An investigation of the role of HIV-1 Gag mutations in failure of protease inhibitors

A thesis submitted to the University College London for the degree of

Doctor of Philosophy (PhD)

By Ana Garcia-Diaz

Centre for Virology

Division of Infection and Immunity

University College Medical School

Royal Free Campus

May 2012

Declaration of Authorship

I declare that the work in the thesis was carried out in accordance with the regulations of the University College London. The work is original, except where indicated by special references in the text, and no part of the thesis has been submitted for any other academic award.

Signed: Ana Garcia-Diaz

Abstract

Classically, the emergence of resistance to protease inhibitors (PIs) in HIV-1 requires the stepwise accumulation of primary and compensatory mutations in the viral protease (PR). In addition, it was demonstrated that mutations occurring on one of the natural substrates of the PR, Gag, could behave as compensatory mutations in the presence of certain primary PR mutations. Furthermore, mutations on the Gag could account for resistance to PIs when appeared in isolation. So far, most studies assessing the effect of Gag mutations on resistance to PIs have focused on two cleavage sites (CS), namely P7/P1 and P1/P6. However, data on the remaining CS and non-cleavage sites is scarce.

In my PhD studies, I developed and optimized an assay for the amplification and sequencing of HIV-1 Gag and PR genes in order to characterize mutations occurring in patients failing PI-based therapy. Initially, I performed a cross-sectional analysis by comparing the Gag and protease sequences from PI-experienced patients and PI-naïve subjects. A number of Gag mutations associated with PI-selective pressure were determined, which were not restricted to P7/P1 and P1/P6 CSs, but present throughout the Gag. Subsequently, I conducted a longitudinal analysis of patients failing a PI-based regimen, which confirmed that under PI-selective pressure the entire Gag evolved along with the PR and that changes were most prominent at P2/P7, P7/P1 and P1/P6 CSs and in the P17 protein outside CSs. Finally, I performed phenotypic characterization of PI susceptibility and replicative capacity studies on patient's viruses and side-directed mutants. As a result of these investigations I found that the evolution of Gag in patients on unsuccessful PI therapy led to increased levels of PI resistance and improved viral replicative capacity. Specifically, I characterized two novel CS mutations (P17/P24: Y132F, and P2/P7: T375A) that conferred resistance in the context of a wild type backbone.

Acknowledgments

I wish to express my deepest gratitude to all who have supported me through all the years that this PhD has taken:

Firstly, I thank to my supervisor Professor Anna Maria Geretti for her excellent guidance and supervision throughout the entire process and my co-supervisor Professor Vincent Emery for his always useful feedback.

Thank you also to Professor Deenan Pillay for allowing me to spend some time with his research group (Adele McCormick and Chris Parry) and gain experience in phenotypic resistance testing.

I am very grateful to Prof Jens Lundgren and Prof Andrew Phillips for insightful discussions about the aims of the project. I would also like to thank Dr Zoe Fox and Dr Colette Smith for their help with database management and patient selection and for being always available for my "multiple questions"

I am very grateful to all the staff in the virology department for their untiring support.

A big Thanks to my friends Sarah, Aysha, Sam and Mauli for always being there and encouraging me.

Thank you to my brothers and sisters for always believing in me.

Lastly, I would like to thank my parents and my partner Alvaro, this work is dedicated to you because I will not be where I am know it wasn't for your love, support and immense generosity.

Table of Contents

Index of Figures

Index of Tables

Abbreviations

ABC	Abacavir
AIDS	Acquired Immunodeficiency Syndrome
APOBEC	Apolipoprotein B mRNA-editing enzyme, catalytic polypeptide-like
APV	Amprenavir
APV/r	Ritonavir-boosted Amprenavir
ARVs	Antiretrovirals
ATV	Atazanavir
ATV/r	Ritonavir-boosted atazanavir
BCO	Biological cut-off
CCO1	Lower clinical cut-off
CCO2	Upper clinical cut-off
CCR5	Chemokine (C-C motif) receptor 5
CDC	Centers for Disease Control and Prevention
CMV	Citomegalovirus
CRF	Circulating recombinant form
CS	Cleavage site
CSMs	Cleavage site mutations
CTS	Central termination sequence
CypA	Cyclophilin A
d4T	Stavudine
ddC	Zalcitabine
ddI	Didadosine

ddNTPs	Di-deoxyribonucleotide triphosphate
DLV	Delavirdine
dNTPs	Deoxyribonucleotide triphosphate
DRV	Darunavir
DRV/r	Ritonavir-boosted darunavir
EFV	Efavirenz
Env */Env*	Envelope protein/gene
ESCRT	Endosomal sorting complex required for transport machinery
ETV	Etravirine
FDA	Food and drug administration
FEN1	Enzyme Flap Endonuclease 1
FIs	Fusion inhibitors
FTC	Emtricitabine
Gag */Gag*	Group specific antigen protein/gene
Gp120	Glipoprotein 120
Gp160	Glipoprotein 160
Gp41	Glipoprotien 41
HAART	Highly active antiretroviral therapy
HF	High fidelity
HIV	Human immunodeficiency virus
HLA-II	Human leucocyte antigen group II
HR1	Hepta repeat 1
HTLV-III	Human T-cell Lymphotropic virus type 3
IC_{50}	Drug concentration to inhibit virus replication by 50 %
IDV	Indinavir

IDV/r	Ritonavir boosted indinavir
IN	Integrase
INIs	Integrase inhibitors
3TC	Lamivudine
LAV	Lymphadenopathy-Associated Virus
LEDGF/p75	Lens Epithelium-Derived Growth Factor/p75
LPV	Lopinavir
LPV/r	Ritonavir boosted lopinavir
LTR	Long terminal repeats
MHC-I	Major histocompatibility complex class I
MSM	Men who have sex with men
MTT	3-(4,5-Dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide
MVC	Maraviroc
Nef	Negative factor
NFV	Nelfinavir
NLS	Nuclear location signal
NNRTIs	Non-nucleo(s)tide reverse transcriptase inhibitors
Non-CSMs	Non-cleavage site mutations
NRTIs	Reverse transcriptase inhibitors
NVP	Nevirapine
ORF	Open reading frame
P17	Matrix protein (MA)
P24	Capsid protein (CA)
P7	Nucleocapsid protein (NC)
PBMCs	Peripheral Blood Mononuclear Cells

PCP	Pneumocystis carinii pneumonia
PBS	Primer binding site
PI/r	Ritonavir boosted protease inhibitor
PIC	Pre-integration complex
PIs	Protease inhibitor
PPT	Polypurine Tract
PR	Protease
pTEFb	Positive Transcription Elongation Factor b
RAL	Raltegravir
RAMs	Resistance associated mutations
RC	Replicative capacity
Rev	Regulatory factor for HIV expression
RNase H	Ribonuclease H
RNP	Ribonucleoprotein
RPV	Ripilvirine
RT	Reverse transcriptase
RTC	Reverse transcription complex
RTIs	Reverse transcriptase inhibitors
RTV	Ritonavir
RTVs	Resistance test vectors
SDM	Side-directed mutagenesis
SIV	Simian immunodeficiency virus
SQV	Saquinavir
SQV/r	Ritonavir boosted saquinavir
T-20	Emfuvirtide

TAMs	Thymidine analogue mutations
TAR	transactivator response element
Tat	Transactivator of HIV gene expression
TDF	Tenofovir
TPV	Tipranavir
TPV/r	Ritonavir boosted tipranavir
URF	Unique recombinant forms
VF	Virological failure
Vif	Viral infectivity factor
Vpr	Viral protein R
Vpu	Viral protein u
WT	Wild-type
ZDV	Zidovudine

1 Chapter one: general introduction

1.1 History of the discovery of HIV-1

On 5th June 1981, the Centers for Disease Control and Prevention (CDC) published a report of *Pneumocystis carinii* pneumonia in five previously healthy young homosexual men in Los Angeles, two of whom had died (CDC Weekly 1982a; CDC Weekly 1982b; CDC Weekly 1982c). Examination of the patients showed a marked reduction in cellular immune response as a common denominator. Owing to the media influence, the disease was originally dubbed Gay-related immune deficiency (GRID). Similar cases were soon reported in Western European countries, all of which were characterized by a profound depression of cell-mediated immunity and the presence of opportunistic infections and rare malignancies previously described only in severely immunocompromised patients. These new cases were not restricted to men who have sex with men (MSM), but affected other population groups, such as blood transfusion patients and injecting drug users. In 1982, the CDC introduced the term acquired immune deficiency syndrome (AIDS) to describe this newly recognized entity.

In January 1983, a potential causative agent was isolated at the Pasteur institute in France by Luc Montagnier and colleagues from cultured T lymphocytes derived from a patient with cervical lymphadenopathy, a sign that was considered a precursor of AIDS. The new retrovirus was named lymphadenopathy-associated virus (LAV), but no proof of causality with AIDS was established at this time (Barre-Sinoussi, 1983). In May 1984, Robert Gallo's group compiled sufficient evidence to convince the medical and scientific communities that the new virus, which was renamed human T-lymphotropic virus type III (HTLV-III), was the etiological agent for the emerging AIDS epidemic (Popovic et al, 1984). In August 1984, Jay

levy confirmed Gallo's findings by reporting they had isolated a retrovirus designated ARV for AIDS-associated retrovirus, in 22 patients with AIDS (Levy et al, 1984).

In February 1985, sequencing of the entire genome from LAV, HTLV-III and ARV viruses demonstrated that they were variants of the same virus (Ratner et al, 1985). In 1986, the Committee on the Taxonomy of Viruses proposed the name human immunodeficiency virus (HIV) to designate the causative agent of AIDS. On the same year, a second retrovirus was isolated from West African patients hospitalized at a Lisbon hospital (Clavel et al, 1986), who had a clinical picture compatible with AIDS but no detectable antibodies against HIV. This new virus was called human immunodeficiency virus type 2 (HIV-2), and the original HIV was renamed as human immunodeficiency virus type 1 (HIV-1).

For the discovery of the infectious agent currently known as HIV-1, Luc Montagnier and Francoise Barré-Sinoussi from the Pasteur Institute were awarded the Nobel Prize in Physiology or Medicine in 2008.

1.2 HIV-1 origin and diversity

HIV-1 is classified in four groups, named M (major), N (non-M, non-O), O (outlier) and P, each of which arose from an independent zoonotic transmission from non-human primates infected with simian immunodeficiency virus (SIV) to humans in Central and West Africa. Human infections probably occurred through cutaneous or mucous membrane exposure to infected ape blood/body fluids during activities such as hunting or butchering of primates. HIV-1 groups M and N originated from two different lineages of SIV_{cpzPtt} that infected common chimpanzees (*Pan troglodytes troglodytes*) inhabiting southeastern and south-

central Cameroon, respectively (Corbet et al, 2000; Gao et al, 1999; Keele et al, 2006). HIV-1 groups O and P are related to SIV_{gor} found in gorillas (*Gorilla gorilla*) living in Cameroon, although SIV_{gor} itself is supposed to be derived from an ancestor of a divergent SIV_{cpz} lineage acquired by cross-species infection from sympatric chimpanzees (Plantier et al, 2009 and Van Heuverswyn et al, 2006).

The timing of cross-species transmissions varies according to particular HIV-1 groups. HIV-1 group M appears to be the oldest lineage in humans with an estimated time to the most recent common ancestor around 1908 (range 1884-1924). The estimated times of the most recent common ancestors of HIV-1 groups O and N are 1920 (1890-1940) and 1963 (1948-1977), respectively (Korber et al, 2000; Wertheim et al, 2009). The time of origin of HIV-1 group P is unknown as only two different sequences are available, precluding relevant phylogenetic analysis.

The current global diversity of HIV-1 is the result of the interaction of several elements, namely the intrinsically high genetic variation of the virus, the selection and evolution of viral strains within the host and the different patterns of propagation of viral strains.

HIV has a huge intrinsic genetic variability. The lack of 3'-5' exonuclease activity (proofreading activity) of the reverse transcriptase enzyme introduces 0.2-2 substitutions per genome per replication cycle (Darke et al, 1993; Roberts et al, 1988). In addition, relatively large insertions and deletions are common occurrences in the viral genome. These high rates of mutation in conjunction with high rates of virus replication (10^{10}-10^{12} virions per day) (Ho et al, 1995) result in the rapid generation of genetically diverse viral populations within each individual, where viral sequences can differ by up to 10%. Moreover, further genetic

variability occurs due to a high recombination rate (7-30 recombination events per genome per replication cycle) (Jetzt et al, 2000) when two or more different strains simultaneously infect an individual.

The enormous genetic diversity of HIV-1 is subject to intra-host evolutionary selective pressures. Among them, the action of the immune system plays a crucial role in shaping the structure of the viral population. Despite the variability of the HIV-1 genome, only 10% undergoes positive selective pressure. The regions affected are those that define critical residues in host-pathogen interaction, half of which are mapped to CTL epitopes located in the Gag protein. The importance of the host immune response in driving viral diversity and evolution was clearly showed by Draenert and colleagues in a study of monozygotic twins infected with the same viral strain (Draenert et al, 2006). They found that the initial CD8+ T cell response targeted 17 epitopes, 15 of which were identical in each twin. Three years post-infection, four responses had declined in both twins, three of which showed mutations at the same Gag residues. Similarly, the antibody responses cross-neutralized the other twin's virus and also showed similar evolutionary changes in the envelope gene. These results illustrated a considerable concordance in cellular and humoral immune response and HIV-1 evolution in the same genomic environment.

In addition to adaptive immune response, mammalian cells express a number of proteins whose function is to suppress viral replication. These have been termed restriction factors and provide an initial line of defense against infection as a component of, or even preceding, innate antiviral responses. The most extensively described host restriction factors include the apolipoprotein B messenger RNA (mRNA)-editing enzyme catalytic polypeptide-like 3 (APOBEC 3) family of proteins, and in particular APOBEC3G, the tetherin/bone marrow stromal cell antigen 2 (BST2)/CD317 termed tetherin, and the tripartite-motif-containing 5α

(TRIM5α). APOBEC3G is a member of a family of vertebrate proteins with polynucleotide cytidine deaminase activity; its action interferes with reverse transcription by inducing numerous deoxycytidine to deoxyuridine mutations on the nascent proviral negative strand, which ultimately results in guanosine-to-adenosine hypermutation at hot spots within the proviral DNA, thus rendering the provirus defective. Its action is counteracted by HIV-1 vif protein, which binds APOBEC3G inducing its proteosomal degradation. TRIM5α is a cytoplasmic protein, whose mechanism of action is unknown, but it has been demonstrated that binds to the viral capsid precipitating its degradation soon after entry and ultimately blocking reverse transcription. Tetherin is a transmembrane protein that is incorporated into the lipid envelope of the HIV-1 particle and causes virions to remain trapped at the surface of the infected cells hindering its dissemination. The HIV-1 Vpu protein antagonizes tetherin action by an uncertain mechanism (Neil et al, 2008). In general, host restriction factors are poor inhibitors of retrovirus that are found naturally in the same host species, but are often active against retroviruses that are found in other species and therefore these proteins are important determinants of host range and cross species transmission. In addition, in order to establish infection a host-pathogen co-evolution is required and consequently, host restriction factors are important determinants of viral evolution. For instance, it has been proposed that APOBEC action is not completely abolished by HIV-1 vif and that it can induce infrequent mutations (as opposed to hypermutation) which may be benefitial to the virus in terms of immune escape or drug resistance and consequently, APOBEC3 appears to be an important contributor to viral diversity and evolution (Wood et al, 2009; Kim et al, 2010 and Sadler et al, 2010).

Furthermore, exogenous pressures, such as anti-retroviral therapy, further delineate the nature of quasispecies present in single individuals.

The patterns of dissemination of HIV-1 groups are substantially different. Whereas groups N, O and P have not spread significantly beyond Central and West Africa, group M is responsible for the HIV pandemic as the vast majority (>95%) of viral strains distributed worldwide belong to this group.

HIV-1 group M is highly heterogeneous on the basis of phylogenetic analysis. Currently, it is divided into nine subtypes or clades (A, B, C, D, F, G, H, J, K) (Geretti, 2006), which show an average intersubtype genetic variability of 15% for the *gag* gene and 20% for the *env* gene (Geretti 2006; Robertson et al, 2000). Moreover, some subtypes are subdivided in sub-subtypes (Geretti, 2006), as is the case for clades A and F, which are separated into A1 and A2 and F1 and F2, respectively. In addition, full-length genome sequencing has revealed the existence of intersubtype recombinants, which are classified either as unique recombinant forms (URFs) or circulating recombinant forms (CRFs). URFs represent recombinant viruses that have been only identified in a single individual or an epidemiologically linked cluster without evidence of epidemic spread, whereas CRFs are recombinant viruses that have been identified in at least three epidemiologically unlinked individuals (Geretti, 2006; Thomson and Najera, 2005). CRFs are labeled with numbers in order of discovery followed by the letters of the two parental subtypes in alphabetical order (e.g., CRF03_AB) In addition, the extension “cpx”, for complex, is given if the CRF consists of contributions from three or more subtypes. At present, a total of 51 CRFs have been described, which constitute around 20% of all HIV infections worldwide (Hemelaar et al, 2012).

