resistance mutations, which may allow it to be used for salvage of patients with virologic failures on the other INSTIs (Underwood et al., 2009).

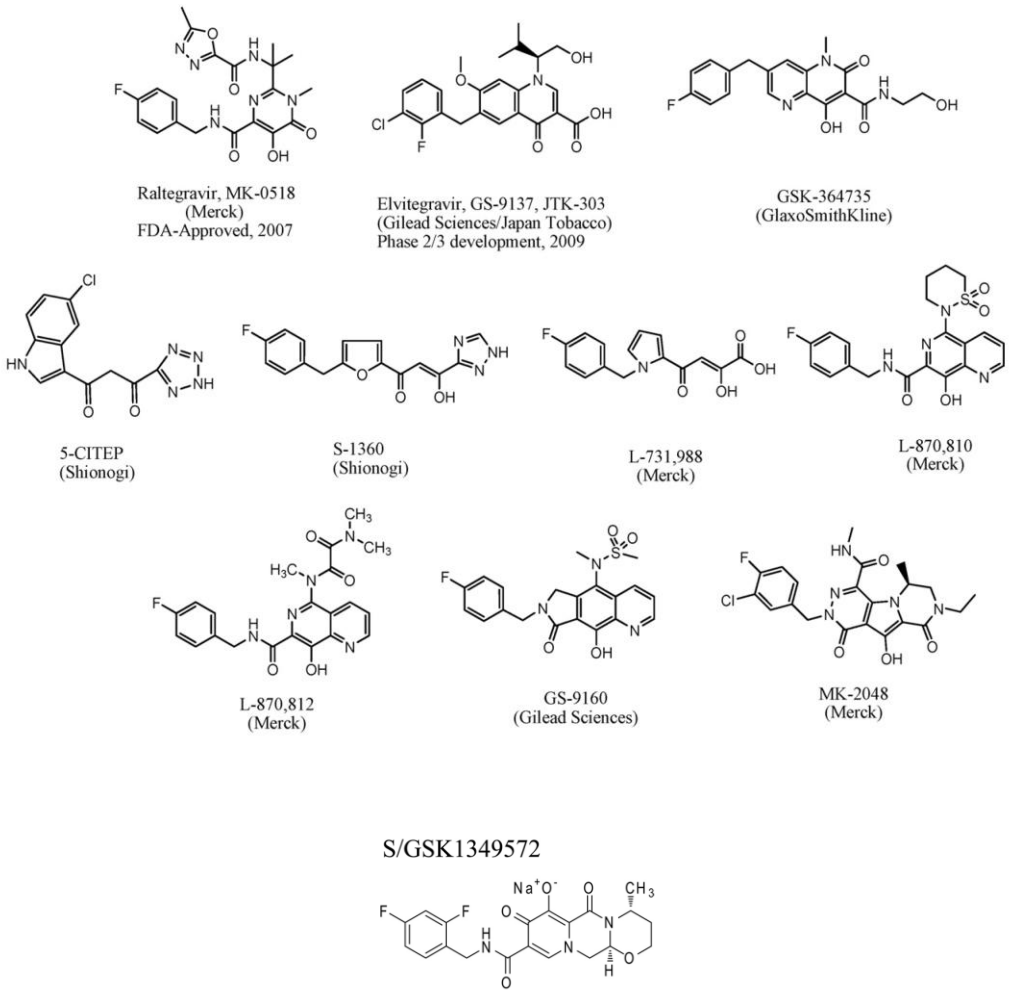


Figure 19: Chemical structures of INSTIs.

The evolution of INSTIs from discovery to clinical trials to approval, including raltegravir (RAL, MK-0518, Isentress®); elvitegravir (in phase 2/3 development); the naphthyridinone GSK-364735; early diketo acid INSTIs including 5-CITEP and S-1360 and L-731,988; the naphthyridines L-870,810 and L-870,812; a tricylic INSTI GS-9160 and a 'second generation' INSTI with an enhanced resistance profile, MK-2048 (Kobayashi et al., 2011; McColl and Chen, 2010).

III.2.6.2. Resistance to INSTIs

As is the case for other antiviral drugs, drug resistance to INSTIs emerges both in vitro and/or in vivo through the selection of mutations in the IN coding region of the *pol* gene that affect the susceptibility of the virus to the inhibitors. In many respects, resistance to INSTIs has features in common with that observed for other antiretroviral drug classes, notably the accumulation of mutations leads to higher levels of resistance (McColl and Chen, 2010). Whereas, one disadvantage of INSTI is that moderate level to high level resistance can follow after only one or two amino acid mutations. About 40 substitutions have been specifically associated with the development of resistance to RAL and/or EVG; some of them were also found in vivo in patients failing such INSTIs (Ceccherini-Silberstein et al., 2009). A summary of these INSTI resistance selections is provided in **Figure 20**.

As currently understood, resistance to INSTIs can be mediated by a number of IN mutations that can be considered 'primary' resistance mutations which are all located in the CCD of IN, close to the INSTI binding site. Numerous additional IN mutations, some of which are natural polymorphisms and most of which are also found in the CCD, comprise secondary resistance mutations which increase phenotypic resistance and/or viral fitness rescue when added to primary mutations. Generally, all secondary mutations (for both RAL and EVG) had little if any effect on drug susceptibility in vitro in the absence of a primary 'signature' mutation. Some secondary resistance mutations have been noted in the CTD of IN including mutations at residues S230 and R263; however, no resistance mutations have been described in the NTD of IN.

Raltegravir-associated mutations	Primary: **Y143R**, **Q148H/K/R**, **N155H**
	Secondary: **L74M**, **E92Q/A**, T97A, **E138K**/A, **G140S/A**, **Y143C/H**, V151I, K156N, E157Q, G163K/R, T206S, D232N
	In vitro: **L68V**, S119R/G, **F121Y**, **N155S**, I203M, **S230R/N**
Elvitegravir-associated mutations	Primary: T66I, **E92Q**, **S147G**, **Q148H/K/R**, **N155H**
	Secondary: **H51Y**, **T66K/A**, L68I/V, S119R/G, **E138K**, **G140S/C**, E157Q, **K160N**, **R166S**, **E170A**, **S230R**, D232N
	In vitro: **E92G**, **Q95K**, **H114Y**, **F121Y**, **T125K**, **Q146P**, V151I, **S153Y**, M154I, **N155S**, **R263K**
Mutations associated with other integrase inhibitors	*In vitro*: V72I, T112I, **A128T**, **Q146K**, S153A, **N155S**, **K160D**, V165I, V201I, **V249I**, **C280Y**

Figure 20: Prevalence of mutations in positions associated with in vivo and/or in vitro resistance to INSTIs.

The amino acid sequence of HIV-1 IN of clade B consensus is shown as a reference. Resistance mutations completely absent (or found only in single isolates among all studies) are shown in black bold; mutations found with ≤ 1% variability are shown in black; mutations with 1.1-10% variability are shown in dark grey bold; mutations with > 10% variability are shown in grey (Ceccherini-Silberstein et al., 2009)

- ***Resistance to RAL***

Resistance to RAL in vivo has been associated, to different degrees, with 14 mutations. But the virologic failure of RAL treatment is predominantly associated with the initial development of two principal independent genetic pathways involving primary mutations of N155 (N155H) and Q148 (Q148 H/K/R), either alone or combination with other IN mutations. The third pathway involving Y143 mutations was observed less frequently and later than N155 and Q148 pathway. The level of phenotypic resistance of HIV-1 IN to RAL associated with the Q148 or Y143 mutation is much higher (>100 times higher) than that associated with N155H (Hatano et al., 2010).

Among these secondary resistance mutations, the most common mutation pattern was G140S/Q148H. The G140S mutation rescues a replication defect resulting from the Q148H primary mutation (Delelis et al., 2009a). A small number of mutations involving residues E92, E157 or T97A/G163R may lead to alternative pathways of resistance. Genotypic

switching among RAL resistance patterns was observed with viruses initially expressing N155H and subsequently switching to the Q148 resistance pathway in many cases.

- ***Resistance to EVG***

In patients failing EVG, the mutations T66I, E92Q, S147G, Q148H/K/R, and N155H have been identified as 'signature' resistance mutations. E92Q mutation was the most common occurring in 50% patients with EVG virologic failure. The most common genotypic switch was from an initial E92Q pattern to a Q148H/K/R pattern, reminiscent of RAL in which initial N155H-containing patterns switched to the Q148 pathway. Genotypic switching or addition of other IN mutations to an initial pattern generally resulted in higher levels of phenotypic resistance to EVG and often, cross-resistance to RAL (Marinello et al., 2008 -a; Mouscadet et al., 2010a).

- ***Resistance to S/GSK1349572***

GSK1349572 seems to have a higher genetic barrier than RAL or EVG., highly resistance mutants were not selected in the presence of GSK1349572, but resistance mutations comprised L101I, T124A, S153F/Y that effected a low fold change in the EC_{50} (up to 4.1 fold) were identified in the vicinity of the integrase active site. GSK-1349572 also showed a different mutation pattern compared with RAL and EVG in *in vitro* selection. It did not exhibit cross-resistance to resistance mutants of RAL and EVG (Vandekerckhove, 2010). GSK1349572 was actively against IN mutants containing the RAL-resistant signature mutations Y143R, Q148K, N155H, and G140S/Q148H (Kobayashi et al., 2011).

III.2.6.3. Mechanism of INSTIs inhibition and resistance

The mode of INSTI binding and action are unlikely to significantly differ, because the core contact points consisting of invariant nucleotide bases and amino acid residues are conserved in HIV-1. Pharmacophore analysis suggests two functional domains: the metal-binding moiety, or catalytic moiety, and an enzyme-binding moiety (Marchand et al., 2002). All INSTIs contain divalent metal ion binding motifs. Unlike the enzyme-binding moiety, the metal-binding moiety of the INSTIs tends to be structurally conserved. The enzyme-binding moiety of INSTIs contains one or more aromatic hydrophobic groups and can accommodate hydrophobic groups of diverse size, and may enable the transport of INSTI across the cellular membrane.

Recently obtained diffracting crystals of the full-length IN from the prototype foamy virus (PFV) revealed that the hydrophobic groups of enzyme-binding moiety fit within a

tight pocket is close to the DDE catalytic triad (Hare et al., 2010b). Within hydrophobic pocket, the drugs make intimate Van der Waals interactions with the bases of the invariant CA dinucleotide, guanine 4 from the nontransferred strand and conserved residues of IN. The hydrophobic pocket forms following a conformation change induced by the binding and 3'-processing of the viral DNA (Chen et al., 2008). Viral DNA may well form part of the inhibitor binding site. Indeed, the diketo acid inhibitor can only bind to the acceptor site after 3'-processing. Crucially, the drug binding results in displacement of the reactive 3' viral DNA end from the active site, which can result in deactivation of the IN (Hare et al., 2010b).

Diffracting crystals of IN from PFV also revealed that metal chelating oxygen atoms of INSTI orient towards the metal cofactors of the active site. Evidence suggests that the catalytic moiety of INSTIs sequesters the divalent metal ions (either Mg^{2+} or Mn^{2+}) in the IN active site and thereby inhibits enzymatic catalysis (**Figure 21)**. Mutations at Q148 or N155 are likely to interfere with coordination of metal cofactors by the active site carboxylates. Conceivably, a slight shift in metal ion cofactor positions might suffice to abrogate drug binding, which relies on its spatially constrained metal chelating groups, albeit at a high price of impaired viral replication fitness due to detuning of the IN active site structure (Hare et al., 2010b). In addition, an overall linear inverse correlation between resistance and catalytic activity was observed for RAL as well as EVG. This correlation is consistent with the specific interaction of the drugs with the IN catalytic site (Marinello et al., 2008 -b).

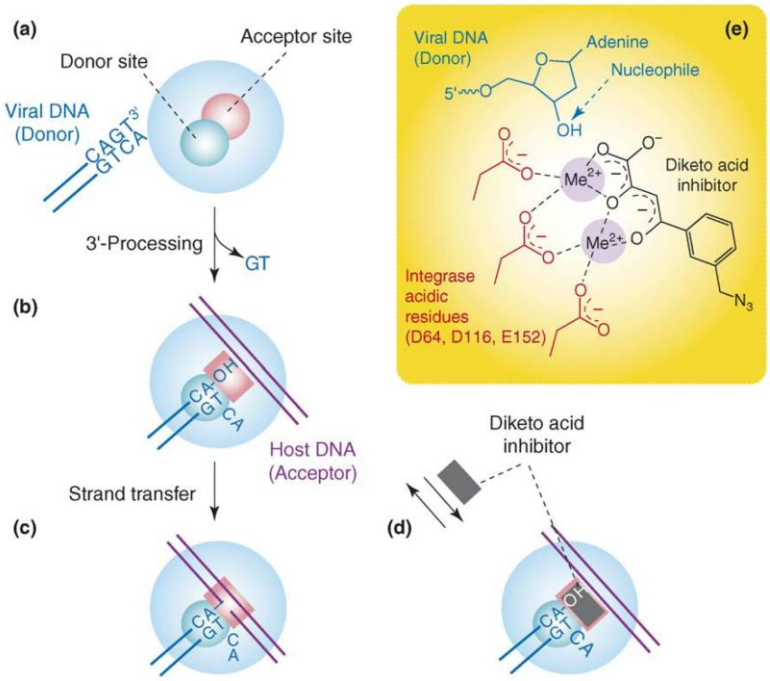


Figure 21: Schematic diagram of inhibition by diketo acid inhibitor (DKA)

(a-c) IN 3'-processing and strand transfer. (d) The DKA inhibitor (gray rectangle) can only bind to the acceptor site after 3'-processing. (e) Details of the hypothetical binding of DKAs at the interface of the IN-DNA-divalent metal complex. The processed viral 3'-DNA ends (in blue) are bound to IN (in red, DDE catalytic residues) ready to attack a host DNA phosphodiester bond. DKA inhibitors have been proposed to chelate the divalent metals in the IN catalytic site (Marchand et al., 2006).

III.3. Current challenges of antiretroviral therapy

Despite intensive research efforts have been devoted to develop effective preventive vaccines and to identify immune-based strategies boosting the depleted immune systems, a cure for HIV infection remains elusive. In addition to these ultimate challenges, there are a number of unmet goals which seem closer and more realistic to study.

The high cost of HAART may be too much to sustain treatments worldwide, since an estimation of 33.3 million adults and children infected at the end of 2009. The broad majority of patients living with HIV in less-developed countries have not yet been reached by current therapy options.

Current antiretroviral treatment is still with several limitations, even if it is available.

Patients are destined to undergo treatment for life, since HIV infection cannot be cured with current therapies. Considering the need for life-long antiviral therapy, drug resistance remains an important limitation, especially when drugs with low genetic barrier for resistance development are part of a treatment regimen. Transmitted resistance can also compromise the efficacy of HAART. Transmitted resistance to at least one antiviral drug has been established in 6 to 16% of patients (Ross et al., 2007). Patients with baseline resistance have been found to have a poorer response to therapy including further selection of drug resistance to other components in the regimen (Johnson et al., 2008a; Little et al., 2002).

Despite the extended suppression of HIV replication below the standard limits of detection, the persistence of latently infected memory CD4+ T lymphocytes precludes the elimination of virus by HAART alone for the lifetime of the patient. Viral relapse from latent reservoir could be caused by discontinuation of HAART. Ongoing viremia is still detectable though at low levels of 1 to 50 copies per milliliter in the majority of patients. Viremia does not appear to compromise the extended success of HAART therapy in the adherent patients. However, emergence of resistance mutations can occur during phases of low level viremia, for instance, in patients receiving INSTI containing antiretroviral regimens (Gallien et al., 2011). Futhermore, chronic consequences could be led due to that virions can use CD4 and chemokine receptors and may activate pathways. Suboptimal penetration of numerous antiretrovirals into the central nervous system might allow low levels of viral replication and/or release from stable viral reservoirs, resulting in neuropathology (Lambotte et al., 2005; Watanabe et al., 2010). Despite the extremely low rates of toxicity of the novel HAART regimens, a lot of these antiretroviral drugs modulate lipid and glucose metabolism. Over many years of treatment, even modest toxicities could have cumulative effects. Moreover, prolonged therapy might exhibit toxicities which were not estimable with animal toxicology assay or several years of clinical surveillance. Moreover, medications are required not only for HIV infection but also for related or unrelated co-morbidities, providing further challenges for healthcare providers and patients alike (Richman et al., 2009).

The foreseeable near-term goals are simplification of therapy, improvement of patient adherence and minimization of adverse events and particularly drug resistance and cross-resistance between drug classes. These goals will very possible also be important milestones towards achieving long-term drug-free virus control and ultimately finding a cure (Estéa and Cihlarb, 2010).

III.4. Proviral latency

Despite the durable suppression of detectable plasma viremia, patient data suggest that about 1 in 10^6 CD4+ T cells are latently infected with HIV. It is considered that the latently infected cells are intermittently activated by antigen recognition or as bystanders in a local inflammatory process in vivo, which results in the release of progeny virions. Activation of latently infected CD4+ T cells is probablely not the unique cause of low-level plasma viremia. These performed researchs propose that a chronically infected clonal reservoir, analogous to a persistently infected stem cell, likely exists (Bailey et al., 2006). Other cellular or tissue sources of virus, such as cells of the monocyte and macrophage lineages, may also contribute to low levels of viremia. Because viral reservoirs express no viral protein, latently-infected reservoir cells are immunologically indistinguishable from uninfected cells and are insensitive to immune clearance and HAART (Colin and Lint, 2009). It remains unclear that how a persistently infected cell population could produce virions at a steady state for years, in the presence of some level of cell-mediated immunity.

Integrated viral genomes can be durably and reversibly repressed by a combination of mechanisms. Firstly, the provirus can be repressed via neighboring cis-acting sequences. Integration can occasionally occur within less favorable chromatin environments, resulting in latency. Secondly, the transcriptional status of HIV is tightly linked to the activation state of its host cell. In resting T cells, nucleosomes adjacent to the HIV promoter, notably one situated at the transcriptional start site (Nuc 1), possess markers of silent heterochromatin. The 5' LTR of HIV harbors sequences that can bind negative regulators in resting T cells. Lastly, latency can be strengthened by posttranscriptional mechanisms, such as impaired HIV mRNA nuclear export, because of low amounts of polypyrimidine tract-binding protein (PTB) and expression of host or viral micro-RNAs (miRNAs) (Trono et al., 2010).

Since latency is likely established and maintained by numerous blocks at multiple steps in the HIV replicative cycle (Williams and Greene, 2007), which potentially complicate eradication strategies.

Then combination approaches of eradication could be required and depend on current or future antiretroviral therapy to reactivate HIV-1 from latently infected cells and completely inhibit all new infection events. Viral eradication might also be achieved by strategic interventions targeting the resistance of infected cells to apoptosis and molecules involved in latency reactivation. Furthermore, according to recent data highlighting the existence of two subsets of memory T cells serving as a reservoir, a combined application of strategic

intervention targeting viral replication (through reinforced HAART) and antiproliferative drugs (e.g., anti-cancer drugs) has been proposed. This treatment has to be introduced very early in the course of infection, as it reduced the constitution of the proliferative reservoir drastically (Redel et al., 2009). Antilatency agents would be given, intermittently and for a limited period of time, to purge the last sanctuaries of HIV infection **(Figure 22)**.

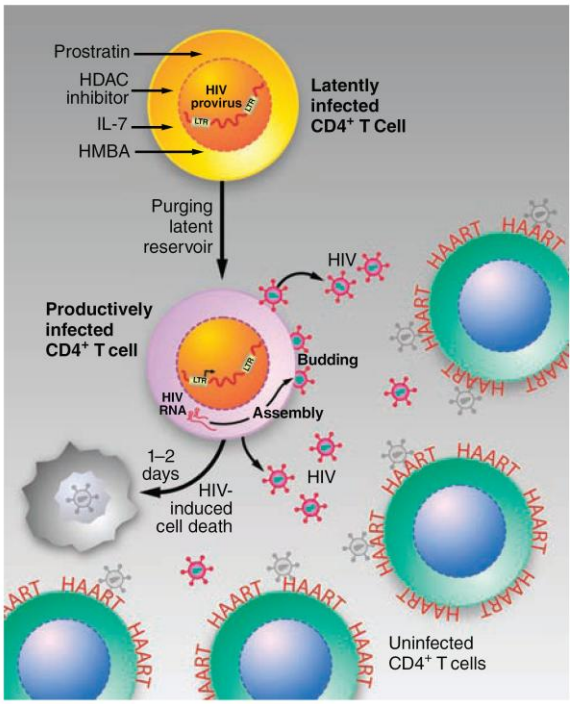


Figure 22: Purging persistent proviral infection.

If targeted approaches, alone or in combination, succeed in activating latent HIV proviruses present in differentiated CD4+ T cells, the life span of these cells should be short. These inductive agents must be used in combination with HAART to prevent further HIV spread to uninfected CD4+ T cells (Richman et al., 2009).

MOTIVATION AND OBJECTIVE

Integrase (IN) is responsible for the integration of viral genome into the host genomic DNA. Owing to its key role in viral replication cycle, and no counterpart in the host cells, IN is an attractive target for antiretroviral therapy as well as the other two viral enzymes, reverse transcriptase (RT) and protease (P). Raltegravir (RAL), the first and to date unique IN strand transfer inhibitor (INSTI) approved by FDA in October, 2007, entered the therapeutic arsenal which led to the establishment of highly active therapies (HAART) that had a profound effect on the morbidity and mortality of HIV-1 infected individuals. Representing a novel drug class, RAL has demonstrated a rapid and sustained antiretroviral effect in both antiretroviral treatment-naïve and experienced patients with a favourable side-effect profile. Due to the distinct mechanism of action, RAL is also efficiently against viruses resistant to other classes of antiretroviral drugs.

However, most of currently available antiretroviral medicines were developed and tested principally using HIV-1 subtype B which represents less than 10% of the total HIV infected population over the world. Few data are available on *in vitro* and *in vivo* evaluation of IN inhibitors on HIV-1 non-B subtypes that are predominant in the world and, particularly for subtype CRF02_AG (circulating recombinant form A/G) which is highly prevalent in West Africa and becoming more frequent in other countries. Previous in silico study suggests that polymorphisms in HIV-1 subtype CRF02_AG IN, especially the significant variations, might have consequential effects on the interaction between the inhibitor and its protein target, and hence modulate IN inhibitor susceptibility. Therefore, the first study was performed to evaluate the susceptibility of naïve patient-derived HIV-1 subtype CRF02_AG INs to the latest INSTIs.

In spite of the apparent diversity of currently available antiretroviral drugs, the alternatives for HIV-2 infected patients are more limited, for instance, non-nucleoside reverse-transcriptase inhibitor (NNRTI) and fusion inhibitor are not active against HIV-2. Therefore, the development of novel treatments based on drug classes with high efficacy against HIV-2 is essential. Despite approximately 40% of heterogeneity between HIV-1 and HIV-2 IN genes, INSTIs are actively against HIV-2 IN with a phenotypic susceptibility similar to that of HIV-1 and represent therefore a promising option for treatment of HIV-2 infected patients. However, as is the case for other antiviral drugs, resistance to INSTIs emerges rapidly through the selection of mutations within the IN coding sequence, dramatically reducing the susceptibility of virus to INSTIs. Recent studies established that HIV-2 resistance to RAL involves one of the three primary resistance mutations N155H, Q148R/H and Y143C already described for the HIV-1 IN coding sequence. While the resistance of HIV-1 IN has been confirmed in vitro by means

of site-directed mutated INs, no such study was performed on HIV-2 IN. Thereby the second part of this study was carried out to determine in vitro resistance to RAL and catalytic properties of HIV-2 IN mutants.

Emergence of resistance to existing drugs serves as a main impetus for the development of new drug classes with distinct inhibition mechanism or agents in classes without cross-resistance, despite the availability of 23 approved drugs for the treatment of HIV/AIDS. Camelid single-domain antibody (sdAb) targeting HIV IN is a promising candidate for the development of antiretroviral agents, in virtue of many attractive properties. SdAbs recognize different epitopes normally inaccessible for the conventional antibodies and exhibit more homology with human. In addition, sdAb can be easily produced and highly stable. It can be correctly folded and retains active in reducing cellular environment such as cytoplasm and nucleus, thus it is possible to obtain the intracellular active sdAbs. Hence, for developing new antiretroviral agents able to block the strand transfer activity of the broadest range of INs, including INs resistant to INSTIs, and new research tool to elucidate the molecular functions of the targeted proteins during different steps of the virus life cycle, specific sdAbs targeting HIV IN were selected and charaterized in the last part of this study.

MATERIALS AND METHODS

I. Purification of IN and in vitro activity tests

I.1. Cloning of IN

The HIV IN used in this study is a recombinant IN expressed in a heterologous bacterial system. All the expression constructions in this study were from the same plasmid pET-15b. INs produced by this system contain a 6 x histidine tag at N-Ter which allows the purification by one step described in the following section.

The IN coding sequence obtained from HIV infected patient was amplified with polymerase chain reaction (PCR) using primers containing Nde I and BamH I restriction enzyme cutting sites, then linked into TA cloning plasmid pGEM-T easy (Promega). Ligation product was transformed into competent bacteria XL1- Blue or XL10-Gold (Stratagene), and selected by ampicillin. After verification with sequencing (Eurofins MWG operon) and digestion by Nde I and BamH I, IN coding sequence was inserted into expression vector pET-15b, and then verified by sequencing.

I.2. Site-directed mutagenesis

In the second part of this thesis, site-directed mutations had been introduced using QuikChange® Lightning mutagenesis kit (Stratagene) into different positions of HIV-2 wild-type IN obtained from patient under RAL-treatment at the time of baseline. These mutants were used to study the effects of single or double mutations on enzymatic activity and the susceptibility of mutant to RAL. This mutagenesis system allows substituting residue by a single step PCR using oligonucleotide primer harboring the desired mutation. The synthesized complementary oligonucleotide sequences for mutagenesisi are as follows:

E92Q (The Glutamic acid at position 92 in HIV-2 wild-type IN was substituted by a Glutamine)
(+) 5’ - GAG GCA GAA GTA ATA CCC CAA **C**AA ACA GGA AGA CAA ACA GCT - 3’
(-) 5’ - AGC TGT TTG TCT TCC TGT TT**G** TTG GGG TAT TAC TTC TGC CTC - 3’

Patient T2 Y143C (The Glutamine at position 92 in HIV-2 IN from Patient T2 E92 Q /Y143C was retrieved with the wild-type residue Glutamic acid)
(+) 5’ - GAG GCA GAA GTA ATA CCC CAA **G**AA ACA GGA AGA CAA ACA GCT -

3’
(-) 5’ - AGC TGT TTG TCT TCC TGT TT**C** TTG GGG TAT TAC TTC TGC CTC - 3’

T97A (The Theronine at position 92 in HIV-2 wild-type IN was substituted by an Alanine)
(+) 5’- CAA GAA ACA GGA AGA CAA **G**CA GCT CTC TTC CTG TTG AAA CTG -3’
(-) 5’- CAG TTT CAA CAG GAA GAG AGC TG**C** TTG TCT TCC TGT TTC TTG -3’

G140S (The Glycine at position 140 in HIV-2 wild-type IN was substituted by a Serine)
(+) 5’ - GGA ATA GAA CAA ACC TTC **TC**G GTA CCC TAT AAC CCA CAA - 3’
(-) 5’ - TTG TGG GTT ATA GGG TAC C**GA** GAA GGT TTG TTC TAT TCC - 3’

Q148R (The Glutamine at position 148 in HIV-2 wild-type IN was substituted by an Arginine)
(+) 5’ - CCC TAT AAC CCA CAA AGT C**G**G GGA GTA GTA GAA GCA ATG - 3’
(-) 5’ - CAT TGC TTC TAC TAC TCC C**C**G ACT TTG TGG GTT ATA GGG - 3’

N155H (The Asparagine at position 155 in HIV-2 wild-type IN was substituted by a Histidine)
(+) 5’ - GGA GTA GTA GAA GCA ATG **C**AC CAT CAC CTA AAA AAC CAG - 3’
(-) 5’ - CTG GTT TTT TAG GTG ATG GT**G** CAT TGC TTC TAC TAC TCC - 3’

The introduced nucleoside mutation in the context of original sequence is in bold and underlined. The following protocol was engaged for each mutation. The 50 µl reaction mixture consists of 5 µl 10 x reaction buffer, 20-50 ng plasmid produced by bacteria to mutate, 125 ng of each complementary oligonucleotide, 1 µl mixture of dNTP, 2.5 units of Pfu Ultra (DNA-dependent DNA polymerase with a very low error rate) and sterile Milli-Q water filled to the final reaction volume.

The PCR cycle is as follows: 2 minutes at 95°C, 18 cycles of 30 seconds at 95°C, 15 seconds at 58°C, and 4.5 minutes at 68°C (30 sec/kb), then 7 minutes at 68°C. Then Dpn I (10 units) was added into amplification products to digest the methylized parental plasmid, but not the new non-methylized DNA product carrying the desired mutation. The reaction proceeded at 37 ℃ for 1 hour. 50 µl competent bacteria XL1 blue or XL10 gold were transformed with 1 µl PCR mixture.

Transformed bacteria were plated on LB plates containing ampicillin and incubated overnight at 37 ℃ or 2 days at room temperature. The clones were obtained after the mini-preparation of plasmids and sequencing to verify the presence of the desired mutation and the absence of any other mutation due to the error of Pfu Ultra.

I.3. Expression of HIV IN

His-tagged INs were produced in Escherichia coli BL21-CodonPlus (DE3)-RIPL (Agilent) and purified under nondenaturing conditions as previously described (Leh et al., 2000). Derived from the BL21-Gold competent cell line, BL21-CodonPlus cells are able to express heterologous proteins efficiently due to extra copies of the ArgU, IleY, and LeuW tRNA genes. A pre-culture was incubated at 37°C overnight. Then the pre-culture was added into 500ml medium with ampicillin, incubated at 37°C, 180 rpm. When DO_{600} was from 0.6 to 0.8, Isopropyl β-D-1-thiogalactopyranoside (IPTG) was added to a final concentration of 0.5 mM to induce the production of IN. The expression can be induced in different temperature according to recombinant IN solubility profile: 30°C for 3 hours if the protein remains soluble in the bacteria; or 25°C for 6 hours if the protein intends to form inclusion bodies. The bacterial culture was centrifuged at 3500 rpm for 20 minutes, at 4 °C. The supernatant was removed and the pellet can be stored at -80°C.

I.4. Purification of HIV IN

All the solution used for IN purification must be filtered and pre-cooled. All the procedures should be performed on ice or at 4 °C. The cell pellet was resuspended in pre-cold buffer A (50 mM Tris-HCl pH 8, 1 M NaCl, 4 mM β-mercaptoethanol), then lysed by French press or sonicator (frequence 60Hz, pulses 9 seconds for 3 minutes, 2 times on ice). After centrifugation (30 minutes at 10000 rpm, 4°C), the supernatant was filtered (pore size 0.45 μm) and incubated with nickel-nitrilotriacetic acid agarose beads (Ni-NTA) (Qiagen) in the presence of 5 mM imidazole under gentle shaking for at least 2 hours or overnight at 4°C. About 2 ml prewashed Ni-NTA beads solution were used for 500 ml bacterial culture. After binding, the beads were washed twice with 10 volumes of buffer A firstly. The presence of protein can be verified with a visual test with Bradford in 96-well plate (10 μl supernatant or beads add into 190 μl Bradford 1 x). When the protein was not present in supernatant (absence of bleu colorization in Bradford test) and present on beads, 10 volumes of buffer A supplemented with 80 mM imidazole were used to wash beads for 5 times. If necessary, one or two times washing with 10 volumes of buffer A containing 100 mM imidazole can be performed. After the last washing, His-tagged recombinant IN was transferred into 2 ml Eppendorf tubes, then eluted with elution buffer (buffer A supplemented with 50 μM $ZnSO_4$ and 1 M imidazole) for 20 minutes at 4°C. IN concentration was estimated by Bradford. During elution, dialysis membrane was cut and washed with Milli-Q water and then with buffer A. The dialysis was performed at 4°C overnight with gentle stirring, in the presence of 20 mM Tris-HCl (pH 8), 1 M NaCl, 4 mM

β-mercaptoethanol, 10% glycerol and 50 µM $ZnSO_4$. IN was recovered and dosed by Bradford. It was aliquoted into 30-50 µl and rapidly frozen with liquid nitrogen or dry ice, then conserved at -80°C.

In the case when IN precipitates during the dialysis, the solubility can be improved by lowering the pH of dialysis buffer to 6.5 and/or using gradient concentration of imidazole (300 mM, 100 mM, and then 0 mM imidazole). Dialysis can also be replaced by gradual buffer change performed with Amicon cut-off of 10 kDa.

I.5. In vitro IN enzymatic activity test

Two sorts of IN activity tests using radio-marked oligonucleotides or fluorescence anisotropy were engaged in this study. ^{32}P marked oligonucleotides were used to test 3'-processing (3'-P) and strand transfer (ST) activity, while fluorescence anisotropy was performed to test substrate DNA-binding and 3'-processing activity of IN.

I.5.1. Activity test using radio-marked oligonucleotides

Oligonucleotides (ODN) mimicked the long terminal repeat of viral genome were synthesized (Eurogentec Belgium) and purified on 16% acrylamide-urea sequencing gel, and then passed on NAP^{TM} 5 column (Amersham) to eliminate the salts. The oligonucleotide is visible under 254 nm UV. The elution was performed by soaking the gel containing oligonucleotices with 1 ml Milli-Q water overnight at room temperature. After quantified with spectrophotometer (absorbance at 260 nm), the oligonucleotides were dried with Speed Vac Plus (SC110A, Savant), and dissolved in 20 mM Tris-HCl pH 7.0 to a concentration of 50 or 100 µM. Purified oligonucleotides were conserved at -20°C.

HIV-1 ODN substrate sequences were as follows:

U5A (21-mer): 5' - ACTGCTAGAGATTTTCCACAC - 3';

U5B (21-mer): 5'- GTGTGGAAAATCTCTAGCAGT - 3';

U5B-2 (19-mer): 5'- GTGTGGAAAATCTCTAGCA - 3' (Leh et al., 2000).

HIV-2 ODN substrate sequences were as follows:

U5A (28-mer): 5' - CCTGCTAGGGATTTTCCTGCCTCGGTTT- 3';

U5B (28-mer): 5' - AAACCGAGGCAGGAAAATCCCTAGCAGG - 3';

U5B-2 (26-mer): 5' - AAACCGAGGCAGGAAAATCCCTAGCA - 3' (Gent et al., 1992).

- ***Radiolabeling oligonucleotide substrates at 5'***

Oligonucleotides were radiolabeled with T4 polynucleotide kinase (Biolabs) and [γ-^{32}P]

ATP (3000 Ci/mmol) (Amersham). The reaction mixture constitutes [γ-^{32}P] ATP (2.5 µcurie/µl final), 2 µl 10 x reaction buffer, 10 units kinase, oligonucleotide (500 nM final concentration), then completed to 20µl with sterile Milli-Q H_2O. The marking reaction was performed at 37°C for 1hour, and stopped by adding 2 µl 0.5 M EDTA. The kinase was inactivated by heating at 65°C for 10 minutes. Double-stranded ODNs were obtained by mixing equimolar amounts of complementary strands (250 nM final) in the presence of 100 mM NaCl. The 40 µl reaction volume was completed with sterile Milli-Q H_2O. The mixture was heated at 90°C for 5 minutes. Then hybridation was carried out at room temperature for 1 hour by slow cooling. The ODN was then purified to eliminate [γ-^{32}P] ATP and EDTA on G-25 column (Amersham). In the end of purification, the presence of radioactivity was verified by checking the collected phase.

- ***3'-processing and strand transfer assays***

3'-P and ST assays were carried out at 37°C in a buffer containing 20 mM HEPES pH 6.8, 1 mM dithiothreitol (DTT), 7.5 mM Mg^{2+}, and 50 mM NaCl in the presence of 12.5 nM either U5A/U5B (3'-P) or U5A/U5B-2 (ST) double-stranded DNA substrates, respectively **(Figure 23)**. The reaction was stoped with 2 µl EDTA 0.5 M. To precipitate ODN, a mixture containing 10 µl NaAc 3 M pH 7, 1 µl glycogene 10 mg/ml and 67 µl TE pH 7 was added in reaction tube. The extraction of phenol/chloroforme allows eliminating proteins. The organic phase in the bottom of the tube was eliminated with the help of the control probe comfirming the presence of radioactivity, and the surprenant aqueous phase was kept. The ODNs were precipited in ethanol (2.5 volume of absolute ethanol/1 volume extract). The tubes were kept at - 20°C for 20 minutes, and then centrifuged at 13400 rpm for 30 minutes. After second washing with 250 µl 70% ethanol, the ODNs were dried and the radioactivity values were counted. The ODNs were dissolved in formamide/EDTA (20 mM final) to have around 5000 to 10000 cpm/µl (the minimum resuspension volume is 5 µl). The ODNs were denatured by heating at 90°C and fast-cooled on ice. 3'-P and ST products were separated by electrophoresis in 16% acrylamide/urea denaturing gel. The migration was carried out at 80 W for 1h30. The gel was dried. The exposure was performed with a phosphorescente screen which is excited by the β radiation of ^{32}P. Gels were analyzed with a Typhoon TRIO variable mode imager (GE Healthcare) which scanned the screen and measured the intensity of emitted phosphorescence. The quantification was preformed using the Image Quant TL software program.

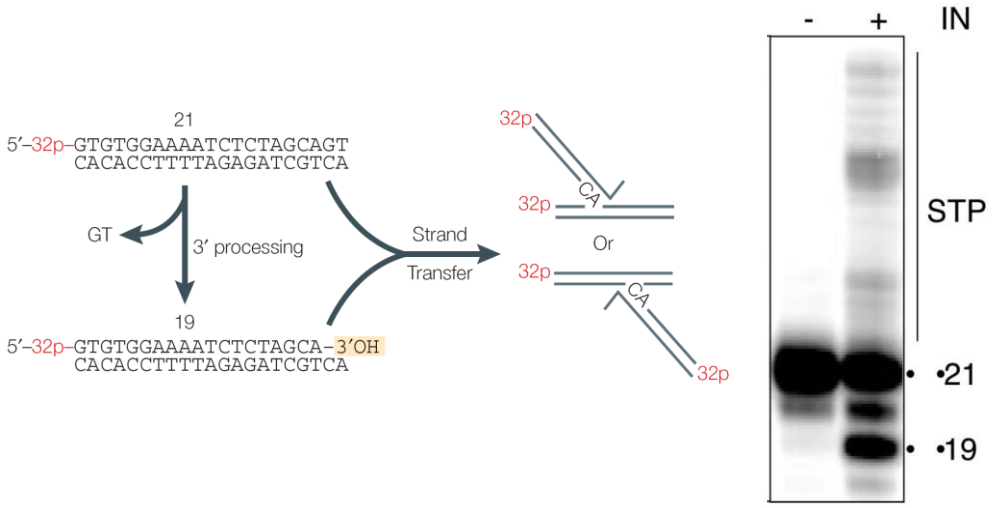


Figure 23: In vitro assay to analyze IN enzymatic activities.

The oligonucleotides substrates derived from the U5 LTR ends is radiolabelled with ^{32}P (in red) at the 5' terminus. Recombinant integrase can use the same oligonucleotide species as both donor and acceptor. The strandtransfer reaction generates a series of products longer than 21 nucleotides (STP) which migrate slower on a denaturing sequencing gel (Marchand et al., 2009; Pommier et al., 2005).

- ***Susceptibility of INs to INSTI***

Susceptibility of INs to RAL was determined in vitro by testing IN ST activity in the presence of increasing concentration of INSTI dissolved in DMSO. The quantification was preformed using the Image Quant TL software program. Fifty percent effective concentration (EC_{50}) calculation was performed using Prism 5.0 software (GraphPad Software, San Diego, CA).

I.5.2. Activity test using static fluorescence anisotropy

The Beacon ®2000 (PanVera, Madison, USA) is an apparatus to measure the polarization or static fluorescence anisotropy of liquid samples with a volume varied from 100 µl to 1 ml (non-polarizing glass cubes). The measuring chamber is thermostated (6°C to 65°C). This device is controlled by computer. It is equipped with excitation and emission filters which are able to work with nanomolar fluorophores, and with the fluorescence which is excited at 495 nm and emits around 515 nm.

This test is based on additive proprieties of fluorescence anisotropy. Marked with the fluoresein at T of the extrimity CAGT, ODN substrate possesses a static anisotropy related

to its size and sequence. The observed fluorescent complex has a higher hydrodynamic volume when IN fix to the ODN substrates, therefore a larger anisotropy. Hence, during the fixation of IN to DNA, an augmentation of static anisotropy of reaction mixture can be observed.

Steady-state anisotropy values were recorded in a cell thermostatically held at 25 °C or 37 °C, for the DNA-binding step and 3′-processing reaction respectively. The cubes were treated with BSA solution (Sigma A0281) at 20 mg/ml for 20 minutes at room temperature and then removed. This pretreatment allows limiting the adsorption of IN on the wall of tubes. The formation of IN/DNA complexes was studied by incubating 3 nM fluorescein-labeled dsODNs with IN in 20 mM Hepes pH 7 10 mM Mg2+, 1 mM DTT and 50 mM NaCl (the sample volume was 200 µl).. An equimolar mixture of two complementary ODN strands marked with fluoresein were hybrided in 20 mM HEPES, pH 7, with 50 mM final concentration of NaCl to form double strand. The final concentration of ODN substrates was 1 µM. The sequences of ODN substrate were EA21F3: 5'- GTG TGG AAA ATC TCT AGC AGT -3'F; EB21NN: 5'- ACT GCT AGA GAT TTT CCA CAC - 3' (Eurogentec).

Anisotropies of IN-bound and free ODN were recorded (no significant concomitant change in fluorescence intensity was observed). Catalytic 3'-P were performed at 37 °C. The reaction solution contained two fluorescent species: the non-processed ODN and the fluorescein-labeled dinucleotide released by the cleavage reaction. The released fluorescent dinucleotide GT has a lower static anisotropy due to its size compared with the substrate and DNA/IN complex. Hence the 3′-P activity can be assessed by quantifying the decrease in anisotropy. The reaction was stopped by adding SDS (0.25% final) to disrupt all the IN/DNA complexes in the sample.

RESULTS

I. Determining the susceptibility of latest integrase strand transfer inhibitors on HIV-1 subtype CRF02_AG in vitro

I.1. Introduction

Currently, intensive researches on HIV have led to further understanding of the mechanism of virus infection and continually advance in antiretroviral treatment. As one of the three essential viral enzymes, integrase (IN) is responsible for the integration of viral DNA into host chromosome. Moreover, it has no counterpart in human cells which also renders IN an attractive candidate for antiretroviral therapy. In October 2007, the first IN strand transfer inhibitor (INSTI) ralvitegravir (RAL) was approved by the U.S. FDA. RAL has demonstrated potent anti HIV activity in both antiretroviral treatment-naïve and experienced patients with a favorable side-effect profile (Dickinson et al., 2009; Nair and Chi, 2007). Another INSTI, elvitegravir (EVG, GS-9137) is currently showing promising efficiency in Phase III clinical studies (Marinello et al., 2008 -b; Shimura and Kodama, 2009). Data from previous clinical trials have demonstrated excellent virological responses and minimal toxicities with EVG (Klibanov, 2009). Another highly active diketo acids, L-731,988, is also one the most potent inhibitors of PIC described (Hazuda et al., 2000).

However, few data are available on in vitro and in vivo evaluation of INSTIs on HIV-1 subtypes other than B which are predominant in the world and, particularly for subtype CRF02_AG (circulating recombinant form A/G) which is highly prevalent in West Africa and becoming more frequently in other countries (Derache et al., 2007; Descamps et al., 2005; Martínez-Cajas et al., 2008; Wensing et al., 2005). Most studies of IN genetic variation and structure were preformed on HIV-1 subtype B strains which are prevalent in Americas, Europe and Australia. Antiretroviral drugs are also developed and primarily tested using HIV-1 subtype B isolates. Nevertheless, subtype B just represents around 11% of the HIV infected population worldwide (Hemelaar et al., 2011).

The knowledge from studies on B subtype should not be directly applied to all subtypes, given the significant sequence differences observed in both the structural and regulatory genes of HIV-1 subtypes. Enzymatic and virological data support the concept that naturally occurring polymorphisms in different non-B subtypes ca may influence the viral properties such as infectivity, transmissibility, the susceptibility to different antiretroviral drugs, the magnitude of resistance conferred by major mutations, and the propensity to acquire some

resistance mutations (Brenner et al., 2011a; Maïga et al., 2009; Martinez-Cajas et al., 2009; Martínez-Cajas et al., 2008). For instance, polymorphisms in some non B subtypes were able to modulate susceptibility to antiretroviral treatments. Recombinant A/G exhibits a lower baseline of susceptibility to PIs (Fleury et al., 2006); and some recombinant A/G isolates have reduced susceptibility to a NRTI, abacavir, potentially related to associated substitutions of RT at positions 123 (D123N) plus 135 (I135V); some CRF02_AG are resistant to atazanavir (Fleury et al., 2006).

Previous analysis of IN coding sequence showed that polymorphism rates equal or above 0.5% were found for 34% of the catalytic core domain positions, 42% of the C-terminal domain positions, and 50% of the N-terminal domain positions of HIV-1 group M including CRFs IN which were obtained from INI naïve patients (Rhee et al., 2008). In silico analysis that compared the IN coding sequence of subtype B and CRF02_AG found 13 significant variations: K14R, V131I, L101I, T112V, T124A, T125A, G134N, I135V, K136T, V201I, T206S, L234I, S283G (Malet et al., 2008c). Those variations are generally regrouped into three main clusters C1 (Leu101, Thr112, Gly134, Ile135, Lys136), C2 (Val31, Thr124, Thr125), and C3 (Val201 and Thr206). C1 and C2 are along the channel formed by a pair of IN monomers and crossed by host DNA. They are also close to and localized between HHCC and DDE metal binding motifs of each IN monomer. C3 is located at the dimeric interfaces. The difference of residues size caused by replacement at position 14 (K to R) may affect the strength of the putative contact between Lys14 and the viral DNA (Wielens et al., 2005). Leucine showing a little preference for α-helix at position 234 is substituted by isoleucine which is more easily accommodated in β-strand. Variations at position 283 (Ser to Gly) lead to a noticeable change from hydrophilic in subtype B to hydrophobic residues in subtype CRF02_AG (Malet et al., 2008c).

Since these important variations are located in highly order domains of IN, they may alter the structure of IN and the interaction between INSTIs and IN as well, consequentially have an impact on the susceptibility of non-B subtype IN to INSTI. To verify this hypothesis, we determined the susceptibility of latest integrase strand transfer inhibitors on HIV-1 subtype CRF02_AG in vitro, and compared with in silico study.

I.2. Manuscript

In silico and *in vitro* comparison of HIV-1 subtypes B and CRF02-AG integrases susceptibility to integrase strand transfer inhibitors

Xiaoju Ni[1,2*], Safwat Abdel-Azeim[1*], Rohit Arora[1], Osamuede Osemwota[1], Elodie Laine[1], Anne-Geneviève Marcellin[3], Vincent Calvez[3], Jean-François Mouscadet[1] and Luba Tchertanov[1$]

[1]LBPA, CNRS, ENS de Cachan, 94235 Cachan, France;

[2]School of Life Science, East China Normal University, 200062 Shanghai, China;

[3]Department of Virology, AP-HP, Hôpital Pitié-Salpêtrière, EA 2387, UPMC Univ Paris 06, Paris, France

Abstract

Background- Most antiretroviral medical treatments were developed and tested principally on HIV-1 B non–recombinant strain, which represents less than 10% of the worldwide HIV-1 infected population. HIV-1 circulating recombinant form CRF02_AG is prevalent in West Africa and is becoming more frequent in other countries. Previous studies suggested that the HIV-1 polymorphisms might be associated to variable susceptibility to antiretrovirals. This study is pointed to compare the susceptibility to integrase inhibitors of HIV-1 subtype CRF02-AG integrase (IN) with HIV-1 B IN.

Methods- Structural models of HIV-1 B and CRF02_AG INs as non-bonded enzymes and in complex with DNA substrate were built by homology modelling. IN strand transfer inhibitors (INSTIs) (Raltegravir-RAL, Elvitegravir-ELV and L-731, 988) were docked onto the different models and their binding affinity for both HIV-1 B and CRF02_AG INs were compared. Moreover, HIV-1 B IN and CRF02_AG IN from plasma of INSTI-naïve infected patients were cloned and expressed. *In vitro* activity and INSTI susceptibility of CRF02_AG IN were similar to the ones of the HIV-1 B IN.

Results- Our *in silico* and *in vitro* studies showed that the sequence variations between the integrases of CRF02_AG and B strains did not lead to an indicative difference in the structural features of the enzyme and did not impact the susceptibility to the integrase inhibitors. The binding modes and affinity of INSTIs to HIV-1 B and CRF02_AG INs were found to be similar. DNA viral end plays a more important role in the target-binding of RAL than ELV and L731, 988.

Conclusions- Although previous studies suggested that several naturally occurring variations of CRF02_AG IN might alter either IN•vDNA interactions or INSTIs binding, our study demonstrates that these variations affect neither IN activity nor its susceptibility to INSTIs.

---

* These authors contributed equally in this work.

$ Corresponding author: Dr Luba Tchertanov; E-mail: Luba.Tchertanov@lbpa.ens-cachan.fr

Introduction

The *pol*-encoded HIV-1 integrase (IN) is a key enzyme in the replication of retroviruses. It catalyses the covalent insertion of viral cDNA (vDNA) into the chromosomes of the infected cells (Brown, 1990). Two reactions are required for the integration of viral DNA. First, IN binds to a short sequence located at either end of the long terminal repeat (LTR) of vDNA and catalyze an endonucleotide cleavage, 3'-processing reaction, resulting in the removing of two nucleotides from each of the 3'-ends of LTR and delivering of hydroxy groups for nucleophilic attacks. The trimmed DNA is then used as a substrate for strand transfer (ST) reaction, leading to the covalent insertion of the DNA into the host genome (Brown, 1990). IN strand transfer inhibitors (INSTIs) constitute a novel family of antiretroviral drugs, with Raltegravir (RAL) at the cape, which is the first approved INTI for AIDS treatment. Other inhibitors in advanced phase of development are Elvitegravir (ELV) and GSK572.