1.3 HIV-1 epidemiology

1.3.1 Global distribution of HIV

Currently, HIV-1 accounts for more than 30 million infections worldwide. Group M viruses constitute the pandemic form of HIV representing over 95% of all HIV-1 infections and has been found in virtually every country on the globe. By contrast, the contribution of HIV-1 groups O, N and P to the pandemic is negligible as these clades have remained confined to specific countries in West and Central Africa. Group O viruses have not spread significantly beyond Cameroon, Gabon and Equatorial Guinea, and represent less than 1% of global HIV-1 infections. Group N infections have been only detected in 13 individuals from Cameroon, whereas group P viruses have so far only been identified in two Cameroonians from Yaounde.

The different group M subtypes and recombinant forms have distinct global distribution patterns (Hemelaar et al, 2012). All HIV-1 group M subtypes and a high proportion of URFs and CRFs are present in West Central Africa, the potential epicentre of the global HIV epidemic. However, in other regions usually one or two genetic forms are predominant, fact that responds to a founder effect whereby the earliest genetic form successfully introduced within a population establishes itself as predominant and gains an initial advantage over other genetic forms arriving later. Nonetheless, this situation is not static and replacement of established genetic variants can occur. The best example is possibly the case occurred in Thailand were the initially introduced subtype B was replaced by CRF01-AE a year later. Similarly, an increase in non-B subtype infections has been reported in several Western European countries, mainly among native individuals.

On a global scale, subtype C is the most successful of the HIV-1 M lineages (48%) (Hemelaar et al, 2012), followed by subtypes A (12%) and B (11%). The most abundant CRF is CRF02_AG, which accounts for 8% of global HIV-1 infections. Other subtypes and recombinant forms represent individually less than 5% of the global total.

Regarding subtype distribution, subtype C is highly prevalent in Southern Africa countries, India and Ethiopia and also circulates as a minor form in Brazil and Russia (Geretti, 2006). Subtype A viruses are predominant in East and West Africa as well as in Eastern Europe and Central Asia (Thomson and Najera, 2005). Subtype B predominates in the Americas, Western Europe, the Middle East and North Africa. However, the prevalence of non-B subtype infections in high-income countries of North America and Western Europe has increased as a consequence of the influx of immigrant population from Africa and Asia (Fox et al, 2010). With regard to CRFs, CRF02-AG is highly prevalent in West Africa (Takebe et al, 2004), whereas CRF01-AE is the major genetic variant in South and East Asia (Thomson and Najera, 2005), where it has replaced subtype B as the predominant form.

Other less prevalent subtypes and CRFs are more locally distributed. For example, subtype D is mainly found in East Africa, subtype F is predominant in Romania, subtype G is mainly spread across West and Central Africa, with the highest prevalence in Nigeria but it is also found in Portugal and northwest Spain (Thomson and Najera, 2005), and CRF12_BF widely circulates in Argentina. In addition, URFs also feature in the HIV epidemic. A diversity of URFs have been reported in areas such as DR Congo, Tanzania, Argentina, Cuba and Galicia (Thomson and Najera, 2005).

1.3.2 Current status of the HIV pandemic

Since the pandemic began in 1981, the cumulative total of individuals infected with HIV ascends to 60 millions, 25 millions of whom have since died (UNAIDS 2010). At the end of 2009, the Joint United Nations Programme on HIV/AIDS (UNAIDS) estimated that there were 33.3 million [31.4-35.3] people living with HIV infection, that 2.6 million [2.3-2.8] people became newly infected with the virus during 2009 and that there were 1.8 million [1.6-2.1] AIDS-related deaths in the same year (UNAIDS 2010).

Worldwide, the rates of annual new infections have been steadily declining since 1997, the year in which the epidemic peaked. Between 2001 and 2009, a more than 25% reduction in HIV incidence was documented in 33 countries, 22 of which were located in sub-Saharan Africa, the global region bearing the highest HIV prevalence. In spite of the number of new infections falling in most parts of the world, the number of people living with HIV has continued to rise due to a significant reduction in AIDS-related deaths as a consequence of expanded access to antiretroviral therapy.

Sub-Saharan Africa is the global region with the highest HIV prevalence (67.5%). The HIV burden is unevenly distributed, with countries in southern Africa most severely affected (South Africa, Zambia, Zimbabwe, Malawi, Mozambique, Namibia, Botswana, Lesotho, Swaziland). In particular, South Africa, which had an estimated 5.6 million [5.4-5.8] infected people in 2009, remains the largest epidemic in the world. Although the rate of new

infections in sub-Saharan Africa decreased from 2.2 million in 2001 to 1.8 million in 2009, the number of people living with HIV has steadily risen during this period reaching a total of 22.5 million [20.9-24.2] at the end of 2009, figure which corresponds to 68% of the global total. Importantly, women account for 60% of infections. The main transmission modes are unprotected sexual intercourse and perinatal infection. In regard to AIDS-related mortality, sub-Saharan Africa continues to be the most important contributor worldwide, with an estimated 1.3 million [1.1-1.5] deaths ascribable to the epidemic during 2009, which equals to 72% of the global total.

South and East Asia is second to sub-Saharan Africa in terms of HIV prevalence (14.3%), with an estimated 4.9 million [4.5-5.5] infected people at the end of 2009. The HIV epidemic in this global region is largely stable as the previous figure is similar to the one from 2004. Between 2001 and 2009, the HIV incidence has fallen by more than 25% in India, Nepal and Thailand, remained stable in Malaysia and Sri Lanka, but increased by more than 25% in Bangladesh and Philippines. The HIV burden is mainly concentrated in intravenous drug users, sex workers and their clients and homosexual men.

The Eastern Europe-Central Asia global region deserves special merit as the number of people living with HIV has more than doubled since 2001 and reached an estimated total of 1.4 million [1.3-1.6] in 2009. The HIV incidence has increased in several countries in this region such as Armenia, Georgia, Kazakhstan, Kyrgyzstan, Tajikistan, The Russian Federation and Ukraine. However, The Russian Federation and Ukraine account for almost 90% of newly reported HIV diagnoses. The HIV epidemics is mainly concentrated among intravenous drug users and their sexual partners, which is leading to an increase of the proportion of HIV infected women, and sex workers and their clients (Mathers et al, 2008).

The Middle East and North Africa global region shows a similar trend to Eastern Europe-Central Asia global region as the number of people living with HIV has leapt from an estimated 180,000 (150,000-200,00) in 2001 to an estimated 460,000 (400,000-530,000) in 2009. In addition, the HIV incidence has more than doubled and AIDS-related deaths have more than tripled during this period. However, reliable data of the epidemic in these regions are still scarce, hindering proper data analysis.

The HIV epidemic in the North America and Western-Central Europe global region appear to have stabilized as the rates of annual new HIV infections have remained constant for at least the past five years. However, the number of people living with HIV has steadily risen between 2001 and 2009 reaching a total of 2.3 million [20.9-2.7] at the end of 2009. Unprotected sex between men continues to dominate patterns of HIV transmission in these areas. However, injecting drug use and unprotected heterosexual paid sex are also important, especially in Mexico and parts of Southern Europe. Worth of mention is the increasing role played by immigrants from countries with generalized epidemics, as they represented almost 17% of people newly infected with HIV in Europe during 2007.

The HIV epidemics in other global regions, including Central and South America, the Caribbean and Oceania are either stable or declining. In South and Central America, about one third of all people living with HIV live in Brazil and most of the epidemic is concentrated in this region in and around networks of MSM. Injecting drug use has been the other main route of transmission, especially in the southern cone of South America. In the Caribbean, unprotected heterosexual sex is believed to be the main mode of transmission in this region and the burden of HIV varies considerably between countries, the exceptionally low prevalence of Cuba (0.1%) contrast with a 3.1% adult HIV prevalence in the Bahamas. The

Caribbean remains the only region, besides sub-Saharan Africa, where women outnumber men among people living with HIV. The HIV epidemic in Oceania is small, the largest and the only one generalized is in Papua New Guinea. The HIV epidemic is mainly driven by sexual transmission. Unprotected heterosexual intercourse is the main mode of transmission in Papua New Guinea, while unprotected sex between men predominates in the smaller Pacific countries and those of Australia and New Zealand.

1.3.3 The status of the HIV epidemic in the UK

To date, 120,000 people have been diagnosed with HIV in the UK, of whom 27,000 have developed AIDS and more than 20,000 have died. An estimated number of 91,500 (85,400-99,000) people were living with HIV in the UK at the end of 2010, of whom approximately a quarter were unaware of their infection (HIV in UK 2011 report). Most individuals acquired their infection heterosexually [47,000 (43,900-50,400)] and through sex between men [40,100 (35,300-46,700)], whereas a minority [2,300 (1,900-2,700)] were injecting drug users.

A total of 6,660 individuals (4,510 men and 2,150 women) were diagnosed with HIV infection during 2010. An estimated 50% (3,350) of newly diagnosed individuals acquired their infection heterosexually. Most of these individuals were black Africans who acquired the infection abroad, mainly in sub-Saharan Africa. However, 33% of heterosexually acquired infections occurred within the UK. An estimated 45% (3,000) of new HIV diagnoses resulted from sex between men. Most of these individuals (81%) acquired the infection in the UK, were of white ethnicity and two thirds were born in the UK. While the number of HIV diagnoses among people infected heterosexually has declined, new diagnoses among MSM have reached an all-time high. The number of new infections among people

who inject drugs remains low (2.5%), mostly due to early and effectual harm-reduction programmes. No case of HIV acquisition through blood transfusion has been documented in the UK since 2002.

Concerning subtype and CRFs distribution, subtype B continues to be the most common subtype among MSM, however an increase of non-B subtypes among this population has been reported with subtypes C, and A followed by CRF01-AE and CRF02-AG recombinant being the most common subtypes found, indicating that risk-group segregation of HIV-1 clades is becoming less distinct (Fox et al, 2010).

1.4 HIV-1 virion characteristics

HIV-1 is a member of the *Lentivirus* genus in the family of *Retroviridae* (Ratner et al, 1985; Wain-Hobson et al, 1985). The HIV-1 virion is spherical in morphology and measures between 100 and 150 nm in diameter.

1.5 HIV-1 genome and proteins

The HIV-1 genome is approximately 9.2 kb in length. The sequence is flanked by the two long terminal repeats (LTRs). The 5' LTR contains the enhancer/promoter sequence for viral transcription and the 3' LTR contains the polyadenylation signal. The viral genome contains nine open reading frames (ORFs). From 5' to 3' it comprises the *gag* gene, which encodes the virion structural components; the *po*l gene, which encodes the viral enzymes and the *env* gene, which encodes the envelope glycoproteins. The HIV-1 genome contains six additional genes: *tat* and *rev* which code for two regulatory proteins and *vif*, *vpr*, *nef* and *vpu* which encode four accessory peptides.

Figure 1.3 Genome organization of HIV-1 (adapted from Sierra-Aragon, 2008).

The three primary HIV-1 translation products (Gag, pol and env) are initially synthesized as polyprotein precursors, which are subsequently processed by viral or cellular proteases into mature proteins. The Gag precursor is cleaved by the viral protease into the matrix (P17), capsid (P24), nucleocapsid (P7) and p6 proteins as well as two small spacer peptides P1 and P2 (Henderson et al, 1992; Mervis et al, 1988). The P17 protein is located along the inner leaflet of the viral lipid envelope, where it directs the incorporation of the envelope glycoproteins into the forming virion (Dorfman et al, 1994). The P24 protein assembles to form the conical core of the virion. The P7 is an RNA binding protein responsible for packaging of the genomic RNA into the virion (Gorelick et al, 1990). The p6 protein appears to be important for viral budding (Accola et al, 2000). Autocatalysis of the Gag-Pol precursor give rise to the retroviral enzymes: the protease (PR), which is an homodimeric protein that is required for the formation of fully mature and infectious viral particles; the

reverse transcriptase (RT), which provides both the RT activity, that allows RNA-dependent RNA polymerization and the RNase H activity, that allows the specific degradation of RNA in RNA-DNA duplexes; and the integrase (IN) which mediates the integration of the viral DNA into the chromosome of the host cell (Farnet et al, 1996). Finally, proteolytic digestion of the envelope glycoprotein (gp160) by cellular serine proteases leads to surface (gp120) and transmembrane (gp41) subunits (Willey et al, 1988). The remaining six HIV-1 encoded proteins (Vif, Vpr, Tat, Rev, Vpu and Nef) are the primary translation products of spliced mRNA. The tat protein is a trans-activating protein that enhances the rate of viral transcription and permits synthesis of full-length transcripts to occur (Laspia et al, 1990; Marciniak et al, 1990). The rev protein mediates the transport of single spliced and unspliced viral RNAs from the nucleus to the cytoplasm (Emerman et al, 1989; Malim et al, 1989). The vif protein acts late in the viral life cycle to facilitate virus release and infectivity, it interacts with the cellular proteins APOBEC3F and APOBEC3G inducing their degradation by proteosomes. In the absence of vif the APOBEC proteins are incorporated into virions where they deaminate cytidine to uridine in the minus strand of the forming cDNA leading to inactivating hypermutation in the HIV-1 genome. Consequently, Vif is required for production of infectious virions (Malim et al 2008; Henriet et al, 2009). The vpr protein is involved in infectivity, apoptosis, cell cycle control, viral transcription and nuclear import of the pre-integration complex (Romani and Engelbrecht, 2009). The nef protein modulates both cellular signal transduction and membrane trafficking. It plays a role in downregulation of CD4 and class I major histocompatibility complex (MHC) from the cell surface preventing immune recognition of infected cells (Garcia et al, 1991). Finally, the vpu protein enhances the release of virions from infected cells (Klimkait et al, 1990; Nomaguchi et al, 2008) and degrades CD4 during virus production preventing and inhibitory effect on infectivity of

progeny virions that occur when the viral receptor interacts with the viral envelope (Levesque et al, 2003).

1.6 The HIV-1 replication cycle

The entire replication cycle of HIV-1, which is completed in approximately 24 hours (Perelson et al, 1996), comprises a series of sequential steps which are shown in figure 1.4 and reviewed below.