Human immunodeficiency virus type 1 (HIV-1) exhibits an exceptional level of genetic variability, which may influence the viral properties such as infectivity, transmissibility or response to antiviral treatment (Rhee *et al.*, 2008). The most prevalent HIV-1 group M genetic forms are subtypes A, B, C and circulating recombinant form CRF02_AG. Enzymatic and virological data support the concept that naturally occurring polymorphisms in different non-B subtypes can affect the susceptibility of HIV-1 to different antiretroviral drugs, the magnitude of resistance conferred by major mutations, and the propensity to acquire some resistance mutations (Martinez-Cajas *et al.*, 2008). The genetic variation between viral isolates retroviral enzymes shows 25-35%; particularly the *pol* gene exhibits high variation, about 10-15 % for reverse transcriptase (RT) and 8-12% for IN (Taylor *et al.*, 2008). IN inhibitors are active *in vivo* against B and non B subtypes. Furthermore, *in vitro* studies suggested that subtype C integrase is equally susceptible to INSTIs (Bar-Magen *et al.*, 2009). Similarly, the analysis of *pol* gene in infected patients showed that highly prevalent polymorphisms have little effect on INSTIs susceptibility (Low *et al.*, 2009). Nevertheless, comparison of IN sequences of B and CRF02_AG strains showed that CRF02-AG consensus differed from the B sequence by 13 residues (K/R14, V/I31, L/I101, T/V112, T/A124, T/A125, G/N134, I/V135, K/T136, V/I201, T/S206, L/I234 and S/G283) (Malet et al., 2008d). According to a model of the HIV-1 B IN/DNA complex (Fenollar-Ferrer *et al.*, 2008), it was suggested that several of these variations K/R14, T/V112, T/A125, G/N134, K/T136, and T/S206 may impact IN interaction with DNA or IN susceptibility to INSTIs. Later we compared genetic barrier between subtype B and CRF02-AG strains; we found that the variability between subtypes impacted the genetic

barrier for G140C/S and V151I with a higher genetic barrier being calculated for subtype CRF02-AG suggesting a great difficulty in selecting these mutations for CR02-AG compared to subtype B (Maiga et al., 2009).

HIV-1 IN is a 288 amino acids enzyme, which consists in three structurally distinct functional domains (Asante-Appiah & Skalka, 1997). Structures reported that HIV-1 IN single or two-domains data allow the generation of biologically relevant models, representing either non- bonded dimeric enzyme or IN complexes with vDNA or/and host DNA (hDNA) (Mouscadet *et al.,* 2010). The X-ray structure of full-length prototype foamy virus (PFV) IN complex with its cognate DNA and INSTIs, (RAL, ELV and the other first- and second-generation INSTIs) was recently reported (Hare *et al.*, 2010a, Hare *et al.*, 2010b). This study used the reported structures for homology modeling of the non-bonded IN and IN in complex with vDNA to compare *in silico* conformations of subtype CRF02_AG and B INs. Further, the constructed models were used for study of the susceptibility of both INs to strand transfer inhibitors, RAL, ELV and L731, 988 (Chart 1). Modeling results were compared with experimental data obtained with B and CRF02_AG INs which were isolated from plasma samples of HIV-1 infected patients and then cloned and expressed *in vitro*.

Materials and methods

Molecular modeling

All calculations were carried out on a Linux station (4x2 cores) running Centos 5.4. The IN models were constructed using Modeller package 9V8 (Eswar *et al.*, 2003). The sequence alignment was performed using CrustalW server (Larkin *et al.*, 2007, Thompson *et al.*, 1994) (http://www.ebi.ac.uk/Tools/clustalw2/index.html). The docking of ST inhibitors, RAL, ELV and L731_988 (Chart 1) onto the IN Models **1** – **4** were performed using two algorithms, GLIDE (Friesner *et al.*, 2004) incorporated in the Schrödinger suite (Schrödinger Inc.) and Autodock 4.2 (Morris *et al.*, 2009). Figures were produced with PyMol (DeLano, 2002) and Chimera (Goddard *et al.*, 2005).

Models of the HIV-1 IN from B and CRF02 IN strains

3D models of the full-length IN homo-dimer, IN^{1-270} (non-bonded IN, or *apo* state respectively to DNA) containing one Mg^{2+} cation in each active site were generated by homology modeling from crystallographic structures of isolated pairs of IN domains. Two

structures of the HIV-1 IN, one containing the N-terminal domain (NTD) and the catalytic core domain (CCD) (IN^{1-210}, PDB code: 1K6Y) (Wang *et al.*, 2001) and the other containing the CCD and the C-terminal domain (CTD) (IN^{56-270}, PDB code: 1EX4) (Chen *et al.*, 2000), were chosen as the initial templates. These structures represent multiple mutants of the HIV-1 subtype B IN, the mutations being W131D/F139D/F185K in 1K6Y and C56S/W131D/F139/F185K/C180S in 1EX4. Both structures were superimposed and CCD domain (IN^{56-210}) of 1EX4, determined at lower resolution (2.8 Å) than 1K6Y (2.4 Å), was deleted. The disordered residues 271-288 were also omitted. Sequences of the WT HIV-1 INs from B and CRF02_AG strains, which differ by 13 amino acids (K/R14, V/I31, L/I101, T/V112, T/A124, T/A125, G/N134, I/V135, K/T136, V/I201, T/S206, L/I234 and S/G283), were aligned to the templates sequences using ClustalW. The missing regions, CCD-NTD linker (47-55 aas) and catalytic site loop (140-148 aas), were constructed by an *ab initio* approach with Modeller 9V8, based on Discrete Optimized Protein Energy (DOPE) scoring function (Shen & Sali, 2006). 100 models were generated for each IN, from B and CRF02_AG strains, and those with the lowest energy were retained. Mg^{2+} cation was inserted into the active site (D64, D116 and E152) as reported in structure 1BI4 (Maignan *et al.*, 1998). We shall refer to these models as model **1** (B strain) and model **2** (CRF02_AG strain).

Models of the HIV-1 IN from B and CRF0_AG strains in complex with vDNA

3D models of the IN•vDNA pre-integration complex (*holo* state respectively to DNA) from B and CRF02_AG strains were generated by homology modeling following a two-step procedure. The coordinates of the recently published crystal structure of the PFV IN•vDNA complexes co-crystallized with RAL (PDB code: 3OYA, resolution of 2.65Å), supersedes: 3L2T, resolution of 2.88Å) Hare et al., 2010a; Hare et al., 2010b) was used as template. The sequence alignment of the HIV-1 IN dimer (B strain) and the PFV IN was performed using ClustalW. The sequence identity between these two INs is 22%. Nevertheless, structure-based alignment of INs from the PFV and HIV-1 demonstrates high conservation of key structural elements and consequently, the PFV IN X-Ray structure provides a good template for the HIV-1 IN model generation. In order to increase the quality of our model, the NED domain (residues 1 to 50), only present in PFV IN was removed from the corresponding sequence. Then, the sequences of the structural domains of HIV-1 and PFV INs were aligned separately, taking into account the conservation of the secondary structure. The obtained sequence alignment was used for homology modeling of the HIV-1 intasome. The inter-domains linkers were constructed using the *ab initio* LOOP module in Modeller (Sali & Blundell, 1993). 100 models were generated for each IN, from B and CRF02 strains,

and those with the lowest energy were retained. We shall refer to these models as model **3** (B strain) and model **4** (CRF02_AG strain).

Refinement of models 1-6 and quality check out

Hydrogen atoms were added by the HBUILD facility in CHARMM (Brooks *et al.*, 2009). The resulting models were slightly minimized while constraining carbon-α to remove clashes. The stereochemical quality of the models was assessed with ProTable Procheck (Laskowski *et al.*, 1993), which showed that more than 97% of the residues in all models had dihedral angles in the most favorable and allowed regions of the Ramachandram plot, consistent with high model quality.

Molecular docking

Initial molecular geometry of ELV and RAL were taken from the X-ray structures 30YA (RAL) and 3L2U (ELV) of PFV IN•vDNA complexes (Hare, Gupta, Valkov, Engelman, & Cherepanov, 2010a). The 3D structure of the compound L-731,988 was generated by ChemBioOffice 2010 (Kerwin, 2010). The models of all inhibitors (Chart 1) in deprotonated form were minimized with Density Functional Fheory (DFT) B3LYP 6-31G* method implemented in *Gaussian03* program (Gaussian, Inc.) (Curtiss *et al.*, 1998). Inhibitors RAL, ELV and L731,988 were docked onto models **1-6** using two algorithms, GLIDE (Friesner, Banks, Murphy, Halgren, Klicic, Mainz, Repasky, Knoll, Shelley, Perry, Shaw, Francis, & Shenkin, 2004) and Autodock 4.2 (Morris, Huey, Lindstrom, Sanner, Belew, Goodsell, & Olson, 2009). The receptor is considered as a rigid body while the ligand is treated wholly flexible.

For AutoDock technique, the graphical user interface (GUI) of AutoDock 4.2 was used for preparation of ligand and receptor files. Grid maps of interaction energies for various atom types was carried out with a grid box of dimension 25×25×25 $Å^3$ centered on the active site. Calculations were performed with a population size of 150, number of energy evaluations of $5x10^6$, maximum number of generations of 27,000 and crossover rate of 0.02 and 0.8 respectively. The number of runs was set to 100 to explore a large number of poses of the highest affinity and the Solis and Wets algorithm was used to relax the best 10 % of the obtained conformations. The active site coordinates were explicitly defined in the input file.

For GLIDE 4.5, coordinates were taken from models **1-6**. Model structures were prepared using Maestro protein preparation wizard (addition of hydrogens, bond order assignment) and the hydrogen bonds network was optimized with the *protassign* module. Receptor grids were generated by Glide within an enclosing box of size 20 Å centered on the active site.

Inhibitors were docked flexibly to these pre-computing grids using standard precision (SP) scoring. For each compounds, the best-scored pose was saved and analyzed.

Cloning of IN gene

IN cDNA derived from naïve HIV-1 subtype CRF02-AG infected patients. Plasmid pET15b- HIV-1subtype B IN (HBX2) was our lab's conservation (Leh et al., 2000). Amplification of IN coding sequence was carried out with specific primers at 94°C for 10 min, then 28 repeat cycles (94°C for 30 s, 55°C for 45 s, and 72°C for 1 min) followed by incubation at 72°C for 10 min. PCR products corresponding to the entire IN sequences were purified and ligated into pGEM-T Easy vector (Promega), and sequenced (Eurofins MWG operon). Then IN gene was inserted into expression vector pET-15b (Novagen) after digested with Nde I and BamH I, and verified by sequencing. Forward primer: 5'-CATATGTTTTTAGATGGCATAGATAAAGCC-3'; backward primer for CRF02-AG 33CR, 49CR: 5' GATCCTAATCCTCATCCTGTCTACCTGC-3'; backward primer for CRF02-AG 52CR Q148K: 5'-GATCCTAATCCTCATCCTGTCCACTTGC-3'; backward primer for CRF02-AG 68CR: 5'-GGATCCTAATCTTCATCCTGTCTACTTGC-3'.

Expression and purification of recombinant INs

His-tagged INs were produced in Escherichia coli BL21-CodonPlus (DE3)-RIPL (Agilent) and purified under non-denaturing conditions as previously described (Leh et al., 2000; Malet et al., 2008a). Protein production was induced at an OD_{600} of 0.6 to 0.8, by adding isopropyl β-D-1-thiogalactopyranoside (IPTG) to a concentration of 0.5 mM. Cultures were incubated for 3 h at 30°C and then centrifuged 20 min at 1100 g, 4°C. Cells were resuspended in buffer A (50 mM Tris-HCl (pH 8), 1 M NaCl, 4 mM β-mercaptoethanol) and lysed by passage through a French press. The lysate was centrifuged (30 min at 12 000 g, 4°C), and the supernatant was filtered (pore size 0.45 µM) and incubated with nickel-nitrilotriacetic acid agarose beads (Qiagen, Venlo, The Netherlands) for at least 2 hours at 4°C. The beads were washed with buffer A and then with buffer A supplemented with 80 mM imidazole. His-tagged proteins were then eluted from the beads in buffer A supplemented with 1 M imidazole and 50 µM zinc sulfate. They were then dialyzed overnight against 20 mM Tris-HCl (pH 8), 1 M NaCl, 4 mM β-mercaptoethanol and 10% glycerol. Aliquots of the purification products were rapidly frozen and stored at -80°C.

Steady-state fluorescence anisotropy-based assay

Steady-state fluorescence anisotropy values were recorded on a Beacon 2000 Instrument (Panvera, Madison, WI), in a cell maintained at 25°C or 37°C under thermostatic control. The principle underlying the anisotropy-based assay was published elsewhere (Agapkina et al., 2006; Smolov et al., 2006a). DNA-binding assay was carried out at 25°C for 20 minutes in a buffer containing 10 mM HEPES pH 6.8, 1mM dithiothreitol, and 7.5 mM magnesium chloride in the presence of 12.5 nM double stranded DNA substrate (21-mer oligodeoxynucleotide mimicking the U5 viral DNA end, fluorescein-labeled at the 3'-terminal GT) and 100, 150, 200 and 250 nM recombinant IN, respectively. In kinetic study, steady-state fluorescence anisotropy-based 3'-processing activity assay was performed in the presence of 50, 100, 200 and 250 nM recombinant IN proteins and 12.5 nM double stranded fluorescein-labeled DNA substrate, at 37°C for 10, 20, 30, 60, 90, 120 and 180 min.

IN 3'-processing and strand transfer activity assay

In vitro 3'-processing and strand transfer activity assay were carried out using the 21/21-mer or 21/19-mer double stranded oligodeoxynucleotides marked with [γ-^{32}P] ATP respectively, as previously described (Malet et al., 2008a). 3'-processing and strand transfer activities assays were carried out for 3 hours at 37°C in a buffer containing 10 mM HEPES pH 6.8, 1mM dithiothreitol, and 7.5 mM magnesium chloride in the presence of 12.5 nM double stranded DNA substrate and 100 nM recombinant IN. The kinetic study was carried out by testing in vitro 3'-processing activity in the presence of 50 nM, 100 nM, 150 nM and 200 nM recombinant IN proteins, at 37°C for 10 min, 20 min, 30 min, 60 min, 90 min, 120 min and 180 min, respectively.

Susceptibility to INSTIs

Susceptibility of INs to INSTI was determined by testing *in vitro* strand transfer activity in the presence of increasing concentration of INSTI in DMSO. Inhibition of the drug was expressed as a fractional product (percentage of the activity of the control without drug). The 50% inhibitory concentration (IC_{50}), defined as the concentration of drug that results in 50% inhibition, was calculated from inhibition curves fitted to experimental data with Prism software, version 5.0 (GraphPad Software, Inc., San Diego, CA).

Results

Analysis of CRF02-AG IN sequences from INSTI-naïve infected patients

The complete sequence of IN coding region of the *pol* gene was amplified and cloned from the plasma samples of four INSTI-naïve CRF02_AG infected patients. The four IN sequences, N^1 to N^4, harbored several variations of amino acid residues among the 13 residues that were shown to be subjected to the polymorphic substitutions between subtype CRF02_AG and B consensus (K/R14, V/I31, L/I101, T/V112, T/A124, T/A125, G/N134, I/V135, K/T136, V/I201, T/S206, L/I234 and S/G283) (**Table 1**) (Malet *et al.*, 2008). Sequence N_1 displayed the seven variations – K14R, T112V, T125A, G134N, K136T, T206S, S283G; N_2 the five variations – T112V, T125A, G134N, K136T, S283G; N_3 the five variations – K14R, T112V, T125A, K136T, T206S; and N_4 the two variations – T125A and T206S, along with the INSTI resistance mutation Q148K. Though Q148K is involved in INSTIs resistance, the patient from whom the IN coding DNA derived was not exposed to the INSTI containing-treatment, thereby indicating that Q148K would be a naturally occurring amino acid substitution.

Compared with HIV-1 subtype B IN, seven important variations are presented at position 14, 112, 125, 134, 136, 206 and 283 of CRF02-AG IN coding sequence N_1; five important variations at position 112, 125, 134, 136 and 283 of sequence N_2; five important variations at position 14, 112, 125, 136 and 206 of sequence N_3; sequence N_4 has two important variations at position 125 and 206, and an INSTI resistant mutation Q148K, the R112 was not considered as an important variation.

Structural comparison of HIV-1 B and CRF02_AG INs exhibited no significant difference

The four CRF02_AG strains differed from subtype B consensus by two to seven residues among the polymorphisms K14R/T112V/T125A/G134N/K136T/T206S/S283G. In order to determine the potential impact of these variations on the protein activity and its susceptibility to INSTIs, we built models of the consensus B sequence and the corresponding variant (CRF02-AG IN coding sequence N_1) containing the six modifications. The 18-aas C-terminal end containing the S283G was omitted since the structure of this domain was not resolved in X-ray analysis and this part of protein is extremely difficult to predict the folding in the apo state, due its essential length and its highly solvent-exposed position.

The structural models representing two different conformational states of the B and CRF02_AG HIV-1 INs were built by homology modeling. Comparative structural analysis and study of the inhibitors binding was based on 6 generated models (**Figure 1**). Models **1** (subtype B) and **2** (subtype CRF02_AG) (**Figure 1A**) represent a full-length wild type (WT) dimer of IN (IN^{1-270}), which depicts the conformational state of the enzyme just before the

3’-processing of vDNA (apo state); models **3** (subtype B) and **4** (subtype CRF02_AG) (**Figure 1B**) represent the IN•vDNA intasome of the strand transfer complex (holo state); models **5** (subtype B) and **6** (subtype CRF02_AG) (not shown) were derived from models **3** and **4** by removing vDNA.

Models **1** and **2** were constructed from the crystallographic structures of HIV-1 IN isolated domains or pairs of domains. Overall, the analysis of the models representing the HIV-1 IN conformational state before 3'-processing (apo state), did not show any significant structural change between the two subtypes (**Figure 1A**). The catalytic site loop conformation differed slightly between the two models. However, this difference might be ascribed first to the fact that it was constructed by *ab initio* methods since structural data for the loop were missing in the crystallographic template structures, 1K6Y and 1EX4, and second that it can adopt various conformations in solution due to its high solvent-exposition of the loop which renders it very flexible as previously showed (Metifiot et al., 2010; Perryman et al., 2010).

Models **3** and **4** were constructed from the crystallographic structure of the IN•vDNA complex of the PFV intasome (Hare, Gupta, Valkov, Engelman, & Cherepanov, 2010a). Although the sequence identity between HIV-1 and PFV INs is low (22%), the structure-based alignment of the two proteins demonstrates high conservation of key secondary structural elements; consequently, the PFV IN X-Ray structure may be a good template for the HIV-1 IN model generation. Moreover, the structure of the PFV intasome provides an appropriate model as it displays a distance between the reactive 3’ends of vDNA that corresponds to the expected distance between the integration sites of HIV-1 IN target DNA (4 base pairs). To obtain a robust alignment, we adjusted the targets (HIV-1 INs from B and CRF02_AG subtypes) and template (PFV IN) sequences manually, considering each structural domain separately, in order to take into account the conservation of the secondary structure. For both subtype B and CRF02_AG models, distance restraints were applied to reproduce key interactions reported in earlier experimental studies (see Material and Methods). Again, models **3** and **4,** representing the IN•vDNA intasomes of both strains, superimposed perfectly and no structural dissimilarity was observed (Figure 1B). Most of the variations are located far from the active sites and the nearest two mutated residues to the active site, at positions 134 and 136, are exposed to the solvent and apparently did not affect significantly the structure.

It is worthy to note large structural and conformational changes observed between the apo (models **1** and **2**) and holo (**3** and **4**) states regarding the relative positions of the IN domains (RMSD of 31 Å, based on C_α). The structural alternation resulted in different contacts

between IN domains. In models **1** and **2** (apo state), no interaction was detected between CTD and CCD whereas the two domains interact tightly in the models **3** and **4** (holo state). The NTD-CCD interface also exhibits substantial changes: in the apo form the NTD-CCD interface belongs to the same monomer subunit whereas in the holo form the interface is from two different subunits. Moreover, IN undergoes important structural reorganization lead to structural re-organization of the catalytic site loop upon vDNA binding; the coiled portion of the loop reduces from 10 residues (140 -149 aas) in the apo form to 5 residues (140-144 aas) in the holo form. This partial folding of the catalytic loop is stabilized through intra-IN domain-domain interactions and interactions with vDNA which probably, contribute in the helix α4 elongation.

In vitro enzymatic comparison of recombinant HIV-1 B and CRF02_AG INs was consistent with the modeling study

To confirm experimentally the absence of divergence between INs from both CRF02_AG and B strains, N_1 to N_4 sequences were expressed and purified (**Figure 2A**) and their enzymatic activities were compared to the one of HxB2 B IN. First, DNA binding activities of recombinant INs were compared using a steady-state fluorescence anisotropy assay (**Figure 2B**) (Guiot et al., 2006). In this assay, the binding of IN to a fluorophore-labeled double strand oligodeoxynucleotide (dsODN) substrate mimicking the end of viral DNA is monitored by the increase of the steady-state anisotropy value, resulting from the restriction of the substrate movements. As shown on figure 2B, no significant difference in DNA binding activity of recombinant B IN and CRF02_AG INs was observed within a range of IN concentrations of 100 nM to 250 nM, thereby indicating that the variations in IN sequence did not affect the binding affinity of enzyme. Then, 3'- processing activity of subtype B IN and CRF02_AG INs was compared *in vitro*. No significant difference in 3'-processing activity of recombinant HIV-1 B IN and CRF02-AG INs was found within a range of IN concentrations of 50 to 400 nM (**Figure 2C**). Similar result was obtained in strand transfer activity assay of both subtype enzymes (**Figure 2D**). In agreement with the modeling result, *in vitro* study confirmed that the enzymatic activities of both INs were comparable. Impaired 3'-processing and strand transfer activity, but not the DNA binding ability of CRF02_AG N_4 Q148K was observed, in agreement with previous study (Delelis et al., 2009a).

In *in silico* study, slight difference was observed in catalytic site loop of both subtypes (**Figure 1**). The difference in the conformation of catalytic hinted to the fact that the enzymes may display difference in 3'-processin kinetics. To address this possibility,

3'-processing kinetics of recombinant HIV-1 B IN and CRF02-AG N_1 IN was assayed in the presence of increasing IN concentrations of 50 nM to 200 nM with an augmenting incubation time, using *in vitro* 3'-processing activity assay (**Figure 3**). Again, no difference could be detected. This result was further confirmed by steady-state anisotropy assay (data not shown).

The Docking of INSTIs

Although B and CRF02_AG INs are structurally similar, residue variations may impact the interaction and subsequent activity of the inhibitors. To address this hypothesis, the three inhibitors RAL, ELV and L-731,688 (**Chart 1**) were docked onto INs by using two different docking algorithms, Glide and Autodock. Models RAL and ELV coordinates were taken from the crystallographic structures of PFV intasome co-complexes (Hare et al., 2010a), L-731,988 was built from scratch (see Materials and Methods). The three compounds were considered in their deprotonated form, as it has been clearly established that DKAs mainly exist in this form in solution (Maurin *et al.*, 2004). The binding energies obtained by Glide and Autodock scoring functions are reported in **Table 2.**

The inhibitors were first docked onto the non-bonded to vDNA, models **1** and **2** with single Mg^{2+} ion within the catalytic site (apo state). For subtype B, values of binding energies obtained with Glide range from -7.60 to -5.46 kcal/mol while those obtained with Autodock range from -8.68 to -7.35 kcal/mol. Similar docking results are obtained for the CRF02_AG subtype with binding energy values from -5.80 to -5.10 kcal/mol for Glide and from -8.69 to -7.69 kcal/mol for Autodock. Considering the problem of direct comparison of the two software scoring functions, no significant difference can be assessed between strains B and CRF02_AG.

All three inhibitors are positioned close to the Mg^{2+} ion, far from the catalytic loop (data not shown). RAL binding is also stabilized by H-bonds with the side chains of T66 and N155 while ELV and L731-988 forms H-bonds with C65. In the absence of the second Mg^{2+} cation and the DNA substrate, the flexible RAL molecule is not capable to be fixed in the active site and it docking poses strongly differ from the observed X-ray conformation (Hare, Vos, Clayton, Thuring, Cummings, & Cherepanov, 2010b). The conformations of ELV and L-731,988 are more conserved between the docking onto the model **1** and **2** due to interaction of the high negative charges on the carboxylate group of ELV and L731, 988 with the Mg^{2+} ion. Docking results (Autodock) shows that the fluorobenzyl moiety of RAL establishes hydrophobic interactions within the crevasse formed by C65, E92, N120 and

F121 residues whereas the 1,3,4-oxadiazol cycle interacts with another pocket formed by Q62, D116, I141, P142 and I151 residues in both INs, from B and CRF02_AG strains. In the best docking poses of ELV and L731, 988, the halogenated aromatic moieties (chlorofluorobenzyl of ELV and the fluorobenzyl of L731, 988) establish hydrophobic interactions with C65, H67, E92, E118, N120 and F121 residues.

The inhibitors were then docked onto models **3** and **4** representing IN•2Mg^{2+}•DNA complexes (**Figure 4**). Docking resulted in a favorable binding for the three inhibitors with significantly better scores (from -11.79 to -10.69 kcal/mol for Glide and from -14.35 to -10.82 kcal/mol for Autodock) than those obtained for the apo IN, in agreement with the mechanism of action of these molecules, the strand transfer inhibitors (Kawasuji *et al.*, 2006). The predicted score rankings (binding energy values) and the poses of the three inhibitors are the same for INs from B and CRF02_AG strains. ELV is found by both programs to have the best affinity, thus confirming the *in vitro* data which pointed to ELV as the most efficient INSTI. Furthermore, the best poses of RAL and ELV predicted by Glide and Autodock are identical to the ones observed in the X-ray structures of IN•2Mg^{2+}•DNA complex from PFV co-crystallized with the RAL or ELV, thus validating the docking methods.

The three compounds are positioned at the catalytic site and bond to the Mg^{2+} ions as it was observed in the X-ray structure of the PFV intasome complex (Hare, Gupta, Valkov, Engelman, & Cherepanov, 2010a, Hare, Vos, Clayton, Thuring, Cummings, & Cherepanov, 2010b). RAL interacts twice with the metal cofactors of the active site, through the alcoholic oxygen atom with the one Mg^{2+} ion and with the second with the keto group of pyrimidine cycle. The large distance from the alcoholic oxygen atom to the second Mg^{2+} ion is not satisfying to consider the alcoholic oxygen as a bridge atom between two Mg^{2+} ions. The distances between acyclic carboxamide oxygen atom and the metal cofactors are larger than the coordination bond length. ELV is coordinated by Mg cofactors by three bonds: the β-ketoenolate chelates the first Mg^{2+} ion as it was detailed in (Tchertanov & Mouscadet, 2007) and the carboxylic oxygen atom coordinates the second Mg^{2+} ion. L-731, 988 molecule binds only one Mg^{2+} ion through the DKA moiety oxygen atoms. The stabilizing non-covalent interactions with Ade17 of the vDNA and the strong coordination bonding of the Mg^{2+} ions with two oxygen atoms of the pyrimidine cycle are presumably the causes for RAL location at the catalytic site. The docking results shows that RAL forms hydrogen bonds with Y143, N144, N117 and D116 and multiple hydrophobic contacts within the pocket formed by D116, N117, G140, I141, N144, S147 and Q148 residues. Particularly,

RAL participates in hydrogen bonding with the Y143 backbone amine and in π-π interactions involving the tyrosine cycle and 1,3,4-oxadiazole, while its fluorobenzyl fragment stacks with pyrimidine (Cyt16) of the vDNA which affect also RAL affinity to the IN•vDNA complex. Such stacking is also observed for ELV and L731. All three compounds are thus oriented as if they were using DNA 3' nucleotide bases AC (Ade-Cyt) as a support for their interaction with IN.

RAL conformation obtained by computing slightly differs from those observed in X-ray structure. The conformational change, a turn of the carbamoyl moiety around its C-N bond stabilized by H-bond between E152 and the NH group of the carbamoyl moiety, allows to minimize steric constrains and electrostatic repulsions at the active site loaded by metal cofactors, DNA substrate and inhibitor. For the other inhibitors, ELV and L-731,988, predicted poses correspond well to those observed by the X-ray analysis (Hare et al., 2010d).

The docking calculations evidenced that (i) the IN•DNA complex represents the best target for the studied inhibitors and (ii) the presence of the vDNA is required to get the creditable docking results (scores and poses) characterizing the inhibitors affinity. To explore further these findings, vDNA was removed from the IN•vDNA complex and inhibitors were docked again on non-bonded IN with a fold corresponding to holo state. The binding energies of the three inhibitors to the INs were differently affected by the DNA removal ; particularly, scores get worse, the value increase being especially striking for RAL (3.3 kcal/mol and 2.3 kcal/mol with Glide and Autodock respectively) and ELV (2.8 kcal/mol with Glide, 1.5 kcal/mol with Autodock).

The docking results obtained by Glide and Autodock are only slightly divergent. For the Glide, Ral best predicted poses in models **5** and **6** display completely flipped orientations compared to the best poses predicted in models **3** and **4**, while the center of mass of the ligand remains in the same position. The fluorobenzyl ring, which was stacking upon AC vDNA bases in models **3** and **4**, is now exposed to the solvent while the methyl group attached to the 1, 3, 4-oxadiazol cycle points toward the removed vDNA. As a result, the Mg^{2+} ions are no properly chelated anymore. By contrast to RAL, the position of ELV and L731 are flipped, while their carboxylic and β-ketoenol groups are remain in place, through which the inhibitors still coordinated by the Mg^{2+} ions. The halogenated benzyl rings of ELV and L-731,988 which were stacking with the AC vDNA bases in models **3** and **4**, position themselves upon the cycle of Y143 from the catalytic loop in models **5** and **6**. Such position enables L-731,988 to form a favorable interaction with the side chain of N117, counterbalancing the loss of DNA-mediated stacking. This explains why L-731,988 is

subjected to a non-significant score increase (less than 1.2 kcal/mol) and becomes the best-ranked ligand upon DNA removal.

The Autodock results also show that vDNA removal has a drastic effect on the binding of RAL and the preferred orientation in the active site. Very similar to the observation with Glide, in the absence of vDNA the fluoro-benzyl moiety is placed in the position of 1, 3, 4-oxadiazol cycle when the DNA is present. The 1, 3, 4-oxadiazol cycle is now anchored in another crevasse formed by C65, E92, G118, N120 and F121. The β-ketoenol motif is now solvent exposed and RAL only bonds one Mg^{2+} ion by the oxygen of the carboxamide group. Interestingly, the ranked 4^{th} pose is very similar to the best pose in presence of the vDNA but it is of less affinity (-7.99 kcal/mol). The best poses of RAL onto models **5** and **6** are quite similar. The other two inhibitors, ELV and L-731,988 are less impacted upon removing the vDNA than RAL. Both inhibitors are less voluminous than RAL and they can adopt versatile conformations in the active site with a small energy difference (1.5 and 0.33 kcal/mol for ELV for L-731,988, respectively).

***In vitro* comparison of the susceptibility of HIV-1 B IN and CRF02-AG INs to three INSTIs confirmed the absence of difference in susceptibility**

To compare the susceptibility of subtypes B and CRF02-AG 33CR INs, the efficiency of INSTIs RAL, ELV and L-731,988 on recombinant INs proteins were determined with *in vitro* strand transfer assay in the presence of increasing concentration of INSTI (see Materials and Methods). As to all the three studied INSTIs, no significant difference in IC_{50} values against recombinant HIV-1 INs from B and CRF02-AG strains was observed (**Table 3**). IC_{50} of RAL, ELV and L-731, 988 against HIV-1 INs from B and CRF02-AG strains are 41.8 nM, 93.4 nM, 855 nM and 13.7 nM, 55.6 nM, 21.2 nM, respectively. The result of in vitro susceptibility assay is consistent with the Docking of INSTIs in modeling study.

Discussion

Molecular modeling approaches were undertaken to investigate the potential effect of the natural variations showed in CRF02-AG strain on the *in vitro* activities of the enzyme and its susceptibility to INSTIs as compared to the ones of the consensus B IN. We found that the structural models of non-bond (apo state) and viral DNA bound (holo state) IN showed very similar structures for the studied strains. A slight difference between apo INs from B and CRF02_AG strains that was detected in the catalytic loop conformation did not impact the

IN activity. The structural models of IN•vDNA complex superimposed perfectly. This similarity was confirmed by comparable strand transfer activity. Consequently, the naturally occurring variations in the HIV-1 IN subtype CRF02-AG, e.g., K14R, T112V, T125A, G134N, K136T, T206S and S283G, which were suggested to modify IN structure do not affect significantly *in vitro* DNA binding activity, neither 3'-processing nor strand transfer reaction. Furthermore, docking results reveal that the modes of binding and docking conformations of three studied inhibitors are identical for B and CRF02_AG strains and the INSTIs possessed similar IN inhibitory activity against B and CRF02-AG HIV-1 strains. Altogether these results confirm the absence of difference in susceptibility previously reported for subtype B and C HIV-1 INs (Bar-Magen T et al., 2009). Thus, in contrast to the lower baseline susceptibilities of recombinant A/G subtype virus to protease inhibitors (PIs) and reduced susceptibility of some A/G isolates to abacavir, INSTIs potentially provide an excellent therapeutic options for the treatment of HIV-1 subtype CRF02-AG infected patients (Martínez-Cajas JL et al., 2008).

In the targets all the three molecules are positioned similarly with diketo acid moiety in orientation encouraged interactions with the two metal cofactors in the active site. Furthermore, independently of the method, the three INSTIs displayed a more favorable binding onto the holo state than on the apo one, in good agreement with their mechanism of action (Kawasuji, Fuji, Yoshinaga, Sato, Fujiwara, & Kiyama, 2006). The observed transition from apo to the holo state of IN led to a distinguished change in the structure and conformation of the catalytic site loop allowing the binding of ELV and L-731,988. Such binding was not altered by the removal of the viral DNA. Conversely, removing vDNA had a significant effect on the docking results RAL, thereby highlighting the role of vDNA for RAL recognition most likely due to the halogenated benzyl moiety that displaces the unpaired 5'adenine and stacks with the Cyt16 through π-π interactions. Although such interaction is thought to be involved in all the IN strand transfer inhibitors examined (Hare et al., 2010a), our results suggest that ELV and L-731,988 binding determinants differed in part from the ones of RAL.

It should be noted that slight differences were observed between the results obtained with Glide and Autodock scores, which can be ascribed to the impact of electrostatic interactions in the studied molecular systems. Indeed Glide uses higher negative charge localized on the two oxygen atoms attached to the pyrymidine cycle of RAL than Autodock (-1.22 and -0.5 vs. -0.183 and -0.265) measurements unit. Also, within Autodock scoring function, the carboxylate charges used for ELV (2x -0.64) and L-731,988 (2x -0.62) are more than two

oxygen atoms attached to the pyrimidine cycle of RAL. To verify this hypothesis, we repeated the docking calculations of ELV and L-731,988 using the founded charges of two oxygen atoms attached to the pyrimidine ring of RAL instead of those assigned Gasteiger charges. The new binding energies of both inhibitors increased from -14.35 and -8.65 to -12.73 and -7.91 kcal/mol for ELV and L-731,988 respectively. Since these atomic charges contribute highly in the binding energy as the atoms coordinate Mg^{2+} ions, they are likely responsible for the discrepancies found between the theoretical binding energies and the experimental IC_{50} values. The experimental ranking of the three inhibitors based on IC_{50} is, RAL > ELV > L-731,988 while it is ELV > L-731,988 > RAL obtained by docking with Autodock and ELV > RAL > L-731,988 with Glide. The high negative charges of the carboxylate oxygen atoms of ELV and L-731,988 might be the obstacle to have effective inhibitory actions on IN, as these charges increase the desolvation free energy and so increase the binding penalty for these inhibitors.

Conclusions

The naturally occurring variations in HIV-1 subtype CRF02-AG IN, such as K14R, T112V, T125A, G134N, K136T, T206S and S283G neither affect indicatively IN structure, nor *in vitro* catalytic activity. Docking results of all the studied inhibitors into the non-bonded IN model show considerably low scores, compared with docking into the preintegration IN•DNA complex. Further, the generated structure of IN•DNA complex is the appropriate model used to explain the inhibition mechanism of the strand transfer inhibitors which is correlated with the earlier studies. All the three studied molecules are polydendate ligands able to wrap around the metal ions in the active site. These results are in perfect agreement with the proposed mechanism of action for INSTIs. Docking results reveal that the modes of binding and docking conformations of the three studied molecules are identical for the HIV-1 IN from B and CRF02_AG strains. The proposed mechanism of the IN inhibition based on considering of different conformational states, non-bonded IN and IN•vDNA complex holds for the two studied strains.

Acknowledgements

This work was funded by the Ecole Normale Supériere de Cachan, the Centre National de la Recherche Scientifique (CNRS) and SIDACTION.

The authors have declared no competing interests.

References

Agapkina, J., Smolov, M., Barbe, S., Zubin, E., Zatsepin, T., Deprez, E., Bret, M.L., Mouscadet, J.-F., and Gottikh, M. (2006). Probing of HIV-1 Integrase/DNA interactions using novel analogs of viral DNA. Journal of biological chemistry *281* 11530-11540.

Asante-Appiah, E. & Skalka, A. M. (1997). Molecular mechanisms in retrovirus DNA integration. *Antiviral Research 36*, 139-156.

Bar-Magen, T., Sloan, R. D., Faltenbacher, V. H., Donahue, D. A., Kuhl, B. D., Oliveira, M., Xu, H. T., & Wainberg, M. A. (2009). Comparative biochemical analysis of HIV-1 subtype B and C integrase enzymes. *Retrovirology 6*, doi:ARTN 103;DOI 10.1186/1742-4690-6-103.

Brooks, B. R., Brooks, C. L., Mackerell, A. D., Nilsson, L., Petrella, R. J., Roux, B., Won, Y., Archontis, G., Bartels, C., Boresch, S., Caflisch, A., Caves, L., Cui, Q., Dinner, A. R., Feig, M., Fischer, S., Gao, J., Hodoscek, M., Im, W., Kuczera, K., Lazaridis, T., Ma, J., Ovchinnikov, V., Paci, E., Pastor, R. W., Post, C. B., Pu, J. Z., Schaefer, M., Tidor, B., Venable, R. M., Woodcock, H. L., Wu, X., Yang, W., York, D. M., & Karplus, M. (2009). CHARMM: The Biomolecular Simulation Program. *Journal of Computational Chemistry 30*, 1545-1614, doi:DOI 10.1002/jcc.21287.

Brown, P.O. (1990). Integration of retroviral DNA. CurrTopMicrobiolImmunol *157:19-48*, 19-48.
Delelis, O., Malet, I., Na, L., Tchertanov, L., Calvez, V., Marcelin, A.-G., Subra, F., Deprez, E., and Mouscadet, J.-F. (2009). The G140S mutation in HIV integrases from raltegravir-resistant patients rescues catalytic defect due to the resistance Q148H mutation. Nucleic Acids Research, 1–9.

Chen, J. C. H., Krucinski, J., Miercke, L. J. W., Finer-Moore, J. S., Tang, A. H., Leavitt, A. D., & Stroud, R. M. (2000). Crystal structure of the HIV-1 integrase catalytic core and C-terminal domains: A model for viral DNA binding. *Proceedings of the National Academy of Sciences of the United States of America 97*, 8233-8238.

Curtiss, L. A., Raghavachari, K., Redfern, P. C., Rassolov, V., & Pople, J. A. (1998). Gaussian-3 (G3) theory for molecules containing first and second-row atoms. *Journal of Chemical Physics 109*, 7764-7776.

Dayam, R. & Neamati, N. (2004). Active site binding modes of the beta-diketoacids: a multi-active site approach in HIV-1 integrase inhibitor design. *Bioorganic & Medicinal Chemistry 12*, 6371-6381, doi:DOI 10.1016/j.bmc.2004.09.035.

Eswar, N., John, B., Mirkovic, N., Fiser, A., Ilyin, V. A., Pieper, U., Stuart, A. C., Marti-Renom, M. A., Madhusudhan, M. S., Yerkovich, B., & Sali, A. (2003). Tools for comparative protein structure modeling and analysis. *Nucleic Acids Research 31*, 3375-3380, doi:DOI 10.1093/nar/gkg543.

Fenollar-Ferrer, C., Carnevale, V., Raugei, S., & Carloni, P. (2008). HIV-1 integrase-DNA interactions investigated by molecular modelling. *Computational and Mathematical Methods in Medicine 9*, 231-243, doi:DOI 10.1080/17486700802167918.

Friesner, R. A., Banks, J. L., Murphy, R. B., Halgren, T. A., Klicic, J. J., Mainz, D. T., Repasky, M. P., Knoll, E. H., Shelley, M., Perry, J. K., Shaw, D. E., Francis, P., & Shenkin, P. S. (2004). Glide: A new approach for rapid, accurate docking and scoring. 1. Method and assessment of docking accuracy. *Journal of Medicinal Chemistry 47*, 1739-1749, doi:DOI 10.1021/jm0306430.

Goddard, T. D., Huang, C. C., & Ferrin, T. E. (2005). Software extensions to UCSF Chimera for interactive visualization of large molecular assemblies. *Structure 13*, 473-482, doi:DOI 10.1016/j.str.2005.01.006.

Grobler, J. A., Stillmock, K., Hu, B. H., Witmer, M., Felock, P., Espeseth, A. S., Wolfe, A., Egbertson, M., Bourgeois, M., Melamed, J., Wai, J. S., Young, S., Vacca, J., & Hazuda, D. J. (2002). Diketo acid inhibitor mechanism and HIV-1 integrase: Implications for metal binding in the active site of phosphotransferase enzymes. *Proceedings of the National Academy of Sciences of the United States of America 99*, 6661-6666, doi:DOI 10.1073/pnas.092056199.

Hare, S., Gupta, S. S., Valkov, E., Engelman, A., & Cherepanov, P. (2010a). Retroviral intasome assembly and inhibition of DNA strand transfer. *Nature 464*, 232-236, doi:DOI 10.1038/nature08784.

Hare, S., Vos, A. M., Clayton, R. F., Thuring, J. W., Cummings, M. D., & Cherepanov, P. (2010b). Molecular mechanisms of retroviral integrase inhibition and the evolution of viral resistance. *Proceedings of the National Academy of Sciences of the United States of America 107*, 20057-20062, doi:DOI 10.1073/pnas.1010246107.

Kawasuji, T., Fuji, M., Yoshinaga, T., Sato, A., Fujiwara, T., & Kiyama, R. (2006). A platform for designing HIV integrase inhibitors. Part 2: A two-metal binding model as a potential mechanism of HIV integrase inhibitors. *Bioorganic & Medicinal Chemistry 14*, 8420-8429, doi:DOI 10.1016/j.bmc.2006.08.043.

Kerwin, S. M. (2010). ChemBioOffice Ultra 2010 Suite. *Journal of the American Chemical Society 132*, 2466-2467, doi:DOI 10.1021/ja1005306.

Leh, H., Brodin, P., Bischerour, J., Deprez, E., Tauc, P., Brochon, J.-C., LeCam, E., Coulaud, D., Auclair, C., and Mouscadet, J.-F. (2000). Determinants of Mg2+-dependent activities of recombinant human immunodeficiency virus type 1 integrase. Biochemistry *39*, 9285-9294.

Larkin, M. A., Blackshields, G., Brown, N. P., Chenna, R., McGettigan, P. A., McWilliam, H., Valentin, F., Wallace, I. M., Wilm, A., Lopez, R., Thompson, J. D., Gibson, T. J., & Higgins, D. G. (2007). Clustal W and clustal X version 2.0. *Bioinformatics 23*, 2947-2948, doi:DOI 10.1093/bioinformatics/btm404.

Laskowski, R. A., Macarthur, M. W., Moss, D. S., & Thornton, J. M. (1993). Procheck - A Program to Check the Stereochemical Quality of Protein Structures
1. *Journal of Applied Crystallography 26*, 283-291.

Leh, H., Brodin, P., Bischerour, J., Deprez, E., Tauc, P., Brochon, J.-C., LeCam, E., Coulaud, D., Auclair, C., and Mouscadet, J.-F. (2000). Determinants of Mg2+-dependent activities of recombinant human immunodeficiency virus type 1 integrase. Biochemistry *39*, 9285-9294.

Low, A., Prada, N., Topper, M., Vaida, F., Castor, D., Mohri, H., Hazuda, D., Muesing, M., & Markowitz, M. (2009). Natural Polymorphisms of Human Immunodeficiency Virus Type 1 Integrase and Inherent Susceptibilities to a Panel of Integrase Inhibitors. *Antimicrobial Agents and Chemotherapy 53*, 4275-4282, doi:DOI 10.1128/AAC.00397-09.

Maiga, A.I., Malet, I., Soulie, C., Derache, A., Koita, V., Amellal, B., Tchertanov, L., Delelis, O., Morand-Joubert, L., Mouscadet, J.F.*, et al.* (2009). Genetic barriers for integrase inhibitor drug resistance in HIV type-1 B and CRF02_AG subtypes. Antivir Ther *14*, 123-129.

Maignan, S., Guilloteau, J. P., Zhou-Liu, Q., Clement-Mella, C., & Mikol, V. (1998). Crystal structures of the catalytic domain of HIV-1 integrase free and complexed with its metal cofactor: High level of similarity of the active site with other viral integrases
1. *Journal of Molecular Biology 282*, 359-368.

Malet, I., Delelis, O., Valantin, M.-A., Montes, B., Soulie, C., Wirden, M., Tchertanov, L., Peytavin, G., Reynes, J., Mouscadet, J.-F.*, et al.* (2008a). Mutations Associated with Failure of Raltegravir Treatment Affect Integrase Sensitivity to the Inhibitor In Vitro. Antimicrobial agents and chemotherapy *52*, 1351-1358.

Malet, I., Soulie, C., Tchertanov, L., Derache, A., Amellal, B., Traore, O., Simon, A., Katlama, C., Mouscadet, J. F., Calvez, V., & Marcelin, A. G. (2008). Structural effects of amino acid variations between B and CRF02-AG HIV-1 integrases. *Journal of Medical Virology 80*, 754-761, doi:DOI 10.1002/jmv.21169.

Martinez-Cajas, J. L., Pant-Pai, N., Klein, M. B., & Wainberg, M. A. (2008). Role of Genetic Diversity amongst HIV-1 Non-B Subtypes in Drug Resistance: A Systematic Review of Virologic and Biochemical Evidence. *Aids Reviews 10*, 212-223.

Maurin, C., Bailly, F., Buisine, E., Vezin, H., Mbemba, G., Mouscadet, J. F., & Cotelle, P. (2004). Spectroscopic studies of diketoacids-metal interactions. A probing tool for the pharmacophoric intermetallic distance in the HIV-1 integrase active site. *Journal of Medicinal Chemistry 47*, 5583-5586, doi:DOI 10.1021/jm0408464.

Metifiot, M., Maddali, K., Naumova, A., Zhang, X., Marchand, C., and Pommier, Y. (2010). Biochemical and Pharmacological Analyses of HIV-1 Integrase Flexible Loop Mutants Resistant To Raltegravir. Biochemistry.

Morris, G. M., Huey, R., Lindstrom, W., Sanner, M. F., Belew, R. K., Goodsell, D. S., & Olson, A. J. (2009). AutoDock4 and AutoDockTools4: Automated Docking with Selective Receptor Flexibility. *Journal of Computational Chemistry 30*, 2785-2791, doi:DOI 10.1002/jcc.21256.

Mouscadet, J. F., Delelis, O., Marcelin, A. G., & Tchertanov, L. (2010). Resistance to HIV-1 integrase inhibitors: A structural perspective. *Drug Resist. Updat. 13*, 139-150, doi:S1368-7646(10)00021-X [pii];10.1016/j.drup.2010.05.001 [doi].

Mouscadet, J. F. & Tchertanov, L. (2009). Raltegravir: molecular basis of its mechanism of action. *Eur. J. Med. Res. 14 Suppl 3*, 5-16.

Perryman, A.L., Forli, S., Morris, G.M., Burt, C., Cheng, Y., Palmer, M.J., Whitby, K., McCammon, J.A., Phillips, C., and Olson, A.J. (2010). A dynamic model of HIV integrase inhibition and drug resistance. J Mol Biol *397*, 600-615.

Rhee, S. Y., Liu, T. F., Kiuchi, M., Zioni, R., Gifford, R. J., Holmes, S. P., & Shafer, R. W. (2008). Natural variation of HIV-1 group M integrase: Implications for a new class of antiretroviral inhibitors. *Retrovirology 5*, doi:ARTN 74;DOI 10.1186/1742-4690-5-74.

Sali, A. & Blundell, T. L. (1993). Comparative Protein Modeling by Satisfaction of Spatial Restraints. *Journal of Molecular Biology 234*, 779-815.

Shen, M. Y. & Sali, A. (2006). Statistical potential for assessment and prediction of protein structures. *Protein Science 15*, 2507-2524, doi:DOI 10.1110/ps.062416606.

Smolov, M., Gottikh, M., Tashlitskii, V., Korolev, S., Demidyuk, I., Brochon, J.-C., Mouscadet, J.-F., and Deprez, E. (2006). Kinetic study of the HIV-1 DNA 3'-end processing single-turnover property of integrase. FEBS Journal *273* 1137-1151.

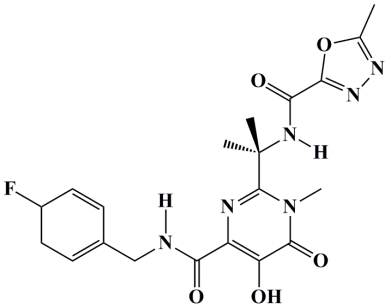
Taylor, B. S., Sobieszczyk, M. E., McCutchan, F. E., & Hammer, S. M. (2008). Medical progress: The challenge of HIV-1 subtype diversity. *New England Journal of Medicine 358*, 1590-1602.

Tchertanov, L. & Mouscadet, J. F. (2007). Target recognition by catechols and beta-ketoenols: Potential contribution of hydrogen bonding and Mn/Mg chelation to HIV-1 integrase inhibition. *Journal of Medicinal Chemistry 50*, 1133-1145, doi:DOI 10.1021/jm061375j.

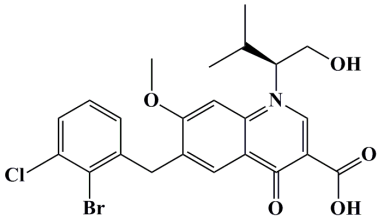
Thompson, J. D., Higgins, D. G., & Gibson, T. J. (1994). Clustal-W - Improving the Sensitivity of Progressive Multiple Sequence Alignment Through Sequence Weighting, Position-Specific Gap Penalties and Weight Matrix Choice. *Nucleic Acids Research 22*, 4673-4680.

Wang, J. Y., Ling, H., Yang, W., & Craigie, R. (2001). Structure of a two-domain fragment of HIV-1 integrase: implications for domain organization in the intact protein. *Embo Journal 20*, 7333-7343.

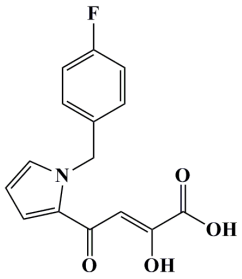
Chart 1



RAL



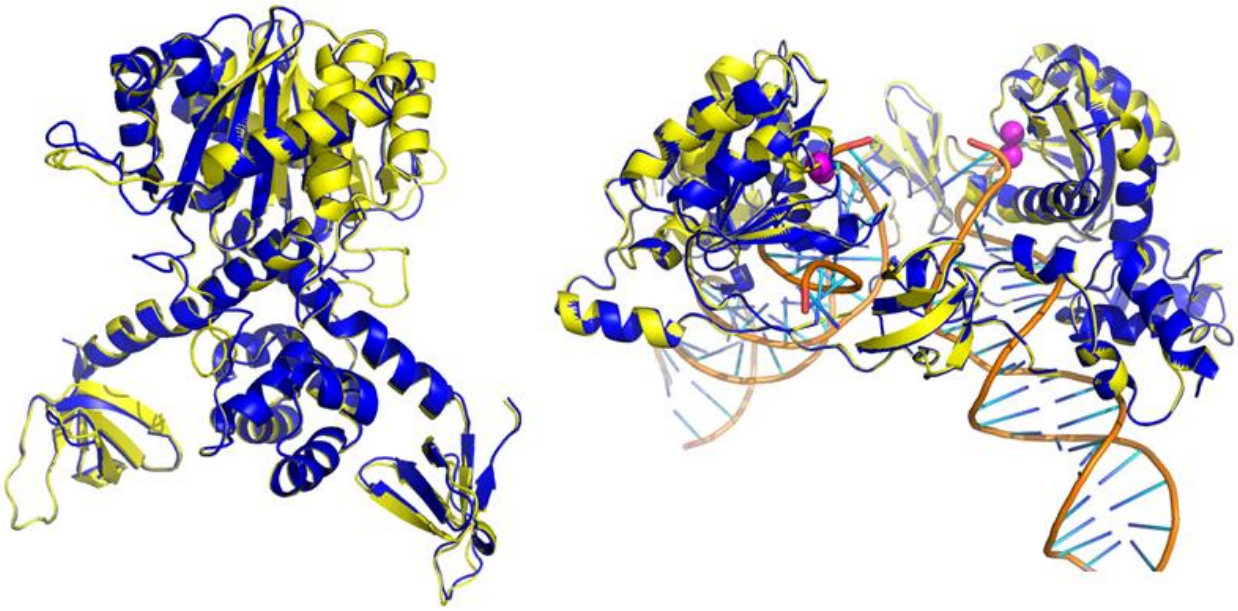
ELV



L-731, 988

Position	B consensus	Subtype CRF02-AG			
		N_1	N_2	N_3	N_4 (Q148K)
14	K	R	K	R	K
112	T	V	V	V	*R*
125	T	A	A	A	A
134	G	N	N	G	G
136	K	T	T	T	K
206	T	S	T	S	S
283	S	G	G	S	S

Table 1. Amino acid variations at the positions putatively affecting the susceptibility to INSTI in 4 isolated HIV-1 subtype CRF02-AG IN coding sequences



A **B**

Figure 1. Structural models of the HIV-1 INs from B and CRF02_AG strains.

(**A**) Superimposition of models **1** and **2**, representing the enzyme of B (in blue) and CRF02_AG (in yellow) subtypes, respectively, before 3'-processing. (**B**) Superimposition of models **3** and **4,** representing the IN•DNA pre-integration complex of B (in blue) and CRF02_AG (in yellow) subtypes. The proteins are shown as cartoons, Mg^{2+} ions as spheres (in magenta).

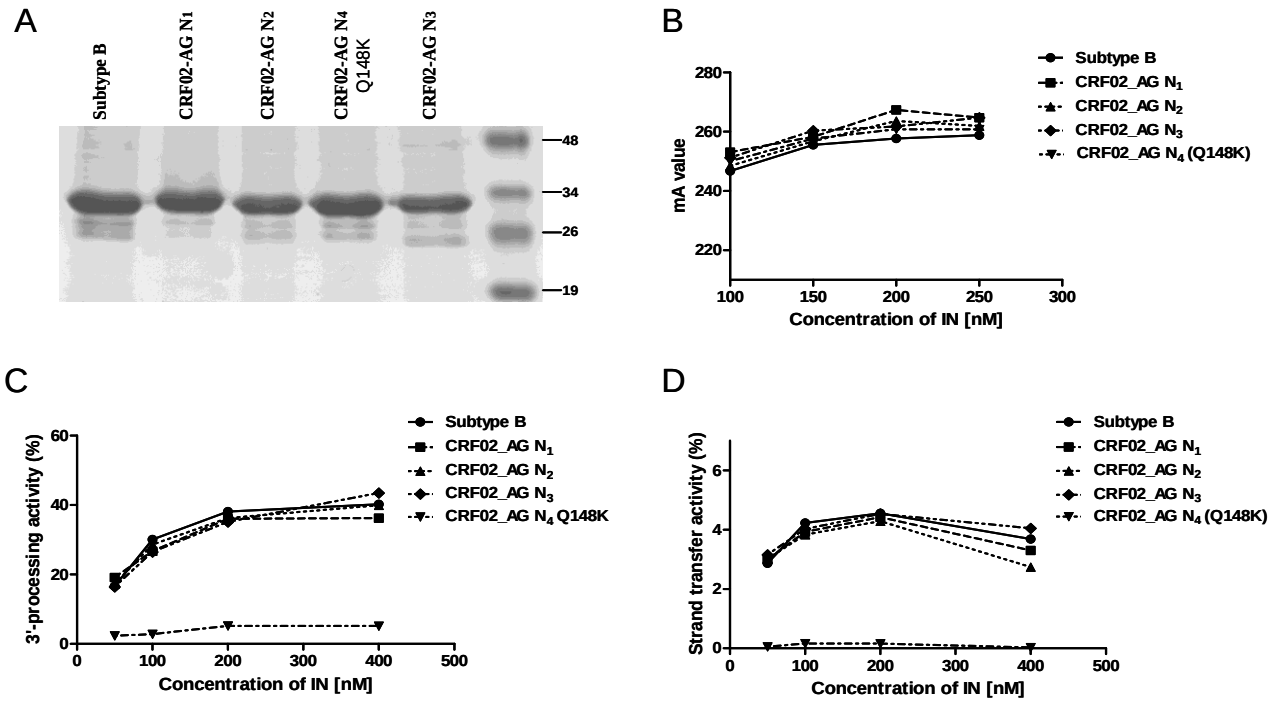


Figure 2. Purification and activity comparisons of recombinant HIV-1 subtype B and CRF02-AG INs.

(A) Purification products of recombinant HIV-1 subtype B and CRF02-AG IN N_1, N_2, N_3 and N_4. (B) DNA binding activity of HIV-1 B IN and CRF02-AG INs as a function of IN concentration. DNA binding activity was measured by steady-state fluorescence anisotropy assay, incubated at 25℃ for 20 minutes, using 21/21-mer 12.5 nM fluorescein-labled DNA substrate mimicking the viral DNA substrate, with 7.5mM MgCl2 as a cofactor in 20mM Hepes (pH 6.8), 1mM DTT and 50mM NaCl. Experimental condition was similar to (C) and (D). (C) *In vitro* 3'-processing activity of HIV-1 B IN and CRF02-AG INs as a function of IN concentration. 3'-processing activity was determined using 21/21-mer 12.5 nM P^{32}-labled DNA substrate, incubated at 37℃ for 3hours and quantified. (D) Strand transfer activity of HIV-1 B IN and CRF02-AG INs as a function of IN concentration. Strand transfer activity was determined using 21/19-mer 12.5 nM P^{32}-labled DNA substrate, incubated at 37℃ for 3hours and quantified.

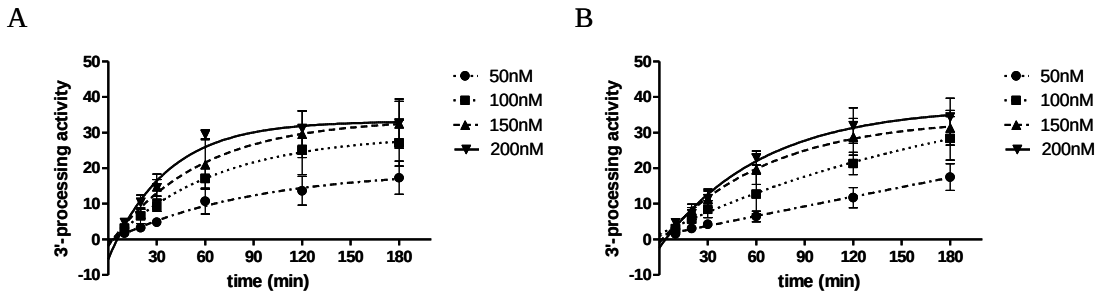


Figure 3. Kinetic comparison of HIV-1 B and CRF02-AG N_1 IN

(A) The kinetic features of recombinant HIV-1 B IN and (B) CRF02-AG N_1 IN were determined in vitro using 3'-processing activity assay, in the presence of 50 nM, 100 nM, 150 nM and 200 nM recombinant IN proteins with an incubation time of 10 min, 20 min, 30 min, 60 min, 90 min, 120 min and 180 min, respectively.

Target	Inhibitor	The free binding energies (kcal/mol)	
		Autodock_Gasteger	Glide
IN B apo	RAL	-8.68	-6.68
	ELV	-8.65	-7.60
	L-731,988	-7.35	-5.46
IN CRF02_AG	RAL	-7.89	-5.8
	ELV	-8.69	-5.3
	L-731,988	-7.69	-5.10
IN•DNA_B holo	RAL	-10.82	-11.43
	ELV	-14.35 (-8.65)	-11.79
	L-731,988	-12.73 (-7.91)	-10.69
IN•DNA CRF02_AG holo	RAL	-11.43	-10.97
	ELV	-13.33	-11.44
	L-731,988	-11.96	-10.74
IN* B holo	RAL	-8.52	-8.17
	ELV	-12.85	-8.99
	L-731,988	-13.06	-9.57
IN* CRF02_AG holo	RAL	-8.39	-8.37
	ELV	-12.97	-8.83
	L-731,988	-12.72	-9.78

Table 2. Docking binding energies (kcal/mol) of the INSTIs on the HIV-1 IN for the B and CRF02-AG strains.

The targets are the IN model with the one Mg^{2+} cation in the active site (apo state) and IN•DNA model with two Mg^{2+} cations (holo state). The inhibitors are RAL, ELV and L-731,988. The values in columns 3 and 4 are the binding energy averaged over the used conformations for the IN in apo state more or less the standard deviation produced by Autodock and Glide, respectively. The values in the brackets of the 3 column are the binding energies of ELV and L-731,988 when the charges were modified for the oxygen atoms of the carboxylate by the charges of both oxygens of the central ring of RAL.

A **B**

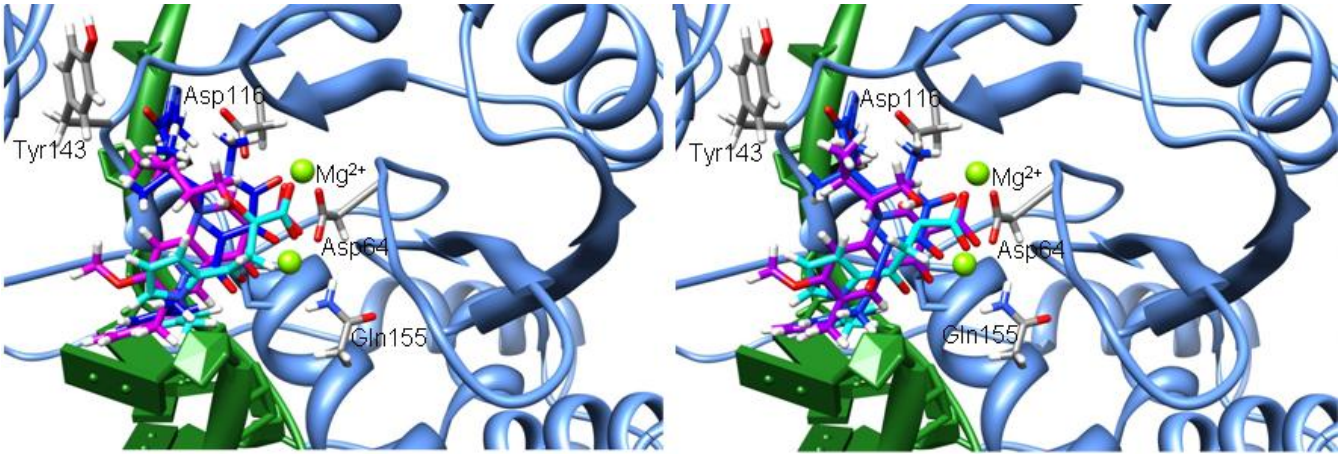


Figure 4. The RAL, ELV and L-731,988 best poses predicted by Autodock in subtype B (A) and CRF02_AG (B).

Protein (in ice blue) and DNA (in green) are shown as cartoon; inhibitors (RAL is in blue, ELV is in magenta; L-731, 988 is in cyan and the active site residues are in a type atoms colors) are shown as balls and sticks.

	IC_{50} (M)		
	Raltegravir	Elvitegravir	L-731,988
Subtype B	4.185e-008	9.340e-008	8.554e-007
CRF02_AG N_1	1.373e-008	5.562e-008	2.115e-007

Table 3. IC_{50} of 3 INSTIs against recombinant HIV-1 B IN and CRF02_AG IN N_1.

I.3. Conclusion

The natural occurring variations in HIV-1 subtype CRF02_AG IN, such as K14R, T112V, T125A, G134N, K136T, T206S and S283G, which were speculated to modify IN structure do not remarkably effect in vitro DNA binding activity, neither 3'- processing nor strand transfer activity. This result is consistent with the analysis of the models representing free and viral DNA bound INs which neither shows any significant structural change between the two subtypes. It is also similar with a previous report in which no function difference between recombinant HIV-1 subtype B and C IN enzymatic activities was detected (Bar-Magen et al., 2009). Most of the significant variations are located far from the active site; and the two variations at positions 134 and 136, near to the active site, are exposed to the solvent and apparently do not significantly affect the structure of CRF02_AG IN. A slight difference in the catalytic loop detected in in silico study do not impact the IN activity, confirmed by in vitro 3'- processing kinetic assay.

Moreover, our result indicates that the residue variations do not impact the interaction of INSTI/IN and the subsequent activity of INSTIs. The three INSTIs tested in this study possessed similar inhibitory activity against both HIV-1 B IN and CRF02_AG INs. This observation is consistent with molecular modeling analysis, which demonstrates that the conformation modes of binding and docking of the three studied INSTIs are identical for HIV-1 subtype B and CRF02_AG strains. The conclusion that the modes of interaction between INSTI and both HIV-1 subtype INs are similar is also supported by the fact that HIV-1 IN residues C65, T66, H67, N117, F121, T122, A128, G149, and K159 that were proposed to be involved in the interaction with RAL are highly conserved among subtypes (Mouscadet et al., 2010b). Altogether these results confirms the lack of difference in susceptibility previously reported for subtype B and C HIV-1 INs (Bar-Magen T et al., 2009). Another in vitro study also demonstrated similar phenotypic susceptibility of HIV-2 clinical isolates obtained from IN inhibitor- naïve patients to INSTI (Roquebert et al., 2008b).

In contrast to the lower baseline susceptibilities of recombinant A/G subtype virus to PIs, and reduced susceptibility of some A/G isolates to abacavir (Fleury et al., 2006), the susceptibilities of CRF02_AG INs in our study are even slightly higher than HIV-1 subtype B. Despite enzymatic and virologic data support the concept that naturally occurring polymorphisms in different non-B subtypes can affect their susceptibility to different antiretroviral drugs (Martínez-Cajas et al., 2008), and IN coding region contiains as many variations as protease, most of the highly prevalent polymorphisms have little

effect on INSTI susceptibility and are rarely associated with high-level resistance to INSTIs tested (Low et al., 2009). Moreover, natural variations and signature polymorphisms which observed at codon positions 140, 148, 151, 157, and 160 among HIV subtypes predicted higher genetic barriers to G140S, G140C and V151I in subtypes CRF02_AG (Brenner et al., 2011b). Therefore, INSTIs potentially provide an excellent therapeutic option for the treatment of HIV-1 subtype CRF02_AG infected patients.

In addition, one isolation CRF02_AG 52CR contains an INSTI resistance mutation Q148K. Since the IN coding sequence was derived from INSTI naive patients, how the substitution was acquired remains unclear. Previous study also found that a INI-resistance mutations Q148H (subtype G) and Q148K (CRF02_AG) were each present in one isolate (Rhee et al., 2008).

This study suggests that the treatment of HIV-1 subtype CRF02_AG with INSTIs employed in this study would probably give comparable results as subtype B. Thus, it potentially provides an additional therapeutic option for the management HIV-1 subtype CRF02_AG infected individuals.

II. In vitro catalytic properties and resistance to raltegravir of HIV-2 integrase clinic isolates and site-direct mutants

II.1. Introduction

Highly active antiretroviral therapy (HAART) has brought a dramatic effect on the morbidity and mortality of HIV-infected individuals. However, the antiretroviral efficiency for HIV-2 infected patients is reduced since the available therapeutic arsenal is developed to against HIV-1. HIV-2 is naturally resistant to all available non-nucleoside reverse transcriptase inhibitors (NNRTI), which has been mechanistically linked to differences in the NNRTI binding pocket of HIV-2 harbouring the natural Y181I polymorphism (Spira et al., 2003), and to the fusion inhibitor enfuvirtide (Poveda et al., 2004; Witvrouw et al., 2004). HIV-2 is also less sensitive to some protease inhibitors (PI) (Damond et al., 2005; Desbois et al., 2008; Rodes et al., 2006). Furthermore, compared with HIV-1, lower genetic barrier to resistance to other PIs of HIV-2 results in more rapid virologic failure (Ntemgwa et al., 2007; Pieniazek et al., 2004). Thereby, the application of new drug classes with high efficacy on HIV-2 is essential since antiretrovial therapeutic options for HIV-2 infected patients are limited compared with HIV-1 patients. A novel drug class, integrase inhibitors (INI), more specifically IN strand transfer inhibitors (INSTIs) represent a promising treatment option for HIV-2 (Bercoff et al., 2010).

As one of the three viral enzymes, IN is an attractive target for antiretroviral therapy due to its essential role in viral replication cycle. Integration catalyzed by IN renders the chromosomal archiving of drug-resistant HIV genomes in long lived cells which could permanently foreclose the application of antiretroviral therapy. Integration is processed in two successive steps (Marchand et al., 2006). The first reaction, 3’-processing (3’-P), takes place within pre-integration complexes (PICs) in the cytoplasm of infected cell following reverse transcription. During 3’-P, IN catalyzes an endonucleolytic cleavage at the 3’-site of the conserved CA of viral long terminal repeat, and generally releases a terminal 5’-GT-3’ dinucleotide. The pair of recessed CA 3’-hydroxyl ends provides the nucleophile groups required for the second reaction, strand transfer (ST). ST occurs in nucleus after the translocation of PIC along the microtube network. Both viral DNA ends are inserted into a host-cell chromosome (full-site integration) by one-step transesterification. ST reaction leaves one single-stranded gap at each junction with a five-base-pair stagger across the DNA major groove. Gap filling and release of the unpaired two-base overhang at the

5’-ends of the viral DNA are carried out in coordination with cellular DNA repair enzymes, although RT has been proposed to be involved in this reaction.

Up to now, raltegravir (RAL) is the unique INSTI approved by FDA in October, 2007. Well tolerated, RAL has demonstrated a rapid, sustained antiretroviral effect in patients with advanced HIV-1 infection (Grinsztejn et al., 2007; Markowitz et al., 2007). Due to the distinct mechanism of action, RAL is also efficiently against viruses resistant to other classes of antiretroviral drugs (Grinsztejn et al., 2007). Moreover, despite approximately 40% of heterogeneity between HIV-1 and HIV-2 IN genes, RAL is actively against wild-type HIV-2 with a phenotypic susceptibility similar to HIV-1 (Roquebert et al., 2008b; Roquebert et al., 2010; Xu et al., 2008).

However, as is the case for other antiviral drugs, resistance to RAL selected through the mutations in the IN-coding region of pol gene emerges both in vitro and/or in vivo with a low genetic barrier, reducing the susceptibility of the virus to inhibitors.

Recent studies established that HIV-2 resistance to RAL involves one of the three primary resistance mutations N155, Q148 and Y143 already described for the HIV-1 IN (Bercoff et al., 2010 ; Charpentier et al., 2010 ; Roquebert et al., 2008b; Salgadoa et al., 2009). The virologic failure of RAL treatment of HIV-1 is predominantly associated with the initial development of N155H and Q148 H/K/R principal independent pathways, either alone or combinated with other resistant mutations. The third pathway involving Y143 mutation was observed less frequently and later than N155 and Q148 pathway (Garrido et al., 2008 ; Reigadas et al., 2010). Secondary resistance mutations, i.e. G140S, with little if any effect on drug susceptibility per se increase phenotypic resistance and/or viral fitness (Delelis et al., 2009a). While the resistance to RAL of HIV-1 IN has been confirmed in vitro with IN site-directed mutants harboring these resistant substitutions (Delelis et al., 2009a; Delelis et al., 2010; Reigadas et al., 2010), no such study was carried out on HIV-2 proteins. In this study we describe in vitro catalytic properties and resistance to RAL of HIV-2 IN clinic isolates harboring resistant mutations and site-directed IN mutants.

Additionally, despite having a same protein structure, the INs of HIV-1 and HIV-2 only share 40% identity at the nucleotide level, and 65% similarity at the amino acid level. Moreover, IN length differs between HIV-1 and HIV-2, and between HIV-2 groups A and B as well (Bercoff et al., 2010), the reason of which remains unclear. Recently study showed the connecting residues of the N-terminal domain, the dimer interface and C-terminal LEDGF binding domain of HIV-2 IN were highly conserved but differed from HIV-1 (Bercoff et al., 2010). Considering all those differences between HIV-1 and HIV-2

INs, comparison study was also carried out to investigate their in vitro enzymatic activity and susceptibility to RAL.

II.2. Manuscript

G140S/Q148R and N155H mutations render HIV-2 integrase resistant to raltegravir whereas Y143C does not

Xiao-Ju Ni[1,2], Olivier Delelis[1], Charlotte Charpentier[3], Alexandre Storto[3], Gilles Collin[3], Florence Damond[3], Diane Descamps[3], and Jean-François Mouscadet[1,*]

LBPA, CNRS, Ecole Normale Supérieure de Cachan, Cachan, France[1]; School of Life Science, East *China Normal University, Shanghai, China2; Laboratoire de Virologie, AP-HP Groupe Hospitalier Bichat-Claude Bernard, HUPNVS, Université Paris-Diderot, Sorbonne Paris Cité, EA4409, 75018, Paris, France3*

*Corresponding author.
LBPA, CNRS UMR8113, Ecole Normale Supérieure de Cachan, 61 Avenue du Président Wilson, 94235 Cachan, France. Phone: (33) 1 47 40 76 75. Fax: (33) 1 47 40 76 84. E-mail: mouscadet@lbpa.ens-cachan.fr.

RUNNING TITLE: HIV-2 IN mutations conferring resistance to raltegravir *in vitro*
KEYWORDS: HIV-2; Integrase; Resistance; Raltegravir

Abstract

Background: HIV-2 is endemic in West Africa and has spread throughout Europe. However, the alternatives for HIV-2-infected patients are more limited than for HIV-1. Raltegravir, an integrase inhibitor, is active against wild-type HIV-2, with a susceptibility to this drug similar to that of HIV-1, and is therefore a promising option for use in the treatment of HIV-2-infected patients. Recent studies have shown that HIV-2 resistance to raltegravir involves one of three resistance mutations, N155H, Q148R/H and Y143C, previously identified as resistance determinants in the HIV-1 integrase coding sequence. The resistance of HIV-1 IN has been confirmed in vitro for mutated enzymes harboring these mutations, but no such confirmation has yet been obtained for HIV-2.

Results: The integrase coding sequence was amplified from plasma samples collected from ten patients infected with HIV-2 viruses, of whom three RAL-naïve and seven on RAL-based treatment at the time of virological failure. The genomes of the resistant strains were cloned and three patterns involving N155H, G140S/Q148R or Y143C mutations were identified. Study of the susceptibility of integrases, either amplified from clinical isolates or obtained by mutagenesis demonstrated that mutations at positions 155 and 148 render the integrase resistant to RAL. The G140S mutation conferred little resistance, but compensated for the catalytic defect due to the Q148R mutation. Conversely, Y143C alone did not confer resistance to RAL unless E92Q is also present. Furthermore, the introduction of the Y143C mutation into the N155H resistant background decreased the resistance level of enzymes containing the N155H mutation.

Conclusion: This study confirms that HIV-2 resistance to RAL is due to the N155H, G140S/Q148R or E92Q/Y143C mutations. The N155H and G140S/Q148R mutations make similar contributions to resistance in both HIV-1 and HIV-2, but Y143C is not sufficient to account for the resistance of HIV-2 genomes harboring this mutation. For Y143C to confer resistance in vitro, it must be accompanied by E92Q, which therefore plays a more important role in the HIV-2 context than in the HIV-1 context. Finally, the Y143C mutation counteracts the resistance conferred by the N155H mutation, probably accounting for the lack of detection of these mutations together in a single genome.

Background

HIV-2 is endemic in West Africa and has spread throughout Europe over the last two decades (1,2). The development of seven different classes of antiretroviral drugs has led to the establishment of highly active treatments that have had a profound effect on the morbidity and mortality of HIV-1-infected individuals. These classes are nucleoside (NRTIs), nucleotide (NtRTIs) and non nucleoside (NNRTIs) reverse transcriptase inhibitors, protease inhibitors (PIs), entry inhibitors, fusion inhibitors and integrase (IN) inhibitors (INIs). Despite this apparent diversity, the alternatives for HIV-2-infected patients are more limited because NNRTIs and fusion inhibitors are not active against HIV-2 (3,4) and HIV-2 is also less sensitive to some PIs (5-7). It has also been suggested that the genetic barrier is weaker in HIV-2, potentially resulting in the more rapid emergence of resistance to other PIs (8,9). The development of novel treatments based on drug classes highly effective against HIV-2 is therefore essential. INIs are active against HIV-2 IN and are therefore a promising option for use in the treatment of HIV-2-infected patients (10,11). IN plays a key role in the viral replication cycle. This makes it an attractive target for antiretroviral therapy, together with two other enzymes: reverse transcriptase (RT) and protease (P). The viral integrase catalyzes two spatially and temporally independent reactions, which eventually lead to covalent insertion of the viral genome into the chromosomal DNA. The first reaction, 3'-processing, is an endonucleolytic cleavage trimming both the 3'-extremities of the viral DNA, whereas the second reaction, strand transfer, results in the concomitant insertion of both ends of the viral DNA into a host-cell chromosome through one-step transesterification. IN strand transfer inhibitors (INSTIs) are specific inhibitors of the strand transfer reaction. The flagship molecule in this class is raltegravir (RAL), the first INSTI to have received approval for clinical use for both treatment-experienced and treatment-naïve patients (12). RAL has a rapid and sustained antiretroviral effect in patients with advanced HIV-1 infection (13,14). As it has a different mechanism of action, RAL is also effective against viruses resistant to other classes of antiretroviral drugs (13). Moreover, although HIV-1 and HIV-2 IN nucleotide sequences are only 40 % identical, RAL is active against wild-type HIV-2, which has a phenotypic susceptibility to this drug similar to that of HIV-1 (11,15).

However, as for other antiviral drugs, resistance to RAL emerges rapidly both *in vitro* and *in vivo*, through the selection of mutations within the IN coding region of the *pol* gene, greatly reducing the susceptibility of the virus to the inhibitor. In HIV-1, three main resistance pathways, involving the residues N155, Q148 and Y143, have been shown to confer resistance to RAL *in vivo*. The virological failure of RAL-based treatment in HIV-1 infection is associated primarily with the initial, independent development of the principal N155H and Q148H/K/R pathways, either alone or together with other resistance mutations. Secondary resistance mutations, such as G140S, which have little or no direct effect on drug susceptibility per se, increase phenotypic resistance or viral fitness (16). More than 60 mutations have been shown to be specifically associated with resistance to INSTIs, but biochemical studies have demonstrated that the mutations affecting residues Y143, Q148 and N155 are sufficient to decrease the susceptibility of IN to the inhibitor in vitro (16-18).The third pathway, involving the Y143R/C mutation, is less frequently observed and was identified after the N155 and Q148 pathways (17,19,20). .

Recent phenotypic studies have established that HIV-2 resistance to RAL may also involve one of the three primary resistance mutations: N155, Q148 and Y143 (10,21,22). However, whereas the resistance of HIV-1 IN to RAL has been confirmed in vitro with IN site-directed mutants harboring these mutations, no such study has yet been carried out for the HIV-2 proteins (16,17,23). We describe here the

in vitro catalytic activity and resistance to RAL of HIV-2 recombinant IN isolated from clinical isolates harboring resistance mutations. By comparing these isolates with IN mutants generated by single-site mutagenesis, we demonstrate that G140S/Q148R and N155H are sufficient to confer resistance to RAL, whereas Y143C mutation is not. We show also that N155H and Y143C mutations have antagonistic effects.

Results

Analysis of HIV-2 IN sequences from clinical isolates before RAL-based treatment

The complete sequence of the HIV-2 IN coding region from clinical isolates N1 to N3 was determined by amplification, cloning and sequencing of the IN coding region of the pol gene from plasma samples obtained at the start of RAL-based treatment. All three isolates were HIV-2 group B, as shown by comparisons with the HIV-2 group B reference sequence EHO, from which they diverged very little (between 3 % and 5 % over the first 293 residues) (Table 1). Substitutions with respect to the HIV-2 EHO sequence were found in all three viruses, at nine residues (N17, R34, I133, T180, T215, R224, N270, M287, V292) and in one or two viruses at eight other residues (F26, I50, D125, D163, V175, I204, Q221, I260). These results are consistent with previous estimates of variation for group B isolates (11). The divergence between the three isolates was even weaker, with only ten residues displaying variation, mostly conservative, in one of the three sequences (F26, I50, D125, D163, V175, T180, I204, T215, Q221, I260), demonstrating a high degree of conservation of the IN sequence (Table 1). None of these substitutions affected a residue previously associated with INSTI resistance in vivo, consistent with the absence of prior exposure to INIs. As expected for group B sequences, the C-terminal domain was of variable length. Thus, the full length of IN was 301 codons for the virus from patient N1, 299 for the virus from patient N2 and 293 for the virus from patient N3. The N1 and N2 sequences had an AQS motif for codons 293 to 295, consistent with the EHO/B reference sequence.

We investigated possible effects of sequence and length variation on IN activity, by producing and purifying the three proteins according to the protocol developed for HIV-1 IN, which favors the physiological Mg2+-dependent activity of the protein (24). The three enzymes, N1, N2 and N3, performed both catalytic activities efficiently (Figure 1). Differences in specific activity were observed, through assessments of the amount of product obtained as a function of enzyme concentration, but they remained within the range of variation for recombinant IN preparations, suggesting that neither divergence at the C-terminal end nor sequence variation had a significant impact on enzyme activity in vitro.

Table 1. Amino acid variation of HIV-2 IN isolates from INI-naive patients

Patient	Group	EHO Reference sequence																
		17	26	34	50	125	133	163	175	180	204	215	221	224	260	270	287	292
		N	F	R	I	D	I	D	V	T	I	T	Q	R	I	N	M	V
N1	B	G	-	K	V	-	V	-	-	V	V	A	-	Q	-	H	R	M
N2	B	G	-	K	-	-	V	-	I	A	-	N	-	Q	V	H	R	M
N3	B	G	Y	K	V	E	V	N	-	A	-	A	K	Q	V	H	R	M

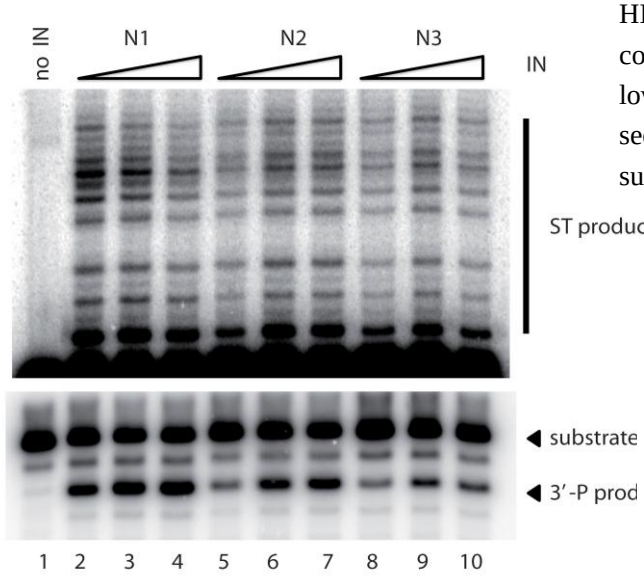


Figure 1
Study of catalytic properties of INs amplified from plasma of three INI-naive patients (N1, N2, N3) infected with HIV-2. Processing activity (bottom panel) and strand transfer activity (top panel) were assayed as a function of IN concentration. Both panels represent a unique gel with the top panel corresponding to a longer exposure. Full assay for catalytic activity was performed as described in Materials and methods section using a 21-mer (U5A/U5B) blunt DNA substrate (12.5 nM), with MgCl2 as a cofactor (7.5 mM). Lane 1. No IN. Lanes 2, 5, 8: 100 nM ; Lanes 3, 6, 9 : 200 nM ; Lanes 4, 7, 10 : 300 nM IN. ST: strand transfer; 3'-P: 3'-processing.

The enzymes from clinical isolates N1 to N3 were obtained from plasma samples collected before treatment with RAL. We confirmed the susceptibility of these INs to this INSTI, by determining IC50 values in vitro in dose-response assays carried out in the presence of various concentrations of inhibitor.

RAL efficiently inhibited the strand transfer activity of the enzyme (Figure 2A), but had no significant effect on 3' processing in vitro at concentrations up to 1 μM (data not shown), as expected for an INSTI. All three enzymes were susceptible to RAL (Figure 2B). The clinical isolates of HIV-2 studied were, therefore, susceptible to RAL before treatment initiation. The IC50 values obtained for enzymes from clinical isolates N1 and N2 were respectively equal to 23 nM and 30 nM, similar to that obtained for HxB2 HIV-1 IN (IC50 = 28 nM) in these experimental conditions, whereas the IN from N3 had a slightly lower susceptibility to RAL (IC50 = 48 nM). Thus, sequence polymorphism may slightly affect IN susceptibility to RAL.

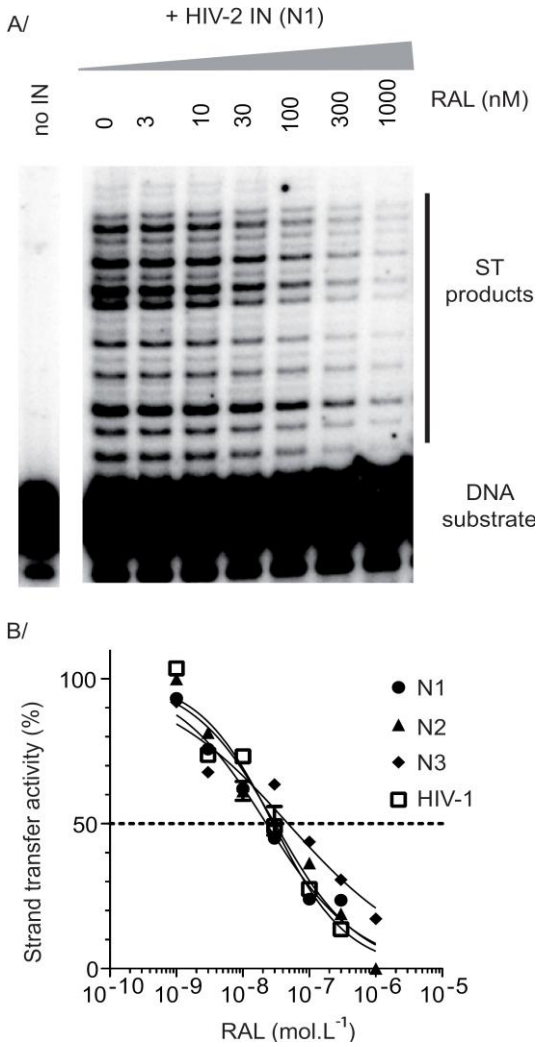


Figure 2
Comparison of in vitro RAL susceptibility of HIV-2 N1-N3 and HIV-1 HxB2 INs. (A) A representative gel obtained for HIV-2 N1 IN-mediated strand transfer reaction in the presence of increasing concentrations of raltegravir. The ST reaction was performed using a 32P-labeled oligonucleotide mimicking the preprocessed substrate. Drug concentrations are indicated above each lane. (B) Susceptibility to RAL of HIV-2 INs. Strand transfer reaction was carried out for three hours in the presence of 200 nM IN and increasing concentrations of RAL. Activity is expressed as a % of control without drug. Experiments were repeated two times.

Table 2. Amino acid substitutions of HIV-2 IN at RAL failure

Sequence	Residue			
	92	97	143	155
EHO (B)	E	T	Y	N
Pat. T1	A	A	-	H
Pat. T2	Q	-	C	-

	140	148
ROD (A)	G	Q
Pat. T3	S	R
Pat. T4	S	R
Pat.T5	A	R
Pat. T6	-	R
Pat. T7	A	R

Identification of mutations associated with RAL resistance

Plasma samples were collected from seven antiretroviral-experienced RAL-treated HIV-2-infected patients (patients T1 to T7), at the time of RAL treatment failure. Complete IN sequences were obtained by amplification, cloning and sequencing of the IN coding region. Patients T1 and T2 were infected with HIV-2 group B and patients T3 to T7 were infected with HIV-2 group A (Table 2). Comparison with the HIV-2 group A reference sequence ROD showed that all five group A sequences displayed the following polymorphisms V19I, R34K, S39T, A41T, T60V, I133V, I180V, I201T, E207D at positions that were previously showed to be subject to variation (11). All group A viruses harbored a Q148R resistance mutation, associated with G140S in two cases (patients T3 and T4) and G140A in two others (patients T5 and T7; Table 2).

The two B viruses differed from the EHO reference sequence by identical variations at the following residues: N17G, R34K, S56A, V72I, I84V, A129V, I133V, E146Q, T180V, I201L, L213F, T215A, R224Q, D240E and N270H. One group B sequence harbored the N155H resistance mutation (patient T1), and another had the Y143C resistance mutation (patient T2) (Table 2). The N155H-mutated virus also harbored the E92A and T97A mutations known to be associated with RAL resistance. Several substitutions, G27E, G70E, G82R, E92Q, and Q124R, were also detected in the Y143C-mutated virus, including one (E92Q) known to be associated with RAL resistance. These data confirm that the three main mutation patterns giving rise to RAL resistance in HIV-1 are also associated with resistance in HIV-2, as suggested by the direct sequencing of viral genomes in plasma samples (15,21).

***In vitro* enzymatic activity of HIV-2 INs**

Biochemical studies have demonstrated that Q148R, N155H and Y143C are primary resistance mutations giving rise to HIV-1 resistance whereas G140S/A and E92Q are secondary resistance mutations increasing resistance and viral fitness (16,17). We determined whether the roles of these mutations were similar in HIV-2, by producing and purifying three recombinant IN sequences corresponding to clinical isolates T1 (N155H; R1), T2 (Y143C; R2) and T3 (G140S/Q148R; R3). The catalytic activities of these enzymes were assessed and compared with that of the wild-type susceptible enzyme from patient N1, used as a control (Figure 3). Both catalytic activities were affected, to various extents, in all three enzymes. The N155H-containing enzyme displayed about 60 % the activity of the control *in vitro*, within the range of variation for wild-type enzymes. The G140S/Q148R double mutant was more strongly impaired, displaying only 30 % control levels of activity, and E92Q/Y143C mutant activity was barely detectable under standard conditions *in vitro*, indicating a functional defect.

The mutant enzymes harboring the E92A/T97A/N155H (T1) and G140S/Q148R (T3) mutations retained sufficient strand transfer activity for tests of their susceptibility to RAL, whereas this was not the case for enzymes with E92Q/Y143C (T2) mutations. The strand transfer activity of HIV-2 IN was measured in the presence of various concentrations of RAL (Figure 4). N155H mutation, in conjunction with secondary

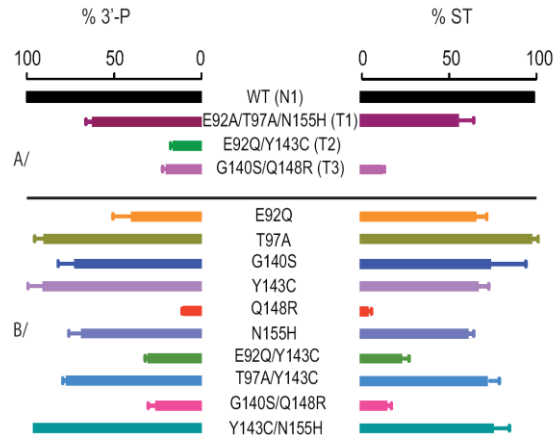


Figure 3
***In vitro* enzymatic activity of HIV-2 INs.** 3'-processing (3'-P) and strand stranfer (ST) activities of the mutants are normalized and represented as percentage of activity wild-type B-type sequence of INI-naïve patient N1 that was taken as a reference for HIV-2 IN activity. A/ INs amplified from HIV-2 infected patients. B/ INs harboring mutations obtained by site-directed mutagenesis in N1 background.

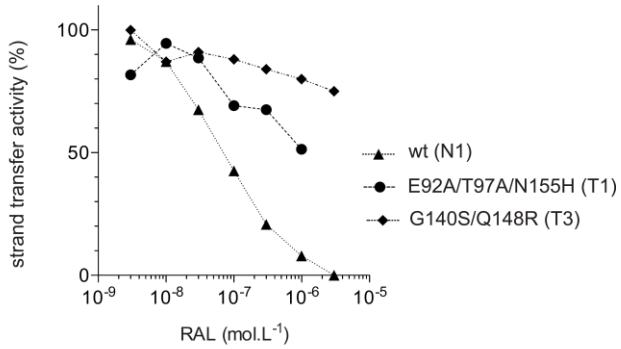


Figure 4
***In vitro* RAL susceptibility of the HIV-2 reference (N1) and T1 and T3 resistant INs amplified form clinical isolates.** Strand transfer reaction was carried using a 32P-labeled oligonucleotide mimicking the preprocessed substrate and 200 nM IN, in the presence of increasing concentrations of RAL at 37°C. Activity is expressed as a % of control without drug. Experiments were performed two times.

substitutions at positions 92 and 97, increased the IC_{50} by a factor of about 50, whereas the IC_{50} was not reached for the G140S/Q148R double mutant, for concentrations up to 1 µM. Thus both the N155H- and G140S/Q148R-containing enzymes were much less susceptible to RAL *in vitro* than the wild-type N1 HIV-2 IN, confirming that these mutations were the cause of viral resistance to RAL.

Effect of single and double mutations on IN activity *in vitro*

We investigated the contribution of each individual mutation to RAL resistance, by introducing G140S, Q148R, N155H and Y143C single mutations and the G140S/Q148R double mutation into the HIV-2 wild-type IN N1 sequence by site-directed mutagenesis. We first assessed the impact of these mutations on enzymatic activity *in vitro*, for both the 3'-processing and strand transfer activities, by comparing the efficiency of IN activities with that of the wild-type reference N1 enzyme. HIV-2 IN harboring the mutation Q148R had a much lower level of catalytic activity (<10% wild-type levels) than the wild-type enzyme (Figure 3). By contrast, the N155H mutation had no significant effect on IN activity. Introduction of the secondary mutation G140S into the Q148R background resulted in the partial recovery (up to 30% of wild-type levels) of IN catalytic activity, which was strongly impaired by the Q148R mutation. This result is similar to that obtained for HIV-1 (16). The recombinant enzymes harboring the N155H, Y143C, G140S and G140S/Q148R mutations were assayed for susceptibility to RAL. The Q148R-containing enzyme had only low levels of activity, precluding precise evaluation of its resistance, but preliminary studies with high protein concentrations suggested that this enzyme was not susceptible to RAL. The G140S mutant retained full activity and was as susceptible to RAL as the wild-type reference N1 enzyme (Figure 5A). By contrast, introduction of the G140S mutation into the Q148R background yielded a protein that was highly resistant to RAL. Thus, the G140S and Q148R mutations play the same role in the resistance of IN to RAL as in the HIV-1 integrase. Introduction of the N155H mutation into the wild-type background also

resulted in a high level of resistance (Figure 5B), with a fold-change with respect to the wild-type enzyme similar to that for the clinical isolate harboring the E92A/T97A/N155H triple mutation, which confirmed the identification of N155H as a primary resistance mutation for HIV-1 IN (22). By contrast, introduction of the Y143C mutation did not lead to significant resistance of the protein *in vitro*, suggesting that Y143C is not a primary RAL resistance mutation in HIV-2 IN.

Population sequencing results suggested that resistant viruses could harbor both the Y143C and 155H mutations. We tested this hypothesis by assaying the double mutant. However, introduction of the Y143C mutation into an N155H background significantly decreased resistance levels, by one order of magnitude (IC_{50} = 290 nM) with respect to the resistant N155H-containing IN (Figure 5B). The solubility of the Y143C/N155H recombinant IN was also lower than that of the other HIV-2 INs, suggesting that proteins carrying both these mutations were unable to adopt an appropriate conformation. Consistent with this hypothesis, the catalytic activity of the Y143C/N155H recombinant mutant was strongly increased by the addition of 10% DMSO (data not shown), probably because DMSO may help to stabilize partially folded conformations of proteins (25). These observations suggest that the Y143C and N155H mutations are mutually exclusive in a context of natural selection.

The two group B recombinant enzymes amplified from clinical isolates T1 and T2 at the time of virological failure contained E92A/Q mutations. Such mutations were previously implicated in the resistance of HIV-1 IN to INSTIs. We investigated the role of the E92 mutation, by preparing the E92Q single mutant and the E92Q/Y143C double mutant in the wild-type N1 background. Both enzymes were active *in vitro*, suggesting that the impairment of the catalytic activity of the T2 IN (E92Q/Y143C) was not directly related to the presence of these mutations (Figure 6A). The

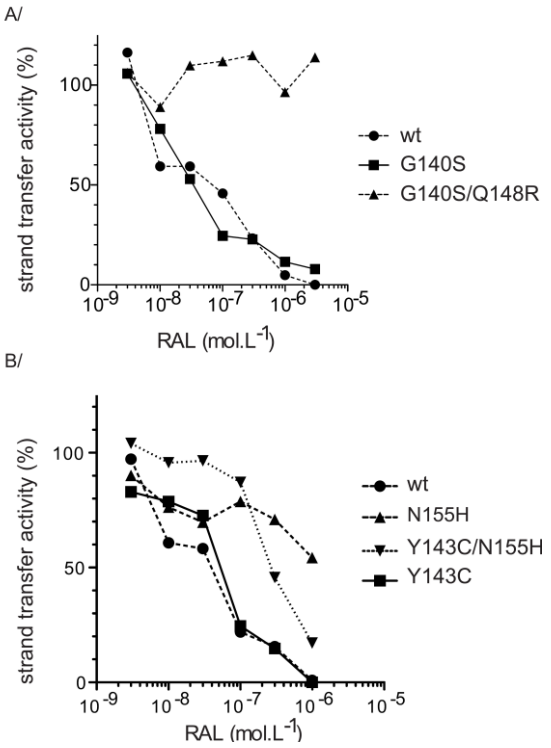


Figure 5
***In vitro* RAL susceptibility of wt and mutated HIV-2 INs. Mutations were introduced in the HIV-2 N1 background by mutagenesis.** (A) Comparison of strand transfer activity in the presence of RAL of wt (circle), G140S (square) and G140S/Q148R (triangle) mutants. (B) Comparison of strand transfer activity in the presence of RAL of wt (circle), N155H (triangle), Y143C (square) and N155H/Y143C (inverted triangle) HIV-2 INs. Strand transfer reaction was carried using a 32P-labeled oligonucleotide mimicking the preprocessed substrate and 200 nM IN, in the presence of increasing concentrations of RAL at 37°C. Activity is expressed as a % of control without drug. Experiments were performed two times.

susceptibility of recombinant INs to RAL was determined by quantifying the inhibition of *in vitro* strand transfer activity in the presence of various concentrations of RAL (Figure 6B). Like Y143C, E92Q alone did not confer significant resistance to RAL *in vitro*. By contrast, the resistance level increased when E92Q was introduced into a Y143C resistant background (IC_{50} = 370 nM), suggesting that the concomitant presence of the two mutations was necessary for this significant increase in resistance.

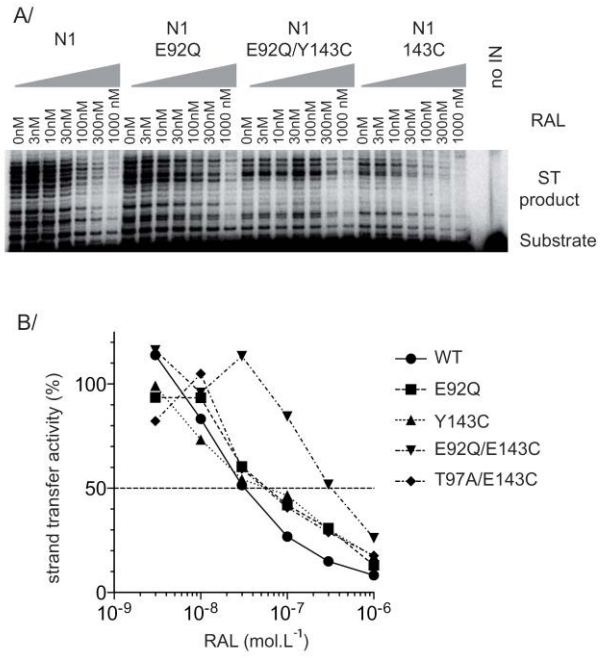


Figure 6
***In vitro* RAL susceptibility of wt and mutated HIV-2 INs. Mutations were introduced in the HIV-2 N1 background by mutagenesis.** (A) Representative gel of the strand transfer activity of recombinant HIV-2 IN mutants in the presence of increasing concentrations of RAL. (B) Comparison of strand transfer activity in the presence of RAL of wt (N1) (circle), E92Q (square), Y143C (triangle), E92Q/Y143C (inverted triangle) and T97A/Y143C (diamond) mutants. Strand transfer reaction was carried using a 32P-labeled oligonucleotide mimicking the preprocessed substrate and 200 nM IN, in the presence of increasing concentrations of RAL at 37°C. Activity is expressed as a % of control without drug. Experiments were performed two times.

Finally, it has been suggested that the T97A mutation is involved in resistance to RAL. We tested this hypothesis, by producing a T97A-containing IN. The introduction of T97A into the Y143C background did not lead to detectable IN resistance *in vitro* (Figure 6B), ruling out a direct role for this mutation in RAL resistance. Nevertheless, IN harboring the T97A substitution has higher levels of activity, suggesting that T97A may increase the fitness of resistant viral mutants.