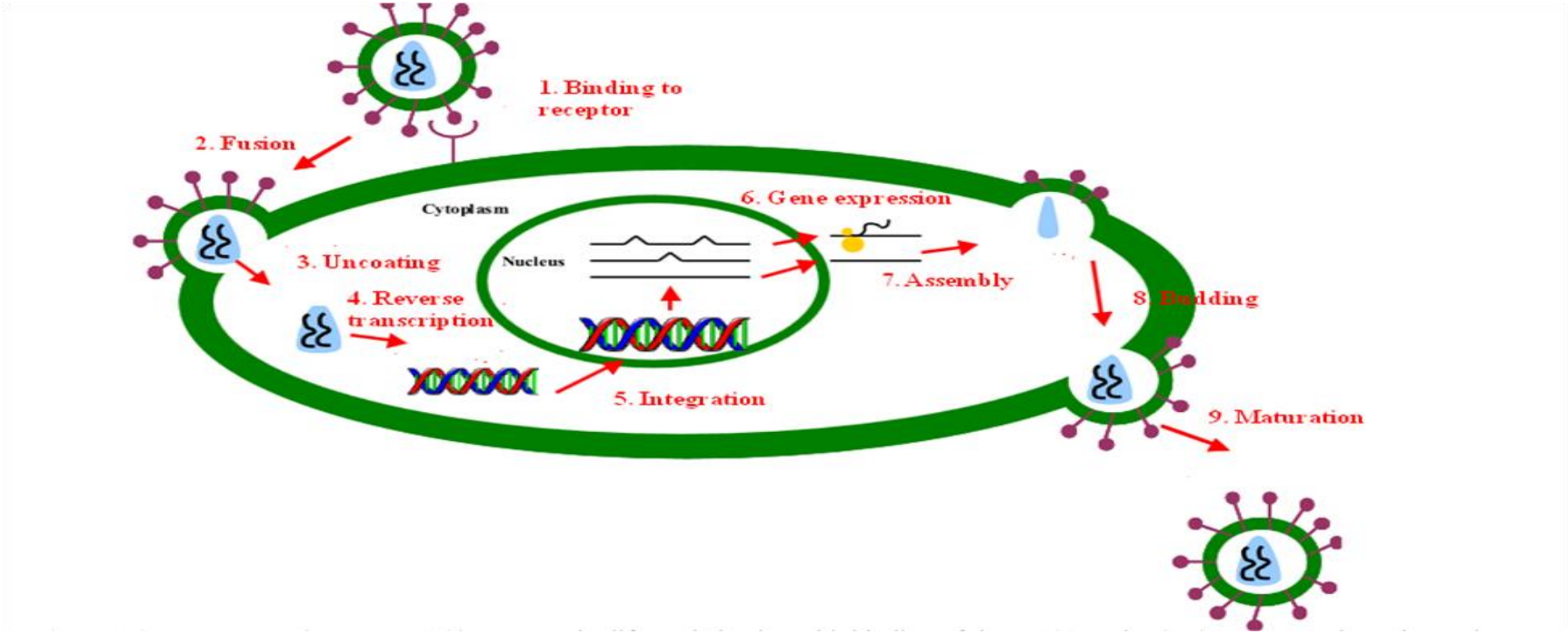


Figure 1.4 Replication cycle of HIV-1

The HIV-1 replication cycle begins with the binding of viral envelope surface glycoprotein (gp120) to the cell surface CD4 receptor and CCR5 or CXCR4 chemokine coreceptors. This is followed by fusion, entry and uncoating of the conical viral core. Reverse transcription converts the single-stranded viral RNA genome into a double-stranded DNA copy. The preintegration complex (PIC) of viral and cellular proteins and proviral DNA is transported to the nucleus, followed by integration into the host chromosomal DNA. Next, the integrated viral DNA is transcribed by cellular RNA Pol II forming spliced and unspliced mRNA templates used for translational synthesis of the accessory factors and polyproteins (Gag and Gag-Pol) encoding structural proteins and functional enzymes. Viral RNA, polyproteins, and envelope localize to the inner face of the plasma membrane where they are packaged into assembling viral particles. Finally, progeny virions bud from the cell surface as immature particles and acquire infectious capacity by proteolysis-induced morphological maturation.

1.6.1 Viral entry

Initially, the virion adsorbs to the target cell (i.e., T helper lymphocytes, macrophages and some populations of dendritic cells) through the interaction of the viral envelope protein (env) or human cell membrane proteins present on the viral envelope with a number of several attachment factors displayed on the target cell membrane (e.g., heparan sulfate proteoglycans, α4β7 integrin), which in turn brings the viral receptor (i.e., CD4 molecule) and co-receptor (i.e., CC or CXC family of chemokine receptors) in close proximity to env, thus increasing the efficiency of infection. Subsequent binding of the gp120 subunit of env to the CD4 molecule causes rearrangements in the V1/V2 loops followed by the V3 loop and leads to the formation of a bridging sheet composed of two double stranded β sheets that are spatially separated in the unliganded state. These conformational changes in gp120 enable its binding to the co-receptor molecule (Sattentau et al, 1991; Wu et al, 1996). Two major co-receptor molecules, the α-chemokine receptor CXCR4 and the β-chemokine receptor CCR5, have been identified. Although a number of other chemokine receptors can act as co-receptors for HIV-1 entry in cultured cells (Berger et al, 1999), there is no compelling evidence that they play an important role *in vivo*. HIV-1 strains can be classified based on their co-receptor usage preference, which is mainly determined by the base sequence of the V3 loop. R5-tropic strains are those that employ CCR5, X4-tropic strains those that use CXCR4 and dual-tropic strains those that can employ both co-receptors (Berger et al, 1998). The importance of viral co-receptors for HIV-1 infection was demonstrated by the discovery of a 32 base-pair deletion in CCR5, termed CCR5Δ32, which has an allelic frequency of 10% in Caucasians. This mutation results in a truncated CCR5 protein that is not expressed on the cell membrane, but retained in the endoplasmic reticulum. Homozygosity for this polymorphism confers profound resistance to HIV-1 infection as homozygous individuals are

only rarely infected despite persistent high risk behavior (Paxton et al, 1996; Samson et al, 1996). Co-receptor binding induces the exposure of a hydrophobic fusion peptide in the amino-terminal ectodomain of gp41, which inserts into the cell membrane. This allows the fusion peptide of each gp41 in the trimer to fold at an angle bringing the amino and carboxy-terminal helical regions from each gp41 subunit together to form a six helix bundle, which in turn brings the viral and cellular lipid bilayers into close apposition, resulting in the formation of a fusion pore (Melikyan et al, 2000) through which the viral contents are delivered into the cytoplasm.

1.6.2 Uncoating

Uncoating is defined as the loss of viral capsid that occurs within the cytoplasm of the infected cell after the virion enters the cell and before the viral genome penetrates the nucleus. It is an obligatory step in the HIV life cycle that accompanies the transition between reverse transcription complexes (RTCs), in which reverse transcription occurs, and pre-integration complexes (PICs), which are able to integrate into the host genome. HIV-1 enters the nucleus through a nuclear pore which diameter falls behind that of the viral capsid. Consequently, uncoating should occur at some point after viral entry and before nuclear import. Nevertheless, the exact time and location of the event remains unclear. Recent studies suggest that uncoating occurs gradually, possibly in response to both cellular and viral signals. Among cellular factors are found: Cyclophilin A (CypA), which has been seen to bind the capsid and assist uncoating (Javanbakht, et al 2007), prolyl isomerases pin 1 (Pin1) which specifically recognised phosphorilated serine-proline residues in the viral capsid and promotes uncoating as demonstrated by the linked between dysfunctional uncoating and depletion of pin1 in cell targets (Misumi et al, 2010) and also cellular factors present in non-resting cells as it has been demonstrated that uncoating requires cell activation (Auewarakul

et al, 2005). Viral factors such as the integrase protein that appears to be required to maintain the interaction between CypA and capsid (Briones et al, 2010) as well as the complexion of reverse transcription and formation of the central DNA flap seems also to play a crucial role in HIV-1 uncoating (Arhel et al, 2007).

1.6.3 Reverse transcription

Reverse transcription is the process whereby the single-stranded RNA viral genome is converted into a linear double-stranded DNA that is the substrate for integration into the host genome. The enzyme that performs this action is the HIV-1 reverse transcriptase (RT), which has two activities: (a) DNA polymerase that can copy either RNA or DNA templates and (b) RNase H that degrades RNA from DNA-RNA duplexes.

Reverse transcription commences with the binding of the host $tRNA^{Lys3}$ to a complementary sequence located approximately 180 nucleotides from the 5' end of the viral genome called the primer binding site (PBS). The RT initiates the synthesis of the minus strand DNA from the 3' end of the tRNA and proceeds towards the 5' end of the viral genome sequentially copying the U5 and R sequences. The RNase H activity of the RT removes the U5 and R RNA sequences from the RNA-DNA structure, exposing the newly synthesized minus DNA strand. As a result of the degradation of RNA-DNA hybrid, the minus DNA strand is exposed facilitating the annealing between the newly synthesized R DNA sequence and a complementary R RNA sequence present at the 3' end of the RNA genome, which in turn leads to the transfer of the minus DNA strand to the direct repeat at the 3' end of the RNA genome. After this transfer, the RT continues to elongate the minus DNA strand towards the

PBS located at the 5’ end of the RNA genome. As DNA synthesis proceeds, so does RNase H degradation. In HIV-1, there are two short purine rich sequences, known as polypurine tracts (PPT), which are resistant to the RNase H activity and serve as primers for the synthesis of the plus DNA strand. The PPT located adjacent to the 3’ end of the viral genome is essential for viral replication, whereas the PPT located near the middle of the viral genome increases the ability of the virus to complete the plus DNA strand, but is not essential. The RT initiates the synthesis of the plus DNA strand from the PPT adjacent to the 3’ end of the viral genome and proceeds towards the 5’ end of the minus DNA strand sequentially copying the U3, R, U5 sequences, but also the first 18 nucleotides of $tRNA^{Lys3}$. Afterwards, the tRNA is partially removed by the RNaseH activity, exposing the PBS sequence in the 3’ end of the plus DNA strand and facilitating its pairing with the complementary PBS sequence located at the 3’ end of the minus DNA strand and consequently leading to a second translocation event. Extension of plus and minus DNA strands by the RT leads to the synthesis of the complete double-stranded linear viral DNA, which is longer than the viral genome as each end is flanked by a long terminal repeat (LTR) containing the U3-R-U5 sequence. Plus-strand synthesis terminates at the end of the minus strand at a sequence known as central termination signal (CTS) (Charneu et al, 1994). The position of the central PPT upstream of the CTS results in the displacement of approximately 100 nucleotides of plus-strand DNA and the formation of a central triplex DNA structure termed central DNA flap which appears to have a role in translocation of the viral DNA into to the nucleus (Zennou et al, 2000) and that is subsequently eliminated by the cellular enzyme flap endonucelase I (FEN1).

1.6.4 Intracellular transport and nuclear entry

Prior to integration, the viral cDNA is translocated into the nucleus as part of a large nucleoprotein complex, the pre-integration complex (PIC), which contains both viral (i.e., P17, P24, RT, integrase and Vpr) (Farnet et al, 1991; Miller et al, 1997) and cellular proteins.

HIV-1 and other lentiviruses have the unusual ability of infecting non-dividing cells, which implies that the PIC must enter the nucleus through an intact nuclear membrane. In contrast, other retroviruses require the disintegration of the nuclear membrane during mitosis to gain access to the nuclear components. The exact mechanism by which HIV-1 enters into the nucleus remains to be established (Fassati 2006). Over the years, a plethora of viral determinants involved in viral translocation through the nuclear pore have been proposed, including several components of the PIC, such as matrix (P17), capsid (P24) and integrase proteins, as well as a triple stranded DNA structure known as the flap. Nuclear import of the PIC is mainly directed by nuclear location signals (NLS) present in the above mentioned viral proteins. Thus, both P17 and integrase proteins are recognised by the NLS-binding site present on the importin α protein. This binding triggers the interaction between importin α and β proteins, which in turn targets the PIC to the nuclear pore by attachment to nucleoporins (Nitahara-Kasahara, et al, 2007; Görlich et al, 1996). Although Vpr does not contain a NLS, it has been shown to promote the translocation of the PIC by tethering to the nuclear pore in a manner analogous to the importin α/β family (Gallay et al, 1996; Vodicka, et al 1998). In addition, the three-stranded DNA flap structure generated at the central polypurine tract also appears to contribute to the nuclear import of the PIC as a ten-fold reduction in the efficiency of this process was documented in its absence (Zennou et al, 2000).

However, the latest evidence suggests that the capsid protein (P24) is the main viral determinant of nuclear import and point mutations in this protein selectively impair viral entry into the nucleus. In addition, genome-wide screens studies have identified several host nuclear pore proteins that HIV-1 can utilize to gain access to the nucleus; among them are transportin 3 (product of the TNPO3 gene) and Nup358 (product of RANBP2 gene) and P24 is responsible for controlling the virus interaction with such host factors (Brass et al, 2008 and Lee et al, 2010).

1.6.5 Integration

Once in the nucleus, the next and second distinguishing feature of the HIV-1 replication cycle is the integration of a copy of the viral cDNA into a cellular chromosome. The integrated viral cDNA, termed provirus, serves as the template for the synthesis of viral RNAs, which may either be translated into viral proteins or act as genomic RNA in progeny virions. In addition, the proviral DNA is replicated along with cellular DNA during cycles of cell division and is maintained as part of the host genome for the lifetime of the infected cell. Integration of the viral cDNA into a cellular chromosome is catalyzed by the integrase and proceeds in a series of coordinated events. Firstly, the integrase catalyses the removal of two nucleotides from each 3' terminus of the linear viral cDNA leading to a pre-integration substrate with 3'-recessed ends that always terminate with the conserved CA-3´ sequence. In the next step, the integrase catalyses a strand transfer reaction where the cellular DNA is cleaved at the integration site and the newly-generated 3' ends of the viral cDNA are covalently bound to the 5´ends of the cellular DNA. Finally, proviral formation is completed by cellular enzymes that remove two unpaired bases at the 5´ends of the viral cDNA, fill in the single-strand gaps between viral and cellular DNA and ligate the 5´ends of viral DNA to

the cellular DNA (Delelis et al, 2008; Jegede et al, 2008). Although viral DNA can integrate into any sequence in the host, it has been shown to preferentially target actively transcribed regions in order to promote efficient gene expression (Marshall et al, 2007). Cell proteins, primarily the transcription activator factor LEDGF/p75, appear to boost the efficiency of integration and mediate targeting to active transcription units. LEDGF/p75 contains a chromatin-binding region, which comprises a PWWP domain at the amino terminus and a pair A/T hook domains, and a carboxy-terminal domain that bounds tightly to IN. This protein binds simultaneously to the integrase and chromatin at active transcription units, thus directing integration to these locations.

1.6.6 Gene expression

Control of gene expression of the HIV-1 provirus is exerted by both *cis* and *trans*-acting viral elements, which orchestrate complex interactions with the cellular transcription, splicing and RNA export and translation apparatus.

Transcription is positively regulated both at initiation and elongation stages. The U3 region located within the 5'-LTR contains two important elements that stimulate transcription initiation: an extremely efficient promoter, which is capable of supporting even higher levels of transcription than the adenovirus major late promoter or the CMV immediate early promoter, and an enhancer that contains two NF-κB binding motifs, which facilitate transcription initiation by removing chromatin restrictions near the promoter through the recruitment of histone acetyltransferases. The Tat protein up-regulates transcription elongation by binding simultaneously to a transactivating-responsive region (TAR) present in the 5' end of nascent viral RNA and to the positive transcriptional elongation factor (pTEFb)

components CDK9 kinase and cyclin T1. These interactions trigger a complex set of phosphorylation reactions that stimulate transcription elongation by inactivating negative elongation factors, such as NELF and DSIF, and by enhancing RNA polymerase II processivity (Dingwall et al, 1990).

Depending on the degree of splicing, viral transcripts produced during the replication cycle are classified into three categories: completely spliced mRNAs that encode the viral proteins Tat, Rev and Nef, incompletely spliced mRNAs that code for Env, Vif, Vpr and Vpu and full-length unspliced transcripts, which act both as the virion genomic RNA and the mRNA for the Gag/Gag-Pro-Pol polyproteins. The efficiency of splicing is regulated by the intrinsic strength of 5' and 3' splice sites and the presence of *cis*-acting elements, such as splicing enhancers and silencers. Immediately after infection, the synthesis of completely spliced mRNAs is predominant. These mRNAs are exported to the cytoplasm via a constitutive endogenous pathway used by cellular mRNAs and translated to yield Tat, Rev and Nef proteins. The Rev protein dictates the fate of incompletely spliced and unspliced viral RNA transcripts. When Rev levels are below a specific threshold these RNAs are either spliced or degraded in the nucleus. However, when Rev levels exceed the threshold, these intron-containing viral RNAs are exported to the cytoplasm due to the interaction of Rev with an elongated stem-loop structure of 351 nt, termed the Rev-responsive element (RRE), present on these viral transcripts. The binding of Rev to the RRE induces its own polymerization and triggers the interaction with a protein of the nuclear pore complex, the karyopherin family member Crm1, which in a GTP-dependent process translocates the viral RNA species into the cytoplasm. As the infection proceeds, incompletely spliced mRNAs and full-length unspliced transcripts are increasingly exported to the cytoplasm (Groom et al, 2009).

Translation is regulated by frameshifting, in which specific *cis*-acting sequences in the RNA cause a reading frame change during translation. For example, the production of Gag-Pro-Pol precursor during the translation of full-length unspliced transcripts occurs as a consequence of a -1 shift in the Gag translational reading frame. This frameshift occurs 5% of the time and results in the production of one Gag-Pro-Pol precursor for every 20 Gag precursors synthesized.

1.6.7 Assembly, budding and maturation

During assembly, viral and cellular components are packaged at nucleation sites leading to the formation of immature virions. Gag polyprotein is largely responsible for viral assembly and its expression is sufficient for the formation of non-infectious spherical viral-like particles. Gag, which is translated from full-length unspliced RNA transcripts in cytoplasmic polysomes, is composed by folded domains separated by flexible linker regions that display HIV-1 protease cleavage sites (CSs) (Figure 1.5).

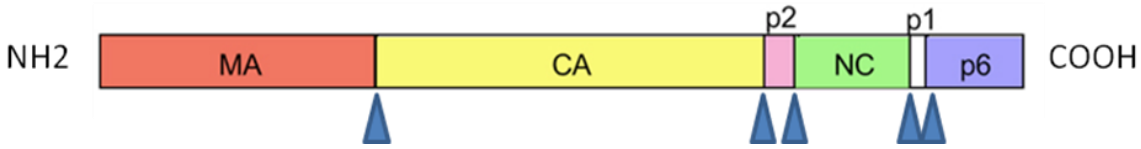


Figure 1.5 Gag polyprotein domains and protease cleavage sites.