Discussion

RAL has been reported to be clinically effective against HIV-2 infection (26,27). However, HIV-1 and HIV-2 IN nucleotide sequences are only 40 % identical and their amino-acid sequences are only 65 % similar, suggesting possible differences in the response of the enzymes to inhibitors and in the mechanisms by which resistance emerges. In this study, we characterized the *in vitro* activity of HIV-2 IN produced from the IN coding region amplified from three plasma samples taken from RAL naive, HIV-2-infected, patients and from seven others after virological failure following exposure to RAL.

Clonal analysis of three IN sequences recovered from the untreated patients identified HIV-2 group B sequences with different length as already observed for groups A and B HIV-2 IN (11). The three IN sequences studied all displayed the N17G, R34K, I133V, R224Q, N270H, M287R and V292M variations, with respect to the EHO group B reference sequence, suggesting that these variations are prevalent and may even form a consensus in HIV-2 group B circulating strains. This is particularly likely for the G17, V133 and Q224 residues, which were also detected in all five group A sequences amplified. The residues involved in resistance are therefore probably less variable than previously thought (11).

The catalytic activities of the three proteins were similar and the specific activities of the enzymes *in vitro* were similar to that of an HIV-1 IN control, indicating a lack of impact of C-terminal variation on enzyme activity. HIV-2 IN-mediated strand transfer was inhibited to the same extent by RAL *in vitro* as HIV-1 IN strand transfer, consistent with the similar phenotypic susceptibilities of HIV-1 and HIV-2 to this class of inhibitor. This result also indicates that RAL interacts with these two INs in similar ways. This conclusion is strongly supported by the complete conservation in HIV-2 of the HIV-1 IN residues C65, T66, H67, N117, F121, T122, A128, G149 and K159, which

have been implicated in the interaction with RAL (28).

Whole-population analysis of viral samples recovered from HIV-2-infected patients who experienced virological failure on RAL treatment indicated that the three main pathways of RAL resistance in HIV-1, involving mutations of residues Q148, N155 and Y143, may be also responsible for the resistance of HIV-2 to INSTIs (26,29,30). Moreover, a previous clonal analysis found that E92Q was associated with both the N155H and Y143C mutations, whereas T97A mutation was associated with Y143C (15). Our clonal analysis based on seven clinical isolates identified three different patterns, E92A/T97A/N155H, G140S/Q148R and E92Q/143C, confirming the probable existence of these three pathways.

The three recombinant INs obtained by cloning and expressing these sequences were indeed strongly resistant to RAL *in vitro*, thereby confirming that the mutated sequence was responsible for resistance. We showed, by single-site mutagenesis in a RAL-susceptible background, that the N155H mutation and the G140S/Q148H double mutation were sufficient to elicit strong resistance to RAL in HIV-2 IN *in vitro*. For the G140S/Q148H pattern, Q148H caused a major catalytic defect while conferring a high level of resistance to RAL, the catalytic defect being rescued *in vitro* by the secondary G140S mutation, which did not itself confer resistance. This result is entirely consistent with our previous observations for HIV-1 *in vitro* (16). The concomitant selection of the Q148H/R/K and G140S mutations in both HIV-1 and HIV-2 RAL-resistant viruses is grounded in the structure of IN, indicating a close interaction between these two residues that are conserved in retroviral IN (31).

A whole-population study suggested a possible association of T97A with both N155H and Y143C mutations in resistant HIV-2 viruses (21,32). For the N155H mutation, the level of resistance conferred by single-site directed mutagenesis was very similar to that of the enzyme amplified from the E92A/T97A/N155H clinical isolate, suggesting that the E92A and T97A secondary mutations provided no additional resistance. This hypothesis is consistent with the absence of resistance observed for INs into which T97A and E92Q were introduced as single mutations. Thus, N155H is the main, if not only determinant of viral resistance in this pathway, consistent with the findings of previous phenotypic studies of viral replication (22). Similarly, we observed no significant effect of T97A in combination with Y143C, on HIV-2 IN susceptibility to RAL. Nevertheless, limited stimulation of IN activity was detected, suggesting that this mutation may improve the fitness of enzymes that would otherwise be catalytically impaired as previously suggested for HIV-1 IN (33).

Y143C/R has been described as a primary mutation for HIV-1 resistance to RAL (19). We previously demonstrated *in vitro* that the susceptibility to RAL of IN was strongly affected by this single mutation (17). Surprisingly, we observed that, although the Y143C-containing HIV-2 IN was amplified from clinical isolate of a patient at time of RAL failure, Y143C mutation alone was not sufficient to confer resistance to IN *in vitro*, ruling out this mutation as a sole determinant of resistance in the HIV-2 context. The IN sequence amplified from the clinical isolate also contained the E92Q mutation. We therefore studied the impact of this mutation within the Y143C background. E92Q is considered to be a secondary mutation within the N155H and Y143C pathways for HIV-1 resistance to RAL. Indeed, E92Q confers resistance to HIV-1 IN *in vitro*, albeit to a lesser extent than N155H or Q148R (18). However, this effect was not confirmed in the HIV-2 context, because the E92Q mutation, introduced through single-site directed mutagenesis, did not significantly increase the resistance of the HIV-2 mutant IN *in vitro*. Nonetheless, the Y143C/E92Q double mutation

obtained by mutagenesis of the wild-type control resulted in resistance *in vitro*. Thus, although neither Y143C nor E92Q is sufficient for significant resistance to RAL, the presence of both these substitutions in the same protein leads to resistance indicating that unlike HIV-1, at least two mutations seem to be required in this pathway to elicit resistance in HIV-2. A recent structural study suggested that RAL establishes contact with the side chain of Y143 (31). We conclude that, in HIV-2, the loss of this contact is not sufficient to impair RAL binding to IN. A second modification to the RAL binding site, which encompasses E92, is probably required. E92Q has been described as a primary resistance mutation for another INSTI, elvitegravir. Moreover, it was also shown that elvitegravir remains fully active against Y143 mutant HIV-1 integrase (34). It would therefore be interesting to determine whether there are structural similarities between elvitegravir binding to HIV-1 IN and RAL binding to HIV-2 IN.

It has been suggested that the emergence of Y143C during HIV-1 infection may result from a late switch from the N155H pathway (19). Our data show that, in HIV-2 infection, the Y143C mutation counteracts the resistance conferred by the N155H mutation, probably precluding the simultaneous selection of both mutations. Thus, although we did not investigate the effect of E92 mutations on the *in vitro* resistance of N155H-containing HIV-2 IN, we suggest that the emergence of E92A/G/Q secondary mutations, facilitated by the single nucleotide change required for all substitutions — E to A, E to G or E to Q transition — may be involved in the switch from the N155H to the Y143C pathway in the HIV-2 context. Under this hypothesis, Y143C-containing resistant viruses would be expected to be more resistant to RAL than N155H-containing viruses. This was not the case here, as the Y143C/E92Q recombinant enzyme was less resistant than the N155H-containing enzyme. There may therefore be other, as yet unidentified determinants. Y143C/E92Q-containing INs from clinical isolates also harbored G27E, G70E, G82R and Q124R variants, which were not found in the HIV-2 group B RAL-susceptible sequences. Moreover, none of these residues has previously been identified as a site of major variation, consistent with selection under RAL pressure in these isolates. However, these residues have also never before been associated with the resistance of HIV-1 or HIV-2 to INSTIs, and their role remains to be determined (35,36).

Conclusion

In conclusion, this study confirms that HIV-2 resistance to RAL is due to the N155H, G140S/Q148R or E92Q/Y143C mutations in the IN coding region. The N155H and G140S/Q148R mutations make similar contributions to resistance in both HIV-1 and HIV-2, but Y143C alone is not sufficient to account for the resistance of HIV-2 genomes harboring this mutation. For Y143C to confer resistance in vitro, it must be accompanied by E92Q, which therefore plays a more important role in the HIV-2 context than in the HIV-1 context.

Methods

Patients

Plasma samples were collected from three RAL-naive patients infected with HIV-2, group B (patients N1 to N3). Two of these patients had previously received antiretroviral treatment whereas the third had never been treated. Plasma samples were also collected from seven different HIV-2-infected patients, two group B and five group A, at the time of virological failure on RAL-based treatment (patients T1 to T7).

Cloning and site-directed mutagenesis

The IN coding sequences from the viruses isolated from the plasma of HIV-2-infected patients N1 to N3 and T1 to T7 were amplified by PCR and sequenced according to a previously described procedure (11). PCR products corresponding to the entire IN sequence were amplified again with the following specific primers containing *Nde*I and *Bam*H I restriction sites at their 5’ ends. Primer 1: N1, N2, T1, T2,

5’-CATATGTTTCTAGAAAAGATAGAACCAG C-3’; N3, 5’-CATATGTTTTTAGAG AACATAGAACCAGC-3’; T3-T7, 5’-CATATGTTCCTGGAAAAGATAGAGCCCG C-3’. Primers 2: N1, 5’-GGATCCCTATGCTTCAGGTACTTGACCA G-3’; N2, 5’-GGATCCCAT CCTGGTATCCTCCACGTCGGC-3’; N3, 5’-GGATCCCTATGCCACCTCTCTAGTCTGC C-3’; T1, T2, 5’-GGATCCTTAATTAGACTGTGCCACCTCT CTAG-3’; T3-T7, 5’-GGATC CCTATGCCACCTCTCCATCCTCCCTG -3’**.** The amplicons were then ligated into the TA cloning plasmid pGEM-T easy (Promega, Madison, USA). Single (E92Q, T97A, G140S, Q148R, Y143C, N155H) and double (G140S/Q148R, Y143C/N155H, E92Q/Y143C and T97A/Y143C) mutations were introduced into the HIV-2 wild-type N1 sequence by mutagenesis using the QuickChange II site-directed mutagenesis kit (Agilent Technologies, Santa Clara, USA) according to the manufacturer instructions. *Nde*I - *Bam*H fragments covering the IN coding sequence were then inserted into the expression vector pET-15b.

Expression and purification of recombinant INs

His-tagged INs were produced in *Escherichia coli* BL21-CodonPlus (DE3)-RIPL (Agilent, Santa Clara, USA) and purified under nondenaturing conditions, as previously described (24). Protein production was induced at an OD_{600} of 0.6 to 0.8, by adding isopropyl β-D-1-thiogalactopyranoside (IPTG) to a concentration of 0.5 mM. Cultures were incubated for 3 h at 30°C and then centrifuged 20 min at 1100 *g*, 4°C. Cells were resuspended in buffer A (50 mM Tris-HCl (pH 8), 1 M NaCl, 4 mM β-mercaptoethanol) and lysed by passage through a French press. The lysate was centrifuged (30 min at 12,000 g, 4°C), and the supernatant was filtered (pore size 0.45 µM) and incubated with nickel-nitrilotriacetic acid agarose beads (Qiagen, Venlo, The Netherlands) for at least 2 hours at 4°C. The beads were washed with buffer A and then with buffer A supplemented with 80 mM imidazole. His-tagged proteins were then eluted from the beads in buffer A supplemented with 1 M imidazole and 50 µM zinc sulfate. They were then dialyzed overnight against 20 mM Tris-HCl (pH 8), 1 M NaCl, 4 mM β-mercaptoethanol and 10% glycerol. Aliquots of the purification products were rapidly frozen and stored at -80°C.

Characterization of IN enzymatic activity *in vitro*

The activity of wild-type and mutated INs was determined *in vitro*, as previously described (24). Briefly, oligonucleotides (ODN) mimicking the end of the U5 long terminal repeat of the viral genome were radiolabeled with T4 polynucleotide kinase (Biolabs, Ipswich, USA) and [γ-^{32}P] ATP (3,000 Ci/mmol) (Amersham, GE Healthcare, USA), then purified on a Sephadex G-10 column (GE Healthcare, USA). Double-stranded ODNs were obtained by mixing equimolar amounts of complementary strands in the presence of 100 mM NaCl. We carried out 3’-processing and strand transfer assays at 37°C in a buffer containing 20 mM HEPES (pH 6.8), 1 mM dithiothreitol (DTT), 7.5 mM Mg^{2+} and 50 mM NaCl in the presence of a 12.5 nM solution of U5A/U5B (3’-processing) or U5A/U5B-2 (strand transfer) double-stranded DNA substrates, respectively. The products were separated by electrophoresis in a 16 % acrylamide/urea denaturing gel. Gels were analyzed with a Typhoon TRIO variable mode imager (GE Healthcare, USA) and quantified with ImageQuant TL software. The susceptibility of INs to RAL was determined *in vitro* by assessing IN activity in the presence of various concentrations of RAL. We obtained 50 % inhibitory concentrations (IC_{50}) with Prism 5.0 software (GraphPad Software, San Diego, CA). The HIV-2 ODN substrate sequences were: UA: 5'-CCTGCTAGGGATTTTCCTGCCTCGGTTT-3’; U5B: 5'-AAACCGAGGCAGGAAAATCCCTAGCAG G-3’; U5B-2: 5'-AAACCGAGGCAGGAAAA TCCCTAGCA-3’.

Competing interests: The authors declare that they have no competing interests.

Authors' contributions: XN carried out the cloning, expression and purification of recombinant integrases, performed the *in vitro* activity studies and contributed to the writing of the manuscript, OD participated in the design of the study and in the enzyme mutagenesis, CC carried out the sequence alignment, contributed to the data interpretation and participated in the writing of the manuscript, AS, GC and FD contributed to the sequence amplifications, analysis and interpretation of the genetic data, DD participated to the conception of the study and to the data interpretation, JFM conceived the study and wrote the manuscript. All authors contributed to the critical reviewing of the manuscript and approved the final manuscript.

Funding

This work was supported by Sidaction, the *Agence Nationale de Recherche sur le Sida et les Hépatites* (ANRS) and the *Centre National pour la Recherche Scientifique* (CNRS).

References

1 Matheron, S., Mendoza-Sassi, G., Simon, F., Olivares, R., Coulaud, J.P. and Brun-Vezinet, F. (1997) HIV-1 and HIV-2 AIDS in African patients living in Paris. *Aids*, **11**, 934-936.

2. Reeves, J.D. and Doms, R.W. (2002) Human immunodeficiency virus type 2. *J Gen Virol*, **83**, 1253-1265.

3. Witvrouw, M., Pannecouque, C., Switzer, W.M., Folks, T.M., De Clercq, E. and Heneine, W. (2004) Susceptibility of HIV-2, SIV and SHIV to various anti-HIV-1 compounds: implications for treatment and postexposure prophylaxis. *Antivir Ther*, **9**, 57-65.

4. Poveda, E., Rodes, B., Toro, C. and Soriano, V. (2004) Are fusion inhibitors active against all HIV variants? *AIDS Res Hum Retroviruses*, **20**, 347-348.

5. Damond, F., Brun-Vezinet, F., Matheron, S., Peytavin, G., Campa, P., Pueyo, S., Mammano, F., Lastere, S., Farfara, I., Simon, F., Chene, G. and Descamps, D. (2005) Polymorphism of the human immunodeficiency virus type 2 (HIV-2) protease gene and selection of drug resistance mutations in HIV-2-infected patients treated with protease inhibitors. *J Clin Microbiol*, **43**, 484-487.

6. Desbois, D., Roquebert, B., Peytavin, G., Damond, F., Collin, G., Benard, A., Campa, P., Matheron, S., Chene, G., Brun-Vezinet, F. and Descamps, D. (2008) In vitro phenotypic susceptibility of human immunodeficiency virus type 2 clinical isolates to protease inhibitors. *Antimicrob Agents Chemother*, **52**, 1545-1548.

7. Rodes, B., Sheldon, J., Toro, C., Jimenez, V., Alvarez, M.A. and Soriano, V. (2006) Susceptibility to protease inhibitors in HIV-2 primary isolates from patients failing antiretroviral therapy. *J Antimicrob Chemother*, **57**, 709-713.

8. Ntemgwa, M., Brenner, B.G., Oliveira, M., Moisi, D. and Wainberg, M.A. (2007) Natural polymorphisms in the human immunodeficiency virus type 2 protease can accelerate time to development of resistance to protease inhibitors. *Antimicrob Agents Chemother*, **51**, 604-610.

9. Pieniazek, D., Rayfield, M., Hu, D.J., Nkengasong, J.N., Soriano, V., Heneine, W., Zeh, C., Agwale, S.M., Wambebe, C., Odama, L. and Wiktor, S. Z. (2004) HIV-2 protease sequences of subtypes A and B harbor multiple mutations associated with protease inhibitor resistance in HIV-1. *Aids*, **18**, 495-502.

10. Bercoff, D.P., Triqueneaux, P., Lambert, C., Oumar, A.A., Ternes, A.M., Dao, S., Goubau, P., Schmit, J.C. and Ruelle, J. (2010) Polymorphisms of HIV-2 integrase and selection of resistance to raltegravir. *Retrovirology*, **7**, 98.

11. Roquebert, B., Damond, F., Collin, G., Matheron, S., Peytavin, G., Benard, A., Campa, P., Chene, G., Brun-Vezinet, F. and Descamps, D. (2008) HIV-2 integrase gene polymorphism and phenotypic susceptibility of HIV-2 clinical isolates to the integrase inhibitors raltegravir and elvitegravir in vitro. *J Antimicrob Chemother*, **62**, 914-920.

12. Zolopa, A.R. (2010) The evolution of HIC treatment guidelines: Current state-of-te-art of ART. *Antiviral Res*, **85**, 241-244.

13. Grinsztejn, B., Nguyen, B.Y., Katlama, C., Gatell, J.M., Lazzarin, A., Vittecoq, D., Gonzalez, C.J., Chen, J., Harvey, C.M. and Isaacs, R.D. (2007) Safety and efficacy of the HIV-1 integrase inhibitor raltegravir (MK-0518) in treatment-experienced patients with multidrug-resistant virus: a phase II randomised controlled trial. *Lancet*, **369**, 1261-1269.

14. Markowitz, M., Nguyen, B.Y., Gotuzzo, E., Mendo, F., Ratanasuwan, W., Kovacs, C., Prada, G., Morales-Ramirez, J.O., Crumpacker, C.S., Isaacs, R.D., Gilde, L. R., Wan, H., Miller, M. D., Wenning, L. A., Teppler, H.(2007) Rapid and durable antiretroviral effect of the HIV-1 Integrase inhibitor raltegravir as part of combination therapy in treatment-naive patients with HIV-1 infection: results of a 48-week controlled study. *J Acquir Immune Defic Syndr*, **46**, 125-133.

15. Xu, L., Anderson, J., Garrett, N., Ferns, B., Wildfire, A., Cook, P., Workman, J., Graham, S. and Smit, E. (2009) Dynamics of raltegravir resistance profile in an HIV type 2-infected patient. *AIDS Res Hum Retroviruses*, **25**, 843-847.

16. Delelis, O., Malet, I., Na, L., Tchertanov, L., Calvez, V., Marcelin, A.G., Subra, F., Deprez, E. and Mouscadet, J.F. (2009) The G140S mutation in HIV integrases from raltegravir-resistant patients rescues catalytic defect due to the resistance Q148H mutation. *Nucleic Acids Res*, **37**, 1193-1201.

17. Delelis, O., Thierry, S., Subra, F., Simon, F., Malet, I., Alloui, C., Sayon, S., Calvez, V., Deprez, E., Marcelin, A.G., Tchertanov, L. and Mouscadet, J. F.

(2010) Impact of Y143 HIV-1 integrase mutations on resistance to raltegravir in vitro and in vivo. *Antimicrob Agents Chemother*, **54**, 491-501.

18. Malet, I., Delelis, O., Valantin, M.A., Montes, B., Soulie, C., Wirden, M., Tchertanov, L., Peytavin, G., Reynes, J., Mouscadet, J.F., Katlama, C., Calvez, V. and Marcelin, A. G. (2008) Mutations associated with failure of raltegravir treatment affect integrase sensitivity to the inhibitor in vitro. *Antimicrob Agents Chemother*, **52**, 1351-1358.
19. Sichtig, N., Sierra, S., Kaiser, R., Daumer, M., Reuter, S., Schulter, E., Altmann, A., Fatkenheuer, G., Dittmer, U., Pfister, H. and Esser, S. (2009) Evolution of raltegravir resistance during therapy. *J Antimicrob Chemother*, **64**, 25-32.
20. Reigadas, S., Anies, G., Masquelier, B., Calmels, C., Stuyver, L.J., Parissi, V., Fleury, H. and Andreola, M.L. (2010) The HIV-1 integrase mutations Y143C/R are an alternative pathway for resistance to Raltegravir and impact the enzyme functions. *PLoS One*, **5**, e10311.
21. Charpentier, C., Roquebert, B., Delelis, O., Larrouy, L., Matheron, S., Tubiana, R., Karmochkine, M., Duval, X., Chene, G., Storto, A., Collin, G., Benard, A., Damond, F., Mouscadet, J. F., Brun-Vezinet, F. and Descamps, D. (2011) Hot Spots of Integrase Genotypic Changes Leading to HIV-2 Resistance to Raltegravir. *Antimicrob Agents Chemother*, **55**, 1293-1295.
22. Salgado, M., Toro, C., Simon, A., Garrido, C., Blanco, F., Soriano, V. and Rodes, B. (2009) Mutation N155H in HIV-2 integrase confers high phenotypic resistance to raltegravir and impairs replication capacity. *J Clin Virol*. **46**, 173-175.
23. Metifiot, M., Maddali, K., Naumova, A., Zhang, X., Marchand, C. and Pommier, Y. (2010) Biochemical and Pharmacological Analyses of HIV-1 Integrase Flexible Loop Mutants Resistant To Raltegravir. *Biochemistry*. **49**, 3715-3722.
24. Leh, H., Brodin, P., Bischerour, J., Deprez, E., Tauc, P., Brochon, J.C., LeCam, E., Coulaud, D., Auclair, C. and Mouscadet, J.F. (2000) Determinants of Mg2+-dependent activities of recombinant human immunodeficiency virus type 1 integrase. *Biochemistry*, **39**, 9285-9294.
25. Bhattacharjya, S. and Balaram, P. (1997) Effects of organic solvents on protein structures: observation of a structured helical core in hen egg-white lysozyme in aqueous dimethylsulfoxide. *Proteins*, **29**, 492-507.
26. Garrett, N., Xu, L., Smit, E., Ferns, B., El-Gadi, S. and Anderson, J. (2008) Raltegravir treatment response in an HIV-2 infected patient: a case report. *Aids*, **22**, 1091-1092.
27. Damond, F., Lariven, S., Roquebert, B., Males, S., Peytavin, G., Morau, G., Toledano, D., Descamps, D., Brun-Vezinet, F. and Matheron, S. (2008) Virological and immunological response to HAART regimen containing integrase inhibitors in HIV-2-infected patients. *Aids*, **22**, 665-666.
28. Mouscadet, J.F., Delelis, O., Marcelin, A.G. and Tchertanov, L. (2010) Resistance to HIV-1 integrase inhibitors: A structural perspective. *Drug Resist Updat*, **13**, 139-150.
29. Roquebert, B., Blum, L., Collin, G., Damond, F., Peytavin, G., Leleu, J., Matheron, S., Chene, G., Brun-Vezinet, F. and Descamps, D. (2008) Selection of the Q148R integrase inhibitor resistance mutation in a failing raltegravir containing regimen. *Aids*, **22**, 2045-2046.
30. Charpentier, C., Karmochkine, M., Laureillard, D., Tisserand, P., Belec, L., Weiss, L., Si-Mohamed, A. and Piketty, C. (2008) Drug resistance profiles for the HIV integrase gene in patients failing raltegravir salvage therapy. *HIV Med*, **9**, 765-770.
31. Hare, S., Vos, A.M., Clayton, R.F., Thuring, J.W., Cummings, M.D. and Cherepanov, P. (2010) Molecular mechanisms of retroviral integrase inhibition and the evolution of viral resistance. *Proc Natl Acad Sci U S A*, **107**, 20057-20062.
32. Charpentier, C., Larrouy, L., Matheron, S., Damond, F., Delelis, O., Mouscadet, J.-F., Campa, P., Chêne, G., Brun-Vezinet, F. and Descamps, D. (2011) Long-Lasting persistence of integrase resistance mutations in HIV-2-infected patients after raltegravir removal. *Antiviral Therapy, in press*.
33. Reigadas, S., Masquelier, B., Calmels, C., Laguerre, M., Lazaro, E., M. Vandenhende, M., Neau, D., Fleury, H. and Andréola, M.L. (2011) Structure analysis of HIV-1 IN Y143C/R Raltegravir resistance mutation in association with the secondary mutation T97A. *Antimicrob Agents Chemother*, **55**, 3187-3194.
34. Métifiot, M., Vandegraaft, N., Maddali, K., Naumova, A., Zhang, X., Rhodes, D., Marchand, C. and Pommier, Y. (2011) Elvitegravir overcomes resistance to raltegravir induced by integrase mutation 143. *Aids*, **25**, 1175-1178.
35. Ceccherini-Silberstein, F., Malet, I., D'Arrigo, R., Antinori, A., Marcelin, A.G. and Perno, C.F. (2009) Characterization and structural analysis of HIV-1 integrase conservation. *AIDS Rev*, **11**, 17-29.
36. Ceccherini-Silberstein, F., Van Baelen, K., Armenia, D., Trignetti, M., Rondelez, E., Fabeni, L., Scopelliti, F., Pollicita, M., Van Wesenbeeck, L., Van Eygen, V., Dori, L., Sarmati, L., Aquaro, S., Palamara, G., Andreoni, M., Stuyver, L. J. and Perno, C. F. (2010) Secondary integrase resistance mutations found in HIV-1 minority quasispecies in integrase therapy-naive patients have little or no effect on susceptibility to integrase inhibitors. *Antimicrob Agents Chemother*, **54**, 3938-3948.

II.3. Supplementary results

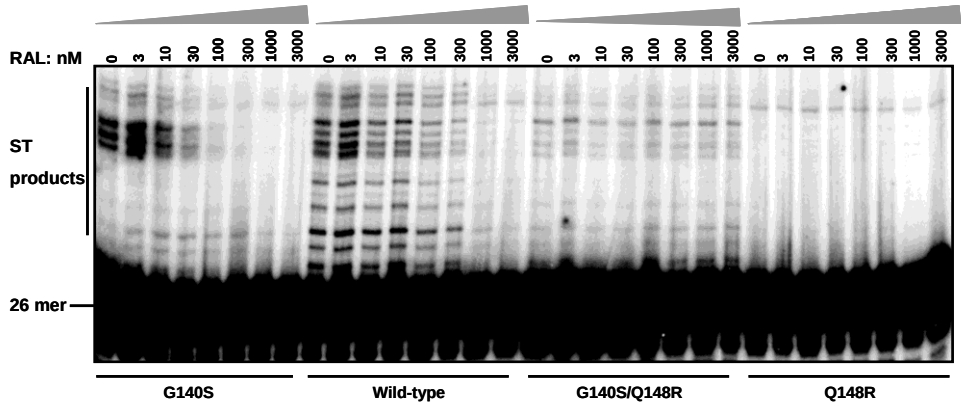


Figure 24: In vitro susceptibility to RAL of wild-type, G140S, Q148R and G140S/Q148R HIV-2 IN mutants.

A representative gel shows RAL resistance with respect to the IN mutations. G140S confers no resistance to RAL. Q148R mutant is highly resistant to RAL but its catalytic activity is severely crippled. The introduction of G140S restores catalytic defect of Q148R mutant while maintaining the high level of resistance. Susceptibility to RAL was determined with strand transfer assay in the presence of increasing concentration of RAL, and 100 nM IN. RAL concentrations are indicated above each lane. The 32P-labeled oligonucleotide substrate mimicking the preprocessed HIV-2 viral substrate is indicated as 26-mer. Products of the strand transfer reaction are indicated (ST products).

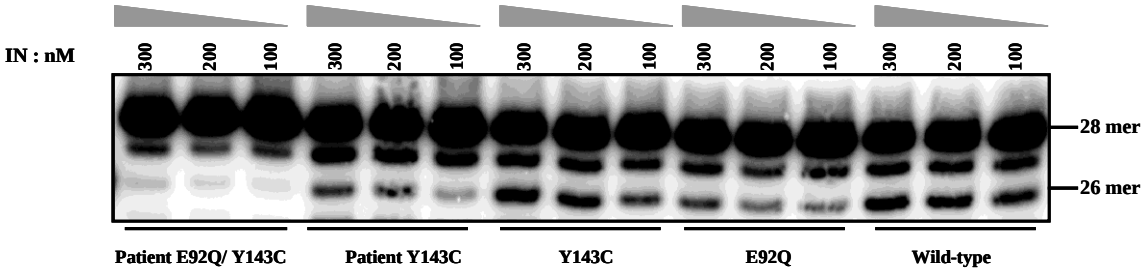


Figure 25: 3'-processing activity is partially rescued by eliminating E92Q mutation in integrase obtained from patient at the time of raltegravir treatment failure.

A representative gel shows 3'-processing activity of INs which were obtained from patient at the time of raltegravir treatment failure, harboring Y143C mutation with or without E92Q mutation; and also site-directed Y143C mutant, E92Q mutant and wild-type IN. 3'-processing activity was determined in the presence of increasing IN concentration indicated above each lane. The ^{32}P-labeled oligonucleotide substrate is indicated as 28-mer. Products of the 3'-processing reaction is indicated as 26-mer.

- **Comparison of the enzymatic activities of HIV-1 and HIV-2 wild-type INs as function of IN concentration.**

The INs of HIV-1 and HIV-2 viruses share 40% identity at the nucleotide level and 65% similarity at the amino acid level. Moreover, the length of HIV-1 and HIV-2 IN differs (Bercoff et al., 2010). In this study, one HIV-2 wild-type IN comprised of 301 amino acids, while HIV-1 IN 288 amino acids. Moreover, it is worthy to note that the recombinant HIV-2 INs were consistently obtained, after purification, at concentration above 1.5 mg/mL when purified in absence of detergent, while the poor solubility of HIV-1 IN generally precludes obtaining such concentrations in non-denaturing conditions (Leh et al., 2000). Considered approximately one third amino acid variations, different length of protein and varied solubility, in vitro enzymatic assays as a function of IN concentration were carried out to compare the catalytic activity of both wild-type INs, using their own oligonucleotide substrates, respectively.

As shown in **Figure 26**, 3'- processing efficiency saturated for concentration of HIV-1 IN above 1 μM. Furthermore, HIV-1-mediated strand transfer displayed a typical bell-curve shape with almost complete inhibition of the activity at concentration above 1μM, a phenomenon previously ascribed to the propensity of HIV-1 IN to form large inactive aggregates (Guiot et al., 2006). In contrast, no saturation was obtained for 3'-processing carried out by HIV-2 IN and the bell-curved shape for strand transfer was less pronounced, confirming the capability of the HIV-2 IN to remain under an active non aggregated form for higher concentrations.

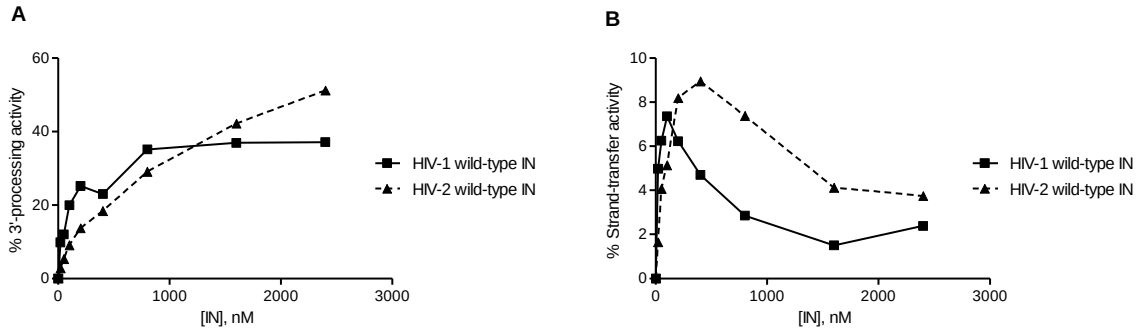


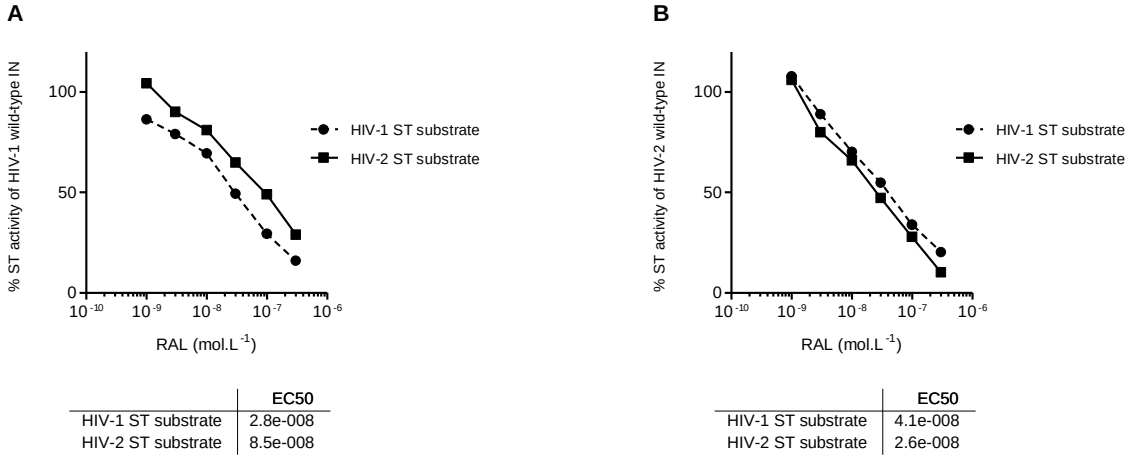
Figure 26: Comparison of the enzymatic activities of HIV-1 and HIV-2 wild-type integrases as a function of integrase concentration using their own substrates, respectively.

(A) 3'-processing activity as a function of IN concentration after incubated at 37 °C for 3 hours, using the respective 3'-processing substrates (12.5 nM) of each wild-type IN with $MgCl_2$ (7.5 mM) as a metallic cofactor in 20 mM HEPES (pH 6.8), 1 mM DTT and 50 mM NaCl. IN concentrations are 20 nM, 50 nM, 100 nM, 200 nM, 400 nM, 800nM, 1600nM and 2400 nM. (B) Stand transfer assay was carried out under the same condition of 3'-processing assay using the respective strand transfer substrates of each wild-type IN.

- **Comparison of the susceptibility to RAL of HIV-1 and HIV-2 wild-type INs.**

To compare the susceptibility to RAL of HIV-1 and HIV-2 wild-type INs, strand transfer reactions were carried out in the presence of increasing concentrations of RAL. The susceptibility of each wild-type IN was determined by using strand transfer substrate of both HIV-1 and HIV-2. It was demonstrated that each IN was slightly more susceptible to RAL when using its corresponding substrate, compared with using the other substrate (**Figure 27**). The EC_{50} of RAL for HIV-1 wild-type IN increased by a factor of about 3 when using HIV-2 substrate, compared with using HIV-1 substrate; and a factor of 1.6 for HIV-2 wild-type IN when using HIV-1 substrate, compared with using HIV-2 substrate.

Interestingly, it was observed that the strand transfer product profiles of the two wild-type INs differed when same substrate was used (either HIV-1 or HIV-2 oligonucleotide substrate). It suggests that HIV-1 and HIV-2 wild-type INs might have preferential site of integration on oligonucleotide substrate due to the variations of amino acids and different length of enzyme. Additionally, compared with HIV-2 wild-type IN, 143C mutant and 143C/155H mutant also exhibited different strand transfer product profiles. Hence residue at some position, e.g., position 143, may impact on the favored site of integration.



	EC50
HIV-1 ST substrate	2.8e-008
HIV-2 ST substrate	8.5e-008

	EC50
HIV-1 ST substrate	4.1e-008
HIV-2 ST substrate	2.6e-008

Figure 27: The susceptibility to raltegravir of HIV-1 and HIV-2 wild-type integrases with respect of two viral substrates in strand transfer assay.

(A) The susceptibility to raltegravir of HIV-1 IN using HIV-1 and HIV-2 viral substrates. (B) The susceptibility to raltegravir of HIV-2 IN using HIV-1 and HIV-2 viral substrates. EC_{50} is shown under each chart. ST reactions were carried out, in the presence of increasing RAL concentrations, at a concentration of 100 nM IN.

II.4. Conclusion

HIV-2 is endemic in West Africa and has spread through Europe during the last two decades (Matheron et al., 1997; Reeves and Doms, 2002; Soriano et al., 2000). Though HIV-2 appears to be significantly less pathogenic than HIV-1 (Spira et al., 2003), HIV-2 infected patients have limited options of antiretroviral treatment. The strand transfer inhibitor of IN, RAL, as a novel class of antiretroviral drug, represents a good therapeutic option for HIV-2 infected patients who obtained a good immunological and virological response (Damond et al., 2008). However, as observed in other classes of antiretroviral agents, the resistance of RAL emergences with a low genetic barrier.

Our study recovers 3 primary pathways of RAL-resistance in HIV-2 by in vitro enzymatic assays, involving N155, G140/Q148 and E92Q/Y143 mutations, respectively. RAL-resistance associated mutations E92Q, T97A, G140S, Q148R, N155H and Y143C were identified by amplifying and sequencing IN coding sequence from HIV-2 infected patients at the initiation of RAL-containing treatment and at time of virological failure.

Mutations at positions 155 and 148 play a similar role for HIV-1 and HIV-2 in conferring resistance to RAL, with a remarkably reduced susceptibility. N155 H resistance pathway in HIV-2 confirmed by our in vitro study using site-directed mutagenesis is consistent to the recently report (Salgadoa et al., 2009), although other data suggest that this mutation pattern might be favored in the IN context of HIV-2 group B strains (Roquebert et al., 2010). Moreover, a sustained low HIV-2 viral load is able to maintain N155H pathway overtime (Salgadoa et al., 2009).

As for HIV-1 (Delelis et al., 2009a), the catalytic defect due to the Q148R mutation can be restored in HIV-2 IN by the secondary G140S mutation conferring little resistance per se, due to the conformational rearrangement compatible with enzymatic activity. It was presumed that G140S mutation converts the inactive conformation form of Q148R harboring mutant into an relatively active form (Hare et al., 2010c), via the physical proximity and interaction between the two mutations (Delelis et al., 2009a), hence increases viral fitness. The G140 S /Q148 H double mutant which is the most common mutation pattern observed in HIV-1 infection confers the highest level of resistance to RAL while remains the capacity of replication (Ceccherini-Silberstein et al., 2009). The advantages of G140 /Q148 pathway rationalize the switch from the N155 to the Q148 pathway observed after a few weeks of RAL treatment (Delelis et al., 2009a).

Conversely, in contrast to HIV-1 (Delelis et al., 2010), Y143C in HIV-2 is not sufficient

by itself to explain the resistance of viral genomes harboring this mutation. Recent report was shown that Y143R mutant of HIV-1 neither confer resistance to another INSTI elvitegravir (Métifiot et al., 2011). Other variations which were observed in Y143C-containing viral genomes isolated from patients, such as E92Q, assist Y143C to confer RAL-resistance. The side chain of the residue at position 143 was thought to interact with bound RAL, hence to be a significant factor to INSTI binding (Hare et al., 2010b). According our experiment data, though being insufficient to impact on drug susceptibility by itself, the substitution of Y143 in HIV-2 IN is able to confer drug resistance together with the mutation at position 92. We assume that polymorphic differences within the HIV-2 IN coding sequences possibly lead to the variations in contribution of mutations, e.g., Y143, to drug resistance. It was speculated that different viral types may favor various mutational pathways, consequently resulting in varying levels of drug resistance among different virus types (Bar-Magen et al., 2010; Wainberg, 2004). The resistance pathway was also suggested to alternate among HIV-1 non-B subtypes (Wensing et al., 2009).

Additionally, the Y143C mutation counteracts the resistance effect of N155H, which may explain why both mutations were not found in single viral genome.

T97A mutation is one of the polymorphic substitutions in HIV-2 IN; and this position is known to be associated with in vitro or in vivo resistance to INI compounds in HIV-1 isolates (Roquebert et al., 2008b). Our in vitro enzymatic assay suggests that T97A doesn't confer the RAL-resistance, nor together with Y143C mutation.

It was observed that HIV-2 wild-type IN showed a better dissolubility and higher activity level on both 3'-processing and strand transfer assays, compared with HIV-1 wild-type IN. We speculate that because the C-terminus of IN is involved in host DNA binding and positioning, its length might lightly effect the interaction of IN and viral substrate. HIV-2 IN may have a better competence on DNA binding and positioning, along with a better dissolubility under in vitro assay condition.

Both HIV-1 wild-type IN and HIV-2 wild-type IN were lightly more susceptible to RAL when using its corresponding substrate, compared with using the other substrate. It was revealed that INSTIs bind to the catalytic metal cations and inactivate the IN by blocking its active site, while the hydrophobic group of INSTI fits within a tight cavity close to DDE catalytic triad of IN (Chen et al., 2008; Hare et al., 2010b), and displaces the reactive viral DNA end from the active site of IN, resulting in deactivation of IN (Hare et al., 2010b). The INSTI binding cavity forms following a conformation change induced by the binding and 3'-processing of the viral DNA (Chen et al., 2008). Viral DNA may well form part of

the inhibitor binding cavity. Indeed, the diketo acid inhibitor can only bind to the acceptor site after 3'- processing. Moreover, a conserved recognition mechanism based on both direct and indirect readout was suggested. Specific binding need conserved amino acids in the α4 helix of IN and conserved nucleotide bases and backbone groups at viral LTR ends (Hobaika et al., 2009). Thus, it is conceivable that binding site of INSTI is better formed when IN binds with its proper, conserved viral DNA substrate, and hence the complex is more susceptible to INSTI.

III. Development of anti-HIV-1 integrase camelid simple-domain antibody

III.1. General introduction of single domain antibody

The field of recombinant antibody technology has rapidly progressed during the last two decades, mainly because of the interest in their human therapeutic use. The ability to select specific human antibodies by display technologies and to improve their affinity, stability, and expression level by molecular evolution has further boosted the field.

As well as a useful research tool, murine monoclonal antibodies opened the way to the development of diagnostics and human therapeutics. However, the application of monoclonal antibodies is constricted from currently technical limitations and drawbacks such as: the high cost of production based on mammalian cell culture systems; inability to optimize the antibody with genetic engineering techniques; and the potential immunogenicity of the mouse antibody to human (Kuus-Reichel et al., 1994). The latter limitation can be overcome by developing chimeric antibodies or humanised antibodies which are available on market as human therapeutic mAbs. The expression of the antibody genes in bacteria could be a way to reduce the production cost and facilitate the genetic engineering for antibody optimization. Certain antibody fragments have been successfully expressed in bacteria, yeast, or fungi, such as Fab, Fv (non-covalently associated heterodimers of VH and VL domains), and single-chain Fvs (scFvs).

ScFv is an engineered format in which the VH and VL domains are joined with a flexible and hydrophilic polypeptide linker preventing dissociation of native Fv fragment **(Figure 28)**. These fragments usually retain the specific, monovalent, antigen-binding affinity of the parent IgG, while showing improved pharmacokinetics for tissue penetration. In addition, protein engineering can modify the functionality and physico-chemical properties of the fragments leading to improved stability, reduced immunogenicity, better tissue distribution and faster blood clearance when used in vivo (Wu et al., 1996). Despite these advances, major technical hurdles remained in implementing the technology on industrial scale. For instance, generating a representative library is tedious; and the genetic constructs are often unstable in the bacterial host. Also, the expression yield and stability of scFv often turn out to be problematic; affinity of scFv decreases compared with parent antibody (Borrebaeck et al., 1992; Mallender et al., 1996); linker induces aggregation and

is easily degraded by proteolysis (Whitlow et al., 1993).

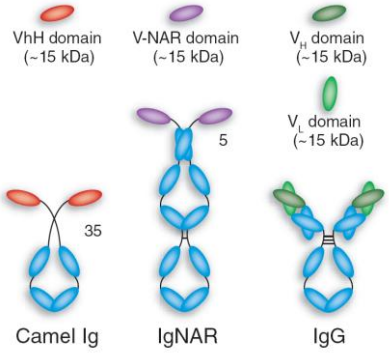
Antigen-binding fragments comprising the single variable domain (VH) of the conventional heavy chains have also been generated (Ward et al., 1989). This approach avoids the introduction of a peptide linker used in the scFv. However, a large hydrophobic surface of the VH exposes to the solvent without the VL domain. Affinity drops by one to three orders of magnitude compared with a scFv (Borrebaeck et al., 1992). Moreover, the insolubility of the isolated VH domains expressed in bacteria posed serious limitations for their use.

Interest was revived when it was discovered that at least two types of organisms, the camelids (camels and llamas) and cartilaginous fish (wobbegong and nurse sharks), have evolved heavy-chain antibodies (HCAb) as an integral and crucial component of their immune system **(Figure 28)** (De Genst et al., 2004; Dooley and Flajnik, 2005; Streltsov, 2004). These peculiar HCAbs lack light chains (and, in the case of camelid antibodies, also the CH1-domain). The high-affinity variable domains of the single heavy chain are called VHH (variable domain of heavy chain of HCAb), and more generally, nanobody or single domain antibody (sdAb). As such, sdAb is the smallest available intact antigen-binding fragment (13 kDa) derived from a functional immunoglobulin.

The selection of the VHH with phage display technology by panning and expression in bacteria is an attractive alternative to obtain small molecular recognition units. As is frequently observed, favorable properties, such as good expression, stability and solubility, are co-selected with binding activity. In addition, improved folding yield can be obtained by selecting V domains that resist aggregation when unfolded on phage.

Only one domain needs to be cloned and expressed to generate an intact in vivo matured antigen-binding fragment, and its intrinsic characteristics, VHH obtained from an immunized camelid have a number of advantages compared with the Fab, Fv or scFv derived from other mammals.

A



B

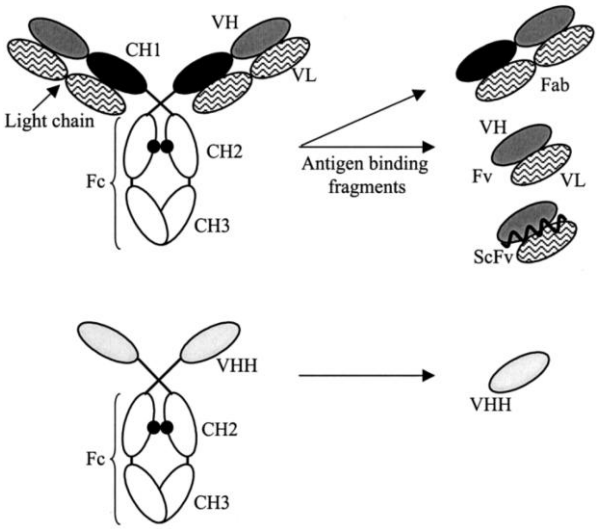


Figure 28: Schematic representation of different antibody formats

(A) Intact 'classic' IgG molecules alongside camelid VHH-Ig and shark Ig-NAR immunoglobulins. Camelid VhH-Ig and shark Ig-NARs comprise a homodimeric pair of two chains of V-like and C-like domains (neither has a light chain), in which the displayed V domains bind target independently. Shark Ig-NARs comprise a homodimer of one variable domain (V-NAR) and five C-like constant domains (C-NAR) (Holliger and Hudson, 2005a). (B) Schematic illustration of the conventional (top) and heavy-chain IgG antibodies (bottom) present in camelid serum. The antigen-binding domains of conventional antibodies obtained after proteolysis (Fab) or after cloning, and expression of the gene VH and VL fragments are shown. A linker introduced between the VH and VL stabilizes and forms scFv. Recombinant VHH is the variable domain of heavy chain of HCAb (Muyldermans, 2001).

III.1.1. VH and VHH sequence differences

Approximately half of the natural serum IgG repertoire of dromedaries consists of HCAbs, llamas having a lower proportion of HCAbs (25 to 45%), and the remaining being conventional antibodies. At least two conventional IgG isotypes and three or four heavy-chain IgG isotypes in dromedary and llama were identified, respectively (Nguyen et al., 2000b). The heavy chain of HCAb is composed of a variable domain (VHH) immediately followed by a hinge, the CH2 and the CH3 domain **(Figure 28B)**. The equivalent of the CH1 domain between the VH and the hinge is missing. CH1 sequences are removed during splicing due to a point mutation at the 5' end of the CH1-hinge intron (Nguyen et al., 1999), and it was shown that no camel-specific trans-acting factors are involved in the unusual splicing (Nguyen et al., 2003).

The genome of the dromedary (and probably llama as well) encodes separate VH and VHH germline genes (Nguyen et al., 2000b). The VH and VHH amino acid sequences share a high degree of identity and are most similar (~80%) to the human VH of family III (Vu et al., 1997), the most common human VH family in terms of both expression and genome complexity (Tomlinson et al., 1992). The amino acids at positions that determine the typical immunoglobulin fold are all well conserved in the VHH. However, four amino acids that are extremely well conserved in all VHs are constitutively substituted in the VHH. These residues (Val37Phe (or Tyr), Gly44Glu (or Gln), Leu45Arg (or Cys) and Trp47Gly (or Ser, Leu, Phe) (Kabat numbering (Kabat et al., 1991), in going from VH to VHH) found in the framework-two region discriminate the conventional VH from the heavy-chain specific VHH. Overall, they render this side of the domain more hydrophilic, presumably lowering propensity of VHH to form a heterodimer with a VL domain.

Besides these hallmark substitutions in the framework, the hypervariable regions of VHH are longer than those of VH generally. The enlarged hypervariable regions of the VHH could increase the actual surface of the antigen binding site and hence compensate for the absence of the antigen-binding surface area provided by the VL. The complementarity determining region 3 (CDR3) is longer in VHH, on average, than in VHs (16 to 18, 12 and 9 amino acids in dromedary VHHs, human VH and mouse VH, respectively) (Vu et al., 1997; Wu et al., 1993). Furthermore, the first hypervariable region CDR1 that normally involves residues 31-35 in VH is enlarged to encompass residues 27-35 in VHH. The hypervariability of residues 27-30 suggests that they are contacting the antigen and that the somatic mutations in this area will be selected during the affinity maturation. A limited number of VHH germline genes (~40) have been identified in the

dromedary genome; this number is approximately half that of functional human VH genes (Cook and Tomlinson, 1995). However, the repertoire of the primary VHH domains is apparently further diversified by active somatic mutation mechanisms (Nguyen et al., 2000a).

The extended CDR3 is frequently stabilized by an additional disulfide bond connecting the CDR3 to the adjacent CDR1 loop or a Cys45 of framework-two (common in camel and shark sdAbs) or to a Cys50 within the CDR2 loop (common in llama sdAbs) **(Figure 29)**. Cys residues at these positions are absent in all VH germline genes including those of the dromedary (Nguyen et al., 2000a). These additional Cys residues form an inter-loop disulphide bond that stabilizes the VHH domain. Furthermore, this bond is expected to impose conformational restraints on the loop flexibility in the absence of antigen so that the entropic penalty upon antigen binding is minimized.

Overall, the CDR1 extended towards the N-terminal end, the somatic hypermutations selected during the affinity maturation process, and a longer CDR3 could increase the diversity of the VHH domains. This might account for the broad antigen-binding repertoire of VHH in the absence of the VH-VL combinatorial diversity.