The majority of Gag traffics in the cytoplasm as soluble monomers or dimers in an auto-inhibited conformation. However, a minor proportion is found complexed with dimers of genomic RNA forming ribonucleoprotein complexes (RNPs). RNPs genesis requires dimerization of genomic RNAs, which occurs in the cytoplasm by virtue of the non-covalent

binding at the dimer initiation sites located within the 5'UTR, followed by the interaction of Gag with genomic RNA dimers, where two retroviral zinc motifs within the P7 domain of Gag recognize highly-organized structures in the 5'UTR. In particular, the ψ sequence, which is located in the 5' LTR region spanning the major splice donor and the Gag initiation codon, plays a central role in this process and is required for efficient genome packaging (Berkowitz et al, 1996).

Nucleation sites are plasma membrane lipid rafts that are enriched in sphingomyelin, cholesterol and plasmalogen-PE and display an increase in saturated fatty acids compared with the cell plasma membrane. Targeting of Gag and RNPs to nucleation sites is mediated by a multipartite membrane-binding signal located within the P17 domain, which consists of a myristic acid covalently attached to the N-terminal Gly in the P17 protein and a patch of basic residues. Binding of the P17 domain to the inner leaflet lipid phosphatidyl inositol (4, 5) biphosphate exposes the myristoyl group stably anchoring Gag/RNPs to the plasma membrane. In addition, electrostatic interactions between the stretch of basic residues and acidic phospholipids on the inner leaflet of the lipid rafts reinforce membrane docking (Hermida-Matsumoto and Resh, 1999; Zhou et al, 1994). The initial presence of RNPs at nucleation sites is important for effective virion assembly. The interaction of newly-arrived Gag monomers or dimers with RNPs through their CA domains converts the auto-inhibited conformation into one that is optimal for Gag polymerization. Multimerization of Gag through its P24-P2 region leads to the formation of an immature lattice, in which membrane-bound Gag molecules are extended and oriented radially with the carboxy-terminal end facing the interior of the particle. During viral assembly, apart from Gag and ribonucleoprotein complexes, other viral components are also incorporated into the immature viral particle. The Gag-Pro-Pol polyprotein, which is translated in the cytosolic polysomes, is

incorporated via the same mechanism as Gag/RNPs. The Env protein is translated in ER-associated ribosomes and then travels through the constitutive cellular secretory pathway where it is glycosylated, olygomerized into trimeric complexes, cleaved by furin to form transmembrane (gp41) and surface (gp120) subunits and transported to the plasma membrane via vesicular transport. The long intracellular tail of gp41 helps to sort the protein into nucleation sites and interacts with the P17 domain of Gag to promote Env virion incorporation. Other viral proteins, such as Vpr, Vif or Nef are incorporated into the immature viral particle by virtue of its interaction with the carboy-terminal P6 domain of Gag. In addition, a number of cellular components have been shown to be packaged into the viral particles. The host $tRNA^{Lys3}$ is recognized by an 18 base-pair sequence (PBS) located within the 5' LTR. In addition, host proteins, such as ICAM-1, HLA-II, actin, cyclophilin A, are incorporated into the virion either passively or by interaction with Gag. The involvement of these human proteins in viral biology is not well established in the majority of cases.

During budding the immature virion is released from the plasma membrane. The process is largely mediated through interactions between P6 protein and the cellular ESCRT (endosomal sorting complexes required for transport) machinery. The ESCRT apparatus usually catalyzes membrane fission reactions that release vesicles into endosomal multivesicular bodies.

The carboxy-terminus of Gag (P6) contains two short sequence motifs (i.e., late assembly domains) that recruit and bind to early-acting ESCRT factors (Katzman et al, 2002). The highly conserved PTAP motif located near the amino terminus of P6 (Huang et al, 1995) binds to the TSG101 component of the ESCRT-I complex, whereas the YPXL motif located downstream of the PTAP motif interacts with the ESCRT-III binding partner ALIX (Usami,

et al 2009; Roxrud et al, 2010; Bieniasz et al, 2009; Strack et al, 2003 and VerPlank et al, 2001). These interactions result in the recruitment of the ESCRT-III proteins of the CHMP 1, 2 and 4 families, which promote closure of the virion neck, and VPS4 ATPases, which complete the fission reaction by hydrolyzing ATP. As a result of these actions the immature viral particle is released from the plasma membrane.

During maturation, the immature virion experiences dramatic morphological changes and becomes infectious. Maturation begins concomitant with or immediately after budding and is driven by the HIV-1 protease cleavage of Gag and Gag-Pro-Pol polyproteins (Kaplan et al, 1993). Construction of mutant virus lacking the viral PR clearly indicated that the proteolytic processing of these polyproteins is required for restructuring of the virion in a fully mature and infectious form (Kohl et al, 1988; Peng et al, 1989).

HIV-1 PR is an aspartyl protease with extensive sequence homology to cellular counterparts, such as pepsin and renin. Like other aspartic acid proteases, HIV-1 PR uses two aspartic acid side chains within a characteristic Asp-Thr-Gly motif to activate a water molecule that catalyses the hydrolysis of the peptide bond. In contrast to cellular proteases, the holoenzyme is a dimer of two identical subunits, each containing 99 amino acid residues (Tozser et al, 2003). As revealed by crystallographic studies, the active site is located in the interior of a long cleft present at the dimer interface and is stabilized by non-covalent interactions between Asp-Thr-Gly motifs and four-stranded mixed β sheets created by the amino and carboxy termini of each subunit. The active site contains two catalytic aspartic acid residues (Asp-25 and Asp-25'), each contributed by a different subunit (Oroszlan et al, 1990; Wlodawer et al, 1993). Substrate access to the active site is regulated by two flexible flaps

that undergo dramatic movements to allow entry of substrates and exit of products. The structure of the HIV-1 PR is shown in figure 1.6.

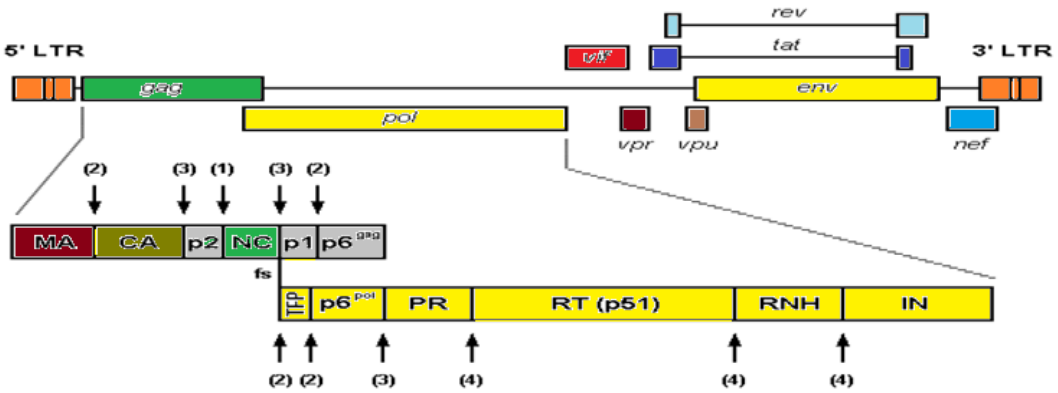
HIV-1 PR binds substrates asymmetrically in an extended anti-parallel β-sheet conformation. Three to four amino acids (P4 to P1 and P1' to P3') located on either side of the cleavable peptide bond (P1↓P1') participate in binding to the substrate cavity of the protease. Substrate specificity studies (Pettit et al, 1991; Poorman et al, 1991; Billich et al, 1988; Griffiths, et al, 1992; Tomasseli et al, 1994; Tozser et al, 1992; Loeb et al, 1989) have reached a series of conclusions: P1 and P1' are usually large hydrophobic residues, P1 never contains a β-branched aliphatic side chain, P2 and P2' are typically hydrophobic or small polar residues, and P4, P3 and P3' can accommodate a variety of residues. However, as Gag and Gag-Pro-Pol cleavage sites vary considerably in amino acid sequence, it has been postulated that the enzyme appears to recognize the shape of the substrate rather than its specific sequence (Prabu-Jeyabalan et al, 2000; Prabu-Jeyabalan et al, 2002).

The Gag-Pro-Pol polyprotein contains three inactive viral enzymes: protease, reverse transcriptase and integrase. The molecular mechanisms that lead to activation and regulation of HIV-1 protease are unclear, but these processes must be tightly regulated, so that the Gag polyprotein is not cleaved before assembly and budding. Experimental findings have suggested that dimerization of Gag-Pro-Pol polyproteins during assembly at nucleation sites is required for initial activation (Franke et al, 1994; Gamble et al, 1997; Gatlin et al, 1998) and an aggregation model has been proposed to explain how this may be regulated. However,

other factors such as drop in pH or an influx of Ca^{2+} ions associated with virion release, may contribute to activation (Skalka, 1989; Vogt, 1996). The partially active HIV-1 PR within the Gag-Pro-Pol dimer appears to be responsible for the initial autocatalysis of Gag-Pro-Pol through intra-molecular mechanisms (Pettit et al, 2004; Tessmer and Krausslich, 1998). Primary processing occurs at the P2↓P7 junction (Pettit et al, 2004) and is followed by secondary and tertiary processing of the transframe region (TFR or P6) at TFR↓$P6^{pol}$ and $P6^{pol}$↓PR, respectively (Phylip et al, 1995). Based on these studies, these initial processing events give rise to p121, p114 and p107 early intermediates as a consequence of cleavage at P2↓P7, TFP↓$P6^{pol}$ and $P6^{pol}$↓PR CSs, respectively. Since the TFR negatively regulates PR function, its removal is concomitant with the appearance of an elevated enzymatic activity characteristic of a mature PR (Louis et al, 1991). Subsequent cleavage events occur by inter-molecular mechanisms and liberate the structural and functional enzymes of the Gag and Pol regions, respectively (Wondrak et al, 1996). There is little information available concerning the sequence of events in processing CSs in HIV-1 pol. Studies support that RT↓IN is processed first and is followed by PR↓RT and the by RTP51↓RTP66 (Tozser et al, 1991). However, a simultaneous processing of all pol CSs resulting in the concurrent release of PR, IN and both RT subunits (P66 and P51) cannot be excluded. As a consequence of the entire autocatalytic process, three fully active protease, reverse transcriptase and integrase are produced (Craven et al, 1991; Swanstrom and Wills, 1997; Xiang et al, 1997; Gross et al, 2000).

The Gag polyprotein contains four structural proteins, matrix (P17), capsid (P24), nucleocapsid (P7) and P6, and two smaller spacer peptides P2 and P1 located between P24/P7 and P7/P6, respectively (Orozslan et al, 1990). Gag is cleaved by the HIV-1 protease

at five different sites (CS). The order of cleavage is highly conserved and is mainly regulated by the intrinsic susceptibility of each site to proteolysis (Tozser et al, 1997). These cleavage sites are classified into three groups according to their rate of processing: rapid (P2 / P7), intermediate (P1 / P6, P17 / P24) and slow (P7 / P1, P24 / P2) (Pettit et al, 1994). Cleavage at each site appears to occur independently (Pettit et al, 1994) and appears to perform a different function. Primary processing occurs at P2 / P7 and gives rise to P43 (P17-P24-P2) and P14 (P7-P1-P6) intermediates. This cleavage is conducive to the activation of the Env protein. Secondary processing involves cleavage at P1 / P6 shortly followed by cleavage at P17 / P24 and leads to P17, P6, and P8 (P7-P1) and P25 (P24-P2) intermediates. Finally, tertiary cleavage at P7 / P1 followed by P24 / P2 gives rise to P7, P1, P24 and P2 final products. These cleavages lead to the condensation and stabilization of the dimeric RNA genome and the assembly of the conical capsid. The dramatic rearrangement of the internal virion components triggered by Gag proteolytic processing is essential for the production of viable infectious virus particles (Kaplan et al, 1993, Pettit et al, 1994, Swanstrom and Wills, 1997; Xiang et al, 1997 and Vogt, 1996).



Site	Amino acid Sequence
Gag	
MA-P17/CA-P24	VSQNY↓PIVQN
CA-P24/P2	KARVL↓AEAMS
P2/NC-P7	SATIM↓MQKGN
NC-P7/P1	ERQAN↓FLGKI
P1/P6	RPGNF↓LQRSP
Pol	
TF/PR	TSFSF↓PQITC
PR/RT	CTLNF↓PISPI
RT (P51/P66)	GAETF↓YVDGA
RT/IN	IRKVL↓FLDGI

Figure 1.7 Schematic representation of the genomic and polyprotein organization of HIV-1.

The upper diagram represents the 9.8 kb provirus genome. The lower diagram represents the HIV-1 Gag and Gag-Pol polyprotein precursors. The proteins encoded in *pol* are synthesized at a frequency of 5-10% by a -1 translational frameshift of the unspliced genomic mRNA template to yield Gag-Pol. At the stage of virion budding and release, Gag and Gag-Pol polyproteins are proteolytically cleaved at domain boundaries by the viral protease to release their constitutive protein species and complete the maturation process. The location of these protease cleavage sites are indicated by the vertical arrows. Primary (1), secondary (2), tertiary (3), and quaternary (4) cleavage events are numbered accordingly. Gag polyproteins are processed into matrix (MA, p17), capsid (CA, p24), nucleocapsid (NC, p7), $p6^{gag}$ and two spacer proteins p2 and p1. Processing of Gag-Pol polyproteins additionally yields the transframe region proteins (TFP and $p6^{pol}$), protease (PR, p10), reverse transcriptase (RT, p66/51) which contains an RNase H (RNH) domain in its larger subunit, and integrase (IN, p32). The HXB2 amino acid sequence for each cleavage site in HIV-1 Gag and Gag-pol precursor is indicated in the table below, the scissile amide bond is indicated by the vertical arrow (↓).

1.7 Antiretroviral drugs and mechanisms of resistance

Currently, there are 24 antiretrovirals (ARVs) approved by the US Food and Drug Administration (FDA) for the treatment of HIV-1 infection (Table 1.1). They are classified into six classes: nucleo(s)tide reverse transcriptase inhibitors (NRTIs), non-nucleoside

reverse transcriptase inhibitors (NNRTIs), protease inhibitors (PIs), fusion inhibitors (FIs), integrase inhibitors (INIs) and co-receptor antagonists (CCR5 antagonists). All, except for the CCR5 antagonists, target essential viral proteins-reverse transcriptase (NRTIs and NNRTIs); protease (PIs), transmembrane protein gp41 (FIs) and integrase (INIs). The CCR5 antagonist, by contrast, is unique among the ARVs in that it targets the host cell chemokine co-receptor CCR5 rather than a viral protein. All ARVs act by inhibiting some of the steps of the viral replication: FIs and CCR5 antagonists inhibit viral entry; NRTIs and NNRTIs inhibit reverse transcription; PIs inhibit the proteolytic processing of Gag and Gag-Pro-Pol polyproteins and INIs inhibit the integration of the viral DNA into the host cell chromosome.

A common feature derived from the use of ARVs, regardless of its class, is the emergence and selection of resistant HIV-1 variants. As a general rule, resistance is conferred by punctual and well-characterized mutations in the target gene. In general, the main drivers of the development of drug resistance are the high level of virus production reaching up to 10^9-10^{12} virus particles per day in untreated patients (Perelson et al, 1996) and the high error rate during reverse transcription, the RT introduces an average of one or two mutations for each viral genome transcribed (Bebenek et al, 1989; Ji and Loeb, 1992). These two characteristics combined ensure that patients have a complex and diverse mixture of viral strains termed "quasispecies", each differing by one or two mutations. Within the quasispecies certain strains dominate. This represents an equilibrium between escape from selective pressure, such as immune system or drug therapy, and preserved ability to replicate and infect. Consequently, variants with reduced susceptibility to ARVs are usually found in the viral population before treatment, but are present at low frequency due to impaired fitness when compared to the wild-type strain. If any of these variants confers a selective advantage to the

virus, such as decreased drug susceptibility, it will become dominant. Emergence of resistance will therefore only occur in patients who have ongoing viral replication in the presence of levels of ARVs that are insufficient to completely abolish viral replication but sufficient to exert a positive selective pressure on variants with decreased susceptibility. The main scenario conducive to emergence of drug resistance is in patients with suboptimal adherence to treatment. However, others circumstances, such as drug interactions, may also lead to sub-therapeutical drug levels and as a result to the development of drug resistance. Many studies have demonstrated that toxicity and side effects are closely associated with adherence to antiretroviral therapy. For instance, in the ICONA study group, patients receiving indinavir-and-ritonavir containing HAART had a higher chance of discontinuing therapy because of toxicity (21%) compared to treatment failure (10%) (D'arminio Monforte, et al, 2000). The development of new, more potent and safer antiretroviral therapies has reduced HAART-related toxicity. However, unfortunately no drug is entirely devoid of secondary effects. Of the NRTIs, tenofovir can cause renal toxic effects and potentially osteopaenia. Abacavir can cause an increased risk of cardiovascular disease and is also associated with a serious hypersensitivity reaction, particularly in HLA-B5701 positive patients. Many of the once popular thymidine analogues, especially stavudine, are currently known to cause profound long-term and probably irreversible side-effects, such as lipoatrophy. Among the NNRTIs, efavirenz is reported to be teratogenic and has also substantial short-term CNS toxic effects. Neviripine can cause severe liver disease and hypersensitivity reactions, particularly in patients starting therapy with high CD4 T-cell counts. Regarding integrase inhibitors, raltegravir appears to be safe, well tolerated and highly effective, but long-term safety data are still missing. Concerning protease inhibitors, most of them increase plasma lipid concentrations and consequently increase cardiovascular risk; in addition, many have clinically relevant drug interactions, usually when boosted with

low dose ritonavir. Furthermore, most PIs are associated with gastrointestinal disturbances. Atazanavir also increases the plasma concentration of unconjugated bilirubin, occasionally causing reversible jaundice. The CCR5 inhibitor maraviroc is generally well tolerated and has no known short or long-term side effects. However, by contrast with all other antiretrovirals, CCR5 inhibitors bind a host target and consequently there are concerns about its long term risk. Lastly, the fusion inhibitor Enfuvirtide is poorly tolerated because of the need for injection twice daily, which is often associated with local pain.

Strict adherence is required to achieve and maintain viral suppression. Suboptimal drug exposure can result in the rapid development of drug resistance. This is especially true for drugs with low genetic barrier to resistance, which are commonly part of some of the most popular first-line regimens, and include NRTIs (e.g., lamivudine or emtricitabine), NNRTIs (e.g., efavirenz and nevirapine) and INIs (e.g., raltegravir). By contrast, PIs have a higher genetic barrier to the development of resistance and consequently long-term exposure is generally required before resistance emerges. In general, it is believed that PIs are more forgiving in terms of non-adherence. The management of HAART side effects is of pivotal importance in the management of HIV-infection. Most treatment modifications of first line HAART are related to toxic effects (Elzi et al, 2010). Similarly, it has been reported that nearly all first regimen virological failure can be attributed to either non-adherence or pre-existing drug resistance (Paredes et al, 2010). Consequently, virological failure should trigger a thorough review of potential causes for non-adherence including drug side-effects, search for other drugs that can be affecting absorption or metabolism and because resistance selection may have occurred, a resistance genotype should also be obtained.

Resistance to ARVs in an individual is not only the result of failure to therapy, but can be due to interpersonal transmission of resistant strains. Importantly, due to the integration of the HIV-1 viral genome into the host cell chromosome during replication, all major quasispecies that have ever been generated within a patient will be archived and replication will favour the form that is fittest under current conditions. However, if conditions change, previously archived variants can rapidly re-emerge.

The mechanisms of resistance differ for ARVs as it does their genetic barrier to resistance defined by the ease of emergence of resistance, which is a function of the number of mutations required to abrogate drug activity. The highest barrier to resistance is observed for ritonavir boosted PIs (PI/r), as the phenotypic impact of individual mutations is generally low. By contrast, the lowest genetic barrier to resistance is documented for first generation NNRTIs, such as efavirenz (EFV) and nevirapine (NVP), as a single mutation is sufficient to confer complete resistance to the drugs.

Below, we describe the mechanism of action and resistance of each class of ARVs:

1.7.1 Fusion inhibitors

Emfuvirtide (T-20) is the only fusion inhibitor currently available in the market. As previously discussed, the entry of the HIV-1 virus into the target cell is a multi-step process that involves attachment, co-receptor binding and fusion of the viral envelope and the cell

membrane. T-20 binds the heptad repeat region 1 (HR1) region in the HIV-1 trans-membrane protein gp41 preventing the formation of the six helical bundle and consequently blocking the fusion between the lipid bilayer of the virus and that of the host cell.

Resistance to T-20 is the consequence of mutations in the HR1 region. The first three substitutions associated with resistance were described in amino acids 36-38 (Derdeyn et al. 2000, 2001 and Rimsky et al, 1998). It was demonstrated that amino acid position 36 played a relevant role in the fusogenic activity of HIV-1 envelope and certain mutations at this residue were associated with increased fusion kinetics, leading to resistance to the drug. On completion of clinical trials in HIV-1 infected patients, the region conferring T-20 resistance was expanded to include amino acid positions 36-45 (Lu et al, 2004; Wei et al, 2002).

1.7.2 Co-receptor antagonists

Maraviroc is so far the sole co-receptor antagonist licensed for the treatment of HIV-1 infection (Macarthur and Novak, 2008). It is a noncompetitive allosteric antagonist of CCR5 (Dorr et al, 2005; Watson et al, 2005) that binds the CCR5 co-receptor and prompts a conformational change that ultimately prevents the interaction between CCR5 and the V3 crown of the surface glycoprotein gp120 (Watson et al, 2005; Dragic et al, 2000) and consequently the entry of the virus into the target cell. Maraviroc binds the CCR5 co-receptor but not the closely related CCR2 chemokine or CXCR4 receptors (Dorr et al, 2005). As a result, maraviroc selectively inhibits the entry of R5-tropic HIV-1 strains into cells, but not that of X4-tropic viruses. Since maraviroc dose not compete with the binding of chemokines to CCR5 it is not expected to affect cellular signaling via CCR5 (Dorr et al, 2005).

Different pathways of resistance to maraviroc have been reported. A first mechanism of resistance involves the use of the co-receptor CXCR4 for entry into the host cell. This requires a change in HIV-1 tropism, which can be acquired by two possible mechanisms: the first is through the *de novo* acquisition of mutations in the *env* gene, which allow the use of the CXCR4 co-receptor (truly co-receptor switch), and the second involves the outgrowth of a pre-existing population of CXCR4-using variants. A number of studies have suggested that the outgrowth of CXCR4-using variants is the most common mechanism of resistance (Westby et al, 2006; Kuhmann and Moore, 2005). Another potential mechanism of resistance is the emergence of mutations in the gp120 envelope protein, primarily in the V3 loop, which will increase the affinity of the protein for the inhibitor-free co-receptor, consequently favouring the binding of the virus *versus*. the antagonist. In addition, mutations in the V3 loop may enable the virus to bind the inhibitor-bound receptor (Westby M, 2007a). Given the extensive variability of the HIV-1 envelope, the patterns of mutations emerging as a consequence of maraviroc selective pressure may considerably differ among patients and as a result the genotypic predictors of this mechanism of resistance have yet to be clearly identified (Westby et al, 2007b).

1.7.3 Reverse transcriptase inhibitors

Inhibition of the reverse transcriptase constitutes the cornerstone of most antiretroviral regimens. Two different classes of drugs are grouped within the reverse transcriptase inhibitors: nucleos(t)ide reverse transcriptase inhibitors (NRTIs) and nonnucleos(t)ide reverse

transcriptase inhibitors (NNRTIs). Although, both classes target the reverse transcription step by inhibiting the activity of the viral reverse transcriptase enzyme, the mechanism of action differed between the two classes.

1.7.3.1 NRTIs

NRTIs are competitive inhibitors of the DNA polymerase activity of the RT. Structurally, NRTIs are nucleos(t)ide analogues that lack the 3'-hydroxy group, which once incorporated into the growing DNA chain caused premature chain termination (Parker et al, 1991, El et al; 2007; Zdanowicz, 2006). There are currently eight NRTIs, comprising seven nucleosides analogues: zidovudine (ZDV), stavudine (d4T), zalcitabine (ddC), didadosine (ddI), lamivudine (3TC), abacavir (ABC) and emtricitabine (FTC) and one nucleotide analogue: tenofovir (TDF). ZDV was in fact the first antiretroviral approved for the treatment of HIV-1 infection in 1987 and since then NRTIs has formed the backbone of antiretroviral therapy.

NRTIs are pro-drugs that require phosphorylation (bi-phosphorylation for nucleotide analogues and tri-phosphorylation for nucleoside analogues). Tissue-dependant cellular kinase activity determines the levels of drug effectiveness (Gao et al, 1993). In addition, NRTIs that rely on the same phosphorylation pathway, as it is the case for ZDV and d4T, can show antagonism when administered in combination (Havlir et al, 2000).

Resistance to NRTIs is the result of the emergence of amino acids changes in the RT enzyme. These changes could be sequential additions (e.g; for ZDV resistance), insertions, or single

amino acid substitutions (e.g; for 3TC resistance). RT mutations confer resistance to NRTIs by two possible mechanisms: the first, which is termed NRTI excision, involves and ATP-dependent pyrophosphorolysis that leads to the selective removal of the NRTIs from the 3'-end of the nascent DNA chain and as a result to reversal of the chain termination (Arion et al, 1998; Boyer et al, 2001; Meyer et al, 1999). Mutations that cause resistance by this mechanism are the thymidine analogue mutations (TAMs), which include a group of amino acids changes selected by ZDV and d4T, such as M41L, D67N, K70R, T215Y/F and K219E/Q (Bacheler et al, 2001; Boucher et al, 1992; Harrigan et al, 1996; Kellam et al, 1992). Emergence of TAMs by NRTI-selective pressure occurs by two different pathways: TAM-1, which includes mutations M41L, L210W, T215Y and occasionally D67N, and TAM-2, which includes D67N, K70R, T215F and K219Q/E. The TAM-1 pathway is associated with greater level of ZDV resistance and NRTI cross resistance than the TAM-2 pathway (Marcelin et al, 2005; Miller, 2004; Hu et al, 2006). TAMs emerge in sequential order and their accumulation over time leads to increasing levels of resistance, mainly to ZDV and d4T but also ABC, ddI and TDF.

The second mechanism of NRTI resistance involves the prevention of NRTI incorporation into the nascent DNA chain. Mutations associated with this mechanism of resistance are M184V/I, K65R, K70E, L74V and Q151M. The M184V/I mutations are selected by 3TC or FTC-containing regimens and associated with high level resistance to both drugs. The mutation is located close to the RT active site and causes steric hindrance that hampers the incorporation of the NRTI (Sarafianos et al, 1999). The K65R mutation is selected by TDF, ABC and ddI and decreases susceptibility to all NRTIs except for ZDV. The mutation favours the incorporation of the natural dNTP substrate over the drug (Deval et al, 2004;

White et al, 2006). K70E mutation is primarily associated with resistance to TDF by prompting a decrease in the incorporation of the inhibitor (Sluis-Cremer et al, 2007). The L74V mutation is selected by ddI and ABC conferring resistance to both drugs; the mutation alters the rate of incorporation of ddI and also favours the incorporation of natural dNTP over ABC (Deval et al, 2004; Winters et al, 1997). Finally, the Q151M mutation is part of multi-drug resistance mutation complex (MDR) along with amino acid changes, such as F116Y, F77L, V75I and A62V. These mutations are typically selected by drug combinations including ZDV and ddI and confer resistance to all NRTIs, albeit less so to 3TC and TDF (Sluis-Cremet et al, 2000). The Q151M mutation interacts with the nitrogen base of the dNTP resulting in the altered recognition and reduced incorporation of the NRTI (Sluis-Cremer et al, 2000).

In addition to the two classic mechanisms of resistance to NRTIs described above, a growing body of evidence has emerged indicating a role in resistance to this drug class for mutations in the connection and the RNase H domains. These mutations have been demonstrated to increase resistance to ZDV by altering the balance between NRTI excision and RNase H activity. Specifically, the mutations reduce the RNase H activity of the RT enzyme allowing more time for the enzyme to excise ZDV from the terminated DNA chain. Mutations in the connection domain conferring resistance to ZDV are E312Q, G335C/D, N348I, A360I/V, A371V, V365I, A376S (Nikolenko et al, 2007; viks-Frankenberry et al, 2007; Viks-Frankenberry et al, 2008 and Yap et al, 2007) and mutations located in the RNase H domain that also contribute to ZDV resistance include H539N, D549N and Q509L (Brehm et al, 2007; Brehm et al, 2008 and Nikolenko et al, 2005).

1.7.3.2 NNRTIs

Currently, there are five NNRTIs licensed for the treatment of HIV-1 infection. They are classified as first generation, which include delavirdine (DLV), efavirenz (EFV) and nevirapine (NVP) and second generation, which comprise etravirine (ETV) and rilpivirine (RPV). They are allosteric noncompetitive inhibitors of the RT.

All NNRTIs exert their action by binding to the HIV-1 RT in a hydrophobic pocket termed the non-nucleoside inhibitor binding pocket with a common butterfly-like binding mode (Kohlstaedt et al, 1992; Ding et al, 1995). The non-nucleoside binding pocket, which exists only in the presence of the NNRTI and is not opened in the unliganded enzyme, consists of hydrophobic residues Y181, Y188, F227, W229 and Y232 and hydrophylic residues, such as K101, K103, S105, D192 and E224 in the p66 subunit of the RT and E138 in the P51 subunit of the enzyme. The NNRTI-binding pocket is close but distinct from the active site and the dNTP binding site of the enzyme and as a consequence it does not prevent the binding of either the dNTP or the RNA template to the RT, but it brings a conformational change that impairs its catalytic activity. The main difference between first and second generation NNRTIs lies in that the former are rather inflexible and their binding to the RT is severely impaired as a consequence of key mutations in the enzyme. By contrast, second generation NNRTIs are much more flexible and consequently can rapidly adapt to changes in the drug binding pocket (Andries 2004; Rodriguez-Barrios et al, 2005).

High level resistance to NNRTIs generally results from the rapid acquisition of single amino acid substitutions located directly in the NNRTI-binding pocket (Tantillo et al, 1994) and as a

result hamper the binding of the inhibitor. Mutations that confer NNRTI resistance cluster in two regions of the HIV-1 RT enzyme: regions adjacent to codons 180-188 and 100-110 (Nunberg et al, 1991; Larder, 1992 and Larder, 1995). Mutations associated with high level resistance to first generation NNRTIs (NVP, EFV and DLV) are K103N, Y181C, G190A/S, Y188L and V106A/M (Wainberg et al, 2003). Other mutations, such as L100I, K101E/P, A98G, V108I, V179D/E, P225H, M230L and K238T/N, cause low level resistance to first generation NNRTIs and they usually occur in combination with the major NNRTI-resistance-associated mutations described above and act synergistically to reduce NVP, EFV and DLV susceptibility (Pelemans et al, 1998; Bacheler et al, 2001; Rhee et al, 2006). In general, resistance mutations are shared between the three first generation NNRTIs, but exceptions are G190A and Y181C, which are frequently selected by NVP-containing regimens but rarely emerge under EFV-selective pressure. By contrast, K103N is by far the most frequently selected mutation by EFV. While K103N and G190A cause high level resistance to all the three agents, Y181C causes high level resistance to NVP and DLV, but only intermediate resistance to EFV (Casado et al, 2000). However, results of treatment with EFV in patients who developed the mutation after NVP-failure has been disappointing (Lecossier et al, 2005) and consequently patients who develop resistance after treatment with any of the three first generations NNRTIs cannot be successfully treated with another of these agents (Delaugerre et al, 2001).

Second generation NNRTIs (ETV and RPV) have a higher genetic barrier to resistance than first generation NNRTIs as accumulation of several mutations is required before significant reduction in susceptibility to the drugs is observed (Seminari et al, 2008). This finding can be explained by the fact that these drugs can bind the RT enzyme in different conformations and

as a result can rapidly adapt to changes in the NNRTI-binding pocket created as a consequence of specific resistance mutations. In general, it is considered that complete resistance to ETV requires accumulation of several of the following mutations V90I, A98G, L100I, K101E/P, V106I, E138A/G/K/Q, V179D/F/T, Y181C/I/V and G190A/SA and M230L. However, studies have shown that mutations at codons 100, 101 and 181 have a greater impact on clinical response to ETV compared to other mutations (Haddad et al, 2010). A total of 15 mutations have been associated with reduced susceptibility to RPV including K101E/P, E138A/G/K/Q/R, V179L, Y181C/I/V, H221Y, F227C and M230L, of them E138K is the most frequently found in patients failing RPV therapy (Rimsky et al, 2012; Haddad et al, 2011).