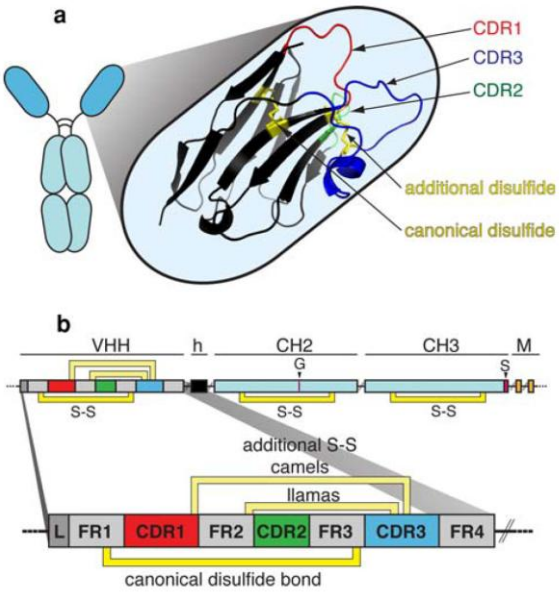


Figure 29: Schematic diagram of the VHH domain of a camelid heavy chain antibody. (a) The three CDRs of the antigen-binding paratope. (b) Many camelid antibodies contain an additional disulfide bond connecting the CDR3 with the end of the CDR1 (camels) or the beginning of the CDR2 (llamas). h Hinge, M transmembrane domain of membrane isoform, G glycosylation site, S stop codon of secretory isoform (Wesolowski et al., 2009).

III.1.2. Structure of VHH

Structure studies showed that VHH scaffold adopts the typical immunoglobulin fold and superimposes on the conventional VH structure (Desmyter et al., 1996). These necessary adaptations took place in the VHH to cope with the absence of a VL domain are concentrated in two areas: the 'side' of the domain that corresponds to the VH-side contacting the VL, and the CDRs themselves. Nine β-strands are folded in two sheets that pack against each other and are stabilized by a conserved disulfide bond. Meanwhile, due to the nature of the VHH-specific amino acid substitutions cluster, the surface formed by the framework-two residues of the VHH is reshaped into a much more hydrophilic region. This hydrophilicity effect is augmented by rotation of the side chains of adjacent residues to expose their most hydrophilic parts to the solvent.

Remarkably, the largest structural difference between a VH and VHH was found at the hypervariable regions. In human and mouse VH, these loops are folded in a limited number of structures, known as canonical structure. In contrast, the antigen-binding loop of VHH possesses apparently more structure deviations on these canonical structures that could not have been easily predicted from their sequence. Accordingly, the antigen-binding loops of VHH exhibit a much larger structural repertoire which is probably due to the additional freedom generating by the absence of VL.

III.1.3. Unique features of sdAb

Besides easy cloning and selection from an in vivo matured library, there are many other unique advantages of sdAb for biotechnological applications. High microbial production is an important advantage of sdAb. Due to the unique hydrophilic substitutions in framework-two, the sdAbs can be efficiently expressed in the form of soluble and non-aggregating recombinant proteins. SdAb fragment can be expressed towards the periplasm of the versatile bacterial expression systems. In the oxidizing environment of periplasm, proteins are folded and the formation of disulfide bond is stimulated, resulting in the production of functional sdAb.

The sdAb domains are naturally soluble because of the substitutions of hydrophobic residues by hydrophilic ones in the framework-two region. In addition, the absence of engineered structures, such as linkers which may lead to aggregation as observed for scFv, promotes the solubility of sdAb. Contrary to conventional antibodies, sdAb also showed a very remarkable stability, a good thermal resistance. This high apparent stability is mainly

attributed to their efficient robust property and refolding after chemical or thermal denaturation and to a lesser extent because of an increased resistance against denaturation (Dumoulin et al., 2002; Ewert et al., 2002; Perez et al., 2001). The augmented apparent stability is also probably in virtue of an increased hydrophilicity of the former VL interface region (Conrath et al., 2005). In addition, the packing of extended CDR3 loops against this former VL interface also contributes to domain stability(Bond et al., 2003).

The sdAbs isolated from the immunized libraries are highly specific for the target antigen with nanomolar affinity and do not cross-react with other non-specific antigens. In contrast with the reduced diversity of conventional antibody libraries due to reshuffling of VL and VH domains during library construction, full functional diversity is retained in VHH libraries generated from immunized camelid. Furthermore, the single domain nature facilitates subsequent molecular manipulation for many applications, e.g., multivalent antibody. Multivalent format antibodies are able to increase functional affinity (avidity) or to produce bispecific antibody fragments binding to different antigens simultaneously. SdAbs are good candidates for multivalent formats because they allow more flexible linker design, which is important for simultaneous binding of antigens, without the domain-mispairing problems in the case of scFv production (Coppieters et al., 2006; Roovers et al., 2007).

SdAbs have demonstrated improved penetration against cryptic (immuno-evasive) target antigens. They recognize unique ‘buried’ epitopes that are currently out of reach for conventional antibodies. Their smaller size together with the enlarged and protrude CDR3-loop particularly accounts for most of the penetration into recessed antigenic sites and the antigen binding interaction (De Genst et al., 2006; Desmyter et al., 1996). Indeed, to escape immunosurveillance, many pathogenic viruses have evolved narrow cavities (canyons) in their surface antigens, which bind their target receptors but are poorly accessible to intact antibodies and are thus largely immunosilent. This ‘blind spot’ of the antibody response is caused by the limited diversity of CDR loop lengths, which constrains the displayed antigen-binding surfaces to mostly flat or concave topologies, depending on the antigen being a protein or hapten, respectively. Meanwhile, grooves and cavities are essential in various biological activities as involved in specific interaction site between two molecules, e.g., receptor-ligand or enzyme-substrate. Antibodies able to specifically bind into such active sites therefore represent powerful tools for biological research.

The high identity degree of human VH and dromedary VHH were demonstrated by sequence comparison. The smaller size and the homology to human antibodies of camel VHH render the ‘humanisation’ more straightforward compared with the murine

counterparts. Based on the structural information, most of the substitutions other than the key camel specific substitutions in framework-two can be replaced by the human sequence without altering the performance of the VHH, and therefore VHH can be good candidate to humanize by veneering. Evidently, the VHH-specific residues in the framework-two should not be modified since they are either involved in the interaction with the enlarged CDR3 or they are solvent exposed and essential for the single-domain nature and the unique physico-chemical features of the isolated VHH.

In addition, in virtue of their small size of about 13 kDa, sdAbs rapidly pass the renal filter, which has a cutoff of about 60 kDa, resulting in their rapid blood clearance. The small size results in a fast tissue penetration. This is advantageous, for instance, targeting of sdAbs coupled to toxic substances to tumors (Cortez-Retamozo et al., 2004).

III.1.4. Biotechnological application of sdAb

Apparently, the unique features of sdAb could lead to enormous applications where it should perform better than other antibody formats. Numerous sdAbs have already proven useful for basic research such as the capturing reagents in immunoaffinity purification (Verheesen et al., 2003), or as improved diagnostic tools and detection units on biosensors (Pleschberger et al., 2004). In vivo studies have underscored the favorable biodistribution of sdAbs, including deep penetration into dense tissues and rapid elimination via the kidney. These features make sdAbs particularly amenable for imaging of tumors and for the delivery of cytotoxic agents. SdAbs should also prove useful for neutralizing soluble extracellular proteins including toxins, cytokines, and blood clotting components. The capacity of sdAbs to block leukocyte ecto-enzymes indicates their potential utility as anti-inflammatory and immune modulating agents (Koch-Nolte et al., 2007). Moreover, several VHHs are now being studied for use in various disease areas, including oncology (Revets et al., 2005) and in infectious, inflammatory, and neurodegenerative diseases.

Concerning on antiretroviral therapy, there are at least two applications for the treatment of HIV infection. Firstly, sdAbs recognizing HIV envelope proteins could constitute a new generation of neutralizing antibodies which block virus replication. By preferentially binding to the ‘buried’ regions which are inaccessible to conventional antibodies, sdAbs access to those epitopes much more conserved than the dominant epitopes normally targeted by conventional antibodies. Secondly, sdAb can be developed as intracellular antibody or intrabodies. Besides the injection of antibodies against the circulating molecules, recombinant antibody can be expressed by the target cells themselves to against

intracellular viral components or be implicated in malignant transformation (Caron de Fromentel et al., 1999; Cochet et al., 1998; Marasco, 1997; Phelan et al., 1998). These intrabodies can neutralize or modify the function of the molecules, inducing the death or the back to normal function of the targeted cells. This approach had led a big hope for the development of anti HIV-1 antibody over 1990s. Unfortunately, enormous studies just led to a few publications concerning three anti-HIV Abs with intracellular anti viral protein activity. In 1995, it was reported that an Ab fragment scFv against Tat could inhibit the transactivation of HIV-1 LTR and the viral infection when the scFv was expressed intracellular (Mhashilkar et al., 1995; Mhashilkar et al., 1999). An intracellular expressed Fab fragment against HIV-1 matrix was able to block the viral replication at the early and late step of viral life cycle (Levin et al., 1997). In addition, a rabbit scFv fragment against Vif showed neutralizing effect on HIV-1 reverse transcription and replication when expressed in the cytoplasm of infected cells (Goncalves et al., 2002). Since then, no characterization study concerning these three anti-HIV-1 fragments was published. In contrast, the vast majority of studies demonstrated the limits of this approach. Except the rare exceptions, these fragments derived from conventional Abs expressed in cytoplasm of eukaryotic cell were unable to form intra-domain disulfide bond which is necessary for the stability and folding. They generally accumulate in the form of inclusion bodies without any biological activity, consequently. The discovery of the surprising properties of camelid sdAb might remove those limitations since most sdAbs are functional without disulfide bond and can be used as a probe for detecting cytoplasmic or nuclear protein, or as an inhibitor to block one or several activities of the targeted proteins (Muyldermans, 2001; Verheesen et al., 2006).

III.2. Development of camelid single-domain antibody targeting HIV-1 integrase

III.2.1. Introduction

Despite the availability of 23 approved drugs for the treatment of HIV/AIDS (Marchand et al., 2009), the great interest in developing new antiretroviral agents maintains due to the following reasons. Firstly, though the application of HAARTs obtains a remarkable improvement of decreasing morbidity and mortality of HIV infected patients, these treatments currently are unable to eradicate virus infection. Secondly, effective vaccine or comparable preventative is unlikely to be available in the near future, and the epidemic will therefore be sustained in endemic areas and increase in new regions where the virus is introduced. Furthermore, emergence of resistance to existing drugs serves as a main impetus for the development of new drug classes with distinct inhibition mechanism or agents in classes without cross-resistance (Flexner, 2007). Finally, better tolerated, more convenient and less expensive treatments should also be taken into account. Therefore, intensive researches leading to the identification of new viral targets and characterization of new technological approaches allowing a direct inhibition are still needed for the development of new antiretroviral strategies.

Integrase (IN), one of the three essential viral enzymes, is responsible for inserting HIV viral DNA into cellular genome which is an indispensable step in viral replication. The process of integration involves two independent reactions: 3'- processing and strand transfer, both catalyzed by IN within a nucleoprotein complex, the pre-integration complex (PIC) (Lewinski and Bushman, 2005). Within PIC, IN interact with other viral and cellular cofactors, such as LEDGF/p75 which facilitates the access to chromatin and ensures the integration in specific sites (Cherepanov et al., 2002; Maertensm et al., 2003). The stability of the PIC and its presence in cell during the most times of pre-integration also make IN an attractive target (Smolov et al., 2006b). This hypothesis is confirmed by the remarkable antiviral activity of IN strand transfer inhibitors (INSTIs) as a new class of antiretroviral drug. The first approved INSTI, raltegravir, which entered the therapeutic arsenal in 2007, demonstrated to be a very efficient inhibitor in vivo. INSTIs act by binding to the viral DNA/ IN complex near to the 3' end of viral DNA and thus block the strand transfer with IC_{50} values in nanomolar range. Nevertheless, INSTIs can not escape the drug-resistance caused by the mutations in IN with a low genetic barrier (Malet et al., 2008b). Therefore,

the research of new inhibitors and especially the compounds which are able to block the strand transfer activity of the broadest range of INs, including INs resistance to INSTIs, remains a primary subject of anti-HIV pharmacology. The neutralizing antibodies of IN, which were developed using monoclonal antibodies (Yi et al., 2002), represent an highly attractive way of research. The accessibility of IN within PIC to intracellularly expressed scFvs was also demonstrated (Levy-Mintz et al., 1996). However those developments were not continued due of the inherent limitations of conventional antibodies. For example, disulfide bond between the tandem VH and VL chains of scFvs is unlikely to form under a reducing environment, such as the cytoplasm and the nucleus, resulting in the lower or no affinity of scFvs (Biocca et al., 1995). Being endowed with many attractive properties, camelid single-domain antibody (sdAb) is a promising candidate for the development of antiretroviral agent.

Particular single-heavy chain antibody devoid of light chain were identified in camelids (Hamers-Casterman et al., 1993). They are able to recognize their antigens via the variable domain of heavy chain (VHH). The VHH domain of 13 kDa can be easily produced as sdAb, a recombinant fragment. SdAbs recognize different epitopes normally inaccessible for the conventional Abs. In virtue of its extended CDR3 loop, sdAb penetrates into the cavities located on antigen surface, and blocks the cryptic sites such as these conserved active sites of enzymes (Desmyter et al., 2002). Additionally, sdAbs exhibit more homology with human VH, especially type III, than with mouse VH (Su et al., 2002; Vu et al., 1997). They may provoke less immunoreaction in human than the murine VH region of chimeric antibody. Moreover, sdAb exhibits a remarkable stability (Dumoulin et al., 2002). Indeed, the remarkable stability of sdAb facilitates the folding in a soluble and active form under a reducing environment. Hence it is more possible to obtain the intracellular active sdAbs (Muyldermans, 2001; Verheesen et al., 2006).

SdAbs targeting different HIV viral components could be developed as new antiretroviral strategies able to block virus replication. These sdAbs also could be efficient tools to elucidate the molecular functions of the targeted proteins during different steps of the virus life cycle. Some sdAbs against HIV viral proteins were successfully developed. For example, sdAb fragments targeting HIV-1 Rev or Vif proteins have been characterized and displayed antiviral activity in cell-culture assays (Aires da Silva et al., 2004; Vercruysse et al., 2010). A sdAb was demonstrated able to specifically bind HIV-1 Nef protein with a high affinity and inhibit critical biologic activities of Nef, recently (Bouchet et al., 2011).

In this study, we developed sdAb targeting HIV-1 IN as a promising candidate for the

development of antiretroviral therapy and for fundamental research.

III. 2. 2. Materials and methods

1. Llama immunization

Most researches on the biotechnological application of single domain antibodies were performed using camelids because of their ease of immunization and handling. There are several options to identify an antigen specific sdAb. When the antigen is available, VHH library could be constructed from animal immunization and the cloning of VHH pool amplified from lymphocytes into a phage display vector from where to select for antigen-specific VHHs by panning with immobilized antigen. The alternative strategy is the generation of a large single pot naïve VHH library from which binders are retrieved after panning (Tanha et al., 2001). In a third approach, a human VH was reshaped in its framework 2 region to mimic the VHH (camelization) (Davies and Riechmann, 1994). The CDR3 codons of such a 'camelized' VH were then randomized and cloned in a phage display vector to obtain a 'synthetic' library from which to retrieve binders (Davies and Riechmann, 1995). The latter two strategies have the advantage that the immunization is avoided, but the disadvantage that lower affinity binders are often obtained which need additional mutagenesis steps to arrive at valuable antigen recognition units. Therefore, it might be more convenient to use 'immune' VHH libraries to arrive at antigen-specific VHH with proper characteristics for applications.

Dromedaries and llamas are easily immunized following standard procedures (Murphy et al., 1989). Immunization with as low as 50µg immunogen in complete and incomplete Freund's adjuvant generates antigen-specific HCAb of high affinity after 5 to 6 weekly boosts. In this study, a llama (Lama glama) was immunized by subcutaneous injections of the purified recombinant HIV-1 wild-type IN protein (HXB2 HIV IN) (Leh et al., 2000), the second llama was immunized by recombinant human foamy virus (HFV) IN. The third was a negative control. Blood samples were taken 15 days after the last injection, and peripheral blood mononuclear cells were isolated as described (Behar et al., 2009). Llama management, inoculation, and sample collection were conducted by trained personnel under the supervision of a veterinarian, in accordance with protocols approved by the Centre National de la Recherche Scientifique's ethical committee of animal welfare.

2. Construction of sdAb phage-display library

Cloning the repertoire of antigen-binding fragments from an immunized animal into a

phage display vector and selection of antigen-specific clones by panning became a routine method to identify specific binders. Moreover, the single-domain nature of the VHH considerably simplifies the work. Obtaining the antigen-binding repertoire of the HCAb of an immunized camelid is less complicated, since the intact antigen-binding fragment is encoded in a single gene fragment.

SdAb libraries are usually generated by amplifying the V-domain repertoire from blood, lymph node, or spleen cDNA obtained from immunized dromedary or llama and cloning into a phage display vector, such as pHEN. Belonging to one single gene family, all VHHs are encoded by a single exon with homologous flank sequences. Hence, the complete in vivo matured VHH repertoire can be amplified by one single set of primers and can be cloned conveniently. Two amplification methods, both based on the absence of VH in HCAb and on the inherent differences between the two antibodies, were developed to eliminate the VH gene fragments which might complicate subsequent pannings due to their possible sticky behaviors from the VHH pool **(Figure 30)**. The first is using hinge specific primers annealing complementarily on the hinge of the HCAb. The hinge harbors the largest sequence differences between the different isotypes, and therefore primers can be designed amplifying the VHH genes specifically. Alternatively, a pan-annealing primer binding to a conserved region of all constant - genes is first employed in combination with a back primer binding at the leader signal sequence of the VH and VHH. Hence, VH and VHH fragments of all IgG isotypes will be amplified. Two kinds of products are yielded. The longer products containing the CH1 exon is derived from the heavy chain of a conventional antibody, and the shorter one lacking the CH1 exon is derived from the heavy chain of the HCAb. VH sequences can be eliminated easily and efficiently just by separating PCR fragments on agarose gel and recovering the shorter fragment. To introduce restriction enzyme sites for cloning and obtaining more products, a secondary PCR using nested primers locating at the two ends of the VHH can be performed.

Then PCR products are cloned into a phagemid vector pHEN, which is a hybrid of phage and plasmid (**Figure 31**). It comprises a M13 replication origin and an Escherichia Coli plasmid replication origin. The inserted anti-HIV IN sdAb gene is located at the downstream of an inducible bacterial promoter (pLac), in-frame behind a periplasm signal leader sequence (PelB) and upstream of two epitope tags (6xHistine and c-Myc), an amber stop codon followed by the coding sequence of the M13 phage head capsid gIII protein (gIIIp). Sfi I and Not I sites were also introduced. The phagemids were transformed into an E. coli strain with amber suppressor genes (e.g., E. coli TG1r) that can read through the amber stop codon. Hence, sdAb can be expressed as a fusion protein with gIII protein and

assembled on the head of M13 phage.

In contrast to scFv libraries where relatively large libraries are required, smaller libraries of only 10^6–10^7 VHH individual clones can routinely result in the isolation of single domain proteins with nanomolar affinity. Semi-synthetic libraries by cassette-mutagenesis of the CDR regions were also constructed successfully. The semi-synthetic libraries offer the advantage of selecting antibodies against toxic or difficult to express antigens.

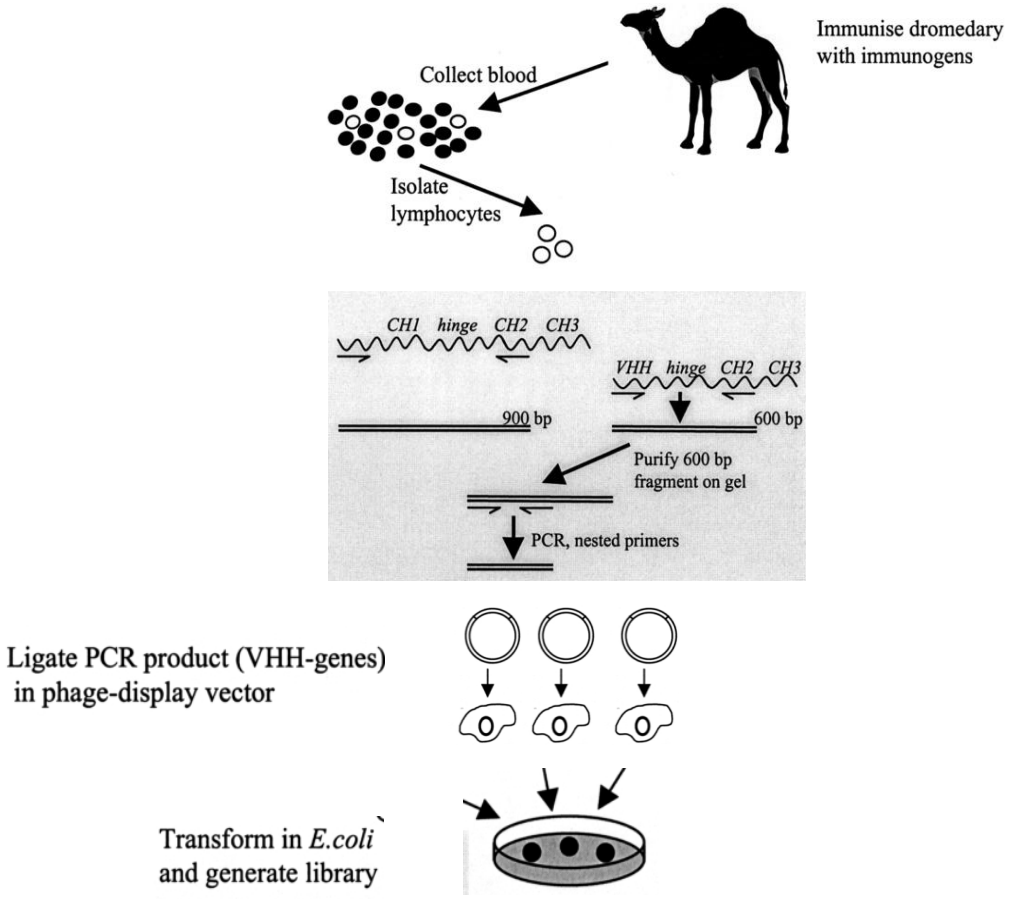


Figure 30: Construction of sdAb phage-display library (Muyldermans, 2001)**.**

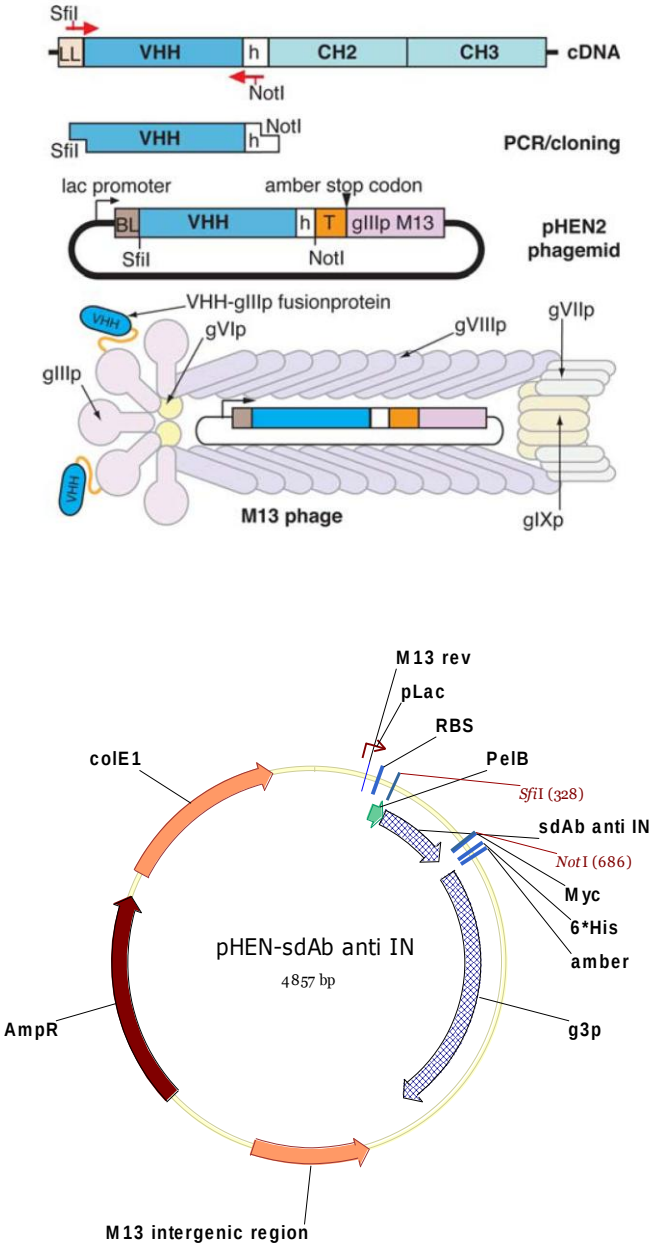


Figure 31: Schematic diagram of the M13 phage display vector used for cloning libraries of the VHH repertoires of immunized camelids.

(A) The coding sequence of the VHH from HcAbs is amplified from cDNA generated from immunized animal. Then PCR products are cloned into the phagemid vector downstream of an inducible bacterial promoter (arrow), in-frame behind a periplasm signal sequence (BL, bacterial leader peptide) and upstream of one or two epitope tags (T), an amber stop codon followed by the coding sequence of the M13 phage head capsid protein gIIIp (Wesolowski et al., 2009). (B) Schematic presentation of phagemid pHEN inserted with anti-HIV IN sdAb coding sequence.

3. Enzyme-linked immunosorbent assay (ELISA) to confirm the immunoreactions in immunized llamas.

50 µl/well of 1 µg/ml ~ 10 µg/ml recombinant HIV-1 wild-type IN or human foamy virus (HFV) IN was coated in 96-well plate overnight at 4℃. Recombinant INs were firstly diluted in PBS (pH ≈ 7-7.5). After blocking non-specific sites with 5% PBSM for 1 hour at room temperature, gradient diluted llama immune sera (1/50, 1/250, 1/1250, 1/5000 in PBS) were incubated for 1 hour at room temperature. Then llama anti-IN antibodies were detected with 1/100 000 diluted goat anti-llama IgG-heavy and light chain Ab-horseradish peroxidase (HRP) (Bethyl lab) using horseradish peroxidase substrate ABTS (Roche), then read at 405 nm.

4. Western blot to confirm the immunoreactions in immunized llamas.

0.05 µg/well of recombinant INs of HIV-1 wild-type, INSTI resistance mutant G140S/Q148H, and four HIV-1 CRF02_AG subtype strains 33CR, 49CR, 52CR Q148K, 68CR were deposed to verify the antigen-recognizing ability of llama immune sera. After incubated with 1/1000 diluted llama sera for 1hour at room temperature, INs were detected with 1/40 000 diluted goat anti-llama IgG-heavy and light chain Ab- horseradish peroxidase (HRP) (Bethyl lab), then developed by western blot detection reagents (Amersham ECL plus).

5. Phage-display library rescue

The phage library must be sufficiently large to encompass the immune repertoire of an animal to maximize the probability of obtaining an antibody of interest. When rescuing a phage-display library (preparing phage stocks from a bacterial stock), it must be aware of the numbers of relevant species (phage, total bacteria, viable bacteria, and infected bacteria) at each stage to ensure a successful rescue. If any of these steps go wrong, the full diversity of the library might not be affected. This is especially critical for large non-immune libraries of about 10^{10} members. For smaller immune libraries (about 10^6-10^7 complexity), it is much easier to retain the full diversity.

The following protocol can be used for rescuing a large non-immune library of about 5 x 10^9 to 1 x 10^{10} complexities. Generally, 1.0 OD_{600} of TG1r cells corresponds to 1.1 x 10^8

bacteria per ml. Inoculate pre-warmed 500 ml (or 250 ml) 2YT medium (supplemented with 100 μg/ml ampicillin and 1% glucose for negative regulation of pLac) with an aliquot of the frozen library (glycerol bacteria stock, about 50 μl) to an OD_{600} of about 0.05. The quantity of stock used is determined by the sufficiency of bacteria quantity to ensure the presence of all members of the library and the controllable total volume. For example, 500 ml of library with OD_{600} of 0.05, that is (0.05) x (1.1×10^8) x (500) = 2.75×10^9 bacteria. If the complexity of library is about 10^8, it means about 20 representatives of each clone present.

OD_{600} of the diluted library should be about 0.05 or less. It ensures the bacteria will at least augment 3.5 times to reach OD_{600} of 0.6. Then check the number of viable bacteria. For example, OD_{600} of 0.05 is 5.5×10^6 bacteria per ml, so dilute 1:100 and spread 10 and 1 μl aliquots on plate, which should be about 550 and 55 colonies respectively. Otherwise, the full library can not be obtained or the initial volume of bacteria stock should be augmented. Incubate the culture at 37 °C and 200 rpm until OD_{600} of 0.6.

Add sufficient helper phage M13KO7 to a final concentration of about 2×10^{12} phages per liter. M13KO7 is an M13 derivative which carries the mutation Met40Ile in gII, with the origin of replication from P15A and the kanamycin resistance gene both inserted within the M13 origin of replication. The mutated gII product interacts less efficiently with its own phage origin than with the phagemid origin of replication. Thus M13KO7 preferentially packages single-stranded plasmid over its phage DNA in the presence of a phagemid bearing a wild-type M13 or f1 origin, and then the new formed phages can be secreted into the culture medium. At OD_{600} of 0.6, there are 6.6×10^{10} bacteria per liter. Hence the multiplicity of infection (MOI) is about 20 (a MOI of 10 is OK). Mix well by shaking, then incubate at 37 °C for 45 minutes without shaking (temperature is critical). Shake at 37°C for 45 minutes. Measure the OD_{600} and calculate how many bacteria are present now. The culture will probably be about OD_{600} of 3.0, for about 3.3×10^8 cells per ml.

Check phage infection by diluting culture 1:100, then 1:100 again for a total dilution of 10^{-4}, (now ~$3.3 \times 10_4$ cells per ml). Plate 100, 10, and 1 μl aliquots on 2YT plates (supplemented with 70 μg/ml kanamycin). There should be about 3000, 300 and 30 colonies on these plates. Calculate the percentage of the successfully infected bacteria, ideally close to 100%. Otherwise, the diversity might be affected.

Spin down bacteria (4500g for 10 minutes at 4 °C). Resuspend in 500 ml of 2YT supplemented with ampicillin and kanamycin (no glucouse for eliminating the negative regulation to pLac, hence the expression of sdAb-gIII fusion protein commence), 200 rpm

overnight at 25 °C.

6. Phage preparation

Centrifuge the culture at 4800g, 4°C for 20 minutes. If necessary, centrifuge again to get a clear supernatant. Add 1/5 culture volume (100 ml) filter sterile PEG/2.5M NaCl to the supernatant to precipitate the phage, then 1 hour on ice. Centrifuge the culture at 10000g, 4°C for 10minutes. Resuspend the pellet in 100 ml of sterile PBS. Add 1/25 culture volume of (20 ml) of PEG/NaCl, 20 minutes on ice. Centrifuge the culture at 4800g, 4°C for 10 minutes. Resuspend the pellet in 1/20 culture volume (25 ml) of sterile PBS. Centrifuge again. Filter sterile the supernatant through 0.45 µm syringe filter (not 0.2 µm!). For short term storage, keep at 4°C. For long term storage, make aliquots of the library in Nunc cryotubes (about 20 tubes). Add 280 µl of 80% sterile glycerol to 1.22 ml phage preparation for a final concentration of 15%, freeze at -80°C.

7. Titration of phage preparation

The expectation of phage is 10^{12} cfu/ml. Plate bacteria: inoculate 50 ml of LB medium, with a single colony of TG1r freshly streaked on minimum medium. Shake at 37°C until OD_{600} is about 0.5-0.7, but no more. Keep the bacteria at 37°C without shaking until titration. Temperature is crucial. Alternatively, make a preculture and inoculate 50 ml LB with 1/100 of the preculture. It will take about 2 hours to reach OD_{600} of 0.5-0.7.

Titration: 10 µl phage preparation into 1 ml 2YT (10^{-2}); 10 µl into 1 ml 2YT (10^{-4}); 10 µl into 1 ml plating TG1r (10^{-6}); incubate at 37°C for 20 - 30 minutes; 10 µl into 1 ml 2YT (10^{-8}); plate 10 µl onto 2YT/Amp/Glu plate (titer = number of cfu x 10^{10}/ml).

8. Selection of the antigen-specific phage-sdAb

Following infection with a helper phage, libraries of recombinant phage particles are harvested from bacterial culture supernatants. Phage particles are produced with the VHH-gIII fusion protein at their tip and also the corresponding VHH gene encoded in the encapsulated phage or phagemid genome. Retrieving the binders by panning is commonly used to the screening of individual colonies, because the panning selects the binders with favorable properties, such as highest affinities, good expression, stability and solubility. Antigen-specific sdAbs are commonly selected on immobilized antigen, e.g., antigen

coated onto the plastic surface of an immuno-tube or magnetic Dynabeads, biotinylated antigens immobilized on Streptavidin beads, or membrane proteins expressed on the surface of cells. Antigen bound phage particles are eluted by pH-shock and regrown through infection of bacteria. Bound phages are subjected to one or several sequential rounds if needed. Phagemids are recuperated from single colonies of infected E. coli and the cDNA insert subjected to sequencing. Soluble VHH can be produced after either subcloning into a soluble expression vector or through the use of bacterial non-suppressor strains and appropriate stop codons between the VHH fusion and the phage gIII protein.

8.1. Panning

- ***Panning with immunotube coated IN***

1ml of 10-50 µg/ml Ag (HIV-1 IN or HFV IN) diluted in PBS was coated on immunotubes (Nunc-Immuno Tube MaxiSorp) at 4°C overnight. Less Ag should be used in the 2^{nd} and 3^{rd} round of panning to select phage-sdAb with higher affinity. Saturate the immunotube with 4 ml 2%PBSM (skimmed milk) for 2 hours at room temperature (fill the tube to the top). After washing with PBS for 3 times, add 1 ml about 10^{11} cfu of the library of phage preparation (i.e. Input which is produced by rescuing phage-display library or amplified from selected phages of previous panning), saturated 2% milk (20mg milk in 1 ml library) for 2 hours at room temperature, shake gently. Then 20 washes with 4 ml 0.1% PBST (Tween) and 10 washes with 4 ml PBS. Elute with 1 ml triethylamine (TEA, prepare just before use, 14 µl in 1 ml water, pH must be 12) for 10 minutes (no more) at room temperature, with gentle shaking. Use a new cap to prevent cross contamination. Transfer the elution into an eppendorf tube. Neutralize with 500 µl 1M Tris pH7. The obtained phages suspension is namely 'Output' which can be amplified by infecting bacteria E. coli TG1r. The total volume of 'Output' should be 1.5 ml. Of note, the phage themselves are inherently sticky which may lead to false positives during screening.

Amplification of Output phage suspension: infect 10 ml plating TG1r (0.5-0.7 OD_{600}) at 37°C for 30 minutes with 750 µl Output (also titer the Input at the same time). Keep the remaining 750 µl at 4°C. Plating 10 µl and 100 µl of 10^{-2} diluted (10 µl in 1 ml) infected TG1r on TAE (or 2YT)/Amp/Glu for titration. The expection contentration of Output is 10^4 -10^6 cfu. Centrifuge the rest infected bacteria at 3000g, 4°C. Resuspend the pellet in 250 µl 2YT. Spread on two large 2YT/Amp/Glu plates.

- ***Panning with immune-tube coated HIV-1 IN/DNA complex***

In the 2nd selection, recombinant HIV-1 IN was coat to immune-tube, and then incubated with oligonucleotide dimer for 1hour at room temperature to form IN/DNA complex. The DNA dimer was formed by annealing U5A/U5B oligonucleotides (mimic LTR of HIV genome). The oligonucleotides used are as follows (Leh et al., 2000):
U5B: 5'-GTGTGGAAAATCTCTAGCAGT-3'
U5A: 5'-ACTGCTAGAGATTTTCCACAC-3'
The following steps refer Materials and methods 7.1.

8.2. Amplification of Output (Rescue)

Scrape bacteria with 5 ml 2YT/Amp/Glu. Freeze the aliquots of the reminding scraped bacteria suspension at -80°C in 15% glycerol (0.6 ml 50% sterile PBS-glycerol for 1.4 ml bacteria suspension). Inoculate 12 ml 2YT/Amp/Glu with 100 µl scraped bacteria suspension (OD_{600} < 0.250), and shake at 37°C for about 1 hour (OD_{600} of 0.9). Add 50 µl helper phage (VCSM13 or M13KO7, 3 x 10^{10} cfu/ml). MOI of 1 is enough. Incubate for 30 minutes at 37°C. Then shake for 30 minutes at 37°C. Plate 1 µl onto small 2YT/Kan plate to check the successful infection. Centrifuge at 4500g at 4°C for 10 minutes. Resuspend pellet in 50 ml 2YT/Amp/Kan, and shake rapidly at 37°C for 30 minutes, then shake rapidly at 30°C for 3 hours, or 25°C overnight. Phage preparation and titeration refer Materials and methods 5. and 6.

8.3. Polyclonal phage ELISA

Polyclonal phage ELISA was performed to ensure the presence of specific phage-sdAbs in the 'Output' phage suspension before selecting monoclones with monoclonal-ELISA. 50 µl/well of 10µg/ml recombinant HIV or HFV IN was coated on 96-well plate at 4°C overnight. Then wash 3-5 times with PBS and block with 200 µl/well 5%PBSM for 2 hours at room temperature. Incubate with 50 µl/well selected and rescued phage suspensions (10^{11}, 10^{10}, 10^{9}, 10^{8}, 10^{7}, 10^{6} cfu/ml in 5% PBSM) for 1hour at room temperature. Negative control was incubated with other phage (M13 phage helper or F12). Wash 5 times with 0.1% PBST, then 3times with PBS. Incubate with 50 µl/well of 1/5000 diluted anti-M13-HRP (GE Healthcare) in 5% PBSM for 1hour at room temperature. Anti-M13-HRP reacts specifically with the bacteriophage M13 major coat protein product

of gene VIII (gpVIII). Then add 100 µl ABTS for 15-30 minutes, read at 405nm.

8.4. Soluble sdAb monoclonal ELISA

Expression of soluble sdAb from E. coli TG1r: The effect of tRNA amber suppressor can be eliminated quickly during the overproduction of sdAb induced by IPTG. Prepare one 96-well U bottom plate (Nunc) for the Output of each selection round (HIV Output 1 and 2; HFV Output 1 and 2); add 120 µl 2YT/1% glucose. Pick 92 clones for culture plus 2 negative clones and 2 wells without clone. Shake at 37°C overnight. These plates are namely master plate. Add sterile glycerol to 15% final concentration (40 µl of 60% glycerol in PBS). Cover the plates with Self-adhesive paper. Stock at -80 °C. Transfer 5 µl overnight cultures from master plate into a plate containing 1ml 2YT/Amp/0.1% glucose. Grow for 3 to 4 hours at 37°C (DO_{600} of 1.0). Then add 100 µl 10 mM IPTG (1mM final concentration) into the cultures. Grow at 25°C overnight. Centrifuge the micro-plate to obtain the supernatant of expression (1500 rpm for 10 minutes at 4°C) which will be used for ELISA. The expressed sdAbs are present in the supernatant.

Test soluble sdAb by ELISA: Coat a 96-well plate with 10 µg/ml antigen solution (50 µl/well) at 4°C overnight. Wash with PBS. Block with 200 µl 2% PBSM for 1 hour at room temperature. Then wash with PBS, add 50 µl supernatant of expression, incubate for 1h30 at room temperature. Wash 3 times with 0.1% PBST then 2 times with PBS. Incubate for 1hour with 50 µl 1/500 diluted anti-c-myc 9E10 antibody (Sigma) which recognizes the amino acid sequence EQKLISEEDL. Wash 3 times with 0.1% PBST then 2 times with PBS. Add 50 µl of 10 µg/ml anti-mouse-HRP (Santa Cruz Biotechnology) in 2% PBSM. Then add 100 µl ABTS for 15-30 minutes, read at 405nm.

9. In vitro expression of recombinant anti-IN sdAb in E.coli HB2151

Phagemids were transformed into E. Coli HB2151. HB2151 is a non-amber suppressor E. coli host strain used for expressing foreign proteins without the gIII fusion moiety. VHH highly enriched in periplasmic can be extracted by a simple osmotic shock (Skerra and Plückthun, 1988). Further purification is conveniently achieved by Ni-NTA batch or IMAC (Immobilized Metal Affinity Chromatography) when the recombinant VHH are fused with a 6xHistine C-terminal tag.

Inoculate monoclone in 2YT (1% Glucose and Carbenicillin). Add 5 ml overnight stock to each 500 ml pre-warmed 2YT (0.1% Glucose, 100 µg/ml Ampecillin). Incubate at 37°C,

200 rpm for approximate 2 hours to OD_{600} of 0.74. Then induce with 0.5 mM IPTG, incubate at 25°C, 200 rpm for 4 hours. Collect cells at 5000 rpm for 15 minutes at 4 °C.

Periplasmic and Osmotic Shock Preparations: Resuspend pellet in PPB buffer with freshly added protease inhibitor (optimal). PPB solution contains 200 mg/ml sucrose, 1 mM EDTA, 30 mM Tris-HCl pH8.0. Use 1/40 of total growth volume of PPB (12.5 ml for 500 ml expression). Keep on ice for 20 minutes. Spin down at 5000 rpm for 15 minutes at 4 °C and collect supernatant into small high-speed centrifuge tubes (Periplasmic Preparations). Resuspend pellet in 5 mM $MgSO_4$ buffer (1/40 of total volume, 12.5 ml per 500 ml expression). Incubate on ice for 20 minutes. Transfer samples to small high-speed centrifuge tubes (Osmotic Shock preparations). Spin both Periplasmic Preparations supernatant and Osmotic Shock preparations at 5000 rpm for 15 minutes at 4 °C. Collect supernatant to dialyze.

Dialysis: Soak dialysis tubing in water. Load supernatant into dialysis tubing and dialyze overnight in 4 liter PBS at 4 °C. The following steps refer Materials and methods: *I.4. Purification of HIV IN.*

10. Recognition of INs by recombinant sdAbs (ELISA assay)

50 µl/well of 10 µg/ml recombinant INs or IN fragments were coated in 96-well plate at 4 °C overnight. Recombinant INs were firstly diluted in PBS. After blocking non-specific sites with 200 µl 5% PBSM for 1 hour at room temperature, 50 µl/well of 30 µg/ml recombinant sdAbs in 5% PBSM were incubated with Antigen for 2 hours at room temperature with gentle shaking. After washing with 0.1% PBST 3-5 times and PBS 3 times, add 50 µl of 10 µg/ml (1/500 diluted in 5% PBSM) anti c-myc 9E10 antibody, incubate at room temperature for 2 hours with gentle shaking. After washing, add 1/10 000 diluted 50 µl anti-mouse-HRP, incubate at room temperature for 1.5 hours with gentle shaking. Then add 100 µl ABTS for 15-30 minutes, read at 405nm.

11. Recognition of INs by recombinant sdAbs (Western blot assay)

1 µg/well of recombinant HIV wild-type INs (HIV-1 IN, HIV-2 IN and HFV IN) were deposed to verify the antigen-recognizing ability of recombinant sdAbs. After incubated with 1.5 µg/ml recombinant sdAbs at room temperature for 2 hours or at 4 °C overnight, INs were detected with 0.4 µg/ml (diluted in 5% PBSM) anti-c-myc 9E10 antibody at room temperature for 2 hours. Then add 1/10 000 diluted anti-mouse IgG-HRP at room

temperature for 1.5 hours for development.

12. The effect of sdAbs on in vitro IN enzymatic activity

The effect of sdAbs on IN enzymatic activity (3'-processing and strand transfer) was determined in vitro by testing IN activity in the presence of increasing concentration of recombinant sdAbs (from 10 nM to 30 μM sdAbs dissolved in PBS) at 37°C for 3 hours. In vitro IN enzymatic activity test refer Materials and methods I. 5. In vitro IN enzymatic activity test. The quantification was preformed using the Image Quant TL software program. Fifty percent effective concentration (EC_{50}) calculation was performed using Prism 5.0 software (GraphPad Software, San Diego, CA).

13. Immunoprecipitation (IP) assay to demonstrate the effect of sdAb on the interaction of IN and LEDGF/p75

Incubate 0.1 μM recombinant HIV-1 IN (0.53 μg) with 0.4 μM sdAb in 200 μl binding buffer rotating at 4℃for 2 hours. For 100 ml binding buffer, use 2 ml of 1M HEPES, 1.25 ml of 4M NaCl, 1 ml of 0.5M EDTA, 100μl NP-40, 100μl of 1M DTT and 5 ml glycerol. Add 0.25 μM recombinant LEDGF/p75 rotating at 4℃ for 2 hours. To precipitate the complex, add 0.05 μM anti-IN IgG, rotating at 4℃ for 2 hours. Then use 30 μl protein G-sepharose to precipitate (prewashed 2 times with binding buffer), rotating at 4℃ for 2 hours. Recover the supernatant for comparison with Western blot. Wash the complex with 1 ml binding buffer rotating at 4℃ for 5 minutes, 3-5 times. Then add 30-50 μl Laemmli buffer, heat at 90℃ for 5 minutes. Depose to SDS-PAGE gel directly. LEDGF/p75 was detected by 0.4 μg/ml goat anit-LEDGF antibody, then 1/5000 diluted donkey anti-goat IgG-HRP.

14. In vivo transitory expression of anti-HIV IN sdAbs

14.1. Expression with pEGFP-N3NN-MH

Plasmid pEGFP-N3NN-MH is derived from pEGFP-N3 (**Figure 32**). pEGFP-N3 encodes a red-shifted variant of wild-type GFP which has been optimized for brighter fluorescence and higher expression in mammalian cells. The MCS in pEGFP-N3 is between the

immediate early promoter of CMV (pCMV IE) and the EGFP coding sequences. SV40 polyadenylation signals downstream of the EGFP gene direct proper processing of the 3' end of the EGFP mRNA. The vector backbone also contains an SV40 origin for replication in mammalian cells expressing the SV40 T-antigen. The pEGFP-N3 backbone also provides a pUC origin of replication for propagation in E. coli and an f1 origin for single-stranded DNA production.

Construction of pEGFP-N3NN-MH-sdAb: The sdAbs coding sequence, 6xHistine and c-Myc tag were digested from pHEN using Not I and Noc I restriction enzymes, then inserted into pEGFP-N3NN-MH vector for cellular expression. The constructions were verified by sequencing (Eurpfins MWG operon).

Calcium phosphate transfection of Hela P4 cell with pEGFP-N3NN-MH-sdAb: The medium for Hela P4 cell contains DMEM GlutaMax, 10% fetal bovine serum (FBS), 100 IU /ml Penicilin and 0.1 mg/ml Streptomycin. One day before transfection, pass Hela P4 cell into 6-well plate at a concentration of 4 x 10^5 cells/well. So the cell confluence for transfection is about 50%-70%. 2 hours before transfection, change the medium for Hela P4 cell. For transfection of 6-well plate (9 cm^2 per well), 100 µl of 0.5 M $CaCl_2$ and 4 µg plasmid were supplemented to 200 µl with pure water, then drop into 200 µl 2xHBS pH7.05 (vortexing to obtain small DNA-$CaPO_4$ precipitation), 15 minutes at room temperature. Then drop the mixture into 2 ml/well cells (mix well). 48 hours after transfection, the transitory expression of anti-HIV IN sdAb-GFP was detected with flow cytometry (FACS) or confocal microscopy.

Preparation cells for FACS: Digest cells were with pre-warmed trypsin (0.5 ml for 9 cm^2 well), then stop the digestion by adding 1.5 ml medium containing 10% FBS. Count the cells with trypan blue. In a FACS tube, add 0.5 ml cells and 3 ml pre-cold FACS buffer (1% FBS in PBS). After spin down, resuspend the cells in 500 µl FACS buffer, keep at 4℃.

HIV infectivity assay: Single-cycle titers of HIV virus can be determined in HeLa-P4 cells. HeLa-P4 cells are HeLa CD4+ LTR-LacZ cells in which lacZ expression is induced by the HIV transactivator protein Tat, making it possible to quantify HIV-1 infectivity precisely from a single cycle of replication. Cells were infected, in triplicate, in 96-well plates, with virus (equivalent of 3 ng of $p24^{gag}$ antigen). The single-cycle titers of viruses were determined 48 h after infection by quantifying β-galactosidase activity in P4 lysates in a colorimetric assay (the CPRG assay) based on the cleavage of chlorophenol red-β-D-galactopyranoside (CPRG) by β-galactosidase.

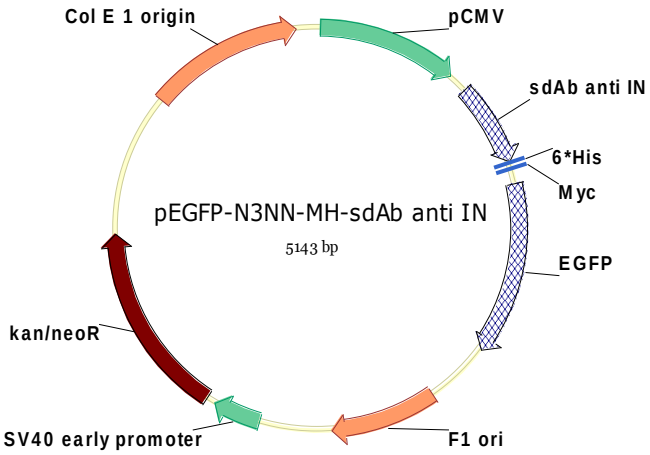


Figure 32: Schematic presentation of plasmid pEGFP-N3NN-MH inserted with anti-HIV IN sdAb coding sequence.

14.2. Expression with pcDNA3.1

pcDNA3.1 is a 5.4 kb vector derived from pcDNA3 and designed for high-level stable and transient expression in mammalian hosts (**Figure 33**). The vector contains the following elements: Human cytomegalovirus immediate-early (CMV) promoter for high-level expression in a wide range of mammalian cells; MCSs to facilitate cloning; Neomycin resistance gene for selection of stable cell lines; and episomal replication in cells lines that are latently infected with SV40 or that express the SV40 large T antigen.