1.7.4 Protease inhibitors

There are currently nine protease inhibitors (PIs) licensed for the treatment of HIV-1 infection: saquinavir (SQV), ritonavir (RTV), indinavir (IDV), nelfinavir (NFV), amprenavir (APV), lopinavir (LPV), atazanavir (ATV), tipranavir (TPV) and darunavir (DRV). All of them are competitive inhibitors that bind to the active site of the viral PR with high affinity and by doing so they prevent the binding of its natural substrate Gag. As a consequence, PIs inhibit the catalytic processing of the Gag and Gag-Pro-Pol polyprotein into their mature components rendering the released virus immature and non-infectious. (Flexner, 1998; Patick and Potts, 1998).

PIs are designed as analogues of the cleavage sites found in the natural substrate Gag of the HIV-1 PR in which the scissile bond has been replaced by a non-cleavable, transition-state motif. All of them, except for TPV, are peptidomimetic inhibitors and therefore they contain

a hydroxyethilene core, which shares structural similarity with the tetrahedral intermediate formed during the hydrolytic cleavage of a peptide bond of the natural substrate (Randolph and Degoy, 2004). By contrast TPV, contains a dihydropyrone ring as a central scaffold, which directly interacts with the flap region of the HIV-1 PR and is structurally unrelated to the tetrahedral intermediate (Chrusciel and Strobach 2004). This unique binding motif and structure of TPV has been reported to increase flexibility allowing the drug to adjust to amino acid changes in the active site of the viral PR (Larder et al, 2000; Turner et al, 1998).

Due to the vital role that HIV-1 PR plays in the viral life cycle and its small size, it was initially believed that resistance to PIs would be infrequent during treatment. However, the protease gene has shown great plasticity with mutations detected in 49 of the 99 amino acids of the HIV-1 protease and more than 20 substitutions associated with resistance to PIs (Shafer et al, 2000). Emergence of PI resistance requires the stepwise accumulation of primary mutations (major mutations) and secondary mutations (minor, compensatory or accessory mutations) (Molla et al, 1996), where each inhibitor usually selects for signature primary mutations and a characteristic pattern of secondary mutations. Thus, multiple substitutions are required for the development of complete PI resistance while maintaining effective virus replication and maturation. In general, primary resistance mutations are located near the active site of the PR at positions involved in inhibitor and substrate binding, some of these are: D30N, G48V, I50V, V82A, and I84V. Often, these mutations have a deleterious effect on the replication capacity of the resistant virus (Nijhuis et al, 2001; Quinones-Mateu, 2001) and such negative effect can be alleviated by the emergence of secondary mutations in the PR. These amino acid changes are generally outside the substrate-binding cavity of the enzyme and promote adaptation to the primary changes observed in the protease and

compensate at least partially the impairment of HIV-1 replication (Eastman et al, 1998; Ho et al, 1994; Mammano et al, 2000 and Nijhuis et al, 1999). Importantly, early studies demonstrated that in addition to mutations in the PR, changes located within the substrate (Gag) Cleavage sites are also selected in the context of PI resistance (Clavel et al, 2000; Doyon et al, 1996; Miller 2001; Nijhuis et al, 2001; Zhang et al, 1997), these mutations have been classified as compensatory mutations similarly to secondary mutations selected in the PR. Primary mutations directly affect the binding of the inhibitor to the protease and by doing this, they confer resistance to the drug. By contrast, the mechanisms in place for secondary mutations located both in the PR and in its natural substrate Gag is more difficult to elucidate. It has been propose that certain secondary PR mutations may alter the active site of the enzyme to adapt to the changes introduced by the active site primary mutations. Similarly, it has been proposed that mutations at cleavage sites provide better substrates for the mutated protease, which partially compensate for the loss of viral fitness displayed by the PI-resistant virus (Clavel et al, 2000; Doyon et al, 1996; Mammano et al, 2000; Nijhuis et al, 2001 and Zennou et al, 1998). However, other studies have shown that in certain instances Gag cleavage site mutations compensate for the replicative capacity of PI-resistant viruses without increasing the rate of cleavage compared to the wild type virus, suggesting that secondary mutations in Gag may exert their action through a variety of mechanism (Mammano et al, 2000). Regardless of the mechanism of action, a common characteristic of secondary mutations in the PR and Gag is that they do not confer significant levels of PI resistance on their own by they are required along with primary PI resistance in order to achieve high level resistance to the inhibitors.

The advent of the first PIs (SQV, RTV, IDV, and NFV) in the early 1990s was a landmark breakthrough in the treatment of HIV-1 infection. They made possible the dual class triple combination therapy that became known as highly active antiretroviral therapy (HAART). HAART not only reduced HIV-1 RNA plasma levels below the level of quantification in most patients (Gulick et al, 1997; Hammer et al, 1997) but also significantly slowed the progression of HIV disease compared with single or dual therapy (Hammer et al, 1997) and as a result it was established as standard of care in all HIV-1 infected patients. However, the clinical utility of early PIs was restricted by their low bioavailability and large pill burden which ultimately reduced adherence and limited long-term viral inhibition. Furthermore, failure to first generation PI-containing therapy often resulted in the development of high levels of PI resistance due to the accumulation of mutations in amino acids generally located in the PR active site, D30, G48, I50, V82, I84V, but also occasionally at non-active site residues such as M46 and L90M. Despite each PI selecting for a characteristic pattern of mutations, cross-resistance was common among first generation PIs as all of them occupy a similar space in the HIV-1 PR.

The next major development in the treatment of HIV-1 came when it was observed that RTV was a strong inhibitor of the CYP34A isoenzyme, which is the main responsible for the catabolism of PIs and consequently co-administration of most PIs with low dose RTV boost the exposure of the PI allowing flexible dosing including once daily dose. In addition, when boosted with RTV, PIs became more effective against PI-resistant viruses by increasing the drug plasma levels and consequently requiring higher levels of resistance to completely abrogate drug activity (Condra et al, 1996). Subsequently, novel PIs were developed which were specifically designed to be active against PI-resistant viruses. While first generation PIs

fixed exactly within the active site of the PR and were designed with large hydrophobic groups to maximize hydrophobic interactions with the HIV-1 PR, novel PIs were developed to fix in the substrate-binding region. In addition, smaller hydrophobic groups were introduced in novel PIs so that the affinity for the inhibitor was not so dramatically affected by mutations in the active site as well as introducing additional polar interaction with main chain atoms which cannot be easily altered by mutations and which provided the inhibitor with enough flexibility as to adapt to changes in the PR active site. The first PI conceived with this approach was APV, this drug was developed to fit predominantly within the substrate envelope (King et al, 2004) and its mutational profile was distinct to the first generation PIs (SQV, RTV and IDV) providing evidence that inhibitors with greater resemblance to the natural substrate will be less affected by mutations selected by first generation PIs. Following the same structural approach, more recently DRV was developed which provide additional interactions with the active site protein backbone which, by contrast with the rest of the PR, is extraordinarily conserved. By introducing additional interactions, DRV provides an impressive resistance profile necessitating up to six active site mutations in the PR before the activity of the drug is completely abolished.

It was soon thought that alternative mechanisms may assist the virus to evade the drug selective pressure. Indeed, it was demonstrated that viral strains resistant to DRV were selected in vitro. These resistant viruses exhibited changes in the Gag CS P7/P1 in the absence of any other changes in the HIV-1 genome, and when introduced in a reference strain, it was demonstrated that such variants conferred 10 fold resistance to the PI (Nijhuis et al, 2007). This was the first evidence that Gag CS mutations could cause resistance by themselves independently of their role as compensatory mutations. Since them, a few studies

have pointed to the role of Gag mutations, not only at CS but also beyond, in resistance to the drugs (Parry et al, 2009; Gupta et al, 2010).

1.7.5 Integrase inhibitors

Raltegravir (RAL) is the only INI currently licensed for the treatment of HIV-1 infection. It is an analogue of the diketo acid class compounds and as a result it shares their β-hydroxy-ketone structural motif (Hazuda et al, 2000; Pommier et al, 2005). As previously specified, the process of integration is catalysed by the viral integrase and is a multistep mechanism that comprises the formation of a pre-integration complex; its nuclear importation; the 3'-processing of the DNA molecule; and strand transfer reaction that results in the attachment of viral and cellular DNA. To exert its action, the HIV-1 integrase requires the presence of divalent cations, such as Mg^{2+} or $Mn^{2+,}$ in its active site. Raltegravir is an inhibitor of the strand transfer event in the integrase process (Hazuda et al, 2000) which, as other diketo-acid compounds, possesses metal-chelating functions. Its mechanism of action is not completely understood. However, it is reported that it binds tightly to the active site of the viral integrase and its chelating properties result in the sequestration of the metal ions present in the active site that are crucial for HIV-1 IN function.

Resistance to Raltegravir is associated with mutations pointed directly to the catalytic site of the HIV-1 integrase enzyme (Pommier et al, 2005) and emerge in at least three different genetic pathways defined by a major mutation at Q148H/K/R, N155H and less often Y143R/H/C and one or more minor mutations which further increase the level of resistance to the inhibitor and/or compensate for the decreased fitness often associated with the presence of major mutations

(Delelis et al, 2010; Marinello et al, 2008; Kobayashi et al, 2008). The Q148 amino acid is a critical part of the active site of the integrase. Not surprisingly, mutations at this codon reduce the susceptibility to raltegravir, but also markedly impair enzyme function. The replication defect associated with the presence of the signature raltegravir mutation Q148 H/K/R is frequently rescued by the compensatory mutations G140S and to a lesser extend E138E/K (Fransen et al, 2009; Delelis et al, 2010). The second most common pathway of raltegravir resistance includes the major mutation N155H, which lies at the base of the catalytic site of the HIV1- integrase interacting with active site residues and directly interfering in enzyme metal binding (McColl et al, 2010). The N155H mutation reduces raltegravir susceptibility without affecting the replicative capacity to the extent of mutations at codon 148. The E92Q minor mutation is frequently found in combination with N115H and in this context further contribute to resistance to the INI, but it does not rescue the replicative capacity of the mutant virus (Fransen et al, 2009). Finally, the third pathway of resistance, which is much less common, comprises the mutation Y143R/C. Raltegravir interacts with Y143 residue during binding to the integrase inhibitor and consequently mutations at this site removes this favourable interaction causing resistance to the drug (Hare et al, 2010 and Delelelis et al, 2010). The T97A minor mutation is commonly seen with Y143R/C and significantly increases raltegravir resistance (Reuman et al, 2010; Fransen et al, 2009).

Name (Abbreviation)	Key mutations implicated in resistance
NRTIs	
Zidovudine (ZDV)	RT: M41L, D67N, K70R, L210W, T215YF, K219EQ
Stavudine (d4T)	RT: M41L, K65R D67N, K70R, L210W, T215YF, K219EQ
Lamivudine (3TC)	RT: K65R, M184VI
Abacavir (ABC)	RT: K65R, L74V, Y115F, M184V
Didadosine (ddI)	RT: K65R, L74V
Emtricitabine (FTC)	RT: K65R, M184VI
Tenofovir (TDF)	RT: K65R, K70E
NNRTIs	
Nevirapine (NVP)	RT: L100I, K101P, K103NS, V106AM, V108I, Y181CI, Y188CLH, G190A
Delaviridine (DLV)	RT: L100I, K101P, K103NS, V106AM, V108I, Y181CI, Y188CLH, G190A
Efavirenz (EFV)	RT: L100I, K101P, K103NS, V106M, V108I, Y181CI, Y188L, G190AS, P225H
Etravirine (ETV)	RT: V90I, A98G, L100I, K101EHP, V106I, E138AGKQ, V179DFT, Y181CIV, G190AS, M230L
Rilpivirine (RPV)	RT: K101EP, E138AGKQR, V179L, Y181CIV, H221Y, F227C, M230IL
PIs	
Saquinavir (SQV)	PR: L10IRV, L24I, G48V, I54VL, I62V, A71VT, G73S, V77I, V82AFTS, I84V, L90M
Ritonavir (RTV)	PR: L10FIRV, K20MR, V32I, L33F, M36I, M46IL, I54LV, A71VT, V77I, V82AFT, I84V, L90M
Indinavir (IDV)	PR: L10IRV, K20MR, L24I, V32I, M36I, M46IL, I54V, A71VT, G73SA, L76V, V77I, V83AFT, I84V, L90M
Nelfinavir (NFV)	PR: L10FI, D30N, M36I, M46IL, A71VT, V77I, V82AFTS, I84V, N88DS, L90M
FosAmprenavir (fAPV)	PR: L10FIRV, V32I, M46IL, I47V, I50V, I54LVM, G73S, L76V, V82AFST, I84V, L90M
Lopinavir (LPV)	PR: L10FIRV, K20MR, L24I, V32I, L33F, M46IL, I47AV, I50V, F53L, I54ALMTSV, L63P, A71TV, G73S, L76V, V82AFTS, I84V, L90M
Atazanavir (ATV)	PR: L10CFIV, G16E, K20IMRTV, L24I, V332I, L33FIV, E34Q, M35ILV, M46IL, G48V, I50L, F53LY, I54ALMVT, D60E, I62V, A71ILTV, G73ACST, V82AFIT, I84V, I85V, N88S, L90M, I93LM
Tipranavir (TPV)	PR: L10V, L33F, M36ILV, K43T, M46L, I47V, I54AMV, Q58E, H69KR, T74P, V82LT, N83D, I84V, L89IMV
Darunavir (DRV)	PR: V11I, V32I, L33F, I47V, I50V, I54LM, T74P, L76V, I84V, L89V
FIs	
Emfuvirtide (T20)	Gp41: G36DS, I37V, V38AEM, Q39R, Q40H, N42T, N43D
INIs	
Raltegravir (RAL)	IN: E92Q, Y143CHR, Q148HKR, N155H
Co-receptor antagonists	
Maraviroc (MVC)	There are not specific mutations described.

Table 1.1 Antiretrovirals approved for HIV-1 treatment.

Abbreviations: NRTIs = nucleoside/nucleotide reverse transcriptase inhibitors; NNRTIs = non-nucleoside/nucleotide reverse transcriptase inhibitors; PIs = protease inhibitors; FIs = fusion inhibitor; INIs = integrase inhibitor; RT = reverse transcriptase; PR = protease; IN = integrase.

1.8 Resistance testing

The development of drug resistance is an important factor leading to treatment failure (Alcorn and Faruki, 2000). Consequently, major guidelines recommend the use of resistance testing at the time of treatment failure in order to decide the optimal regimen (US Department of Health and Human Services guidelines, 2010; Hirsh et al, 2008; The British HIV Association guidelines, 2012). Several randomized studies have demonstrated that this practice leads to superior virological response (Durant et al, 1999; Baxter et al, 2000; Cohen et al, 2002). There are two main types of resistance testing: genotypic and phenotypic. Although both approaches are appropriate to evaluate drug resistance in the context of treatment failure, genotypic testing is preferred in clinical practice due to its faster turnaround time and simplicity.

To perform genotypic resistance testing, the viral RNA is isolated from plasma, the genomic region of interest (HIV-1 RT, PR, IN, Env) is amplified by RT-PCR and the amplicon is sequenced by the Sanger method. This sequencing method uses dideoxynucleotides (ddNTPs), which lack the 3' hydroxyl group required for the formation of a phosphodiester bond between two nucleotides and consequently they act as chain terminators. Briefly, the amplicon is denatured and a primer annealed to one of the template strands. The four deoxynucleotides triphosphate (dNTPs) are added together with the four ddNTPs, each marked with a different colour dye, at a 100:1 ratio. As the DNA polymerase polymerizes, dNTPs are added to the growing chain. However, on occasion a ddNTP rather than a dNTP is incorporated resulting in a chain terminating event. Chain of different lengths are produced and detected via fluorescence.

Direct sequencing from amplicons is known as viral population sequencing, and generates a consensus nucleotide sequence from the patient's most prevalent viral quasispecies, which is further translated into its corresponding amino acid sequence. Point mutations are identified by alignment with a reference wild-type strain sequence, typically HXB2, and those variants with known effects on drug resistance are scored in the interpretation. Several online genotypic interpretation systems are available, which translate a specific mutation pattern into the predicted level of susceptibility. The three most commonly employed in clinical practice are ANRS, Rega and Stanford HIV db (Frentz et al, 2010).

To perform phenotypic testing, in its initial format the isolation of patient's virus and PBMCs were required. These were subsequently cultivated for 2-8 weeks in the presence of increasing concentrations of drugs and the readout was generally the production of P24 in the supernatant. This method has the advantage of incorporating both cells and virus from the patient, thus mimicking "*in vivo*" conditions closely. However, this approach had major disadvantages, particularly the long time required to complete the assay. In addition, due to biological variations, reproducibility was also a problem (Schmidt et al, 2002; Japour et al, 1993). As a result, simpler phenotypic resistance testing methods based on recombinant virus techniques were developed. The initial steps are the same as for genotypic testing and involved the isolation of the patient's RNA and the amplification of the gene of interest, typically RT, PR and/or IN. In phenotypic tests, the amplified gene is then transferred into a laboratory strain of HIV that lacks the gene of interest, producing a recombinant virus. This mosaic virus contains patient-derived sequences that can be tested for drug susceptibility in two possible formats; these are:

- Single cycle phenotypic assay: this method is based on a single round of infection. The assay uses restriction, digestion and ligation to clone the patient's HIV genomic region of interest into an HIV expression vector that lacks this region, to form Resistance Test Vectors (RTVs). The RTVs are replication defective as they are deprived of the envelope gene. Cells are co-transfected with three plasmids: the resistance test vector, which contains the patient's derived sequences, a reporter vector, which contains the HIV packaging sequence and also expresses luciferase that is employed as a marker of virus production, and a vector expressing the vesicular stomatitis G protein, which provides the envelope to the pseudovirus,. Following exposure to increasing drug concentrations, the amount of pseudotyped virus is measured by quantifying the amount of luciferase production. Single cycle assays also provide a measure of the replicative capacity of the virus, which is defined by its ability to replicate in the absence of drugs and reported as a proportion (%) of the replication observed with wild-type virus.

- Multiple cycle phenotypic assay: In this format the tested virus undergoes multiple rounds of infection. The assay uses homologous recombination in cell culture to insert the HIV genomic region of interest into a molecular HIV-1 clone, usually HXB2, which lacks the same region. This yields infectious recombinant viral particles capable of multiple rounds of replication. The virus is grown in the presence of increasing drug concentrations and replication is measured by expression of a reporter gene such as 3-(4, 5-dimethylthiazol)-2, 5-diphenyltetrazolium bromide (MTT).

The main advantage of multiple cycle assays is that the virus undergoes multiple rounds of infection and as a result mimics more closely "*in vivo*" conditions. However, these types of assays are time consuming, processing requires high level of containment, and retain the potential that the virus may mutate during passages. By contrast, single cycle phenotypic testing can typically be performed within a week and the results obtained are generally highly reproducible. In addition, the format of the assay eliminates the possibility of selection of viral populations that may not accurately represent the original virus and therefore may be the method of choice when the effect of specific mutations needs to be characterized. Regardless of the format of phenotypic assay employed, both assays report drug susceptibility results as a fold change for each drug, which is determined by the ratio of the IC_{50} (the concentration of a drug that is required for 50% viral inhibition) from the patient's chimeric virus divided by the IC_{50} of a wild-type reference virus. The interpretation of phenotypic tests is based on defined cut-offs. The technical cut off is generated by measuring the variation seen with repeat testing of the same samples and provides a measurement of the reproducibility of the system. The biological cut off (BCO) represents the phenotypic variability observed within treatment naïve patients and provides a measurement of the normal variation in fold changes observed in wild type viruses. The clinical cut off (CCO) represents the value that discriminates between treatment responders and non-responders among treatment experienced patients. Two clinical cut-offs are defined for each drug: the lower cutoff (CCO1) would define when the susceptibility begins to decline, but the drug still has partial activity, and the upper cutoff (CCO2) represents the fold change at which all drug activity is lost. From a clinical perspective, clinical cut-offs provide the most useful information, but if these clinical cutoffs have not been defined for a given drug, then biological cutoffs are applied.

Both standard phenotypic and genotypic resistance testing have limitations. Firstly, they are unable to detect minority variants, which are present at frequencies below 20%. Alternative methods capable of detecting low frequency variants include allele specific PCR (Charpentier et al, 2004), single genome analysis (Palmer et al, 2005) and ultra-deep sequencing assays (Wang et al, 2007). The allele specific PCR employs genetic probes to detect specific resistance mutations. It can detect mutations at frequencies between 0.5-1 % depending on the specific mutation. In single genome analysis, the complementary viral DNA is synthesized from the patient's plasma and diluted to one copy; then each viral copy is amplified and sequenced. Finally, Ultra-deep sequencing uses large scale parallel pyrosequencing and is able to detect viral quasispecies present at very low levels. The role of minority variants in failure of antiretroviral remains to be elucidated. In drug naïve patients, studies have shown an association between the presence of minority drug resistant variants and subsequent virological failure, particularly when NNRTI-based regimens are administered (Johnson et al, 2008a; Metzner et al, 2009 and Geretti et al, 2009). However, definitive cut-offs remain to be established. There is likely a threshold above which minority variants will lead to therapy failure and this threshold will depend on the barrier to resistance of the affected drug, the antiviral activity of other drugs in the regimen as well as patient's characteristics, such as baseline viral load and adherence patterns. Another important restriction of current phenotypic and genotypic resistance testing methods is that they do not include the entire viral sequence. As a result, there may be important mutations outside the standard area of interest that will be missed by these techniques. For instance, to evaluate resistance to RTIs, the connection and RNase H domains are not routinely sequenced or included in the recombinant viruses used for phenotypic assays. Mutations in these areas have been reported to be selected by NRTI therapy and increase the level of resistance to these drugs (Nikolenko et al, 2007). Similarly, most phenotypic and genotypic tests only include

the PR region of HIV-1, those that include regions of Gag, the main substrate for protease, are restricted to the two CS located at the C-terminal site (P7/P1/P6). However, mutations in Gag outside this region have been recently demonstrated to contribute to resistance to PIs (Parry et al, 2009). Importantly, mutations in the RNase H and connection domains have been reported to emerge in the context of multiple TAMs. Therefore, their presence in already highly resistant virus questions its clinical utility. However, mutations in Gag have been found to confer resistance in the absence of typical PI-resistance-associated mutations (Nijhuis et al, 2007; Gupta et al, 2010).

1.9 Viral fitness and replicative capacity

The term "viral fitness" refers to the ability of a virus to replicate and produce a progeny in a given environment (Domingo et al, 1997a). There are two stages to describe the evolution of viral fitness whilst on therapy. The first stage is characterised by the selection of primary, also referred to as major, resistance mutations resulting in viruses that not only have reduced drug susceptibility but often an impaired replicative capacity relative to the wild type virus. During the second stage, additional mutations arise that alone do not confer drug resistance but in combination with the primary mutations enhance the replicative capacity of the virus. These mutations are known as accessory mutations or also termed minor or secondary mutations (Nijhuis et al, 1999). It can be said that the natural evolution of HIV under drug pressure is towards increasing levels of resistance, cross-resistance and fitness. There are different modes for the assessment of viral fitness and these methods can be broadly classified as *in vivo* and *in vitro* methods. *In vivo* methods assess viral fitness by comparing the amount of wild type and mutant virus detected in "*in vivo*" populations (Devereux et al, 2001). These methods closely mimic the natural setting of the natural host. However, the

entire individual offers a variety of cell types and microenvironments to the infecting HIV-1 with conflicting selective pressure and consequently, *in vivo* methods are difficult to extrapolate to different situations.

In vitro methods employ HIV-1 isolates or more frequently recombinant viruses and can be very useful as models for determining the effect of drug resistant variants on replication in a fixed environment. There are several methods for the evaluation of viral fitness *in vitro*. In general, they are all grouped into two main categories: monoinfection assays (or viral growth kinetics) and growth competition experiments. In monoinfection assays, the replicative capacity of different HIV isolates or recombinant viruses is tested individually and measured by measuring the amount of specific viral proteins, usually P24, or the activity of viral enzyme, such as RT. Virus replicative capacity can also be measured in monoinfection assays by using a reporter gene, such as luciferase, in a single cycle assay (Dykes and Demeter, 2007) and comparing luciferase production by a mutant virus and WT reference virus.

In growth competition experiments, two phenotypically distinguishable viruses are mixed at similar or different proportions and the outgrowth of one of the population is measured (Domingo et al, 1997a; Domingo and Holland et al, 1997b). By doing this, the fitness of both viral strains can be directly compared as two viral populations in cell culture compete with each other until one outgrows the other one. In general, cells are infected with the mixture of viruses and after several passages, the proportion of both viruses is measured and compared with their proportions in the initial mixture (Holland et al, 1991 and Domingo et al, 1997a).

Overall, growth competition experiments are more accurate and sensitive for the determination of small differences in fitness than monoinfection assays. The single cycle assay offers a fast and reproducible method to measure the replicative capacity of mutant virus that can be compared and expressed as a percentage of that observed for a WT reference strain. As it is a monoinfection assay, it cannot accurately determined small differences in replicative capacity. However, it can be of use for the characterization of novel mutations since growth competition experiments increase the potential for mutations to occur in the different passages and divert form the population of interest.

2 Chapter two: materials and methods

2.1 Materials

2.1.1 Bacteria

- TOP10: F^- *mcr*A Δ (*mrr-hsd*RMS-*mcr*BC) Φ80*lac*ZΔM15 Δ*lac*X74 *rec*A1 *ara*D139 Δ (*ara-leu*) 7697 *gal*U *gal*K *rps*L (Str^R) *end*A1 *nup*G.

- HB101: F-, *thi*-1, *hsdS*20 (r_B-, m_B-), *sup*E44, *rec*A13, *ara*-14, *leu*B6, *pro*A2, *lac*Y1, *gal*K2, *rps*L20 (str^r), *xyl*-5, *mtl*-1.

- XL1 blue supercompetent cells: *recA1 endA1 gyrA96 thi-1 hsdR17 supE44 relA1 lac* [F´ *proAB* $lacI^qZ\Delta M15$ Tn*10* (Tet^r)].

2.1.2 Mammalian cell lines

Human Embryonic Kidney 293 cells (HEK 293 Cells): HEK 293 cells were generated in the 70s by transformation of cultures of normal human embryonic kidney cells with sheared adenovirus 5 DNA (Graham FL, et al, 1977).

2.1.3 Bacterial culture media

Luria-Bertani (LB) broth: 10 g tryptone, 5 g of yeast extract, 10 g NaCl and deionised water to a final volume of 1 litre. Ampicillin was added at a concentration of 50 mg/ml.

- LB agar plates: LB medium was prepared as indicated above and 15g/l of agar was added before autoclaving and supplemented with 50 mg/ml of ampicillin after autoclaving.

- LB agar X-Gal plates: ampicillin-containing LB agar plates were prepared as above and spread with 40 µl of 40 mg/ml X-gal 10 minutes before use.

- Super Optimal Broth with catabolite repression (SOC): SOC medium was purchased from Invitrogene and contained 20 g tryptone, 5 g yeast extract, 10 mM NaCl, 2.5 mM KCl, 10 mM $MgCl_2$ and 20 mM glucose.

- NZY^+ broth: 10 g of NZ amine (casein hydrolysate), 5 g of yeast extract, 5 g of NaCl, 12.5 ml of 1M $MgCl_2$, 12.5 ml of 1M MgSO4, 10 ml of 2M glucose and deionised water to a final volume of 1 litre. The pH was adjusted to 7.5 using NaOH.

2.1.4 Cell culture media

Cells were grown in Dulbecco's Modified Eagle Medium (DMEM) (GIBCO, Invitrogene, Paisley, UK) supplemented with 100 U/ml of penicillin, 100 mg/ml of streptomycin and 15% of Fetal Calf Serum (FCS) (Biosera, UK).

Cell transfection was conducted in Opti-MEM® Reduced Serum Medium with GlutaMAX (Opti-MEM) (GIBCO, Invitrogene, Paisley, UK).

2.1.5 Transfection reagent

FuGENE® 6 Transfection Reagent (Roche Diagnostics, US), which is a proprietary blend of lipids and other components, was employed for cell transfection.

2.1.6 Antiretrovirals

Protease inhibitors (PIs): Amprenavir (APV), Atazanavir (ATV), Darunavir (DRV), Indinavir (IDV), Lopinavir (LPV) and Saquinavir (SQV). All antiretrovirals were obtained through the AIDS Research and Reference Program, Division of AIDS, NIAID, NIH.

2.1.7 Primers

Sequences for the primers employed in the study are detailed in tables 2.1 to 2.4.

2.1.8 Vectors

- P8.9NSX: it is a modified form of Gag-Pro-Pol expression plasmid pCMV-ΔR8.2 (Naldini et al, 1996). The vector encodes HIV-1 Gag-pol and the virulence genes (*vif, vpu and nef*); accessory genes (*Rev and Tat*) as well as *env* gene have been deleted. In addition, the vector contains numerous restriction sites such as NotI site upstream of Gag, ApaI and SpeI sites within Gag and XmaI in the integrase allowing foreign Gag-PR and RT sequences to be cloned. We further modified the P8.9NSX vector to allow the independent cloning of external full-length Gag and PR genes. Modifications were performed by SDM and consisted of introduction of a BglII site at the end of Gag and beginning of PR and a BamHI site at the end of the protease. In addition, an additional BglII and a BamHI present in the vector downstream of RT were blocked by "*in vitro"* mutagenesis so that the final sequence contains a unique NotI site upstream of Gag, a unique BglII at the end of Gag and beginning of PR and a unique BamHI at the end of the PR. Upon completion of the "*in vitro"* mutagenesis, the HIV backbone sequence of the P8.9NSX vector was confirmed by sanger sequencing and the presence of the restriction sites by digestion with the corresponding restriction endonucleases.

- pCSFLW: it is a modified form of the pHR-SIN-CSGW vector in which the gene encoding green fluorescent protein has been replaced with firefly luciferase (FL) (Demaison et al, 2002). Therefore, as pHR-SIN-CSGW, pCSFLW is a self-inactivating vector (SIN vector) in which the U3 region of the 5' LTR of HIV has been replaced with the CMV promoter and that has a deletion in the U3 region of the 3' LTR which includes segments encoding the enhancer and promoter functions. The deletion will be transferred to the 5'LTR after reverse transcription and as a result the transcriptional unit from the LTRs in the provirus is eliminated providing a SIN vector. The pCSFLW encoded firefly luciferase which is employed as the reporter gene in the drug susceptibility and replicative capacity assays and also contains the HIV packaging sequence.

- PMDG: it expressed the Vesicular estomatitis virus G envelope protein under regulation of the CMV promoter and is used for the pseudotyping of the HIV and luciferase expressing vectors.

- PCR2.1®TOPO: it is a commercial vector (Invitrogene, Pasley, UK) which is supplied linearized with a 3'-thymidine (T) overhang and Topoisomerase covalently bound to the vector. This vector was employed for TA cloning of PCR products which would be used for clonal analysis of samples which did not render interpretable population sequencing results or in which clonal analysis was required for the study of linkage of mutations.

- PGEM8.9: this vector was obtained by cloning the P8.9NSX HIV backbone containing restriction sites into the commercial vector PgemEasyTvector (Promega, UK). PGEM8.9 was employed as a template for SDMs.

2.2 Methods

2.2.1 Patients, samples and sequences

- In the validation of the HIV-1 Gag and protease (Gag-PR) amplification and sequencing assay, we employed plasma samples from patients attending the HIV services at Royal Free Hospital (RFH). The Virology Department at Royal Free Hospital holds a database in which all HIV-1 infected patients that attend the hospital for drug resistance testing have their *pol* gene sequences and genotypic profile entered. The database also contains information on treatment status, treatment regimen, plasma HIV-1 RNA load and HIV-1 subtype. In 2011, the database contained approximately 10000 *pol* sequences. Plasma samples were selected representing a wide range of HIV-1 subtypes and circulating recombinant forms (CRFs).

- In the cross-sectional comparison of PI-naïve and PI-experienced patients, we included patients from different cohorts in order to increase the number of sequences analyzed. Firstly, we selected both PI-naïve and PI-experienced patients from the RFH (n= 52 and n=50, respectively). Secondly, we analyzed sequences from PI-experienced patients attending the Cologne University Hospital (n = 128). Lastly, we included PI-experienced

patients from the MaxCmin 1, MaxCmin2 and COLATE clinical trials (n = 13). . In addition, we also downloaded 148 drug-naïve Gag-PR sequences from the HIV Los Alamos Database. In total, we compared 200 PI-naïve and 191 PI-experienced sequences. All sequences obtained from PI-experienced patients contained at least one major PR resistance mutations and all sequences were retrieved from subjects with long lasting subtype B HIV-1 infection (i.e.; 3-10 years).

- In the longitudinal analysis of Gag-PR in patients failing PI-based regimens as well as for the assessment of the effect of Gag mutations on PI susceptibility and replicative capacity (RC), we selected patients from the MaxCmin1, MaxCmin2 or COLATE trials, who had matched pre-treatment and treatment failure plasma samples. In addition, we also studied patients from the HIV services at Royal Free Hospital who had long term on-going viraemia while on PI-based regimen and showed evidence of PR evolution on their routine HIV resistance genotypic tests.