The sdAb coding sequences with 6xHistine and c-Myc were amplified from phagemid pHEN by specific primers:

Forward: 5' - GAA TTC GCC ACC **ATG** CAG GTG CAG CTG GTG CAG TCT G - 3'

(EcoR I site = GAA TTC, Kozac = GCC ACC, Start codon = **ATG**)

Reverse: 5' - GGA TCC **TCA** GCT CCC ATG GTG ATG GTG -3'

(BamH I site = GGA TCC, Stop codon = **TCA**)

The following protocol was engaged to amplify the sdAb coding sequences and 6xHistine and c-Myc tag. The 50 µl reaction mixture consists of 5 µl 10 x HF buffer, 2 mM final concentration of $MgSO_4$, 100 ng plasmid template, 400 nM final concentration of

each primer, 0.2 mM of each dNTP, 1-2 units of TaqHF and sterile Milli-Q water filled to the final reaction volume (Invitrogen platinum).

The PCR cycle is as follows: 5 minutes at 94 °C, 30 cycles of 30 seconds at 94 °C, 30 seconds at 55 °C and 40 seconds at 68 °C, then 10 minutes at 68°C. The PCR products were cloned into pGEM-T Easy (Promega). After verified by sequencing (Eurofins MWG Operon), the sdAb coding sequences with 6xHistine and c-Myc were inserted into pcDNA3.1. The constructions were confirmed by sequencing.

HeLa cells were transitory co-transfected by calcium phosphate with pAS1B HA-IN and pcDNA3.1 anti-IN sdAb of 1:1 ratio. Cover slips were stained 24h after transfection. The HA tag was stained with anti-HA 3F10 (Roche) and anti-Rat-A488 (Molecular Device). c-Myc was stained with 9E10 antibody (Roche) and anti-mouse-A555.

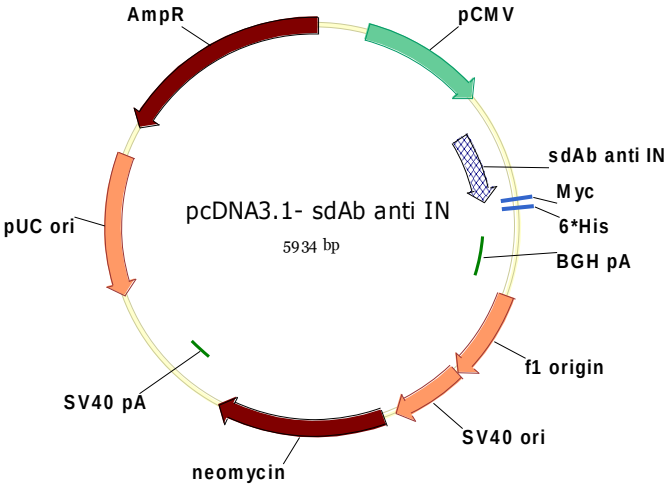


Figure 33: Schematic presentation of plasmid pcDNA3.1 inserted with anti-HIV IN sdAb coding sequence.

15. Selection of anti HIV-1 IN sdAb with Epoxy bead coupled HIV-1 IN

48 hours before selection, prepare the Epoxy bead (Invitrogen, Dynabeads M-450 Epoxy). Dynabeads M-450 Epoxy are hydrophobic and covered with surface epoxy groups. The epoxide chemistry immobilizes ligands containing amino, thiol and hydroxyl functional groups. Firstly, wash the beads with 1ml PBS, then with 100 µl PBS. 25 µl bead is able to couple with 3-5µg of antigen. Selection volume is either 100 µl or 1 ml. When antigen is enough, it is better to work with 1ml. Otherwise, 100 µl selection volume also works very well. The final concentration of antigen in selection system should be 100 nM. For example: 1ml selection volume x 100 nmol/l x 32000 g/mol (IN MW) = 3.2 µg IN; then use about 32 µl beads. Try to keep the bead in the same volume as the starting bead volume during the coupling (coupling in 32 µl, do not dilute too much). Couple for 48 hours at 4 °C, rotating. Then wash the beads with 1 ml 0.1% PBST 3 times and 1 ml PBS 3 times using a magnetic rack. Keep 2.5 µl beads to verify the coupling of antigen. Block with 1ml 2% PBSM, rotating at room temperature for 1 hour. Then add recombinant anti-HIV IN sdAb-1 (≥100 µg/ml) to saturate its epitope for selecting sdAbs with different epitopes of IN, rotating at room temperature for 1.5 hours. The saturation volume is 100 µl. The following selection steps refer Materials and methods 6.1. Panning with immunotube coated IN.

Soluble sdAb monoclonal ELISA with Epoxy bead: 16 µg of HIV-1 IN was coupled to about 150 µl Epoxy beads (for 100 wells). After washing and blocking, the beads in 1 ml PBS were put into PCR strips or 96-well PCR style plate (about 1.5 µl beads in 10 µl suspension per well). To wash the 96-well plate, use a DynaMag-96 side magnetic plate (Invitrogen). The following steps refer Materials and methods *9. Soluble sdAb monoclonal ELISA***.**

16. Surface plasmon resonance imagery (SPRi) assay to develop sdAb as targeting reagents for nanosensors

Surface Plasmon Resonance (SPR) is an optical detection process that occurs when a polarized light hits a prism covered by a thin gold layer. Under certain conditions (wavelength, polarization and incidence angle), free electrons at the surface of the biochip absorb incident light photons and convert them into surface plasmon waves. A dip in reflectivity of the light is seen under these SPR conditions.

Perturbations at the gold surface of the biochip, such as an interaction between antibody immobilized on the chip and captured target proteins, induce a modification of resonance conditions which are in turn seen as a change in reflectivity and which can be measured.

The SPR imaging technology takes SPR analysis a step further. It is a sensitive label-free method of visualizing the whole of the biochip via a video CCD camera. This design enables the biochips to be prepared in an array format on which each active spot provides SPR information simultaneously.

With the SPRi range of equipment, SPR imaging is provided in a robust and sensitive manner. A broad-beam monochromatic polarized light at specific wavelength from a laser diode illuminates the whole functionalized area of the biochip surface which is mounted within the chamber of detection instrument. A high resolution CCD camera provides real-time different images for the array of active spots. It captures all of the local changes at the surface of the biochip providing detailed information on molecular binding, biomolecular interactions and kinetic processes.

The surface plasmon resonance imagery (SPRi) is previously described (Nogues et al., 2010). Different sdAbs (anti-HIV IN sdAb-1, anti-HIV IN sdAb-4, anti-HIV IN sdAb-A7, anti-HIV IN sdAb-A9, anti-HIV IN sdAb-A12, anti-HIV IN sdAb-B5, anti-HIV IN sdAb-C12, anti-HIV IN sdAb-G11) were deposed on the gold surface of a prism. Then the protein solutions, non-sepecific protein fused with histine tag (500 nM LacI-his and 200 nM Prh-his), 200 nM BSA and 200 nM integrase in sample buffer were injected on the surface.

III.2.3. Results

1. Immunization of llamas and construction of phage display library

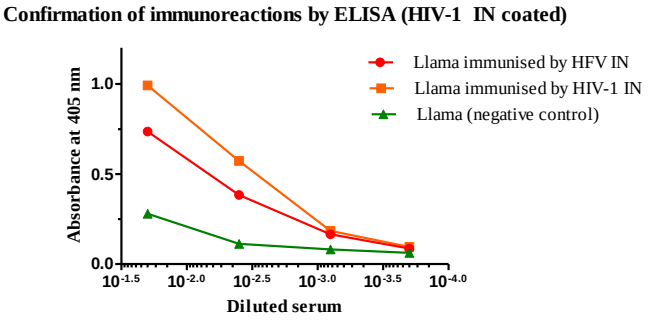
Immunization and construction of phage display library was performed by the collaborative team of D. Batty. Two llamas were immunized with recombinant HIV-1 IN or human foamy virus (HFV) IN, respectively. The third llama was a negative control. After amplified the VHH coding sequence with specific primers, inserted in phagemid pHEN vector, and transformed E.Coli TG, phage display libraries of anti-HIV IN sdAb and anti-HFV IN sdAb were constructed with a diversity of 10^7 to 10^8. It is also interesting to develop specific sdAbs targeting HFV IN because HFV IN is a convenient proxy for HIV structural studies due to the high level of conservation between retroviral INs, in particular within their active sites (Engelman and Craigie, 1992; Kulkosky et al., 1992; Valkov et al., 2009). Especially, a diffracting crystals of the full-length IN from the prototype foamy virus was successfully obtained in recent (Hare et al., 2010b).

2. Confirmation of immunoreactions induced by HIV IN and HFV IN

2.1. ELISA assay to confirm the immunoreactions

Before selecting the specific anti-IN sdAb, we confirmed if the immunoreaction was successfully induced in the llamas. ELISA assay was firstly performed. The coated HIV IN was well recognized by anti-HIV IN serum and anti -HFV IN serum (**Figure 34A**). The coated HFV IN was also well recognized by both immune sera (**Figure 34B**). ELISA tests showed that immunoreaction was very well provoked in the two llamas immunized using recombinant HIV IN or HFV IN. The observed cross-reaction of the two anti-IN sera was probably due to the similar immunogen characteristics shared by HIV IN and HFV IN.

A



B

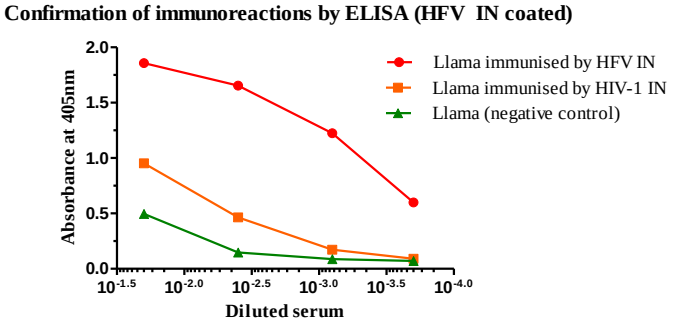


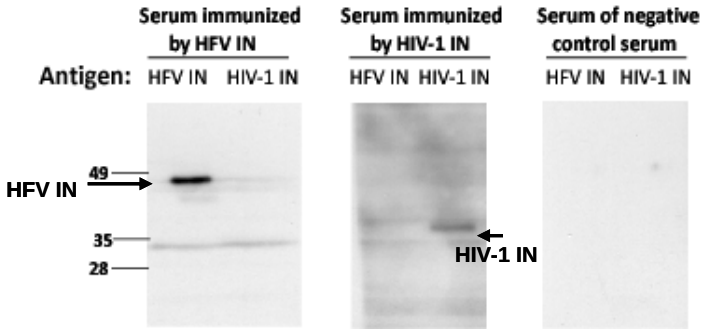
Figure 34: Confirmation of immunoreactions induced by HIV-1 IN (A) and HFV IN (B) with ELISA assay.

Two llamas were immunized with recombinant HIV-1 IN or human fomay virus (HFV) IN, respectively. The third llama was taken as a negative control. HIV-1 IN or human fomay virus (HFV) IN was coated, then incubated with gradient diluted llama immune sera (1/50, 1/250, 1/1250, 1/5000 in PBS). Llama anti-IN antibodies were detected with goat anti-llama IgG-heavy and light chain Ab-HRP.

2. 2. Western blot assay to confirm the immunoreactions

Western blot assay was also preformed to confirm the immunoreactions in llamas. Anti-HIV IN serum was able to recognize coated HIV IN, and anti-HFV IN serum also can recognize coated HFV IN (**Figure 35A**). The cross-reaction was not observed with western blot assay, probably because western blot assay is relatively less sensitive than ELISA test. Remarkably, the anti-HIV IN serum exhibited the ability of recognizing a broad range of INs, including INSTI resistance mutant HIV-1 IN G140S/Q148H, as well as HIV-1 CRF02_AG strains INs (**Figure 35B**). This capacity was not observed in anti-HFV IN serum.

A



B

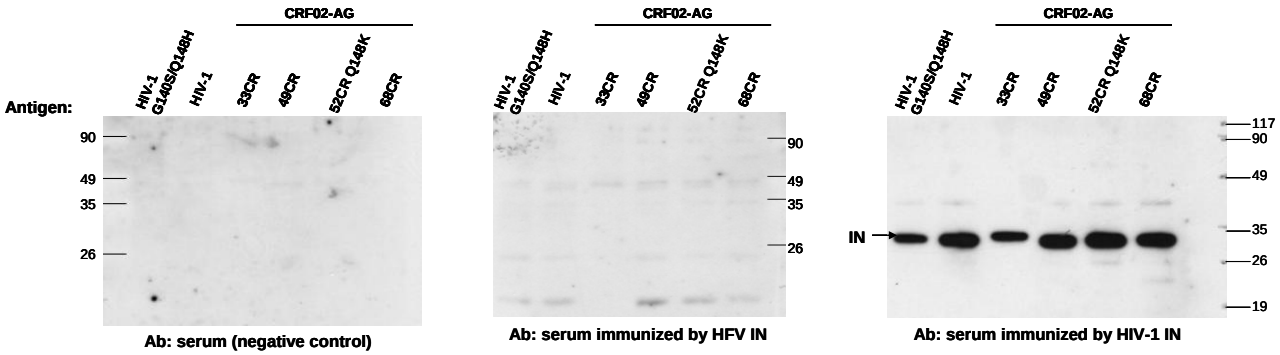


Figure 35: Western blot aasy to confirm the immunoreactions induced by HIV-1 IN or HFV IN.

(A) HFV IN (left lane) and HIV-1 IN (right lane) were deposed and incubated with indicated llama immune serum. HIV-1 IN is about 32 kDa; HFV IN is about 42 kDa. (B) HIV-1 INSTI resistance mutant G140S/Q148H, HIV-1 IN, and four HIV-1 CRF02_AG subtype strains (33CR, 49CR, 52CR Q148K, 68CR) were deposed (from left to right), and incubated with indicated llama immune serum. Then the Abs in serum were detected with goat anti-llama IgG-heavy and light chain Ab-HRP.

3. 1st selection of anti-HIV IN sdAb and anti-HFV IN sdAb (selected with immune-tube coated INs)

3.1. Concentration of specific phage-sdAbs with two rounds of panning

After confirmed that immunoreaction was successfully induced, two rounds of panning for the selection of specific phage-sdAbs were performed. In each round of panning, recombinant phage-sdAbs were selected by their ability to recognize the coated targets: either HIV-1 IN or HFV IN. Each round of panning increases the percentage of positive phage-sdAb in the entire phage population (**Table 4**). Meanwhile, the diversity of phage-sdAb could decrease with rounds of panning. Enrichment of binders is the ratio of Output / Input value of second round to that of first round. The enrichments were about 10^4 for the two rounds of panning

Anti-HIV IN	Input (cfu)	Output (cfu)	Output/ Input	Enrichment
Round 1	1.00E+11	6.00E+03	6.00E-08	
Round 2	1.00E+11	5.70E+07	5.70E-04	9.50E+03

Anti-HFV IN	Input (cfu)	Output (cfu)	Output/ Input	Enrichment
Round 1	1.00E+11	1.20E+04	1.20E-07	
Round 2	1.00E+11	3.75E+08	3.75E-03	3.13E+04

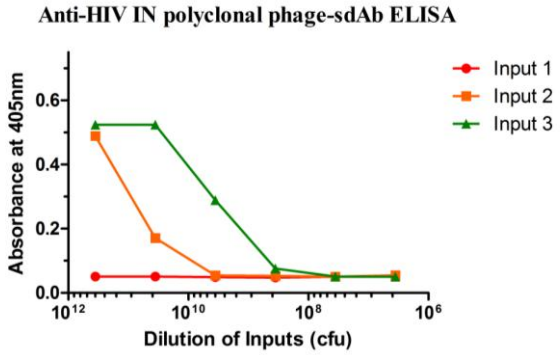
Table 4: Titrations and binder enrichment after two rounds of pannings with immunotube coated INs.

3.2. Confirmation of the presence of specific phage-sdAbs in the population obtained from panning with polyclonal phage-sdAbs ELISA

After concentrated the specific phage-sdAbs, polyclonal phage-sdAbs ELISA was engaged to confirm the presence of positive phage-sdAbs in the suspension of selected phages.

Three Inputs of anti-HIV IN phage-sdAbs and anti-HFV IN phage-sdAbs were gradually diluted, and their target-recognizing ability was tested using coated recombinant HIV-1 IN and HFV IN, respectively **(Figure 36)**. The signal of ELISA augmented with the round of panning, which confirmed that the positive phage-sdAbs were present and concentrated after panning. The signal generally began to disappear when phage-sdAbs were diluted to 10^8 cfu/ml.

A



B

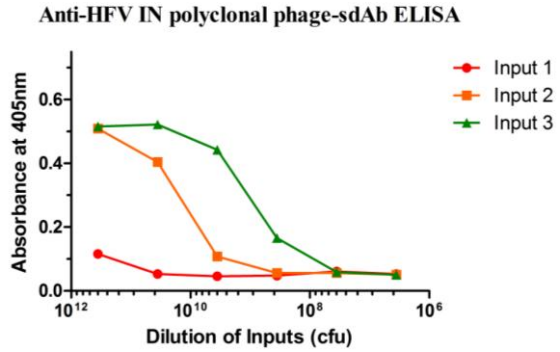


Figure 36: The presence of positive phage-sdAbs in selected phages was verified with polyclonal phage-sdAb ELISA.

Target-recognizing capacity of three Inputs of anti-HIV IN phage-sdAbs (A) or anti-HFV IN phage-sdAbs (B) was verified. Input1 is phage-sdAbs rescued from phage display library. Input 2 is the amplification of selected phage-sdAbs from the first round of panning. Input 3 is from second round of panning.

3.3. Isolation of specific anti-IN sdAb strains with soluble sdAb monoclonal ELISA and sequencing

To obtain the specific anti-IN sdAbs isolates, clones of E.coli TG1r were precultured and induced to produce sdAbs. Soluble sdAbs in supernatant of each miniculture were then identified by their target-recognizing capacity. Six anti-HIV IN sdAbs and eight anti- HFV IN sdAbs were isolated using soluble sdAb monoclonal ELISA.

According to the sequencing result, three different anti-HIV IN sdAbs and two anti-HFV IN sdAbs were obtained (**Figure 37**). The last 8 clones in the list of alignment are identical. Anti-HIV IN sdAb-4 and sdAb-5 are also same. The sequence identity of anti-HIV IN sdAbs and anti-HFV IN sdAbs confirms the epitope similarity of both INs, which is due to their conserved structures. Moreover, it was observed that approximately half of the sequenced sdAbs are identical with anti-HIV IN sdAb-1.

According to the hallmark substitutions in framework II, most of the sequenced sdAbs are VHHs, except anti-HFV IN sdAb-2 which seems to be a VH (V at 37, W at 47 according the positions in **Figure 37**).

4. 2nd selection of anti-HIV IN sdAb (selected with immune-tube coated HIV-1 IN/DNA complex)

In vivo, the tetramer of IN interacts with a pair of viral DNA end within PIC. The conformation of IN could be modulated after the binding of viral DNA. The exhibited epitopes of IN may be changed consequently. The epitopes exposed on the recombinant HIV IN/DNA complex in vitro could represent a better simulation of PIC. Thereby, we preformed the second selection with immune-tube coated recombinant HIV-1 IN/DNA complex.

However, sequencing results showed that the sequences of all the eight isolated anti-HIV IN/DNA sdAbs were almost identical with the anti-HIV IN sdAb-1 from first selection using immune-tube coated HIV-1 IN, while the three CDRs were exactly identical (**Figure 38**). It suggests that the binding site of anti-HIV IN sdAb-1 might be the dominant epitope exposed on IN/DNA complex.

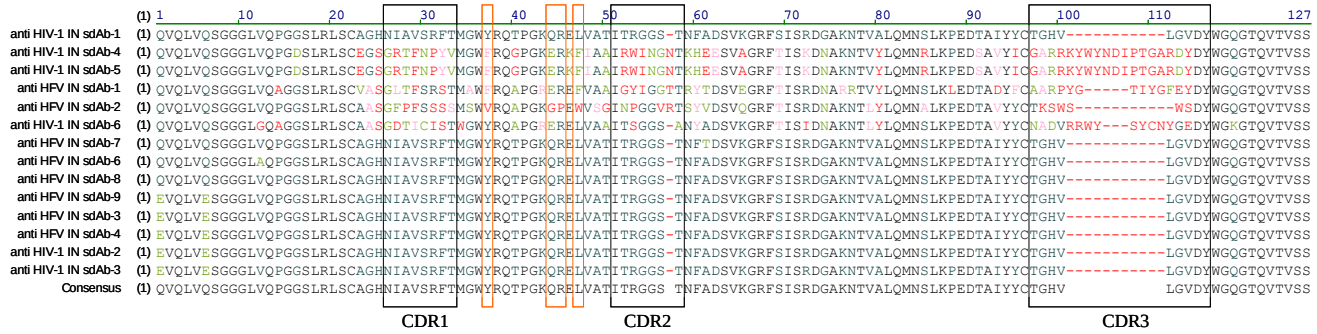


Figure 37: Alignment of amino acid sequences of anti-IN sdAbs selected with immune-tube coated INs.

The four hallmark amino acids differing between VH and VHH in framework (Val37Phe/Tyr, Gly44Glu/Gln, Leu45Arg/Cys and Trp47Gly/Ser/Leu/Phe, in going from VH to VHH) are marked with orange boxes. Three CDRs are marked with black boxes.

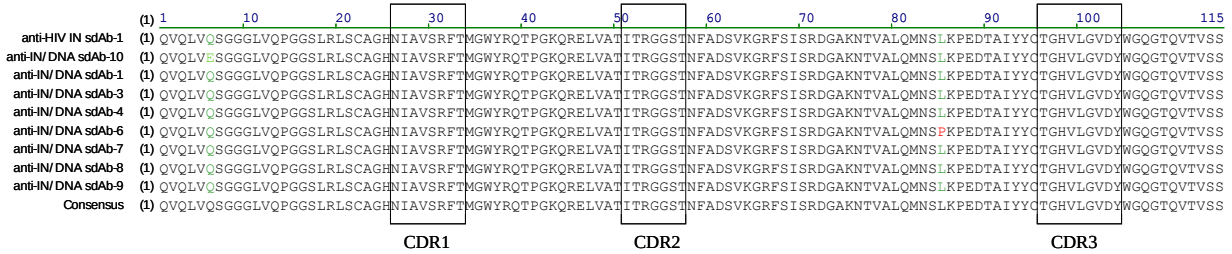


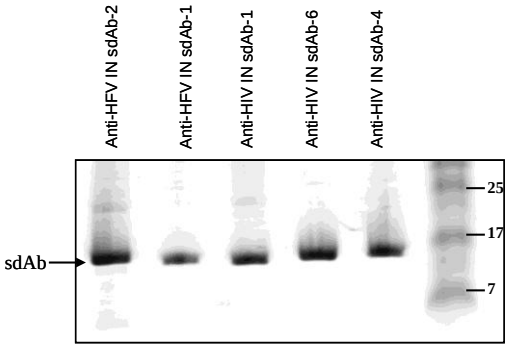
Figure 38: Alignment of amino acid sequences of anti-IN sdAbs selected with immune-tube coated HIV-1 IN/DNA complex.

Three CDRs are marked with black boxes.

5. In vitro activity of recombinant anti-IN sdAbs

5.1. Expression of recombinant anti-IN sdAbs

To investigate the in vitro activity of the five identified anti-IN sdAbs, recombinant sdAbs were expressed without the gIII fusion moiety using a non-amber suppressor E. coli host strain HB2151 and purified in vitro. Downstream of a periplasm signal leader sequence, sdAbs were secreted into periplasm of bacteria, facilitating the purification of protein. The sdAbs purified from 500 ml culture were at the concentration of 0.33 to 1.28 mg/ml in a volume of 0.5 to 1.5 ml (**Figure 39**).



sdAb	Concentration (mg/ml)	Concentration (μ M)
Anti-HIV IN sdAb-1	1.16	68.2
Anti-HIV IN sdAb-4	1.28	67.4
Anti-HIV IN sdAb-6	0.8	43.8
Anti-HFV IN sdAb-1	0.71	41.7
Anti-HFV IN sdAb-2	0.33	19.5

Figure 39: In vitro expression and purification of recombinant anti-IN sdAbs.

5.2. INs including an INSTI resistance mutant can be recognized by recombinant sdAbs with ELISA assay

Firstly, we confirmed the traget-recognizing ability of recombinant sdAbs with ELISA assay. The coated antigens include a HIV-1 IN Δ N fragment in which C280S mutation was introduced for an higher solubility, and HIV-1 IN Δ C fragment in which F185K mutation was introduced for the same reason. Among these recombinant anti-IN sdAbs, anti-HIV IN sdAb-1 showed the a good affinity to HIV-1 IN, HFV IN, INSTI resistance mutant HIV-1 IN G140S/Q148H and HIV-1 IN Δ N fragment, but not HIV-1 CRF02_AG IN (**Table 5**). Anti-HFV IN sdAb-1 also showed a modest affinity to HIV-1 IN, HFV IN, HIV-1 IN Δ N fragment and some affinity to resistance mutant HIV-1 IN G140S/Q148H. Anti-HIV IN sdAb-4 has a relatively lower affinity to HIV-1 IN, resistance mutant HIV-1 IN G140S/Q148H and HIV-1 IN Δ N fragment. None of the three sdAbs binds to HIV-1 IN Δ C fragment or catalytic core fragment with good affinity. Although the identified anti-HIV IN sdAbs were not finely mapped to precise peptide domains in this study, we assume that IN N-terminal domain is not essential for the binding of the three sdAbs. Remarkably, the IN strand transfer resistance mutant HIV-1 IN G140S/Q148H was well recognized by sdAbs, which suggests the double resistant mutations might not be involved in the binding sites of these identified sdAbs.

	BSA (C-)	HIV-1 IN	HIV-1 CRF02AG 33CR IN	HIV-1 IN G140S/Q148H	HFV IN	HIV-1 IN Δ N	HIV-1 IN Δ C	HIV-1 IN Catalitic core
Anti-HIV sdAb-1	0.050	1.296	0.055	0.707	1.220	1.670	0.056	0.216
Anti-HIV sdAb-4	0.050	0.229	0.066	0.396	0.095	0.343	0.076	0.064
Anti-HIV sdAb-6	0.051	0.076	0.052	0.071	0.059	0.083	0.051	0.051
Anti-HFV sdAb-1	0.07	0.617	0.064	0.182	0.735	1.546	0.068	0.094
Anti-HFV sdAb-2	0.055	0.063	0.049	0.061	0.060	0.065	0.054	0.054

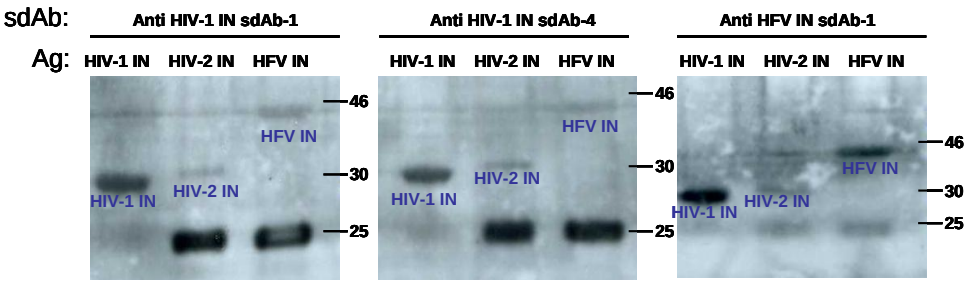
Table 5: Recognition of INs by recombinant sdAbs with ELISA assay.

Different INs or IN fragments were deposed and then incubated with recombinant anti-IN sdAbs. The stronger signal of ELISA is marked in red.

5.3. INs including an INSTI resistance mutant can be recognized by recombinant sdAbs with western blot assay

Target-recognizing ability of the three recombinant sdAbs (anti-HIV IN sdAb-1, anti-HIV IN sdAb-4 and anti-HFV IN sdAb-1) with good affinity in ELISA assay was confirmed by western blot sequentially. Both anti-HIV IN sdAb-1 and anti-HIV IN sdAb-4 were able to recognize coated HIV-1 IN in western blot **(Figure 40A)**. They can also recognize HIV-2 IN and HFV IN though with a lower affinity. Anti-HFV IN sdAb-1 recognized all the three INs. Stirkingly, the degraded form of recombinant IN also can be recognized by these sdAbs. Anti-HIV IN sdAb-1 and anti-HIV IN sdAb-4 even showed a better affinity to the degraded form of IN, presumably due to the better exposed epitopes. Consistent with ELISA assay, anti-HIV IN sdAb-1 and anti-HIV IN sdAb-4 exhibited the target-recognizing ability to INSTI resistance mutant HIV-1 IN G140S/Q148H **(Figure 40B)**. In addition, we found that recombinant sdAbs can be directly detected using anti-llama IgG Ab-HRP with western blot, though with less sensitivity (data not shown).

A



B

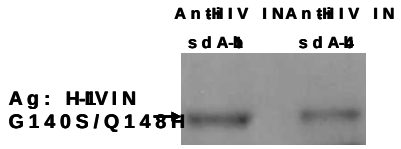


Figure 40: Recognition of INs by recombinant sdAbs (Western blot).

(A) Recombinant HIV-1 IN, HIV-2 IN and HFV IN (from left to right) were deposed, incubated with anti-IN sdAbs, and then detected by mouse anti-cmy Ab and anti-mouse IgG Ab-HRP. The recombinant HIV-1 IN and HIV-2 IN are about 32 kDa, HFV IN is about 42 kDa. The proteins of about 25 kDa were degraded fragments of INs. (B) Recognization of INSTI resistance mutant HIV-1 IN G140S/Q148H by anti-HIV IN sdAb-1 and anti-HIV IN sdAb-4.

5.4. Isolated recombinant sdAbs are not able to inhibit IN enzymatic activity (3'-processing and strand transfer) in vitro

Because sdAbs have shown the target-recognizing ability to resistance mutant HIV-1 IN G140S/Q148H, it is highly interesting to verify the inhibitory activity of sdAb to IN. We performed IN enzymatic activity assay (3'-processing and strand transfer) in vitro using ^{32}P-labeled substrate in the presence of increasing concentration of sdAbs. Nevertheless, the inhibitory activity of anti-HIV IN sdAbs and anti-HFV IN sdAbs were not very efficient (estimated IC_{50}＞30 μM) **(Figure 41A)**. Moreover, a considerable amount of non-specific cleaved products were observed in the control where IN was not present **(Figure 41B)**, which suggested the existence of nuclease in the preparation of recombinant sdAb. The effect of nuclease was more severe especially when a relatively large volume of recombinant sdAb was used in order to inhibit the activity of IN, due to the low IC_{50} of sdAb. The non-specific cleaved products in turn can affect the evaluation of IC_{50} value.

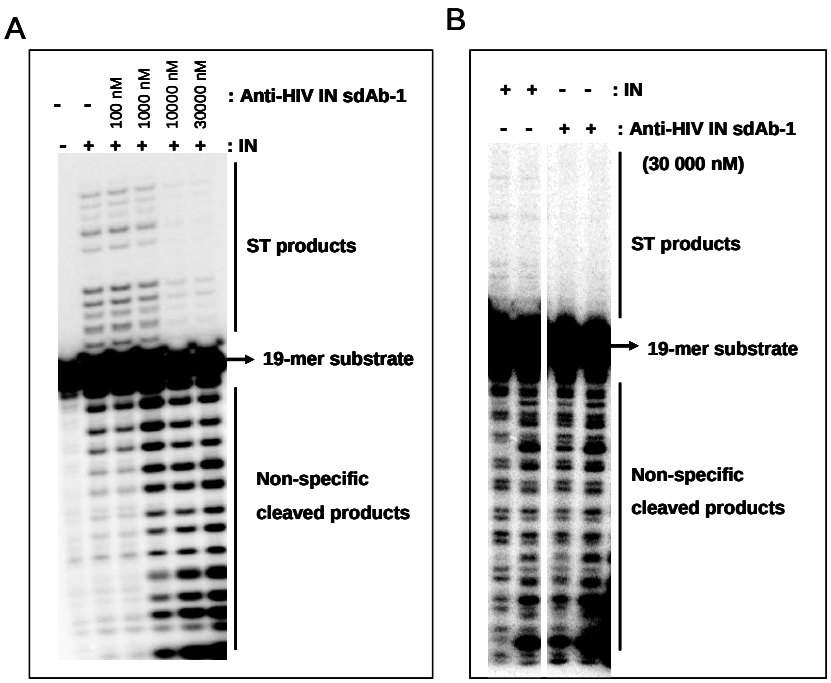


Figure 41: The effect of sdAbs on enzymatic activity of IN in vitro.

(A) A representative gel shows strand transfer assay in the presence of increasing concentration of sdAb. Concentrations of recombinant sdAb are indicated above each lane. The ^{32}P-labeled oligonucleotide substrate mimicking the preprocessed HIV-1 viral substrate is indicated as 19-mer. Products of the strand transfer reaction and non-specific cleaved products are indicated. (B) A considerable amount of non-specific cleaved products were observed in the control where IN was not present.

5.5. Isolated recombinant sdAbs can not interfer the interaction of IN and LEDGF/p75

Cellular co-factor LEDGF/p75 is involved in nuclear import and chromosome tethering of PIC, and increases the efficiency of concerted integration. Moreover, LDEGF/p75 strongly influences the genome-wide pattern of integration, in which active transcription units are favored. Within PIC, a pair of IN tetramers and two subunits of LEDGF/p75 comprise a symmetrical complex. Hence, it is interesting to study if the fixation of sdAb can interfere the interaction of IN and LEDGF/p75 in vitro. This effect of anti-HIV IN sdAb-1 and anti-HFV IN sdAb-1, both with good affinity, were determined using immunoprecipitation assay. Firstly, recombinant HIV-1 IN was incubated with sdAb. Then the complex of IN/sdAb was incubated with recombinant LEDGFp/75 and precipitated with conventional anti-HIV-1 IN antibody and protein G. Then LEDGF/p75 within the precipitated complex was detected with anti-LEDGF antibody with western blot assay. It was showed that the quantity of LEDGF/p75 bound to IN/sdAb complex was approximately equal to the control without sdAb (**Figure 42)**. It suggested that the two sdAbs were not able to interfere the interaction of IN and LEDGF/p75 in vitro. The protein fragment of 44 kDa detected by anti-LEDGF antibody is a degraded form of recombinant LEDGF/p75.

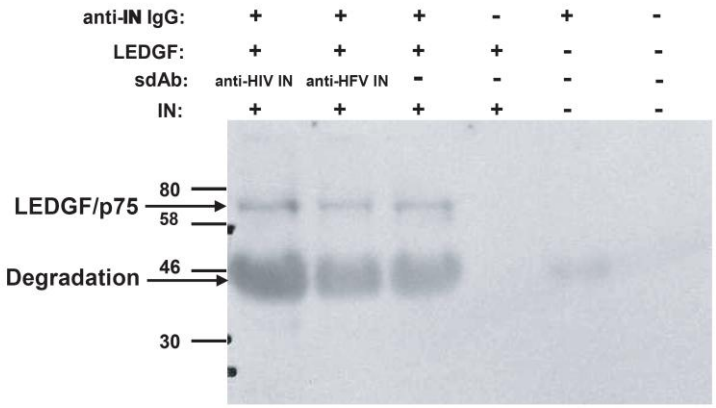


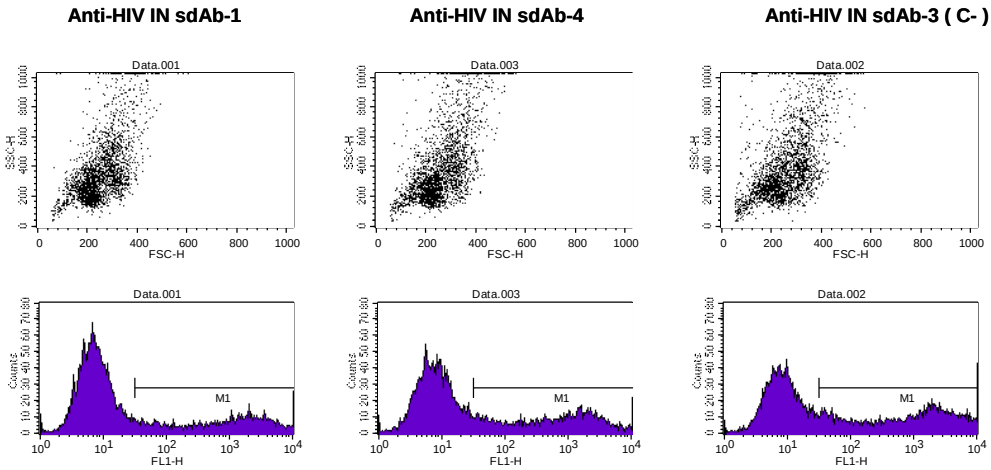
Figure 42: The effect of sdAb fixation on the in vitro interaction of IN and LEDGF/p75.

The components of immunoprecipitation assay were indicated (+ : with the component; - : without the component). In the first two lanes from left, anti-HIV IN sdAb-1 and anti-HFV IN sdAb-1 were incubated with HIV-1 IN, respectivly. The third lane from left is a positive control without sdAb. The last 3 lanes on the right are negative controls. The band indicated as 'Degradation' (molecular weight around 45 kDa) is a degraded LEDGF/p75 from purification product.

5.6. Anti-HIV IN sdAbs expressed ex vivo cannot block the replication of HIV-1.

Though the inhibitory activity of recombinant sdAbs were not found in vitro, it is still interesting to know if ex vivo expressed sdAbs are able to block the infection of HIV-1 virus. Hence, anti-HIV IN sdAb-1, anti-HIV IN sdAb-4 and a negative control anti-HIV IN sdAb-3 were transitory expressed as sdAb-EGFP fusion proteins in Hela P4 cells. Anti-HIV sdAb-1 and sdAb-4 have shown IN-recognizing ability in vitro. Negative control sdAb-3 is a nonspecific sdAb. We observed that all the sdAbs were well expressed as sdAb-EGFP fusion proteins **(Figure 43A).** Confocal microscopy also showed that the sdAb-EGFP fusion proteins were highly expressed in vivo **(Figure 43B)**. But the blocking of HIV-1 replication by sdAb was not observed in HIV infectivity assay in which single-cycle titers of the virus were determined in HeLa-P4 cells.

A



B

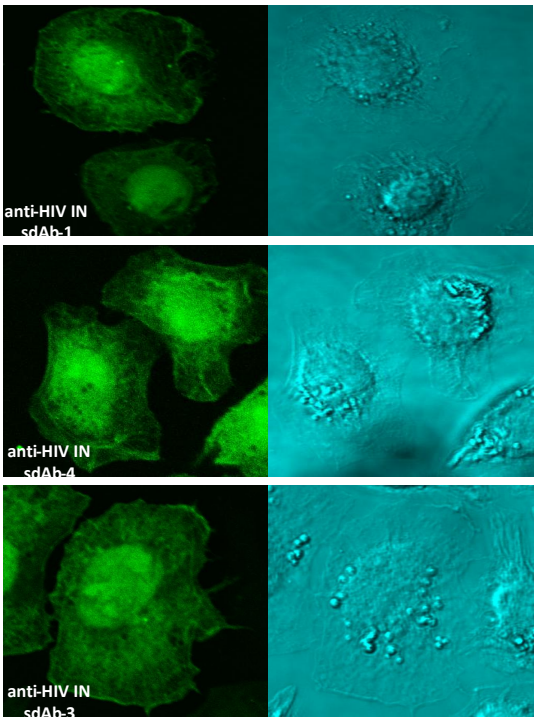


Figure 43: Ex vivo transitory expression of sdAb-EGFP fusion proteins.

(A) A representative FACS assay shows the transitory expression of sdAb-EGFP fusion proteins (48 hours after transfection). (B) Representative confocal microscopy photos show the transitory expression of sdAb-EGFP fusion proteins in Hela P4 cells (48 hours after transfection).

5.7. Transitory co-expressed anti-HIV IN sdAbs and IN-HA are not co-localized ex vivo.

To determine if sdAbs are able to recognize HIV-1 IN in vivo, we also expressed the sdAbs fused with tags 6 x histine and c-Myc. HeLa cells were transitory co-transfected with HA-IN and either anti-HIV IN sdAb-1 or anti-HIV IN sdAb-4. According to immunofluorescence assay, IN and sdAbs were well expressed **(Figure 44)**. However, the co-localization of IN and sdAb was not observed, which suggests that in vivo expressed sdAbs can not bind to INs. IN-HA was preferentially expressed in nucleus, while sdAbs were preferentially expressed in cytoplasm.

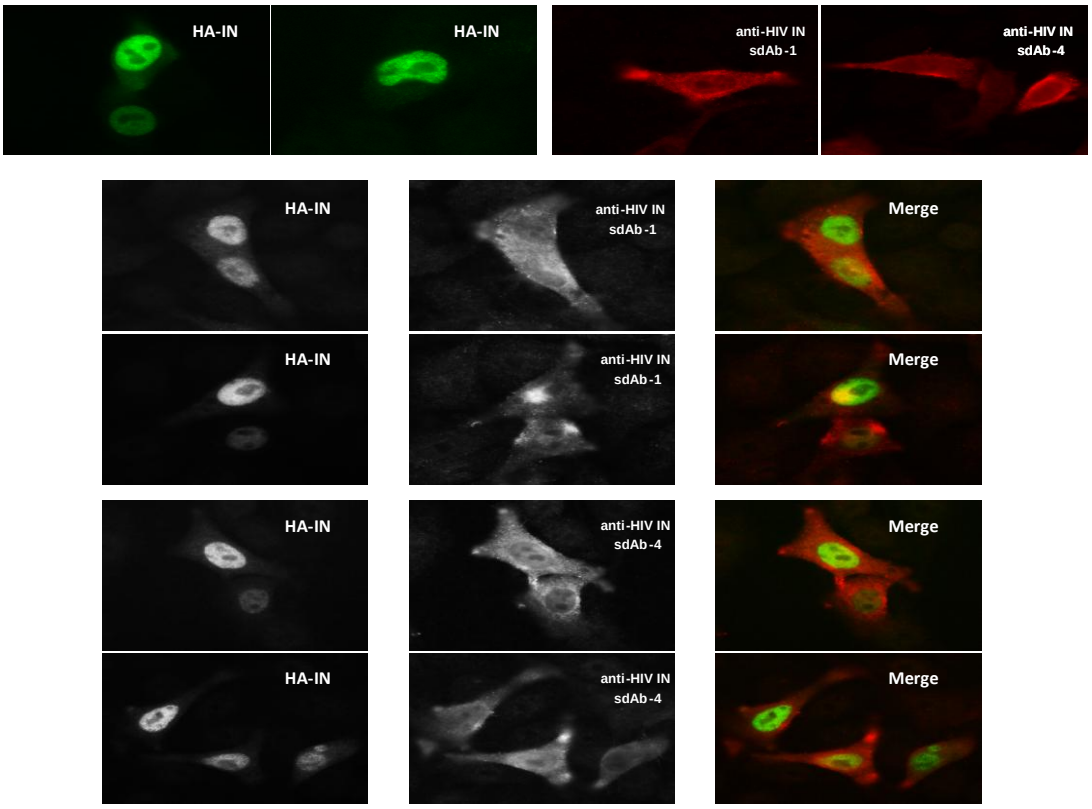


Figure 44: Transitory co-expression of anti HIV-1 IN sdAbs and IN-HA in Hela cells (By Bouchet, J., Institut Cochin).

6. Selection of anti-HIV IN sdAb with Epoxy bead coupled HIV-1 IN

Phage antibody library selection is usually carried out using antigens directly coated onto a plastic surface (e.g., immunotubes). This straightforward method is easy to perform and has been shown to be successful for a diverse set of Ags. However, the limitation of selection with plastic surface is the relatively poor coating efficiency and the altered availability of epitopes on plastic-coated antigens. Compared with being coated on immunotube, the conformations of IN coupled on magnetic bead may be more similar to in vivo situation. Moreover, to select sdAbs which recognize distinct epitopes from anti-HIV sdAb-1, recombinant anti-HIV sdAb-1 was used to saturate its epitope before the selection of specific phage-sdAb. Sequencing result shows that most of sdAbs selected with this method are quite different from anti-HIV sdAb-1 **(Figure 45)**. All these recombinant anti-HIV sdAbs exhibited IN-recognizing ability in western blot assay **(Figure 46)**. Further work should be engaged to verify their in vitro and in vivo inhibitory activity on IN.

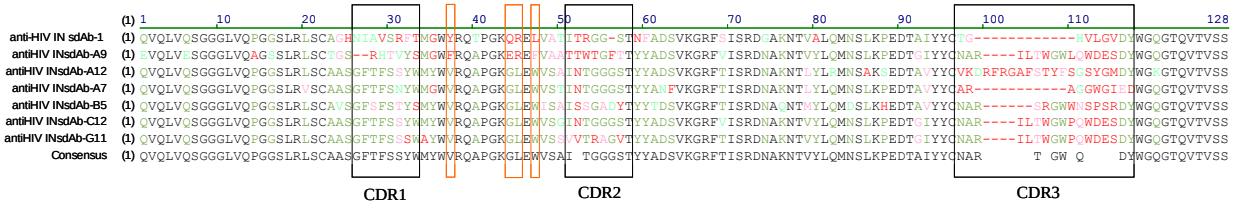


Figure 45: Align of sdAbs selected by Epoxy bead coupled HIV-1 IN.

Three CDRs are marked with black boxes. The four hallmark amino acids differing between VH and VHH in framework (Val37Phe/Tyr, Gly44Glu/Gln, Leu45Arg/Cys and Trp47Gly/Ser/Leu/Phe, in going from VH to VHH) are marked with orang boxs.

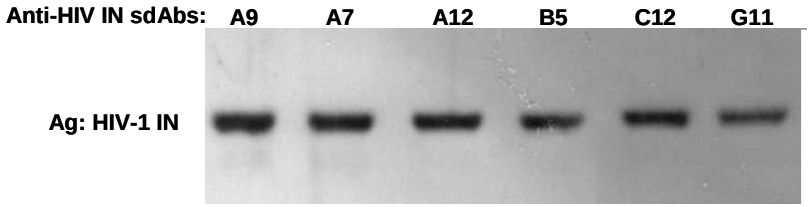


Figure 46: Recognition of HIV-1 IN by recombinant sdAbs.

7. Developing sdAb as targeting reagents for nanosensors

Though the identified anti-IN sdAbs are not able to block HIV infection, presumably due to the location of its binding sites or insufficient affinity, the anti-IN sdAbs could be developed as targeting reagents for surface plasmon resonance (SPR), as a platform of nanosensor. Perturbations at the gold surface of a biochip, e.g., an interaction between the antibody immobilized on the chip and captured target protein, induce a modification of resonance which is in turn observed as a measurable change in reflectivity. SPR imaging (SPRi) technology is a sensitive label-free method of visualizing the entire biochip via a high resolution CCD camera which provides real-time images for the array of active spots on the chip. By capturing the real-time changes on the surface of the biochip, SPRi provides detailed information on molecular binding, biomolecular interactions and kinetic processes.

In our study, linked to the gold surface of a prism, the anit-IN sdAbs which are small and able to recognize IN specifically can be used as a supporter of IN while IN exhibits a natural conformation and remains its enzymatic activity. This platform can be developed as a platform for high-throughput screening of novel IN inhibitors and drug/IN interaction mechanistic studies.

In the preliminary assay of SPRi, different sdAbs were deposed on gold surface of a prism. Then the protein solution was injected on the surface. One sdAb, anti-HIV IN A7, was demonstrated to be an anti-his tag Ab by recognizing non-specific proteins fused with histine tag (LacI-his and Prh-his) **(Figure 47A and C)**. The other sdAbs deposed can specifically recognize recombinant IN **(Figure 47B and C)**. Further work will be devoted to improve this platform, such as the coupling of sdAb to gold surface.

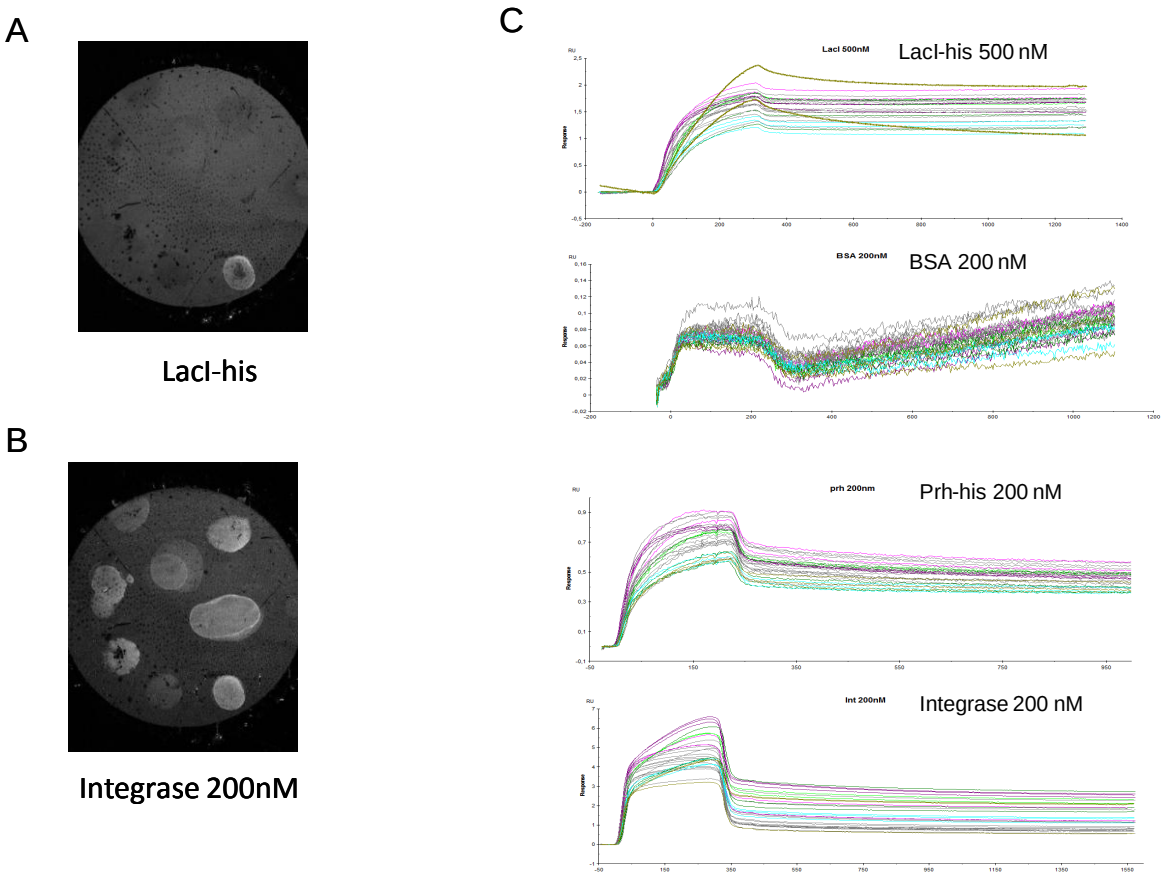


Figure 47: Developing sdAb as targeting reagents for nanosensors (preliminary result of surface plasmon resonance)

Different sdAbs were deposed on gold surface of a prism. Then the protein solutions, non-sepecific protein fused with histine tag (LacI-his and Prh-his), BSA and integrase were loaded on the surface, respectively.

III. 4. Discussion

Engineered antibodies have apparently come of age as biopharmaceuticals, with 18 monoclonal antibody products currently on the market and more than 100 in clinical trials (Holliger and Hudson, 2005b). Smaller recombinant antibody (e.g., scFv) and engineered variants (e.g., diabodies) were developed as alternatives of monoclonal antibody, while retaining the targeting specificity and being produced more economically. Many scFvs have shown their potency for inhibition of their targets in cells. For instance, scFvs targeting HIV-1 IN can neutralize IN activity prior to integration, resulting in resistance to productive HIV-1 infection (Levy-Mintz et al., 1996). However, disulfide bonds of scFvs are unlikely to form in reducing compartments of eukaryotic cells such as cytoplasm and nucleus, resulting in the decrease of scFvs affinity. Furthermore, anti-IN peptides were also designed to inhibit the activity of IN. Using the protein-protein interaction of IN multimer or IN/cellular cofactor, anti-IN peptides derived from Rev, LEDGF/p75, N terminal domain or NLS domain of IN were successfully developed and demonstrated the inhibition of IN activity as well as HIV replication and integration in infected cells (Hayouka et al., 2007; Levin et al., 2009; Maes et al., 2009; Rosenbluh et al., 2007). However, peptides generally do not make good drug candidate due to the shortcomings such as rapid enzymatic degradation, poor bioavailability and aqueous solubility, sensitivity to heat, pH and oxidation. Hence, offering added-value over conventional antibody fragments, sdAbs could overcome these limitations and be promising candidate for a range of diagnostic and therapeutic applications.

In virtue of the essential role of IN in HIV replication cycle, anti-IN sdAbs were developed in this study. Several specific anti-IN sdAbs were successful identified after selection. The recombinant anti-IN sdAbs showed specific target-recognizing ability in different in vitro assays. Remarkably, the anti-IN sdAbs exhibit a broad IN binding repertoire. As well as binding HIV-1 IN, anti-IN sdAbs are able to recognize HIV-2 IN and HFV IN. It is not surprising since the level of conservation among retroviral INs is high, in particular within their active sites (Engelman and Craigie, 1992; Kulkosky et al., 1992; Valkov et al., 2009). HFV IN was served as a convenient proxy for HIV structural studies in a recent research (Hare et al., 2010b). Encouragingly, the anti-IN sdAbs also exhibit the specific target-recognizing ability to INSTI resistance mutant HIV-1 IN G140S/Q148R. Despite the INSTIs, e.g., raltegravir and elvitegravir, are highly efficient in antiretroviral therapy, IN resistance mutants emerge from the preliminary clinic application, due to the low genetic barrier. G140S/Q148R is one of the primary pathways conferring the

resistance to INSTI, leading a highly reduced susceptibility (more than 1000 fold) to both raltegravir and elvitegravir (McColl and Chen, 2010). In virtue of the distinct inhibitory mechanism of sdAb from existing INSTI, anti-HIV IN sdAb is an attractive candidate for antiretroviral treatment.

Cellular co-factor LEDGF/p75, preferentially binding the N-terminal domain of IN within PIC, is essential for nuclear import and chromosome tethering of the PIC, and influence the genome-wide pattern of HIV integration (Engelman and Cherepanov, 2008; Poeschla, 2008). It is interesting to investigate if the fixation of sdAb is able to interfere the interaction of IN and LEDGF/p75. However, in vitro assay showed that the recombinant sdAbs do not interfere the formation of IN/LEDGF/p75 complex. We speculate that the binding-sites of these identified anti-IN sdAbs might not be located on the interaction surface of IN/LEDGF/p75 complex.

As shown in IN enzymatic activity assay, the identified anti-IN sdAbs are not able to efficiently block the catalytic activity of IN in vitro. The estimated IC_{50} is more than 30 μM. These intracellular anti-IN sdAbs are unable to block HIV replication when the Hela P4 cells transitorily expressing anti-IN sdAb were infected by HIV-1 virus. It was suggested that successful inhibition of HIV-1 replication depends upon neutralizing relevant functional domains within the target protein (Mhashilkar et al., 1995; Shaheen et al., 1996), hence we presume that either the binding sites of these sdAbs may not be located in active sites of IN, or the affinity of these sdAbs is insufficient to block IN activity and virus replication, although the effects of anti-HIV Ab do not correspond to the binding affinities of anti-HIV-1 IN antibodies (Bizub-Bender et al., 1994). Moreover, the co-localization of IN and sdAb was not observed in vivo. IN is preferentially expressed in nucleus, while the expression of anti-IN sdAb favors cytoplasm. This is consistent to the previous study in which the anti-IN scFv were located mainly in cytoplasmic with nuclear exclusion, while the anti-IN scFv introduced a Tat NLS domain were localized within the nucleus, specially surrounding the inner nuclear membrane (Levy-Mintz et al., 1996).

Remarkably, the intracellular expression of anti-IN sdAb as EGFP fusion protein is quite well. Probably because the small fragment containing a single domain does not require to be paired with other chain or domain neither the presence of binding peptide, and can be fold efficiently in active form under the reducing environment such as the cytoplasm of eukaryotic cells.

In addition, our result shows that the selection of sdAb with Epoxy bead, on which the coated antigen exhibits the epitopes better, is a good alternative for immunotube selection and able to identify binders with higher diversity.

Though the identified anti-IN sdAbs are not able to block HIV infection, presumably due to the location of its binding sites or insufficient affinity, the anti-IN sdAbs can specifically recognize IN in vitro without blocking the enzymatic activity. Moreover, due to the small size of sdAb (about 13 kDa), sdAb may not interfere the conformation of IN. Hence, as a suitable supporter of IN while remaining IN activity, sdAb could be developed as a targeting reagent of IN for SPRi (nanosensor) where good structural, mechanistic and immunological data are available. SPRi can be used to characterise the kinetics of the antibody-target interaction in greater detail than ELISA methods, and also well suited to a multiplexed immunoassay using sdAbs.

The qualities of sdAbs, such as good solubility, well expression in vivo and the distinct inhibition mechanism from exiting INSTIs, constantly render sdAbs to be a promising candidate for developing antiretroviral agent. In future, selection of anti-IN sdAbs with high affinity and binding sites at active domains of IN remains attractive for obtaining efficient IN inhibitor which is potential for antiretroviral therapy.

GENERAL CONCLUSIONS

- The vast majority of cases of HIV-1-infected people worldwide harbor non-subtype B variants of HIV-1. The HIV-1 has been classified into subtypes, as well as circulating and unique recombinant forms (CRF and URF respectively), because of significant natural genetic variation. Enzymatic and virologic data demonstrate that naturally occurring polymorphisms in different HIV-1 non-B subtypes can affect their susceptibility to some antiretroviral drugs. Recombinant A/G subtype virus showed lower baseline susceptibilities to PIs, and some A/G isolates have reduced susceptibility to a NRTI, abacavir (Martínez-Cajas et al., 2008).

This first part of this study evaluated the susceptibility of CRF02_AG, a predominant non-B HIV-1 subtype in West Africa, to INSTI which represents a novel antiretroviral drug class. Previous in silico study revealed some significant nature variations, such as K/R14, T/V112, T/A125, G/N134, K/T136, T/S206, and S/G283, located in IN of non-B HIV-1 CRF02_AG subtype, compared with HIV-1 subype B. Firstly, our in vitro study showed that these nature variations do not significantly effect the enzymatic activity of IN. The DNA binding activity, as well as 3'-processing and strand transfer activity of CRF02_AG subtype INs, is quite comparable with the HIV-1 B subtype IN. This conclusion is consistent with the molecular modeling of free and viral DNA bound INs which shows no significant structural difference existing between the two subtypes. In silico study also shows that most of the significant variations are located far from the active site; and two variations at positions 134 and 136, near to the active site, are exposed to the solvent and do not significantly affect the structure of CRF02_AG IN.

The susceptibilities of both CRF02_AG subtype and HIV-1 B subtype to the three employed INSTIs are similar. Molecular modelling analysis also confirmed that the conformation modes of binding and docking of the three studied INSTIs are identical for HIV-1 subtype B and CRF02_AG strains. Similar inhibitory activity were observed from studies carried out either on subtypes B and C IN (Bar-Magen et al., 2009), or on HIV-2 clinical isolates obtained from IN inhibitor- naïve patients (Roquebert et al., 2008b). Though the natural variations possessed by IN (proportion of positions with $\geq$ 0.5% variability is about 34.7%) are almost as many as those by protease (about 37.2%) (Rhee et al., 2008), they rarely associated with susceptibility to INSTIs, in contrast to the lower baseline susceptibilities of recombinant A/G subtype to PIs.

The core contact sites of INSTI normally consist conserved amino acid residues in IN and invariant viral nucleotide bases, which make the INSTIs share a similar binding and action mode. Therefore, we presume that the structure modifications caused by the amino acid variations of CRF02_AG IN might not involved in INSTI binding sites, thereby can

not affect DNA binding and catalytic activity of IN in virtue of certain enzyme conformation complementarities.

Our study suggests that the application of INSTIs against HIV-1 subtype CRF02_AG would probably result in comparable outcomes to those obtained against subtype B infections.

- As the first and to date unique INSTI approved by FDA, RAL has exhibited high clinical efficacy of treating HIV-2 infected patients which have limited antiretroviral therapy options. Only 40% nucleotide identity and 65% amino acid similarity between HIV-1 and HIV-2 INs, the response of HIV-2 IN to RAL and particularly the mechanisms of resistance emergence might eventual differ from HIV-1 IN, because differences in codon sequences at positions associated with drug resistance mutations might predispose viruses of different types to encode varied amino acid substitutions. The activity profiles and susceptibility to RAL of HIV-2 INs from plasma samples of HIV-2 patients were characterized.

Firstly, the variable lengths of HIV-2 subtype B INs were observed. These variations were caused by a difference in the length of the nucleotide sequence, instead of the in-frame stop codons within the sequence. Despite the diverse IN lengths for HIV-2 subtype B, IN enzymatic activity is not impacted, since these INs exhibited similar in vitro catalytic activity which is comparable to the activity of an HIV-1 IN as well. Consistent with the comparable phenotypic sensibility of HIV-1 and HIV-2 to INSTI, HIV-2 IN demonstrated the same extend of susceptibility to RAL as HIV-1 IN, which suggest the similar interaction mode between RAL and both INs. It is not surprisingly since the residues of IN involved in the interaction with RAL are completely conserved in HIV-2 (Mouscadet et al., 2010b).

HIV-2 population sequencing analysis revealed that three primary pathways of RAL resistance involving mutations on residues Q148, N155 and Y143, as for HIV-1(Charpentier et al., 2008; Garrett et al., 2008; Roquebert et al., 2008a). Moreover, E92Q and T97A are associated to one or two of the primary genetic pathways (Xu et al., 2009). Three different patterns E92A/T97A/N155H, G140S/Q148R and E92Q/143C were found by the clonal analysis on clinic isolates, confirming the possibility of these three possible pathways. Indeed, recombinant INs from clinic isolates containing the three different patterns were strongly resistant to RAL in vitro.

Then, by introducing single-site mutations in a RAL-sensitive background of HIV-2 IN, our study confirmed that N155H mutation and G140S/Q148H double mutations were sufficient to confer a strong resistance to RAL in vitro. Consistent with our previous

observation on HIV-1 (Delelis et al., 2009b), Q148H induced a dramatic catalytic defect while conferring a high level of resistance to RAL. Secondary G140S mutation restores the catalytic activity though it does not confer any resistance. The concomitant selection of Q148H and G140S mutations indicates that their close interaction is conserved in retroviral INs (Hare et al., 2010d).

Though Y143C/R has been described as a primary mutation for HIV-1 resistance to RAL (Sichtig et al., 2009b), Y143C alone in a RAL-sensitive background was not sufficient to confer resistance to IN, ruling out this mutation as the unique determinant for the resistance in an HIV-2 context. However, when E92Q was introduced simultaneously, the double mutant Y143C/E92Q elicits RAL-resistance in vitro, though single E92Q mutation is neither sufficient to confer the resistance to RAL. Hence, the concomitant presence of Y143C/E92Q is likely to constitute another main pathway toward resistance. As for HIV-2, losing the contact with the side chain of Y143 which is established by RAL is not sufficient to impair RAL binding to IN. A second modification of the RAL binding site, e.g., E92Q, is probably required, thus plays a more important role in the HIV-2 context than in the HIV-1 context. What' more, our data also suggests that the Y143C mutation counteracts the resistance effect of the N155H mutation, presumably precluding the simultaneous selection of both mutations.

Additionally, though whole population study suggested that T97A might be associated to both N155H and Y143C mutations in HIV-2 resistant virus (Charpentier et al., 2011; Charpentier et al., 2010), the significant effect of T97A on HIV-2 IN susceptibility to RAL, neither alone nor in combination with Y143C, was not observed in vitro.

- Emergence of resistance to antiretroviral drugs is a main impetus for the development of novel drug classes with distinct blocking mechanism or new agents in classes without cross-resistance (Flexner, 2007), though 23 approved drugs are available for the treatment of HIV/AIDS currently. Despite the high efficiency of INSTI to varied HIV-1 subtypes and HIV-2 INs as well, resistance generates with low genetic barrier. Specific neutralizing antibodies represent an outstanding alternative strategy for antiretroviral treatment. Though many scFvs have shown their potency for inhibition of their targets in cells, the formation of disulfide bonds in reducing compartments of eukaryotic cells is an inherent limitation. Various anti-IN peptides which are able to block HIV replication also have a lot of intrinsic drawbacks, such as rapid degradation (Hayouka et al., 2007; Rosenbluh et al., 2007). Hence, sdAb composing only one domain, together with other unique advantage, is

an exceptional candidate for developing specific inhibitor for IN and a useful research tool for fundamental study of IN. The last part of our study is to develop camelid sdAb targeting HIV-1 IN.

Several specific anti-IN sdAbs were successfully identified after selections with recombinant IN. These recombinant anti-IN sdAbs showed specific target-recognizing ability in vitro. The anti-IN sdAbs also possess the capacity to recognize INs of diverse retroviruses, such as HIV-2 IN and HFV IN, which is likely owing to the high level of conservation among retroviral INs, particularly their active sites. Encouragingly, these recombinant sdAbs exhibit a similar target-recognizing ability for HIV-1 IN mutant G140S/Q148R which is one of the primary INSTI resistance mutants and causes a highly reduced susceptibility to INSTI (more than 1000 fold) (McColl and Chen, 2010). This evidence suggests that anti-HIV IN sdAb is promising to be developed as an antiretroviral agent by virtue of the distinct inhibition mechanism from existing INSTI compounds.

However, in vitro assay showed that the recombinant sdAbs do not interfere the formation of IN and cellular co-factor LEDGF/p75 complex, putatively because their binding-sites may be out of the interaction sites of IN/LEDG/p75 complex.

The identified anti-IN sdAbs are not able to efficiently block the catalytic activity of IN in vitro, with an estimated IC_{50} greater than 30 μM. These intracellular expressed anti-IN sdAbs can neither block HIV replication. Since the neutralization of relevant functional domains within the target protein is determining for the successful inhibition of HIV-1 replication (Mhashilkar et al., 1995; Shaheen et al., 1996), we presume that either the binding sites of these sdAbs may not be located in active sites of IN; or the affinity of these sdAbs is insufficient to block IN activity and virus replication.

IN is preferentially expressed in nucleus. In contrast, anti-IN sdAbs are mainly expressed in cytoplasm, which is consistent with the previous study in which anti-IN scFv located mainly in cytoplasmic with nuclear exclusion (Levy-Mintz et al., 1996). Moreover, the intracellular expression of anti-IN sdAb is quite well, because the single domain fragment does not require to be paired with other chain or domain neither the linker, and can be efficiently fold in active form under the reducing compartment such as the cytoplasm of eukaryotic cells. Additionally, our result confirms that the method of sdAb selection with antigen coupled on Epoxy bead is able to identify highly diverse binders.

Despite that the identified anti-IN sdAbs are not able to block HIV infection, presumably due to the location of its binding sites or insufficient affinity, the anti-IN sdAb could be developed as targeting reagent for the platform of nanosensor. The anit-IN sdAb, which is able to recognize IN specifically and may not disturb the conformation of IN due

to its small size of about 13 kDa, can be developed as a supporter of IN on the gold surface of biochip in SPRi, rendering IN to exhibit its natural conformation and remain enzymatic activity. SPRi can be used to characterise the kinetics of the antibody-target interaction in greater detail and also well suited to a multiplexed immunoassay using sdAbs. This platform can be applied for high-throughput screening of novel IN inhibitors and drug/IN interaction mechanistic studies. Our preliminary assay of SPRi showed that the recombinant sdAbs can recognize HIV IN specifically. Further work will be performed on such as the optimization of deposing sdAb on the gold surface of prism.

The unique qualities of sdAbs, e.g., the distinct inhibition mechanism from exiting INSTIs, make sdAbs an attractive candidate for the development of antiretroviral agent. Selection of anti-IN sdAbs with high affinity and binding sites involved in the active domains of IN remains promising for obtaining efficient neutralizing antibody of IN.

REFERENCES

References

Adamson, C.S., and Freed, E.O. (2008). Recent progress in antiretrovirals-lessons from resistance. Drug Discovery Today *13*, 424-432.

Agapkina, J., Smolov, M., Barbe, S., Zubin, E., Zatsepin, T., Deprez, E., Bret, M.L., Mouscadet, J.-F., and Gottikh, M. (2006). Probing of HIV-1 Integrase/DNA interactions using novel analogs of viral DNA. Journal of biological chemistry *281* 11530-11540.

Aires da Silva, F., Santa-Marta, M., Freitas-Vieira, A., Mascarenhas, P., Barahona, I., Moniz-Pereira, J., Gabuzda, D., and Goncalves, J. (2004). Camelized rabbit-derived VH single-domain intrabodies against Vif strongly neutralize HIV-1infectivity. J Mol Biol *340*, 525-542.

An, W., and Telesnitsky, A. (2002). HIV-1 genetic recombination: experimental approaches and observations. AIDS Rev *4*, 195-212.

Ao, Z., Huang, G., Yao, H., Xu, Z., Labine, M., Cochrane, A.W., and Yao, X. (2007). Interaction of human immunodeficiency virus type 1 integrase with cellular nuclear import receptor importin 7 and its impact on viral replication. . J Biol Chem *282*, 13456-13467.

Archer, R.H., Dykes, C., Gerondelis, P., Lloyd, A., Fay, P., Reichman, R.C., Bambara, R.A., and Demeter, L.M. (2000). Mutants of human immunodeficiency virus type 1 (HIV-1) reverse transcriptase resistant to nonnucleoside reverse transcriptase inhibitors demonstrate altered rates of RNase H leavage that correlate with HIV-1 replication fitness in cell culture. J Virol *74*, 8390-8401.

Arhel, N.J., Souquere-Besse, S., Munier, S., Souque, P., Guadagnini, S., Rutherford, S., Prévost, M.-C., Allen, T.D., and Charneau, P. (2007). HIV-1 DNA Flap formation promotes uncoating of the pre-integration complex at the nuclear pore. EMBO J *26*, 3025-3037.

Arthur, L.O., Bess, J.W.J., Sowder, R.C., Benveniste, R.E., Mann, D.L., Chermann, J.C., and Henderson, L.E. (1992). Cellular proteins bound to immunodeficiency viruses: implications for pathogenesis and vaccines. Science *258*, 1935-1938.

Bailey, J.R., Sedaghat, A.R., Kieffer, T., Brennan, T., Lee, P.K., Wind-Rotolo, M., Haggerty, C.M., Kamireddi, A.R., Liu, Y., , , Lee, J., *et al.* (2006). Residual human immunodeficiency virus type 1 viremia in some patients on antiretroviral therapy is dominated by a small number of invariant clones rarely found in circulating CD4+ T cells. J Virol *80*, 6441-6457.

Bar-Magen T, Sloan RD, Faltenbacher VH, Donahue DA, Kuhl BD, Oliveira M, Xu H, and MA, W. (2009). Comparative biochemical analysis of HIV-1 subtype B and C integrase enzymes. Retrovirology *6*.

Bar-Magen, T., Donahue. D.A., McDonough. E.I., Kuhl. B.D., Faltenbacher, V.H., Xu, H., Michaud, V., Sloan, R.D., and Wainberg, M.A. (2010). HIV-1 subtype B and C integrase enzymes exhibit differential patterns of resistance to integrase inhibitors in biochemical assays. AIDS *24*, 2171-2179.

Bar-Magen, T., Sloan, R.D., Faltenbacher, V.H., Donahue, D.A., Kuhl, B.D., Oliveira, M., Xu, H., and Wainberg, M.A. (2009). Comparative biochemical analysis of HIV-1 subtype B and C integrase enzymes. Retrovirology *6*.

Barboric, M., Nissen, R.M., Kanazawa, S., Jabrane-Ferrat, N., and Peterlin, B.M. (2001). NF-┤ B binds P-TEFb to stimulate transcriptional elongation by RNA polymerase II. Mol Cell Biol *8*, 327-337.

Barré-Sinoussi, F., Chermann, J.C., Rey, F., Nugeyre, M.T., Chamaret, S., Gruest, J., Dauguet, C., Axler-Blin. C., Vézinet-Brun, F., Rouzioux, C., *et al.* (1983). Isolation of a T-lymphotropic retrovirus from a patient at risk for acquired immune deficiency syndrome (AIDS). Science *220*, 868–871

Behar, G., Chames, P., and Teulon, I. (2009). Llama single domain antibodies directed against

nonconventional epitopes of tumor-associated carcinoembryonic antigen absent from nonspecific crossreacting antigen. . FEBS J *276*, 3881-3893.

Beitzel, B., and Bushman, F. (2003). Construction and analysis of cells lacking the HMGA gene family. Nucleic Acids Res *31*, 5025-5032.

Bercoff, D.P., Triqueneaux, P., Lambert, C., Oumar, A.A., Ternes, A.M., Dao, S., Goubau, P., Schmit, J.C., and Ruelle, J. (2010). Polymorphisms of HIV-2 integrase and selection of resistance to raltegravir. Retrovirology *7*, 98.

Béthune, M.-P.d. (2010). Non-nucleoside reverse transcriptase inhibitors (NNRTIs), their discovery, development, and use in the treatment of HIV-1 infection: A review of the last 20 years (1989-2009). Antiviral Research *85*, 75-90.

Biocca, S., Ruberti, F., Tafani, M., Pierandrei-Amaldi, P., and Cattaneo, A. (1995). Redox state of single chain Fv fragments targeted to the endoplasmic reticulum, cytosol and mitochondria. Biotechnology *13*, 1110-1115.

Bizub-Bender, D., Kulkosky, J., and Skalka, A.M. (1994). Monoclonal antibodies against HIV type 1 integrase: clues to molecular structure. . AIDS Res HumRetroviruses *10*, 1105-1115.

Bond, C.J., Marsters, J.C., and Sidhu, S.S. (2003). Contributions of CDR3 to VHH domain stability and the design of monobody scaffolds for naive antibody libraries. J Mol Biol *332*, 643-655.

Borrebaeck, C.A.K., Malmborg, A.C., and Furebring, C.e.a. (1992). Kinetic analysis of recombinant antibody-antigen interactions: relation between structural domains and antigen binding. Biotechnology *10*, 697-698.

Bosco, D.A., Eisenmesser, E.Z., Pochapsky, S., Sundquist, W.I., and Kern, D. (2002). Catalysis of cis/trans isomerization in native HIV-1 capsid by human cyclophilin A. Proc Natl Acad Sci U S A *99*, 5247-5252.

Bouchet, J., Basmaciogullari, S.E., Chrobak, P., Stolp, B., Bouchard, N., Fackler, O.T., Chames, P., Jolicoeur, P., Benichou, S., and Baty, D. (2011). Inhibition of the Nef regulatory protein of HIV-1 by a single-domain antibody. Blood *117*, 3559-3568.

Bouyac-Bertoia, M., Dvorin, J.D., Fouchier, R.A., Jenkins, Y., Meyer, B.E., Wu, L.I., Emerman, M., and Malim, M.H. (2001). HIV-1 infection requires a functional integrase NLS. Mol Cell *7*, 1025-1035.

Brass, A.L., Dykxhoorn, D.M., Benita, Y., Yan, N., Engelman, A., Xavier, R.J., Lieberman, J., and Elledge, S.J. (2008). Identification of host proteins required for HIV infection through a functional genomic screen. Science *319*, 921-926.

Brenner, B.G., Lowe, M., Moisi, D., Hardy, I., Gagnon, S., Charest, H., Baril, J.G., Wainberg, M.A., and Roger, M. (2011a). Subtype diversity associated with the development of HIV-1 resistance to integrase inhibitors. J Med Virol.

Brenner, B.G., Lowe, M., Moisi, D., Hardy, I., Gagnon, S., Charest, H., Baril, J.G., Wainberg, M.A., and Roger, M. (2011b). Subtype diversity associated with the development of HIV-1 resistance to integrase inhibitors. J Med Virol *83*.

Brown, P.O. (1990). Integration of retroviral DNA. CurrTopMicrobiolImmunol *157:19-48*, 19-48.

Brussel, A., and Sonigo, P. (2004). Evidence for gene expression by unintegrated human immunodeficiency virus type 1 DNA species. J Virol *78*, 11263-11127.

Bukrinsky, M.I., Sharova, N., Dempsey, M.P., Stanwick, T.L., Bukrinskaya, A.G., Haggerty, S., and Stevenson, M. (1992). Active nuclear import of human immunodeficiency virus type 1 preintegration complexes. Proc Natl Acad Sci U S A *89*, 6580-6584.

Cane, P.A. (2009). New developments in HIV drug resistance. J Antimicrob Chemother *64*, i37-40.

Caron de Fromentel, C., Gruel, N., Venot, C., Debussche, L., Conseiller, E., Dureuil, C., Teillaud, J.L., Tocque, B., and Bracco, L. (1999). Restoration of transcriptional activity of p53 mutants in human tumour cells by intracellular expression of anti-p53 single chain Fv fragments. Oncogene *18*, 551-557.

Carteau, S., Gorelick, R.J., and Bushman, F.D. (1999). Coupled integration of human immunodeficiency virus type 1 cDNA ends by purified integrase in vitro: stimulation by the viral nucleocapsid protein. J Virol *73*, 6670-6679.

Ceccherini-Silberstein, F., Malet, I., D'Arrigo, R., Antinori, A., Marcelin, A.G., and Perno, C.F. (2009). Characterization and structural analysis of HIV-1 integrase conservation. AIDS Rev *11*, 17-29.

Cereseto, A., Manganaro, L., Gutierrez, M.I., Terreni, M., Fittipaldi, A., Lusic, M., Marcello, A., and Giacca, M. (2005a). Acetylation of HIV-1 integrase by p300 regulates viral integration. EMBO J *24*, 3070-3081.

Cereseto, A., Manganaro, L., Gutierrez, M.I., Terreni, M., Fittipaldi, A., Lusic, M., Marcello, A., and Giacca, M. (2005b). Acetylation of HIV-1 integrase by p300 regulates viral integration. EMBO J *24*.

Charpentier, C., Karmochkine, M., Laureillard, D., Tisserand, P., Belec, L., Weiss, L., Si-Mohamed, A., and Piketty, C. (2008). Drug resistance profiles for the HIV integrase gene in patients failing raltegravir salvage therapy. HIV Med *9*, 765-770.

Charpentier, C., Larrouy, L., Matheron, S., Damond, F., Delelis, O., Mouscadet, J.-F., Campa, P., Chêne, G., Brun-Vezinet, F., and Descamps, D. (2011). Long-Lasting persistence of integrase resistance mutations in HIV-2-infected patients after raletgravir removal. Antiviral Therapy *in press*.

Charpentier, C., Roquebert, B., Delelis, O., Larrouy, L., Matheron, S., Tubiana, R., Karmochkine, M., Duval, X., Chene, G., Storto, A., *et al.* (2010). Hot Spots of Integrase Genotypic Changes Leading to HIV-2 Resistance to Raltegravir. Antimicrob Agents Chemother *55*, 1293-1295.

Charpentier, C., Roquebert, B., Delelis, O., and VIH-2), t.F.A.H.-C.A.C. (2010). Hot spots of integrase genotypic changes leading to HIV-2 resistance to raltegravir. Antimicrob Agents Chemother.

Chen, X., Tsiang, M., Yu, F., Hung, M., Jones, G.S., Zeynalzadegan, A., Qi, X., Jin, H., Kim, C.U., Swaminathan, S., *et al.* (2008). Modeling, analysis, and validation of a novel HIV integrase structure provide insights into the binding modes of potent integrase inhibitors. J Mol Biol *380*, 504-519.

Cherepanov, P., Maertens, G., Proost, P., Devreese, B., Van Beeumen, J., Engelborghs, Y., De Clercq, E., and Debyser, Z. (2002). HIV-1 integrase forms stable tetramers and associates with LEDGF/p75 protein in human cells. J Biol Chem *278*.

Chiu, T.K., and Davies, D.R. (2004). Structure and function of HIV-1 integrase. Curr Top Med Chem *4*, 965-977.

Chow, S.A., Vincent, K.A., Ellison, V., and Brown, P.O. (1992). Reversal of integration and DNA splicing mediated by integrase of human immunodeficiency virus. Science *255*, 723-726.

Christ, F., Thys, W., De Rijck, J., Gijsbers, R., Albanese, A., Arosio, D., Emiliani, S., Rain, J.C., Benarous, R., Cereseto, A., *et al.* (2008). Transportin-SR2 imports HIV into the nucleus. Curr Biol *18*, 1192-1202.

Cihlar, T., and Ray, A.S. (2010). Nucleoside and nucleotide HIV reverse transcriptase inhibitors: 25 years after zidovudine. Antiviral Research *85*, 39-58.

Clavel, F., Brun-Vézinet, F., Guétard, D., Chamaret, S., Laurent, A., Rouzioux, C., Rey, M., Katlama, C., Rey, F., Champelinaud, J.L., *et al.* (1986). LAV type II: a second retrovirus associated with AIDS in West Africa. C R Acad Sci III *302*, 485-488.

Cochet, O., Kenigsberg, M., Delumeau, I., Virone-Oddos, A., Multon, M.C., Fridman, W.H., Schweighoffer, F., Teillaud, J.L., and B., T. (1998). Intracellular expression of an antibody

fragment-neutralizing p21 ras promotes tumor regression. Cancer Res *58*, 1170-1176.

Colin, L., and Lint, C.V. (2009). Molecular control of HIV-1 postintegration latency: implications for the development of new therapeutic strategies. Retrovirology.

Conrath, K.E., Vincke, C., Stijlemans, B., Schymkowitz, J., Decanniere, K., Wyns, L., Muyldermans, S., and Loris, R. (2005). Antigen binding and solubility effects upon the veneering of a camel VHH in framework-2 to mimic a VH. J Mol Biol *350*, 112-125.

Cook, G.P., and Tomlinson, I.M. (1995). The human immunoglobulin VH repertoire. Immunol Today *16*, 237-242.

Coppieters, K., Dreier, T., Silence, K., de Haard, H., Lauwereys, M., Casteels, P., Beirnaert, E., Jonckheere, H., Van de Wiele, C., Staelens, L., *et al.* (2006). Formatted anti-tumor necrosis factor alpha VHH proteins derived from camelids show superior potency and targeting to inflamed joints in a murine model of collagen-induced arthritis. . Arthritis Rheum *54*, 1856-1866.

Cortez-Retamozo, V., Backmann, N., Senter, P.D., Wernery, U., De Baetselier, P., Muyldermans, S., and Revets, H. (2004). Efficient cancer therapy with a nanobody-based conjugate. Cancer Res *64*, 2853-2857.

Damond, F., Brun-Vezinet, F., Matheron, S., Peytavin, G., Campa, P., Pueyo, S., Mammano, F., Lastere, S., Farfara, I., Simon, F., *et al.* (2005). Polymorphism of the human immunodeficiency virus type 2 (HIV-2) protease gene and selection of drug resistance mutations in HIV-2-infected patients treated with protease inhibitors. J Clin Microbiol *43*, 484-487.

Damond, F., Lariven, S., Roquebert, B., Males, S., Peytavin, G., and Morau, G. (2008). Virological and immunological response to HAART regimen containing integrase inhibitors in HIV-2-infected patients. AIDS *22*, 665-666.

Dau, B., and Holodniy, M. (2009). Novel Targets for Antiretroviral Therapy: Clinical Progress to Date. Drugs *69*, 31-50.

Davies, J., and Riechmann, L. (1994). Camelising human antibody fragments: NMR studies on VH domains. . FEBS Lett *339*, 285-290.

Davies, J., and Riechmann, L. (1995). Antibody VH domains as small recognition units. Biol Technol *13*, 475-479.

De Genst, E., Handelberg, F., Van Meirhaeghe, A., Vynck, S., Loris, R., Wyns, L., and Muyldermans, S. (2004). Chemical basis for the affinity maturation of a camel single domain antibody. J Biol Chem *279*, 53593-53601.

De Genst, E., Silence, K., Decanniere, K., Conrath, K., Loris, R., Kinne, J., Muyldermans, S., and Wyns, L. (2006). Molecular basis for the preferential cleft recognition by dromedary heavy-chain antibodies. Proc Natl Acad Sci USA *103*, 4586-4591.

DeJesus, E., Berger, D., Markowitz, M., Cohen, C., Hawkins, T., Ruane, P., Elion, R., Farthing, C., Zhong, L., Cheng, A.K., *et al.* (2006b). Antiviral activity,pharmacokinetics, and dose response of the HIV-1 integrase inhibitor GS-9137 (JTK-303) in treatment-naive and treatment-experienced patients. Journal of Acquired Immune Deficiency Syndromes (JAIDS) *43*, 1-5.

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

Delelis, O., Malet, I., Na, L., Tchertanov, L., Calvez, V., Marcelin, A.-G., Subra, F., Deprez, E., and Mouscadet, J.-F. (2009a). The G140S mutation in HIV integrases from raltegravir-resistant patients rescues catalytic defect due to the resistance Q148H mutation. Nucleic Acids Research, 1–9.

Delelis, O., Malet, I., Na, L., Tchertanov, L., Calvez, V., Marcelin, A.G., Subra, F., Deprez, E., and

Mouscadet, J.F. (2009b). The G140S mutation in HIV integrases from raltegravir-resistant patients rescues catalytic defect due to the resistance Q148H mutation. Nucleic Acids Res *37*, 1193-1201.

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

Delelis, O., Thierry, S., Subra, F., Simon, F., Malet, I., Alloui, C., Sayon, S., Calvez, V., Deprez, E., Marcelin, A.G., *et al.* (2010). Impact of Y143 HIV-1 integrase mutations on resistance to raltegravir *in vitro* and *in vivo*. Antimicrob Agents Chemother *54*, 491-501.

Deng, H., Liu, R., Ellmeier, W., Choe, S., Unutmaz, D., Burkhart, M., Di Marzio, P., Marmon, S., Sutton, R.E., Hill, C.M., *et al.* (1996). Identification of a major co-receptor for primary isolates of HIV-1. Nature *381*, 661-666.

Depienne, C., Mousnie, A., Leh, H., Rouzici, E.L., Dormont, D., Benichoui, S., and Dargemont, C. (2001). Characterization of the nuclear import pathway for HIV-1 integrase. J Biol Chem *276*, 18102-18107.

Derache, A., Traore, O., Koita, V., Sylla, A., Tubiana, R., Simon, A., Canestri, A., Carcelain, G., Katlama, C., Calvez, V., *et al.* (2007). Genetic diversity and drug resistance mutations in human immunodeficiency virus type 1 from untreated patients in Bamako,Mali. Antivir Ther *12*, 123-129.

Desbois, D., Roquebert, B., Peytavin, G., Damond, F., Collin, G., Benard, A., Campa, P., Matheron, S., , , Chene, G., Brun-Vezinet, F., *et al.* (2008). *In vitro* phenotypic susceptibility of human immunodeficiency virus type 2 clinical isolates to protease inhibitors. Antimicrob Agents Chemother *52*, 1545-1548.

Descamps, D., Chaix, M.L., Andre, P., Brodard, V., Cottalorda, J., Deveau, C., Harzic, M., Ingrand, D., Izopet, J., Kohli, E., *et al.* (2005). French national sentinel survey of antiretroviral drug resistance in patients with HIV-1 primary infection and in antiretroviral-naive chronically infected patients in 2001-2002. J Acquir Immune Defic Syndr *38*, 545-552.

Desmyter, A., Spinelli, S., Payan, F., Lauwereys, M., Wyns, L., Muyldermans, S., and Cambillau, C. (2002). Three camelid VHH domains in complex with porcine pancreatic alpha-amylase. Inhibition and versatility of binding topology. J Biol Chem *277*, 23645-23650.

Desmyter, A., Transue, T.R., Ghahroudi, M.A., Thi, M.H., Poortmans, F., Hamers, R., Muyldermans, S., and Wyns, L. (1996). Crystal structure of a camel single-domain VH antibody fragment in complex with lysozyme. Nat Struct Biol *3*, 803-811.

Dickinson, L., Khoo, S., and Back, D. (2009). Pharmacokinetics and drug-drug interactions of antiretrovirals: an update. Antiviral Research.

Doms, R.W., and Trono, D. (2000). The plasma membrane as a combat zone in the HIV battlefield. . Genes Dev *14*, 2677-2688

Dooley, H., and Flajnik, M.F. (2005). Shark immunity bites back: affinity maturation and memory response in the nurse shark, Ginglymostoma cirratum. Eur J Immunol *35*, 936-945.

Doublie, S., Tabor, S., Long, A.M., Richardson, C.C., and Ellenberger, T. (1998). Crystal structure of a bacteriophage T7 DNA replication complex at 2.2 A resolution. Nature *391*, 251-258.

Dulude, D., Baril, M., and Brakier-Gingras, L. (2002). Characterization of the frameshift stimulatory signal controlling a programmed -1 ribosomal frameshift in the human immunodeficiency virus type 1. Nucleic Acids Res *30*, 5094-5102.

Dumoulin, M., Conrath, K., Van Meirhaeghe, A., Meersman, F., Heremans, K., Frenken, L.G., Muyldermans, S., Wyns, L., and Matagne, A. (2002). Single-domain antibody fragments with high

conformational stability. Protein Sci *11*, 500-515.

Dvorin, J.D., Bell, P., Maul, G.G., Yamashita, M., Emerman, M., and Malim, M.H. (2002). Reassessment of the roles of integrase and the central DNA flap in human immunodeficiency virus type 1 nuclear import. J Virol *76*, 12087-12096.

Dyda, F., Hickman, A., Jenkins, T., Engelman, A., Craigie, R., and Davies, D.R. (1994a). Crystal structure of the catalytic domain of HIV-1 integrase: similarity to other polynucleotidyl transferases. Science *266*, 1981-1986.

Dyda, F., Hickman AB, Jenkins TM, Engelman A, Craigie R, and DR., D. (1994b). Crystal structure of the catalytic domain of HIV-1 integrase: similarity to other polynucleotidyl transferases. Science *266*, 1981-1986.

Ebina, H., Aoki, J., Hatta, S., Yoshida, T., and Koyanagi, Y. (2004). Role of Nup98 in nuclear entry of human immunodeficiency virus type 1 cDNA. Microbes Infect *6*, 715-724.

Eijkelenboom, A.P., Sprangers, R., Hard, K., Puras, L.R.A., Plasterk, R.H., Boelens, R., and Kaptein, R. (1999). Refined solution structure of the C-terminal DNAbinding domain of human immunovirus-1 integrase. Proteins *36*, 556-564.

Eijkelenboom, A.P., van den Ent F.M, Wechselberger, R., Plasterk, R.H., Kaptein, R., and Boelens, R. (2000). Refined solution structure of the dimeric N-terminal HHCC domain of HIV-2 integrase. J Biomol NMR *18*, 119-128.

Engelman, A. (1999). In vivo analysis of retroviral integrase structure and function. Adv Virus Res *52*, 411-426.

Engelman, A., and Cherepanov, P. (2008). The lentiviral integrase binding protein LEDGF/p75 and HIV-1 replication. PLoS Pathog *4*.

Engelman, A., and Craigie, R. (1992). Identification of conserved amino acid residues critical for human immunodeficiency virus type 1 integrase function in vitro. J Virol *66*, 6361-6369.

Esser, M.T.e.a. (2001). Differential incorporation of CD45, CD80 (B7-1), CD86 (B7-2) and major histocompatibility complex class I and II molecules into human immunodeficiency virus type 1 virions and microvesicles: implications for viral pathogenesis and immune regulation. . J Virol *75*, 6173-6182

Estéa, J.A., and Cihlarb, T. (2010). Current status and challenges of antiretroviral research and therapy. Antiviral Research *85*, 25-33.

Ewert, S., Cambillau, C., Conrath, K., and Pluckthun, A. (2002). Biophysical properties of camelid V(HH) domains compared to those of human V(H)3 domains. Biochemistry *41*, 3628-3636.

Fassati, A., Görlich, D., Harrison, I., Zaytseva, L., and Mingot, J.M. (2003). Nuclear import of HIV-1 intracellular reverse transcription complexes is mediated by importin 7. EMBO J *22*, 3675-3685.

Faure, A., Calmels, C., Desjobert, C., Castroviejo, M., Caumont-Sarcos, A., Tarrago-Litvak, L., Litvak, S., and Parissi, V. (2005). HIV-1 integrase crosslinked oligomers are active *in vitro*. Nucleic Acids Res *33*, 977-986.

Feng, S., and Holland, E.C. (1988). HIV-1 tat trans-activation requires the loop sequence within tar. Nature *334*, 165-167.

Fleury, H.J., Toni, T., Lan, N.T., Hung, P.V., Deshpande, A., Recordon-Pinson, P., Boucher, S., Lazaro, E., Jauvin, V., Lavignolle-Aurillac, V., *et al.* (2006). Susceptibility to antiretroviral drugs of CRF01_AE, CRF02_AG, and subtype C viruses from untreated patients of Africa and Asia: comparative genotypic and phenotypic data. AIDS Res Hum Retroviruses *22*, 357-366.

Flexner, C. (2007). HIV drug development: the next 25 years. Nat Rev Drug Discov *6*, 959-966.

Gallay, P., Stitt, V., Mundy, C., Oettinger, M., and Trono, D. (1996). Role of the karyopherin pathway

in human immunodeficiency virus type 1 nuclear import. J Virol *70*, 1027-1032.

Gallien, S., Delaugerre, C., Charreau, I., Braun, J., Boulet, T., Barrail-Tran, A., de Castro, N., Molina, J.M., and Kuritzkes, D.R. (2011). Emerging integrase inhibitor resistance mutations in raltegravir-treated HIV-1-infected patients with low-level viremia. AIDS *25*, 665-669.

Gallo, R.C., Salahuddin, S.Z., Popovic, M., Shearer, G.M., Kaplan, M., Haynes, B.F., Palker, T.J., Redfield, R., Oleske, J., and Safai, B. (1984). Frequent detection and isolation of cytopathic retroviruses (HTLV-III) from patients with AIDS and at risk for AIDS. Science *May 4*, 500-503.

Gao, F., Bailes, E., Robertson, D.L., Chen, Y., Rodenburg, C.M., Michael, S.F., Cummins, L.B., Arthur, L.O., Peeters, M., Shaw, G.M., *et al.* (1999). Origin of HIV-1 in the chimpanzee Pan troglodytes troglodytes. Nature *397*, 436-441.

Gao, F., Yue, L., White, A.T., Pappas, P.G., Barchue, J., Hanson, A.P., Greene, B.M., Sharp, P.M., Shaw, G.M., and Hahn, B.H. (1992). Human infection by genetically diverse SIVSM-related HIV-2 in west Africa. Nature *358*, 495-499.

Gao, K., Gorelick, R.J., Johnson, D.G., and Bushman, F. (2003). Cofactors for human immunodeficiency virus type 1 cDNA integration in vitro. J Virol *77*, 1598-1603.

Garrett, N., Xu, L., Smit, E., Ferns, B., El-Gadi, S., and Anderson, J. (2008). Raltegravir treatment response in an HIV-2 infected patient: a case report. Aids *22*, 1091-1092.

Garrido, C., de Mendoza, C., and Soriano, V. (2008). Resistance to integrase inhibitors. Enferm Infecc Microbiol Clin *26 Suppl* 40-46.

Gelderblom, H.C., Vatakis, D.N., Burke, S.A., Lawrie, S.D., Bristol, G.C., and Levy, D.N. (2008). Viral complementation allows HIV-1 replication without integration. Retrovirology *5*.

Gent, D.C.V., Groeneger, A.A.M.O., and Plasterk, R.H.A. (1992). Mutational analysis of the integrase protein of human immunodeficiency virus type 2. Proc Natl Acad Sci USA *89*, 9598-9602.

Goncalves, J., Silva, F., Freitas-Vieira, A., Santa-Marta, M., Malhó, R., Yang, X., Gabuzda, D., and Barbas, C. (2002). Functional neutralization of HIV-1 Vif protein by intracellular immunization inhibits reverse transcription and viral replication. J Biol Chem *277*, 32036-32045.

Greenberg, M., Cammack, N., Salgo, M., and Smiley, L. (2004). HIV fusion and its inhibition in antiretroviral therapy. Rev Med Virol *14*, 321-337.

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

Grinsztejn, B., Nguyen, B.Y., Katlama, C., Gatell, J.M., Lazzarin, A., Vittecoq, D., Gonzalez, C.J., Chen, J., Harvey, C.M., and Isaacs, R.D. (2007). Safety and efficacy of the HIV-1 integrase inhibitor raltegravir (MK-0518) in treatment-experienced patients with multidrug-resistant virus: a phase II randomised controlled trial. Lancet *369*, 1261-1269.

Guiot, E., Carayon, K., Delelis, O., Simon, F., Tauc, P., Zubin, E., Gottikh, M., Mouscadet, J.F., Brochon, J.C., and Deprez, E. (2006). Relationship between the oligomeric status of HIV-1 integrase on DNA and enzymatic activity. J Biol Chem *281*, 22707-22719.

Hamers-Casterman, C., Atarhouch, T., Muyldermans, S., Robinson, G., Hamers, C., Songa, E.B., Bendahman, N., and Hamers, R. (1993). Naturally occurring antibodies devoid of light chains. Nature *363*, 446-448.

Hare, S., Gupta, S.S., Valkov, E., Engelman, A., and Cherepanov, P. (2010a). Retroviral intasome assembly and inhibition of DNA strand transfer. Nature *464*, 232-236.

Hare, S., Gupta, S.S., Valkov, E., Engelman, A., and Cherepanov, P. (2010b). Retroviral intasome assembly and inhibition of DNA strand transfer. Nature *464*, 232-236.

Hare, S., Vos, A.M., Clayton, R.F., Thuring, J.W., Cummings, M.D., and Cherepanov, P. (2010c). Molecular mechanisms of retroviral integrase inhibition and the evolution of viral resistance. Proc Natl Acad Sci U S A *107*.

Hare, S., Vos, A.M., Clayton, R.F., Thuring, J.W., Cummings, M.D., and Cherepanov, P. (2010d). Molecular mechanisms of retroviral integrase inhibition and the evolution of viral resistance. Proc Natl Acad Sci U S A *107*, 20057-20062.

Hatano, H., Lampiris, H., Fransen, S., Gupta, S., Huang, W., Hoh, R., Martin, J.N., Lalezari, J., Bangsberg, D., and Petropoulos, C. (2010). Evolution of integrase resistance during failure of integrase inhibitor-based antiretroviral therapy. JAcquir Immune Defic Syndr.

Hayouka, Z., Rosenbluh, J., Levin, A., Loya, S., Lebendiker, M., Veprintsev, D., Kotler, M., Hizi, A., Loyter, A., and Friedler, A. (2007). Inhibiting HIV-1 integrase by shifting its oligomerization equilibrium. Proc Natl Acad Sci U S A *104*, 8316-8321.

Hazuda, D.J., Felock, P., Witmer, M., Wolfe, A., Stillmock, K., Grobler, J.A., Espeseth, A., Gabryelski, L., Schleif, W., Blau, C., *et al.* (2000). Inhibitors of strand transfer that prevent integration and inhibit HIV-1 replication in cells. Science *287*, 646-650.

Heinzinger, N.K., Bukinsky, M.I., Haggerty, S.A., Ragland, A.M., Kewalramani, V., Lee, M.A., Gendelman, H.E., Ratner, L., Stevenson, M., and Emerman, M. (1994). The Vpr protein of human immunodeficiency virus type 1 influences nuclear localization of viral nucleic acids in nondividing host cells. Proc Natl Acad Sci U S A *91*, 7311-7315.

Hemelaar, J., Gouws, E., Ghys, P.D., Osmanov, S., and Characterisation., W.-U.N.f.H.I.a. (2011). Global trends in molecular epidemiology of HIV-1 during 2000-2007. AIDS *25*, 679-689.

Hindmarsh, P., Ridky, T., Reeves, R., Andrake, M., Skalka, A.M., and Leis, J. (1999). HMG protein family members stimulate human immunodeficiency virus type 1 and avian sarcoma virus concerted DNA integration in vitro. J Virol *73*, 2994-3003.

Hobaika, Z., Zargarian, L., Boulard, Y., Maroun, R.G., Mauffret, O., and Fermandjian, S. (2009). Specificity of LTR DNA recognition by a peptide mimicking the HIV-1 integrase {alpha}4 helix. Nucleic Acids Res *37*, 7691-7700.

Holliger, P., and Hudson, P.J. (2005a). Engineered antibody fragments and the rise of single domains. Nat Biotechnol *23*, 1126-1136.

Holliger, P., and Hudson, P.J. (2005b). Engineered antibody fragments and the rise of single domains. Nat Biotechnol *23*, 1126-1136.

Huang, H., Chopra, R., Verdine, G.L., and Harrison, S.C. (1998). Structure of a covalently trapped catalytic complex of HIV-1 reverse transcriptase: implications for drug resistance. Science *282*, 1669-1675.

Hwang, S.S., Boyle, T.J., Lyerly, H.K., and Cullen, B.R. (1991). Identification of the envelope V3 loop as the primary determinant of cell tropism in HIV-1. Science *253*, 71-74.

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

Johnson, A.A., Santos, W., Pais, G.C., Marchand, C., Amin, R., Burke, T.R.J., Verdine, G., and Pommier, Y. (2006). Integration Requires a Specific Interaction of the Donor DNA Terminal 5'-Cytosine with Glutamine 148 of the HIV-1 Integrase Flexible Loop. J Biol Chem *281*, 461-467.

Johnson, J.A., Li, J.F., Wei, X., Lipscomb, J., Irlbeck, D., Craig, C., Smith, A., Bennett, D.E., Monsour, M., Sandstrom, P., *et al.* (2008a). Minority HIV-1 drug resistance mutations are present in antiretroviral treatment-naive populations and associate with reduced treatment efficacy. PLoS Med *5*,

e158.
Johnson, V.A., Brun-Vezinet, F., Clotet, B., Gunthard, H.F., Kuritzkes, D.R., Pillay, D., Schapiro, J.M., and Richman, D.D. (2008b). Update of the drug resistance mutations in HIV-1. Top HIV Med *16*, 138-145.
Jones, K.A., and Peterlin, B.M. (1994). Control of RNA initiation and elongation at the HIV-1 promoter. Annu Rev Biochem *63*, 717-743.
Joyce, C.M., and Benkovic, S.J. (2004). DNA polymerase fidelity: kinetics, structure, and checkpoints. Biochemistry *43*, 14317-14324.
Kabat, E., Wu, T.T., Perry, H.M., Gottesman, K.S., and Foeller, C. (1991). Sequence of proteins of immunological interest. US Public Health Services, NIH Bethesda, MD, Publication *No. 91-3242*.
Kalpana, G.V., Marmon, S., Wang, W., Crabtree, G.R., and Goff, S.P. (1994). Binding and stimulation of HIV-1 integrase by a human homolog of yeast transcription factor SNF5. Science *266*, 2002-2006.
Kao, S.Y., Calman, A.F., Luciw, P.A., and Peterlin, B.M. (1987). Anti-termination of transcription within the long terminal repeat of HIV-1 by tat gene product. Nature *330*, 489-493.
Kepler, T.B., and Perelson, A.S. (1998). Drug concentration heterogeneity facilitates the evolution of drug resistance. Proc Natl Acad Sci USA *95*, 11514-11519.
Kim, E.Y., Winters, M.A., Kagan, R.M., and Merigan, T.C. (2001). Functional correlates of insertion mutations in the protease gene of humanimmunodeficiency virus type 1 isolates from patients. J Virol *75*, 11227-11233.
Kim, S.Y., Byrn, R., Groopman, J., and Baltimore, D. (1989). Temporal aspects of DNA and RNA synthesis during human immunodeficiency virus infection: evidence for differential gene expression. . J Virol *63*, 3708-3713.
Klibanov, O. (2009). Elvitegravir, an oral HIV integrase inhibitor, for the potential treatment of HIV infection. Curr Opin Investig Drugs *10*, 190-200.
Kobayashi, M., Yoshinaga, T., Seki, T., Wakasa-Morimoto, C., Brown, K.W., Ferris, R., Foster, S.A., Hazen, R.J., Miki, S., Suyama-Kagitani, A., *et al.* (2011). In Vitro antiretroviral properties of S/GSK1349572, a next-generation HIV integrase inhibitor. Antimicrob Agents Chemother *55*, 813-821.
Koch-Nolte, F., Reyelt, J., Schössow, B., Schwarz, N., Scheuplein, F., Rothenburg, S., Haag, F., Alzogaray, V., , , Cauerh, V.A., and Goldbaum, F.A. (2007). Single domain antibodies from llama eVectively and speciWcally block T cell ecto-ADP-ribosyltransferase ART2.2 in vivo. FASEB J *21*, 3490-3498.
Kogan, M., and Rappaport, J. (2011). HIV-1 Accessory Protein Vpr: Relevance in the pathogenesis of HIV and potential for therapeutic intervention. Retrovirology *8*.
Kohlstaedt, L.A., Wang, J., Friedman, J.M., Rice, P.A., and Steitz, T.A. (1992). Crystal structure at 3.5 A resolution of HIV-1 reverse transcriptase complexed with an inhibitor. Science *256*, 1783-1790.
König, R., Zhou, Y., Elleder, D., Diamond, T.L., Bonamy, G.M., Irelan, J.T., Chiang, C.Y., Tu, B.P., De Jesus, P.D., Lilley, C.E., *et al.* (2008). Global analysis of host-pathogen interactions that regulate early-stage HIV-1 replication. Cell *135*, 49-60.
Kozísek, M., Sasková, K.G., Rezácová, P., Brynda, J., van Maarseveen, N.M., De Jong, D., Boucher, C.A., Kagan, R.M., Nijhuis, M., and Konvalinka, J. (2008). Ninety-nine is not enough: molecular characterization of inhibitor-resistant human immunodeficiency virus type 1 protease mutants with insertions in the flap region. J Virol *82*, 5869-5878.
Kulkosky, J., Jones, K.S., Katz, R.A., Mack, J.P., and Skalka, A.M. (1992). Residues critical for

retroviral integrative recombination in a region that is highly conserved among retroviral/retrotransposon integrases and bacterial insertion sequence ransposases. Mol Cell Biol *12*, 2331-2338.

Kuus-Reichel, K., Grauer, L.S., Karavodin, L.M., Knott, C., Krusemeier, M., and Kay, N.E. (1994). Will immunogenicity limit the use, efficacy and future development of therapeutic monoclonal antibodies? . Clin Diagn Lab Immunol *1*, 365-372.

Lambotte, O., Chaix, M.L., Gasnault, J., Goujard, C., Lebras, P., Delfraissy, J.F., and Taoufik, Y. (2005). Persistence of replication-competent HIV in the central nervous system despite long-term effective highly active antiretroviral therapy. AIDS *19*, 217-218.

Lecossier, D., Bouchonnet, F., Clavel, F., and Hance, A.J. (2003). Hypermutation of HIV-1 DNA in the Absence of the Vif Protein. Science *300*, 1112.

Lee, S.P., Xiao, J., Knutson, J.R., Lewis, M.S., and Han, M.K. (1997). Zn2+ promotes the self-association of human immunodeficiency virus type-1 integrase in vitro. Biochemistry *36*, 173-180.

Leh, H., Brodin, P., Bischerour, J., Deprez, E., Tauc, P., Brochon, J.-C., LeCam, E., Coulaud, D., Auclair, C., and Mouscadet, J.-F. (2000). Determinants of Mg2+-dependent activities of recombinant human immunodeficiency virus type 1 integrase. Biochemistry *39*, 9285-9294.

Lehmann-Che, J., Giron, M.-L., Delelis, O., Löchelt, M., Bittoun, P., Tobaly-Tapiero, J., Thé, H.d., and Saïb, A. (2005). Protease-dependent uncoating of a complex retrovirus. J Virol *79*, 9244-9253.

Lenz, J.C., and Rockstroh, J.K. (2010). Vicriviroc, a new CC-chemokine receptor 5 inhibitor for treatment of HIV: properties, promises and challenges. Expert Opin Drug Metab Toxicol *6*, 1139-1150.

Levin, A., Armon-Omer, A., Rosenbluh, J., Melamed-Book, N., Graessmann, A., Waigmann, E., and Loyter, A. (2009). Inhibition of HIV-1 integrase nuclear import and replication by a peptide bearing integrase putative nuclear localization signal. Retrovirology *6*, 112.

Levin, R., Mhashilkar, A.M., Dorfman, T., Bukovsky, A., Zani, C., Bagley, J., Hinkula, J., Niedrig, M., Albert, J., Wahren, B., *et al.* (1997). Inhibition of early and late events of the HIV-1 replication cycle by cytoplasmic Fab intrabodies against the matrix protein, p17. Mol Med *3*, 96-110.

Levy-Mintz, P., Duan, L., Zhang, H., Hu, B., Dornadula, G., Zhu, M., Kulkosky, J., Bizub-Bender, D., Skalka, A.M., and Pomerantz, R.J. (1996). Intracellular expression of single-chain variable fragments to inhibit early stages of the viral life cycle by targeting human immunodeficiency virus type 1 integrase. J Virol *70*, 8821-8832.

Levy, J.A., Hoffman, A.D., Kramer, S.M., Landis, J.A., Shimabukuro, J.M., and Oshiro, L.S. (1984). Isolation of lymphocytopathic retroviruses from San Francisco patients with AIDS. Science *225*, 840-842.

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

Lewinski, M.K., Yamashita, M., Emerman, M., Ciuffi, A., Marshall, H., Crawford, G., F.Collins, Shinn, P., Leipzig, J., Hannenhalli, S., *et al.* (2006). Retroviral DNA integration: viral and cellular determinants of target-site selection. PLoS Pathog *2* e60.

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

Li, M., and Craigie, R. (2005). Processing of viral DNA ends channels the HIV-1 integration reaction to concerted integration. J Biol Chem *280*, 29334-29339.

Li, M., Mizuuchi, M., Burke T.R., J., and Craigie, R. (2006). Retroviral DNA integration:reaction

pathway and critical intermediates. EMBO J *25*, 1295-1304.
Limón, A., Devroe, E., Lu, R., Ghory, H.Z., Silver, P.A., and Engelman, A. (2002). Nuclear localization of human immunodeficiency virus type 1 preintegration complexes (PICs): V165A and R166A are pleiotropic integrase mutants primarily defective for integration, not PIC nuclear import.
. J Virol *76*, 10598-10607.
Lin, C.W., and Engelman, A. (2003). The barrier-to-autointegration factor is a component of functional human immunodeficiency virus type 1 preintegration complexes. J Virol *77*, 5030-5036.
Little, S.J., Holte, S., Routy, J.P., Daar, E.S., Markowitz, M., Collier, A.C., Koup, R.A., Mellors, J.W., Connick, E., Conway, B., *et al.* (2002). Antiretroviral-drug resistance among patients recently infected with HIV. N Engl J Med *347*, 385-394.
Low, A., Prada, N., Topper, M., Vaida, F., Castor, D., Mohri, H., Hazuda, D., Muesing, M., and Markowitz, M. (2009). Natural Polymorphisms of Human Immunodeficiency Virus Type 1 Integrase and Inherent Susceptibilities to a Panel of Integrase Inhibitors. Antimicrobial agents and chemotherapy *53*, 4275-4282.
Machado, L.F., Ishak, M.O., Vallinoto, A.C., Lemos, J.A., Azevedo, V.N., Moreira, M.R., Souza, M.I., Fernandes, L.M., Souza, L.L., and Ishak, R. (2009). Molecular epidemiology of HIV type 1 in northern Brazil: identification of subtypes C and D and the introduction of CRF02_AG in the Amazon region of Brazil. AIDS Res Hum Retroviruses *25*, 961-966.
Maele, B.V., Busschots, K., Vandekerckhove, L., Christ, F., and Debyser, Z. (2006). Cellular co-factors of HIV-1 integration. Trends in Biochemical Sciences *31*, 98-105.
Maertens, G., Cherepanov, P., Debyser, Z., Engelborghs, Y., and Engelman, A. (2004a). Identification and characterization of a functional nuclear localization signal in the HIV-1 integrase (IN) interactor LEDGF/p75. J Biol Chem *279*, 33421-33429.
Maertens, G., Cherepanov, P., Debyser, Z., Engelborghs, Y., and Engelman, A. (2004b). Identification and characterization of a functional nuclear localization signal in the HIV-1 integrase interactor LEDGF/p75. J Biol Chem *279*, 33421-33429.
Maertens, G., Cherepanov, P., Pluymers, W., Busschots, K., De Clercq, E., Debyser, Z., and Engelborghs, Y. (2003). LEDGF/p75 is essential for nuclear and chromosomal targeting of HIV-1 integrase in human cells. J Biol Chem *278*, 33528-33539.
Maertensm, G., Cherepanov, P., Pluymers, W., Busschots, K., De Clercq, E., Debyser, Z., and Engelborghs, Y. (2003). LEDGF/p75 is essential for nuclear and chromosomal targeting of HIV-1 integrase in human cells. J Biol Chem *278*, 33528-33539.
Maes, M., Levin, A., Hayouka, Z., Shalev, D.E., Loyter, A., and Friedler, A. (2009). Peptide inhibitors of HIV-1 integrase: from mechanistic studies to improved lead compounds. Bioorg Med Chem *17*, 7635-7642.
Maiga, A.I., Malet, I., Soulie, C., Derache, A., Koita, V., Amellal, B., Tchertanov, L., Delelis, O., Morand-Joubert, L., Mouscadet, J.F., *et al.* (2009). Genetic barriers for integrase inhibitor drug resistance in HIV type-1 B and CRF02_AG subtypes. Antivir Ther *14*, 123-129.
Maïga, A.I., Malet, I., Soulie, C., Derache, A., Koita, V., Amellal, B., Tchertanov, L., Delelis, O., Morand-Joubert, L., Mouscadet, J.F., *et al.* (2009). Genetic barriers for integrase inhibitor drug resistance in HIV type-1 B and CRF02_AG subtypes. Antivir Ther *14*, 123-129.
Malet, I., Delelis, O., Valantin, M.-A., Montes, B., Soulie, C., Wirden, M., Tchertanov, L., Peytavin, G., Reynes, J., Mouscadet, J.-F., *et al.* (2008a). Mutations Associated with Failure of Raltegravir Treatment Affect Integrase Sensitivity to the Inhibitor In Vitro. Antimicrobial agents and

chemotherapy *52*, 1351-1358.

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

Malet, I., Soulie, C., Tchertanov, L., Derache, A., Amellal, B., Traore, O., Simon, A., Katlama, C., Mouscadet, J.-F., Calvez, V., *et al.* (2008c). Structural Effects of Amino Acid Variations Between B and CRF02-AG HIV-1 Integrases. Journal of Medical Virology *80*, 754-761.

Malet, I., Soulie, C., Tchertanov, L., Derache, A., Amellal, B., Traore, O., Simon, A., Katlama, C., Mouscadet, J.F., Calvez, V., *et al.* (2008d). Structural effects of amino acid variations between B and CRF02-AG HIV-1 integrases. J Med Virol *80*, 754-761.

Mallender, W.D., Carrero, J., and Voss, E.W. (1996). Comparative properties of the single chain antibody and Fv derivatives of mAb 4-4-20: relationship between interdomain interactions and high affinity for fluourescein ligand. J BiolChem *271*, 5338-5346.

Mammano, F., Trouplin, V., Zennou, V., and Clavel, F. (2000). Retracing the evolutionary pathways of human immunodeficiency virus type 1 resistance to protease inhibitors: virus fitness in the absence and in the presence of drug. J Virol *74*, 8524-8531.

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

Marasco, W.A. (1997). Intrabodies: turning the humoral immune system outside in for intracellular immunization. Gene Ther *4*, 11-15.

Marchand, C., Johnson, A.A., Semenova, E., and Pommier, Y. (2006). Mechanisms and inhibition of HIV integration. Drug Discov Today Dis Mech *3*, 253-260.

Marchand, C., Maddali, K., Metifiot, M., and Pommie, Y. (2009). HIV-1 IN Inhibitors: 2010 Update and Perspectives. Curr Top Med Chem *9*, 1016-1037.

Marchand, C., Zhang, X., Pais, G.C., Cowansage, K., Neamati, N., Burke Jr., T.R., and Pommier, Y. (2002). Structural determinants for HIV-1 integrase inhibition by beta-diketo acids. Journal of Biological Chemistry *277*, 12596-12603.

Marinello, J., Marchand, C., Mott, B., Bain, A., Thomas, C., and Pommier, Y. (2008 -a). Comparison of raltegravir and elvitegravir on HIV-1 integrase catalytic reactions and on a series of drug-resistant integrase mutants. Biochemistry *47*, 9345-9354.

Marinello, J., Marchand, C., Mott, B., Bain, A., Thomas, C.J., and Pommier, Y. (2008 -b). Comparison of Raltegravir and Elvitagravir on HIV-1 Integrase catalytic reactions and on a series of drug-resistant Integrase mutants. Biochemistry *47*, 9345-9354.

Markowitz, M., B. Y. Nguyen, E. Gotuzzo, F. Mendo, W. Ratanasuwan, C.Kovacs, G. Prada, J. O. Morales-Ramirez, C. S. Crumpacker, R. D. Isaacs, *et al.* (2007). Rapid and durable antiretroviral effect of the HIV-1 integrase inhibitor raltegravir as part of combination therapy in treatment-naive patients with HIV-1 infection: results of a 48-week controlled study. J Acquir Immune Defic Syndr *46*, 125-133.

Maroun, M., Delelis, O., G, C.C., Bader, T., Segeral, E., Mbemba, G., Petit, C., P.Sonigo, Rain, J.C., Mouscadet, J.F., *et al.* (2006). Inhibition of early steps of HIV-1 replication by SNF5/Ini1. J Biol Chem *281*, 22736-22743.

Marshall, H.M., Ronen, K., Berry, C., Llano, M., Sutherland, H., Saenz, D., Bickmore, W., Poeschla, E., and Bushman, F.D. (2007). Role of PSIP1/LEDGF/p75 in lentiviral infectivity and integration

targeting. PLoS ONE *2*, e1340.

Martínez-Cajas JL, Pant-Pai N, Klein MB, and MA, W. (2008). Role of genetic diversity amongst HIV-1 non-B subtypes in drug resistance: a systematic review of virologic and biochemical evidence. AIDS Rev *10*, 212-223.

Martinez-Cajas, J.L., Pai, N.P., Klein, M.B., and Wainberg, M.A. (2009). Differences in resistance mutations among HIV-1 non-subtype B infections: a systematic review of evidence (1996–2008). Journal of the International AIDS Society *12*.

Martínez-Cajas, J.L., Pant-Pai, N., Klein, M.B., and Wainberg, M.A. (2008). Role of genetic diversity amongst HIV-1 non-B subtypes in drug resistance: a systematic review of virologic and biochemical evidence. AIDS Rev *10*, 212-223.

Masson, C., Bury-Mone, S., Guiot, E., Saez-Cirion, A., Schoevaert-Brossault, D., C.Brachet-Ducos, Delelis, O., Subra, F., Jeanson-Leh, L., and Mouscadet, J.F. (2007). Ku80 participates in the targeting of retroviral transgenes to the chromatin of CHO cells. J Virol *81*, 7924-7932.

Matheron, S., Mendoza-Sassi, G., and Simon, F., et al (1997). HIV-1 and HIV-2 AIDS in African patients living in Paris. AIDS *11*, 934-936.

McColl, D.J., and Chen, X. (2010). Strand transfer inhibitors of HIV-1 integrase: Bringing IN a new era of antiretroviral therapy. Antiviral Research *85* 101-118.

McGillick, B.E., Balius, T.E., Mukherjee, S., and Rizzo, R.C. (2010). Origins of resistance to the HIVgp41 viral entry inhibitor T20. Biochemistry *49*, 3575-3592.

Meehan, A.M., and Poeschla, E.M. (2010). Chromatin tethering and retroviral integration: recent discoveries and parallels with DNA viruses. Biochim Biophys Acta *1799*, 182-191.

Melek M, Jones JM, O'Dea MH, Pais G, Burke TR Jr, Pommier Y, Neamati N, and M, G. (2001). Effect of HIV integrase inhibitors on the RAG1/2 recombinase. Proc Natl Acad Sci U S A *99*, 134-137.

Mellors, J.W., Rinaldo, C.R.J., Gupta, P., White, R.M., Todd, J.A., and Kingsley, L.A. (1996). Prognosis in HIV-1 infection predicted by the quantity of virus in plasma. Science *272*, 1167-1170.

Metifiot, M., Maddali, K., Naumova, A., Zhang, X., Marchand, C., and Pommier, Y. (2010). Biochemical and Pharmacological Analyses of HIV-1 Integrase Flexible Loop Mutants Resistant To Raltegravir. Biochemistry.

Métifiot, M., Vandegraaff, N., Maddali, K., Naumova, A., Zhang, X., Rhodes, D., Marchand, C., and Pommier, Y. (2011). Elvitegravir overcomes resistance to raltegravir induced by integrase mutation Y143. AIDS, 1175-1178.

Mhashilkar, A.M., Bagley, J., Chen, S.Y., Szilvay, A.M., Helland, D.G., and Marasco, W.A. (1995). Inhibition of HIV-1 Tat-mediated LTR transactivation and HIV-1 infection by anti-Tat single chain intrabodies. EMBO J *14*, 1542-1551.

Mhashilkar, A.M., LaVecchio, J., Eberhardt, B., Porter-Brooks, J., Boisot, S., Dove, J.H., Pumphrey, C., Li , X., Weissmahr, R.N., Ring, D.B., *et al.* (1999). Inhibition of human immunodeficiency virus type 1 replication in vitro in acutely and persistently infected human CD4+ mononuclear cells expressing murine and humanized anti-human immunodeficiency virus type 1 Tat single-chain variable fragment intrabodies. Hum Gene Ther *10*, 1453-1467.

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

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

Mitsuya, H., Weinhold, K.J., Furman, P.A., St Clair, M.H., Lehrman, S.N., Gallo, R.C., Bolognesi, D.,

Barry, D.W., and Broder, S. (1985). 3'-Azido-3'-deoxythymidine (BWA509U): an antiviral agent that inhibits the infectivity and cytopathic effect of human T-lymphotropic virus type III/lymphadenopathy-associated virus in vitro. . USA Proc Natl Acad Sci *82*, 7096-7100.

Molla, A., Korneyeva, M., Gao, Q., Vasavanonda, S., Schipper, P.J., Mo, H.M., Markowitz, M., Chernyavskiy, T., Niu, P., Lyons, N., *et al.* (1996). Ordered accumulation of mutations in HIV protease confers resistance to ritonavir. Nat Med, 760-766.

Morellet, N., Jullian, N., Rocquigny, H.D., Maigret, B., J L Darlix, and Roques, B.P. (1992). Determination of the structure of the nucleocapsid protein NCp7 from the human immunodeficiency virus type 1 by 1H NMR. EMBO J *11*, 3059-3065.

Mouscadet, J.-F., Delelis, O., Marcelin, A.-G., and Tchertanova, L. (2010a). Resistance to HIV-1 integrase inhibitors: A structural perspective. Drug Resistance Updates *13* 139-150.

Mouscadet, J.F., Delelis, O., Marcelin, A.G., and Tchertanov, L. (2010b). Resistance to HIV-1 integrase inhibitors: A structural perspective. Drug Resist Updat.

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

Mulder, L.C., Harari, A., and Simon, V. (2008). Cytidine deamination induced HIV-1 drug resistance. Proc Natl Acad Sci USA *105*, 5501-5506.

Murphy, F.A., Fauquet, C.M., Biship, D.H.L., Ghabrial, S.A., Javis, A.W., Martelli, G.P., Mayo, M.A., and Summers, M.D. (1994). Virus taxonomy: The classification and nomenclature of viruses, Retroviridae, Springer-Verlag, Vienna.

Murphy, J.P., Gershwin, L.J., Tatcher, E.F., Fowler, M.E., and Habig, W.H. (1989). Immune response of the llama (lama glama) to tetanus toxoid vaccination. . Am J Vet Res *50*, 1279-1281.

Muyldermans, S. (2001). Single domain camel antibodies: current status. Reviews in Molecular Biotechnology *74*, 277-302.

Nair, V., and Chi, G. (2007). HIV integrase inhibitors as therapeutic agents in AIDS. Reviews in Medical Virology *17*, 277 - 295.

Neil, S.J., Zang, T., and Bieniasz, P.D. (2008). Tetherin inhibits retrovirus release and is antagonized by HIV-1 Vpu. Nature *451*, 425-430.

Nguyen, D.H., and Hildreth, J.E. (2000). Evidence for budding of human immunodeficiency virus type 1 selectively from glycolipid-enriched membrane lipid rafts. . J Virol *74*, 3264-3272.

Nguyen, V.K., Hamers, R., Wyns, L., and Muyldermans, S. (1999). Loss of splice consensus signal is responsible for the removal of the entire CH1 domain of the functional camel IgG2A heavy chain antibodies. Mol Immunol *36*, 515-524.

Nguyen, V.K., Hamers, R., Wyns, L., and Muyldermans, S. (2000a). Camel heavy-chain antibodies: diverse germline V(H)H and specific mechanisms enlarge the antigen-binding repertoire. EMBO J *19*, 921-930.

Nguyen, V.K., Hamers, R., Wyns, L., and Muyldermans, S. (2000b). Camel heavy-chain antibodies: diverse germline VHH and specific mechanisms enlarge the antigen-binding repertoire. EMBO J *19*, 921-931.

Nguyen, V.K., Zou, X., Lauwereys, M., Brys, L., Bruggemann, M., and Muyldermans, S. (2003). Heavy-chain only antibodies derived from dromedary are secreted and displayed by mouse B cells. Immunology *109*, 93-101.

Nijhuis, M., Schuurman, R., de Jong, D., Erickson, J., Gustchina, E., Albert, J., Schipper, P., Gulnik, S.,

and Boucher, C.A.B. (1999). Increased fitness of drug resistant HIV-1protease as a result of acquisition of compensatory mutations during suboptimal therapy. AIDS *13*, 2349-2359.

Njai, H.F., Gali, Y., Vanham, G., Clybergh, C., Jennes, W., Vidal, N., Butel, C., Mpoudi-Ngolle, E., Peeters, M., and Ariën, K.K. (2006). The predominance of Human Immunodeficiency Virus type 1 (HIV-1) circulating recombinant form 02 (CRF02_AG) in West Central Africa may be related to its replicative fitness. Retrovirology *3*.

Nogues, C., Leh, H., Langendorf, C.G., Law, R.H., Buckle, A.M., and Buckle, M. (2010). Characterisation of peptide microarrays for studying antibody-antigen binding using surface plasmon resonance imagery. PLoS One *5*.

Ntemgwa, M., Brenner, B.G., Oliveira, M., Moisi, D., and Wainberg, M.A. (2007). Natural polymorphisms in the human immunodeficiency virus type 2 protease can accelerate time to development of resistance to protease inhibitors. Antimicrob Agents Chemother *51*, 604-610.

Pan, C., Liu, S., and Jiang, S. (2010). HIV-1 gp41 Fusion Intermediate: A Target for HIV Therapeutics. J Formos Med Assoc *109*, 94-105.

Parissi, V., Calmels, C., De Soultrait, V.R., Caumont, A., Fournier, M., Chaignepain, S., and Litvak, S. (2001). Functional interactions of human immunodeficiency virus type 1 integrase with human and yeast HSP60. J Virol *75*, 11344-11353.

Perelson, A.S., Essunger, P., Cao, Y., Vesanen, M., Hurley, A., Saksela, K., Markowitz, M., and Ho, D.D. (1997). Decay characteristics of HIV-1-infected compartments during combination therapy. Nature *387*, 188-191.

Perelson, A.S., Neumann, A.U., Markowitz, M., Leonard, J.M., and Ho, D.D. (1996). HIV-1 dynamics in vivo: virion clearance rate, infected cell life-span, and viral generation time. Science *271*, 1582-1586.

Perez, J.M., Renisio, J.G., Prompers, J.J., van Platerink, C.J., Cambillau, C., Darbon, H., and Frenken, L.G. (2001). Thermal unfolding of a llama antibody fragment: a two-state reversible process. . Biochemistry *40*, 74-83.

Perryman, A.L., Forli, S., Morris, G.M., Burt, C., Cheng, Y., Palmer, M.J., Whitby, K., McCammon, J.A., Phillips, C., and Olson, A.J. (2010). A dynamic model of HIV integrase inhibition and drug resistance. J Mol Biol *397*, 600-615.

Peterlin, B.M., and Trono, D. (2003). Hide, shield and strike back: how HIV-infected cells avoid immune eradication. Nat Rev Immunol *3*, 97-107.

Phelan, A., Elliott, G., and O'Hare, P. (1998). Intercellular delivery of functional p53 by the herpesvirus protein VP22. Nat Biotechnol *16*, 440-443.

Pieniazek, D., Rayfield, M., Hu, D.J., Nkengasong, J.N., Soriano, V., Heneine, W., Zeh, C., Agwale, S.M., Wambebe, C., Odama, L., *et al.* (2004). HIV-2 protease sequences of subtypes A and B harbor multiple mutations associated with protease inhibitor resistance in HIV-1. AIDS *18*, 495-502.

Pleschberger, M., Saerens, D., Weigert, S., Sleytr, U.B., Muyldermans, S., Sara, M., and Egelseer, E.M. (2004). An S-layer heavy chain camel antibody fusion protein for generation of a nanopatterned sensing layer to detect the prostate-specific antigen by surface plasmon resonance echnology. Bioconjug Chem *15*, 664-671.

Poeschla, E.M. (2008). Integrase, LEDGF/p75 and HIV replication. Cellular and Molecular Life Sciences *65*, 1403-1424.

Pomerantz, R.J., Seshamma, T. & Trono, D. (1992). Efficient replication of human immunodeficiency virus type 1 requires a threshold level of Rev: potential implications for latency. J Virol *66*, 1809-1813

Pommier, Y., Johnson, A.A., and Marchand, C. (2005). Integrase inhibitors to treat HIV/AIDS. Nat Rev Drug Discov *4*, 236-248.

Poveda, E., Rodes, B., Toro, C., and Soriano, V. (2004). Are fusion inhibitors active against all HIV variants? AIDS Res Hum Retroviruses *20*, 347-348.

Price, D.H. (2000). P-TEFb, a cyclin-dependent kinase controlling elongation by RNA polymerase II. . Mol Cell Biol *20*, 2629-2634.

Rambaut, A., Posada, D., Crandall, K.A., and C.Holmes, E. (2004). The causes and consequences of HIV evolution. Nature reviews *5*, 52-61.

Redel, L., Le Douce, V., Cherrier, T., Marban, C., Janossy, A., Aunis, D., Van Lint, C., Rohr, O., and Schwartz, C. (2009). HIV-1 regulation of latency in the monocyte-macrophage lineage and in CD4+ T lymphocytes. J Leukoc Biol *87*, 575-588.

Reeves, J.D., and Doms, R.W. (2002). Human immunodeficiency virus type 2. J Gen Virol *83*, 1253-1265.

Regoes, R.R., and Bonhoeffer, S. (2005). The HIV coreceptor switch: a population dynamical perspective. Trends Microbiology *13*, 269-277.

Reigadas, S., Anies, G., Masquelier, B., Calmels, C., Stuyver, L.J., Parissi, V., Fleury, H., and Andreola, M.L. (2010). The HIV-1 integrase mutations Y143C/R are an alternative pathway for resistance to Raltegravir and impact the enzyme functions. PLoS One *5*.

Revets, H., De Baetselier, P., and Muyldermans, S. (2005). Nanobodies as novel agents for cancer therapy. . Expert Opin Biol Ther *5*, 111-124.

Rhee, S.Y., Gonzales, M., J,, , Kantor, R., Betts, B.J., Ravela, J., and Shafer, R.W. (2003). Human immunodeficiency virus reverse transcriptase and protease sequence database. Nucleic Acids Res *31*, 298-303.

Rhee, S.Y., Liu, T.F., Kiuchi, M., Zioni, R., Gifford, R.J., Holmes, S.P., and Shafer, R.W. (2008). Natural variation of HIV-1 group M integrase: implications for a new class of antiretroviral inhibitors. Retrovirology *5*.

Rice, P.A., and Baker, T.A. (2001). Comparative architecture of transposase and integrase complexes. Nat Struct Biol *8*, 302-307.

Richman, D.D., Margolis, D.M., Delaney, M., Greene, W.C., Hazuda, D., and Pomerantz, R.J. (2009). The challenge of finding a cure for HIV infection. Science *323*, 1304-1307.

Rivière, L., Darlix, J.-L., and Cimarelli, A. (2010). Analysis of the Viral Elements Required in the Nuclear Import of HIV-1 DNA. J Virol *84*, 729-739.

Rodes, B., Sheldon, J., Toro, C., Jimenez, V., Alvarez, M.A., and Soriano, V. (2006). Susceptibility to protease inhibitors in HIV-2 primary isolates from patients failing antiretroviral therapy. J Antimicrob Chemother *57*, 709-713.

Roovers, R.C., Laeremans, T., Huang, L., De Taeye, S., Verkleij, A.J., Revets, H., De Haard, H.J., and Van Bergen en Henegouwen, P.M. (2007). Efficient inhibition of EGFR signaling and of tumour growth by antagonistic anti-EFGR Nanobodies. Cancer Immunol Immunother *56*, 303-317.

Roquebert, B., Blum, L., Collin, G., Damond, F., Peytavin, G., Leleu, J., Matheron, S., Chene, G., Brun-Vezinet, F., and Descamps, D. (2008a). Selection of the Q148R integrase inhibitor resistance mutation in a failing raltegravir containing regimen. Aids *22*, 2045-2046.

Roquebert, B., Damond, F., Collin, G., Matheron, S., Peytavin, G., Bénard, A., Campa, P., Chêne, G., Brun-Vézinet, F., and Descamps, D. (2008b). HIV-2 integrase gene polymorphism and phenotypic susceptibility of HIV-2 clinical isolates to the integrase inhibitors raltegravir and elvitegravir in vitro. J

Antimicrob Chemother *62*, 914-920.

Roquebert, B., Matheron, S., Benard, A., Leleu, J., Tubiana, R., Karmochkine, M., Chene, G., Damond, F., Brun-Vezinet, F., and Descamps, D. (2010). Raltegravir genetic resistance patterns in HIV-2 infected patients failing raltegravir-containing regimen. In CROI (San Fransisco http://retroconference.org/AbstractSearch/Default.aspx?Conf=19).

Rosenbluh, J., Hayouka, Z., Loya, S., Levin, A., Armon-Omer, A., Britan, E., Hizi A, Kotler M, Friedler A, and A, L. (2007). Interaction between HIV-1 Rev and integrase proteins: a basis for the development of anti-HIV peptides. J Biol Chem *282*, 15743-15753.

Ross, L., Lim, M.L., Liao, Q., Wine, B., Rodriguez, A.E., Weinberg, W., and Shaefer, M. (2007). Prevalence of antiretroviral drug resistance and resistance-associated mutations in antiretroviral therapy-naive HIV-infected individuals from 40 United States cities. HIV Clin Trials *8*, 1-8.

Salgadoa, M., Torob, C., Simóna, A., Garridoa, C., Blancoa, F., Sorianoa, V., and Rodésa, B. (2009). Mutation N155H in HIV-2 integrase confers high phenotypic resistance to raltegravir and impairs replication capacity. Journal of Clinical Virology *46* 173-175.

Santiago, M.L., Rodenburg, C.M., Kamenya, S., Bibollet-Ruche, F., Gao, F., Bailes, E., Meleth, S., Soong, S.J., Kilby, J.M., Moldoveanu, Z., *et al.* (2002). SIVcpz in wild chimpanzees. Science *295*, 465.

Shaheen, F., Duan, L., Zhu, M., Bagasra, O., and Pomerantz, R.J. (1996). Targeting human immunodeficiency virus type 1 reverse transcriptase by intracellular expression of single-chain variable fragments to inhibit early stages of the viral life cycle. J Virol *70*, 3392-3400.

Shaw-Reid, C.A., Munshi, V., Graham, P., Wolfe, A., Witmer, M., Danzeisen, R., Olsen, D.B., Carroll, S.S., Embrey, M., Wai, J.S., *et al.* (2003). Inhibition of HIV-1 ribonuclease H by a novel diketo acid, 4-[5-(benzoylamino)thien-2-yl]-2,4-dioxobutanoic acid. J Biol Chem *278*, 2777-2780.

Sheehy, A.M., Gaddis, N.C., Choi, J.D., and Malim, M.H. (2002). Isolation of a human gene that inhibits HIV-1 infection and is suppressed by the viral Vif protein. Nature *418*, 646-650.

Shimura, K., and Kodama, E. (2009). Elvitegravir: a new HIV integrase inhibitor. Antivir Chem Chemother *20*, 79-85.

Shimura, K., Kodama, E., Sakagami, Y., Matsuzaki, Y., Watanabe, W., Yamataka, K., Watanabe, Y., Ohata, Y., Doi, S., Sato, M., *et al.* (2008). Broad antiretroviral activity and resistance profile of the novel human immunodeficiency virus integrase inhibitor elvitegravir (JTK-303/GS-9137). Journal of Virology *82*, 764-774.

Shun, M.-C., Daigle, J.E., Vandegraaff, N., and Engelman, A. (2007). Wild-Type Levels of Human Immunodeficiency Virus Type 1 Infectivity in the Absence of Cellular Emerin Protein. Journal of Virology *81*, 166-172.

Sichtig, N., Sierra, S., Kaiser, R., Däumer, M., Reuter, S., Schülter, E., Altmann, A., Fätkenheuer, G., Dittmer, U., Pfister, H., *et al.* (2009a). Evolution of raltegravir resistance during therapy. J Antimicrob Chemother *64*, 25-32.

Sichtig, N., Sierra, S., Kaiser, R., Daumer, M., Reuter, S., Schulter, E., Altmann, A., Fatkenheuer, G., Dittmer, U., Pfister, H., *et al.* (2009b). Evolution of raltegravir resistance during therapy. J Antimicrob Chemother *64*, 25-32.

Skerra, A., and Plückthun, A. (1988). Assembly of functional immunoglobulin Fv fragment in Escherichia coli. Science *240*, 1038-1041.

Sluis-Cremer, N., Temiz, N.A., and Bahar, I. (2004). Conformational changes in HIV-1 reverse transcriptase induced by nonnucleoside reverse transcriptase inhibitor binding. Curr HIV Res *2*,

323-332.
Smolov, M., Gottikh, M., Tashlitskii, V., Korolev, S., Demidyuk, I., Brochon, J.-C., Mouscadet, J.-F., and Deprez, E. (2006a). Kinetic study of the HIV-1 DNA 3'-end processing single-turnover property of integrase. FEBS Journal *273* 1137-1151.
Smolov, M., Gottikh, M., Tashlitskii, V., Korolev, S., Demidyuk, I., Brochon, J.C., Mouscadet, J.F., and Deprez, E. (2006b). Kinetic study of the HIV-1 DNA 3'-end processing. FEBS J *273*, 1137-1151.
Sokolskaja, E., and Luban, J. (2006). Cyclophilin, TRIM5, and innate immunity to HIV-1. Curr Opin Microbiol *9*.
Soriano, V., and de Mendoza, C. (2002). Genetic mechanisms of resistance to NRTI and NNRTI. HIV Clin Trials *3*, 237-248.
Soriano, V., Gomes, P., and Heneine, W. (2000). Human immunodeficiency virus type 2 (HIV-2) in Portugal: clinical spectrum, circulating subtypes, virus isolation, and plasma viral load. J Med Virol *61*, 111-116.
Spira, S., Wainberg, M.A., Loemba, H., Turner, D., and Brenner, B.G. (2003). Impact of clade diversity on HIV-1 virulence, antiretroviral drug sensitivity and drug resistance. J Antimicrob Chemother *51*, 229-240.
Steitz, T.A. (1999). DNA polymerases: structural diversity and common mechanisms. J Biol Chem *274*, 17395-17398.
Strebel, K., Daugherty, D., Clouse, K., Cohen, D., Folks T., and Martin, M.A. (1987). The HIV 'A' (sor) gene product is essential for virus infectivity. Nature *328*, 728-730.
Streltsov, V.A. (2004). Structural evidence for evolution of shark Ig new antigen receptor variable domain antibodies from a cell-surface receptor. Proc Natl Acad Sci USA *101*, 12444-12449.
Stremlau, M., Owens, C.M., Perron, M.J., Kiessling, M., Autissier, P., and Sodroski, J. (2004). The cytoplasmic body component TRIM5alpha restricts HIV-1 infection in Old World monkeys. Nature.
Strizki, J.M., Tremblay, C., Xu, S., Wojcik, L., Wagner, N., Gonsiorek, W., Hipkin, R.W., Chou, C.C., Pugliese-Sivo, C., Xiao, Y., *et al.* (2005). Discovery and characterization of vicriviroc (SCH 417690), a CCR5 antagonist with potent activity against human immunodeficiency virus type 1. Antimicrob Agents Chemother *49*, 4911-4919.
Su, C., Nguyen, V.K., and Nei, M. (2002). Adaptive evolution of variable region genes encoding an unusual type of immunoglobulin in camelids. Mol Biol Evol *19*, 205-215.
Subbramanian, R.A., and Cohen, E.A. (1994). Molecular biology of the human immunodeficiency virus accessory proteins. J Virol *68*, 6831-6835.
Suzuki, Y., and Craigie, R. (2007). The road to chromatin - nuclear entry of retroviruses. Nat Rev Microbiol *5*, 187-196.
Tanha, J., Xu, P., Chen, Z., Ni, F., Kaplan, H., and Narang, S.A. (2001). Optimal design features of camelized human single domain antibody libraires. J Biol Chem *276*, 24774-24780.
Taube, R., Fujinaga, K., Wimmer, J., Barboric, M., and Peterlin, B.M. (1999). Tat transactivation: a model for the regulation of eukaryotic transcriptional elongation. . Virology *264*, 245-253

Taylor, B.S., Sobieszczyk, M.E., McCutchan, F.E., and Hammer, S.M. (2008). The Challenge of HIV-1 Subtype Diversity. The new england journal o f medicine *35*, 1590-1602.
Temiz, N.A., and Bahar, I. (2002). Inhibitor binding alters the directions of domain motions in HIV-1 reverse transcriptase. Proteins *49*, 61-70.
Tomlinson, I.M., Walter, G., Marks, J.D., Llewelyn, M.B., and Winter, G. (1992). The repertoire of

human germline VH sequences reveals about fifty groups of VH segments with different hypervariable loops. J Mol Biol *227*, 776-798.

Towers, G.J. (2007). The control of viral infection by tripartite motif proteins and cyclophilin A. Retrovirology *4*.

Trono, D., Lint, C.V., Rouzioux, C., Verdin, E., Barré-Sinoussi, F., Chun, T.-W., and Chomont, N. (2010). HIV Persistence and the Prospect of Long-Term Drug-Free Remissions for HIV-Infected Individuals. Science *329*, 174-180.

Underwood, M., Johns, B., Sato, A., Fujiwara, T., and Spreen, W. (2009). S/GSK1349572: a next generation integrase inhibitor with activity against integrase inhibitorresistant clinical isolates from patients experiencing virologic failure while on raltegravir therapy (Poster WEPEA098). In: IAS 2009-2005th Conference on HIV Pathogenesis, Treatment and Prevention, Cape Town, South Africa.

Valkov, E., Gupta, S.S., Hare, S., Helander, A., Roversi, P., McClure, M., and Cherepanov, P. (2009). Functional and structural characterization of the integrase from the prototype foamy virus. . Nucleic Acids Res *37*, 243-255.

Van Damme, N., Goff, D., Katsura, C., Jorgenson, R.L., Mitchell, R., Johnson, M.C., Stephens, E.B., and Guatelli, J. (2008). The interferon-induced protein BST-2 restricts HIV-1 release and is downregulated from the cell surface by the viral Vpu protein. Cell Host Microbe *3*, 245-252.

Van Rompay, A.R., Johansson, M., and Karlsson, A. (2000). Phosphorylation of nucleosides and nucleoside analogs by mammalian nucleoside monophosphate kinases. Pharmacol Ther *87*, 189-198.

Vandekerckhove, L. (2010). GSK-1349572, a novel integrase inhibitor for the treatment of HIV infection. Curr Opin Investig Drugs *11*, 203-212.

Vanegas, M., Llano, M., Delgado, S., Thompson, D., Peretz, M., and Poeschla, E. (2005). Identification of the LEDGF/p75 HIV-1 integrase-interaction domain and NLS reveals NLS-independent chromatin tethering. J Cell Sci *118*, 1733-1743.

Vercruysse, T., Pardon, E., Vanstreels, E., Steyaert, J., and Daelemans, D. (2010). An intrabody based on a llama single-domain antibody targeting the N-terminal alpha-helical multimerization domain of HIV-1 rev prevents viral production. . J Biol Chem *285*, 21768-21780.

Verheesen, P., de Kluijver, A., van Koningsbruggen, S., de Brij, M., de Haard, H.J., van Ommen, G.J., van der Maarel, S.M., and Verrips, C.T. (2006). Prevention of oculopharyngeal muscular dystrophy-associated aggregation of nuclear polyA-binding protein with a single-domain intracellular antibody. Hum Mol Genet *15*, 105-111.

Verheesen, P., ten Haaf, t.M.R., Lindner, N., Verrips, C.T., and de Haard, J.J. (2003). Beneficial properties of single-domain antibody fragments for application in immunoaffinity purification and immunoperfusion chromatography. Biochim Biophys Acta *1624*, 21-28.

Violot, S., Hong, S.S., Rakotobe, D., Petit, C., Gay, B., Moreau, K., Billaud, G., Priet, S., Sire, J., Schwartz, O., *et al.* (2003). The human polycomb group EED protein interacts with the integrase of human immunodeficiency virus type 1. J Virol *23*, 12507-12522.

Votteler, J., Neumann, L., Hahn, S., Hahn, F., Rauch, P., Schmidt, K., Studtrucker, N., Solbak, S.M., Fossen, T., Henklein, P., *et al.* (2011). Highly conserved serine residue 40 in HIV-1 p6 regulates capsid processing and virus core assembly. Retrovirology *8*.

Vu, K.B., Ghahroudi, M.A., Wyns, L., and Muyldermans, S. (1997). Comparison of llama VH sequences from conventional and heavy chain antibodies. Mol Immunol *34*, 1121-1131.

Wainberg, M. (2004). HIV-1 subtype distribution and the problem of drug resistance. AIDS *Suppl 3*, S63-68.

Wang, J., Ling, H., Yang, W., and Craigie, R. (2001). Structure of a two-domain fragment of HIV-1 integrase: implications for domain organization in the intact protein. EMBO J *20*, 7333-7343.

Ward, E.S., Güssow, D., Griffiths, A.D., Jones, P.T., and Winter, G. (1989). Binding activities of a repertoire of single immunoglobulin variable domains secreted from Escherichia coli. Nature *341*, 544-546.

Watanabe, K., Honda, M., Watanabe, T., Tsukada, K., Teruya, K., Kikuchi, Y., Oka, S., and Gatanaga, H. (2010). Emergence of raltegravir-resistant HIV-1 in the central nervous system. Int J STD AIDS *21*, 840-841.

Waters, J.M., O'Neal, W., White, K.L., Wakeford, C., Lansdon, E.B., Harris, J., Svarovskaia, E.S., Miller, M.D., and Borroto-Esoda, K. (2009). Mutations in the thumb connection and RNase H domain of HIV type-1 reverse transcriptase of antiretroviral treatment-experienced patients. Antiviral Ther *14*, 231-239.

Wensing, A.M., van de Vijver, D.A., Angarano, G., Asjo, B., Balotta, C., Boeri, E., Camacho, R., Chaix, M.L., Costagliola D, De Luca A, *et al.* (2005). Prevalence of drugresistant HIV-1 variants in untreated individuals in Europe:Implications for clinical management. J Infect Dis *192*, 958-966.

Wensing, A.M., van Maarseveen, N.M., and Nijhuis, M. (2009). Fifteen years of HIV Protease Inhibitors: raising the barrier to resistance. Antiviral Res *85*, 59-74.

Wesolowski, J., Alzogaray, V., Reyelt, J., Unger, M., Juarez, K., Urrutia, M., CauerhV, A., Danquah, W., Rissiek, B., Scheuplein, F., *et al.* (2009). Single domain antibodies: promising experimental and therapeutic tools in infection and immunity. Med Microbiol Immunol *198*, 157-174.

Whitlow, M., Bell, B.A., and Feng, S.-L. (1993). An improved linker for scFv with reduced aggregation and enhanced proteolytic stability. Protein Eng *6*, 989-995.

Wielens, J., Crosby, I.T., and Chalmers, D.K. (2005). A three-dimensional model of the human immunodeficiency virus type 1 integration complex. J Comput Aided Mol Des *19*, 301-317.

Wild, C.T., Shugars, D.C., Greenwell, T.K., McDanal, C.B., and Matthews, T.J. (1994). Peptides corresponding to a predictive alpha-helical domain of human immunodeficiency virus type 1 gp41 are potent inhibitors of virus infection. Proc Natl Acad Sci U S A *91*, 9770-9774.

Willey, R.L., Maldarelli, F., Martin, M.A., and Strebel, K. (1992). Human immunodeficiency virus type 1 Vpu protein induces rapid degradation of CD4. . J Virol *66*, 7193-7200.

Williams, S.A., and Greene, W.C. (2007). Regulation of HIV-1 latency by T-cell activation. Cytokine *39*, 63-74.

Witvrouw, M., Pannecouque, C., Switzer, W.M., Folks, T.M., De Clercq, E., and Heneine, W. (2004). Susceptibility of HIV-2, SIV and SHIV to various anti-HIV-1 compounds: implications for treatment and postexposure prophylaxis. Antivir Ther *9*, 57-65.

Woodward, C.L., Prakobwanakit, S., Mosessian, S., and Chow, S.A. (2009). Integrase interacts with nucleoporin NUP153 to mediate the nuclear import of human immunodeficiency virus type 1. J Virol *83*, 6522-6533.

Wu, A.M., Chen, W., and Raubitschek, A.A.e.a. (1996). Tumor localisation of anti-CEA single chain Fvs: Improved targeting by non-covalent dimers. . Immunotechnology *2*, 21-36.

Wu, T.T., Johnson, G., and Kabat, E.A. (1993). Length distribution of CDR H3 in antibodies. . Proteins Struct Funct Genet *16*, 1-7.

Xia, Q., Radzio, J., Anderson, K.S., and Sluis-Cremer, N. (2007). Probing nonnucleoside inhibitor-induced active-site distortion in HIV-1 reverse transcriptase by transient kinetic analyses. Protein Sci *16*, 1728-1737.

Xu, L., Anderson, J., Ferns, B., Cook, P., Wildfire, A., and Workman, J. (2008). Genetic diversity of integrase (IN) sequences in antiretroviral treatment-naive and treatment experienced HIV type 2 patients. AIDS Res Hum Retroviruses *24*, 1003-1007.

Xu, L., Anderson, J., Garrett, N., Ferns, B., Wildfire, A., Cook, P., Workman, J., Graham, S., and Smit, E. (2009). Dynamics of raltegravir resistance profile in an HIV type 2-infected patient. AIDS Res Hum Retroviruses *25*, 843-847.

Ye. J., Yu, S.Q., Lu, H., Wang, W., Xin, R.L., and Zeng, Y. (2011). Genetic diversity of HIV-1 isolated from newly diagnosed subjects (2006-2007) in Beijing, China. AIDS Res Hum Retroviruses.

Yi, J., Cheng, H., Andrake, M.D., Dunbrack, R.L.J., Roder, H., and Skalka, A.M. (2002). Mapping the epitope of an inhibitory monoclonal antibody to the C-terminal DNA-binding domain of HIV-1 integrase. J Biol Chem *277*, 12164-12174.

Yoder, K.E., and Bushman, F.D. (2000). Repair of gaps in retroviral DNA integration intermediates. J Virol *74*, 11191-11200.

Yu, X., Yu, Y., Liu, B., Luo, K., Kong, W., Mao, P., and Yu, X.F. (2003). Induction of APOBEC3G ubiquitination and degradation by an HIV-1 Vif-Cul5-SCF complex. Science *302*, 1056-1060.

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

Zhang, H., Dornadula, G., Orenstein, J., and Pomerantz, R.J. (2000). Morphologic changes in human immunodeficiency virus type 1 virions secondary to intravirion reverse transcription: evidence indicating that reverse transcription may not take place within the intact viral core. J Hum Virol *3* 165-172.

Zhou, Y., Zhang, H., Siliciano, J.D., and Siliciano, R.F. (2005). Kinetics of human immunodeficiency virus type 1 decay following entry into resting CD4+ T cells. J Virol *79*, 2199-2210.

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

Zhu, T., Korber, B., Nahmias, A.J., Hooper, E., Sharp, P.M., and Ho, D.D. (1998). An African HIV-1 sequence from 1959 and implications for the origin of the epidemic. . Nature *391*
594-597.