2.2.2 General molecular biology techniques

Standard molecular biology techniques as described in Molecular Cloning: a Laboratory Manual, (Maniatis et al, 1986) were used throughout.

2.2.2.1 RNA extraction

We employed Qiagen QIAamp Viral RNA Kit (Qiagen, Crawley, UK) and EasyMag automated extractor (Nuclisens, Biomerieux, Boxtel, Netherlands) for manual and semi-automated RNA extraction, respectively. Both methods are based on the nucleic acid extraction protocol developed by Boom and colleagues (Boom et al, 1990) which employed the lysing and nuclease-inactivating properties of the chaotropic agent guanidinium thiocynate together with the nucleic acid-binding properties of silica particles in the presence of this agent. One millilitre of plasma was centrifugated for 1 hour at 4°C to concentrate the virus. The supernatant was then removed and the pellet re-suspended to a final volume of 280 µl. The re-suspended pellet was employed for nucleic acid extraction with either Qiagen QIAamp Viral RNA Kit (Qiagen, Crawley, UK) if the plasma HIV viral load was below 1,000 copies/ml or the automated extractor EasyMag (Nuclisens, Biomerieux, Boxtel, Netherlands) if the HIV plasma viral load was above that threshold. In both cases nucleic acid extraction was performed following the manufacturer's instructions.

2.2.2.2 Amplification of HIV-1 Gag-protease region

A RT-nested PCR protocol was designed for reverse transcription followed by specific amplification of full-length Gag-PR. Primers employed for amplification are shown in table 2.1. Three different commercial Kits were evaluated for amplification of the HIV-1 Gag-PR region. Two of the methods performed reverse transcription and amplification in a single tube followed by a second PCR round on the initial PCR product: Qiagen one step RT-PCR Kit/Qiagen HotStar®Taq DNA polymerase Kit (Qiagen, Crawley, UK), namely Qiagen Gag-PR amplification protocol and Invitrogene SuperScript®III One-Step RT-PCR System/

Platinum® Taq DNA polymerase HF Kit (Invitrogene, UK), namely Invitrogene Gag-PR amplification protocol. The third method (AccuScript™ High Fidelity (HF) PCR/ PfuUltra HF DNA Polymerase) (Stratagene, Netherlands), namely Stratagene Gag-PR amplification protocol, separates both reverse transcription and both rounds of amplification in three independent steps.

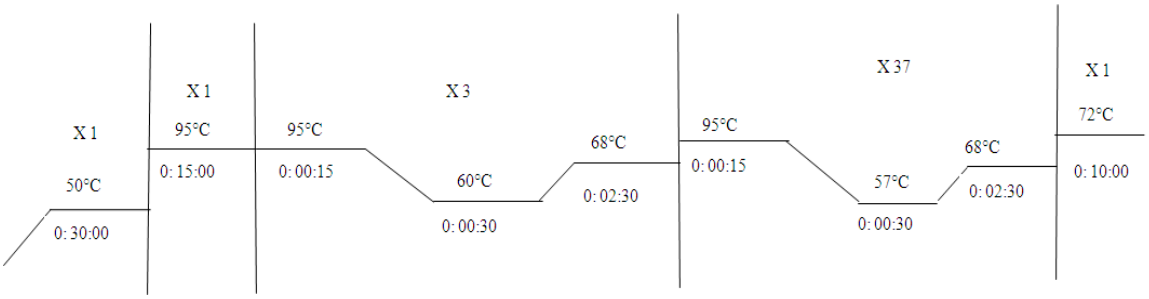
Qiagen incorporates non-proofreading enzymes for both reverse transcription and amplification. Invitrogene system uses a blend of proof-reading and non-proofreading enzymes for reverse transcription and PCR. In addition, the reverse transcriptase included in Invitrogene system displayed a reduced RNase H activity facilitating complete synthesis of cDNA strands. Lastly, Stratagene kits contained proof-reading enzymes for both reverse transcription and amplification.

The three systems were optimized and the selection of one or another method for amplification will depend on the specific purpose i.e.: population sequencing *vs*. clonal analysis.

2.2.2.2.1 Qiagen Gag-protease amplification protocol

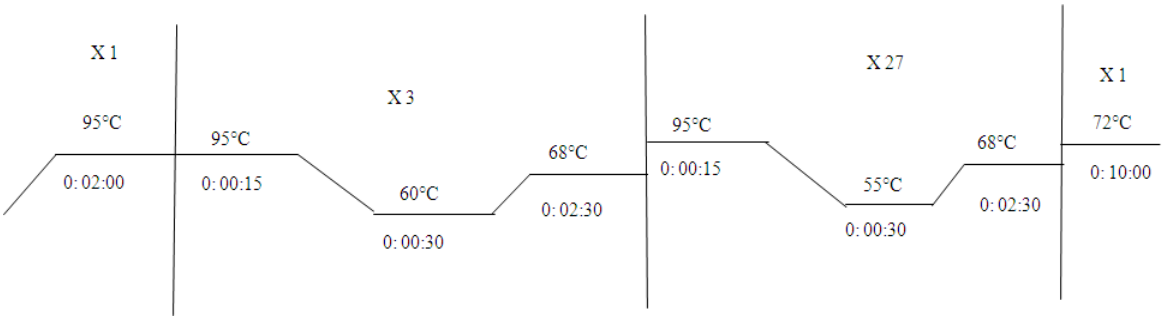
One step RT-PCR:

- Prepare master mix containing: 1 x Qiagen RT-PCR buffer, 400 µM dNTPs, 0.4 µM of forward and reverse primers and 2.5 Units of Qiagen RT/PCR enzyme mix.
- Add 40 µl of the above mix to 10 µl of RNA extract.
- Perform RT-PCR as follows:



2nd round PCR:

- Prepare master mix containing: 1 x Qiagen PCR buffer, 400 µM dNTPs, 0.4 µM of forward and reverse primers and 2 Units of Qiagen Hotstar Taq DNA polymerase.
- Add 48 µl of the above mix to 2 µl of 1st round PCR product.
- Perform PCR as follows:



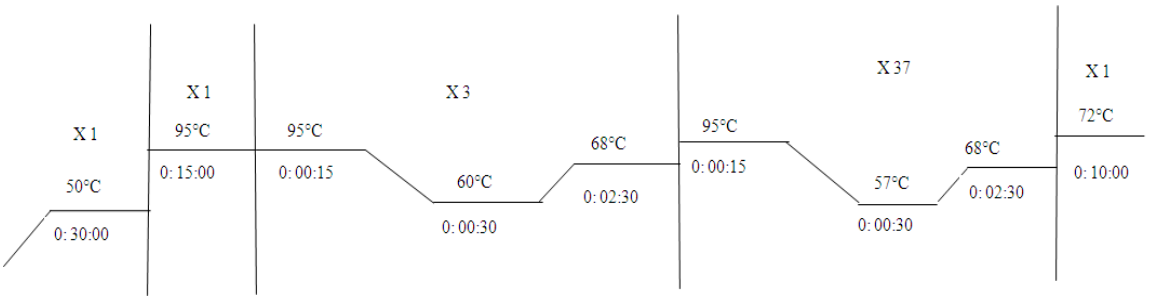
2.2.2.2.2 Invitrogene Gag-protease amplification protocol:

One step RT-PCR:

- Prepare master mix containing: 1 x Invitrogene buffer supplemented with 1.2 Mm Mg^{2+} and 200 µM dNTPs, 0.4 µM of forward and reverse primers and 2 Units of RT/PCR enzyme mix.

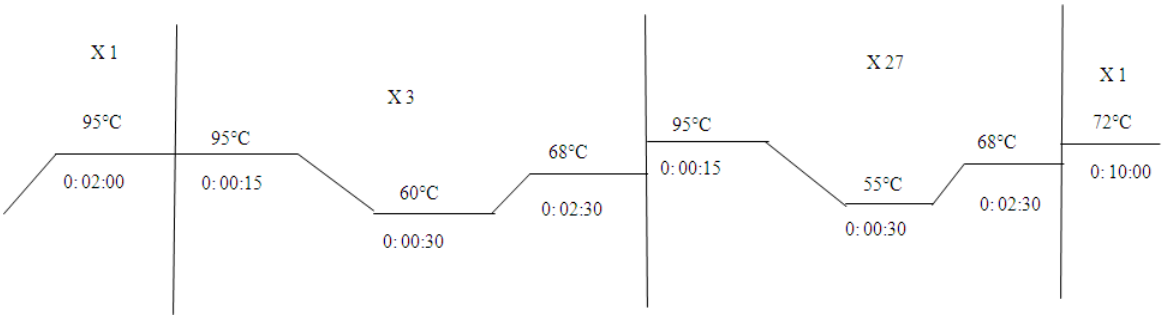
- Add 40 µl of the above mix to 10 µl of RNA extract.

- Perform RT-PCR as follows:



2nd round PCR:

- Prepare master mix containing: 1 x Invitrogene PCR buffer, 1.5 mM Mg^{2+} , 200 µM dNTPs, 0.4 µM of forward and reverse primers and 2.5 Units of Platinum Taq HF Taq DNA polymerase.
- Add 48 µl of the above mix to 2 µl of 1st round PCR product.
- Perform PCR as follows:



2.2.2.2.3 Stratagene Gag-protease amplification protocol:

Reverse transcription (RT):

- Prepare master mix containing: 1 x AccuSript RT buffer, 10 Mm Dithiothreitol (DTT), 1mM dNTPs, 2µM of outer reverse primers and 20 Units of RNase inhibitors.
- Add 39 µl of the above mix to 10 µl of RNA extract

- Incubate the above mix at 65°C for 5 minutes and subsequently cool down reaction to room temperature.

- Add 2 units of AccuScript RT enzyme

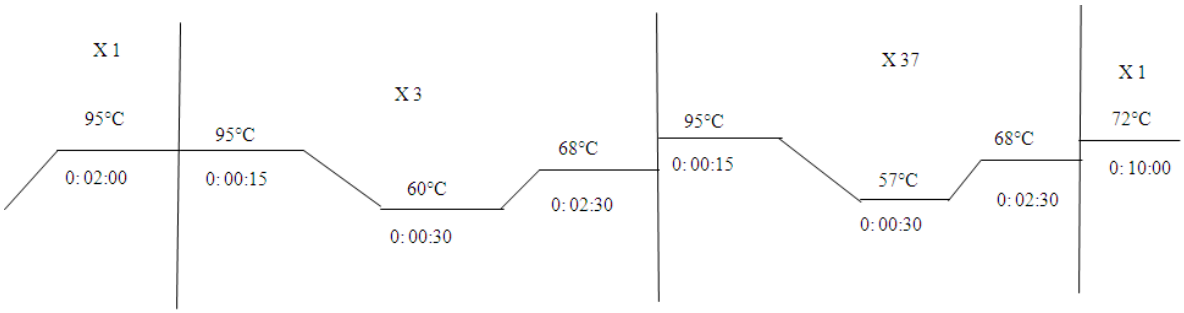
- Perform RT for 1 hour at 42°C.

1st round PCR:

- Prepare master mix containing: 1 x Pfu Ultra HF buffer containing 2 mM Mg^{2+}, 200 µM dNTPs, 0.4 µM of forward and reverse primers and 2.5 Units of Pfu Ultra HF DNA polymerase.

- Add 45 µl of the above mix to 5 µl of cDNA

- Perform PCR as follows:

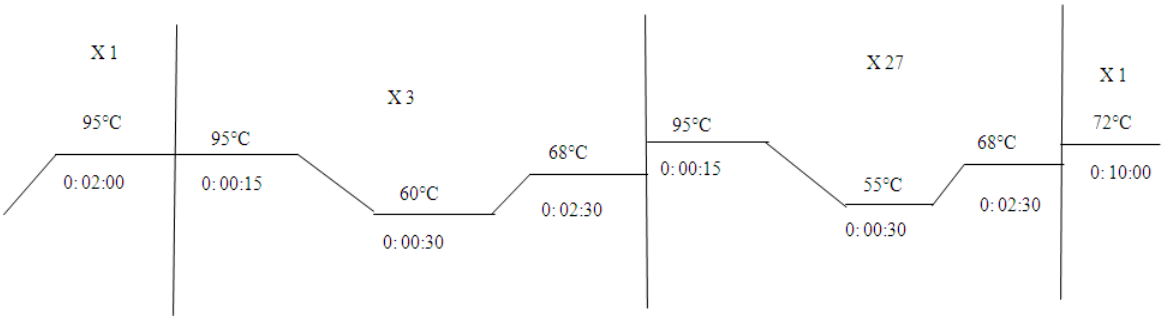


2nd round PCR:

- Prepare master mix containing: 1 x Pfu Ultra HF buffer , 4 mM Mg^{2+} , 200 µM dNTPs, 0.4 µM of forward and reverse primers and 2.5 Units of Pfu Ultra HF DNA polymerase.

- Add 48 µl of the above mix to 2 µl of 1st round PCR product.

- Perform PCR as follows:



2.2.2.3 Agarose gel electrophoresis

Agarose gel electrophoresis was used to confirm the correct size of the PCR product. 1.5 g of agarose was dissolved in TAE buffer in a microwave and once cooled; 10mg/ml ethidium bromide was added. PCR products were mixed with 5X loading dye and loaded onto the gel with a DNA mass ladder. Gels were run for approximately 1-2 hours at 100 volts, depending on the size of the band expected. All gels were visualised using a UV transilluminator (Biorad, Hertfordshire, UK).

2.2.2.4 Purification of PCR products

PCR products were visualised using ethidium bromide staining and agarose gel electrophoresis. Products of the expected size were either excised and purified using the QIAQuick Gel Extraction Kit (Qiagen, Crawley, UK) if the PCR product was going to be subsequently cloned, or directly purified from the PCR mixed using the QIAQuick PCR Purification kit (Qiagen, Crawley, UK). Both purification methods were performed in accordance with the manufacturer's instructions. Briefly, the PCR product was added to the buffer containing the chaotropic agent guanidine thiocyanate and bound to the silica membrane in the QIAquick spin column through centrifugation for 1 minute at 13,000rpm. Impurities and contaminants were removed through washing with an ethanol containing buffer and DNA was eluted into a low salt and pH containing buffer. If gel excision was employed, the gel slice was previously dissolved in 3 volumes of buffer QG and subsequently the DNA purified as indicated above for a PCR product.

2.2.2.5 A-tailing

To enable efficient TA cloning of products generated with high fidelity polymerases, such as those incorporated into Invitrogene or Stratagene systems, purified PCR products were added to a mix of 250 uM dATP, 10Xbuffer containing 2.5 mM $MgCl_2$, and 2.5 units of Amplitaq Gold Polymerase. The mix was heated to 95^{o}C for 10 minutes followed by 20 minutes at 72^{o}C.

2.2.2.6 PCR TA cloning into pCR®2.1-TOPO

The TOPO TA cloning kit (Invitrogen, Paisley, UK) was used to clone the HIV Gag-PR genes. TOPO TA cloning uses Topoisomerase I to ligate the PCR product with the vector. All PCR products produced with conventional Taq polymerases contain a deoxyadenosine (A) overhang to the 3'end, as a consequence of the non-template-dependent terminal transferase displayed by the enzymes. In the case that proof-reading enzymes were employed for PCR, the A-overhang was incorporated into the PCR product by the A-tailing procedure described above. The linearized vector supplied in the kit (pCR®2.1-TOPO) has a deoxythimidine (T) overhang. Ligation will occur between A-overhang in the PCR product and T-overhang in the vector. A number of cloning reaction, varying amount of PCR products and incubation times, were set-up in order to identify the optimal condition for the cloning of HIV-1 Gag-PR genes into the pCR®2.1-TOPO vector. Final ligation reaction consisted of: 4 µl PCR product, 1 µl salt solution and 1 µl pCR®2.1-TOPO vector. Ligation reactions were incubated at room temperature for 30 minutes and then placed on ice ready for transformation. 50 µl of TOP10 *E. coli* cells were transformed with 2 µl of the ligation reaction. Transformation was carried out by incubating cells and ligation reaction for 30 minutes on ice followed by heat shock at 42ºC for 45 seconds and cooled on ice for 2 minutes. 250 µl of SOC medium (Invitrogen, Paisley, UK) was added to the cells and incubated in a 37ºC orbital shaker for 1 hour. Cells were then plated onto LB agar plates containing 50 mg/ml ampicillin and 40 mg/ml of X-Gal and incubated overnight at 37ºC.

2.2.2.7 Plasmid DNA purification

2.2.2.7.1 Minipreps

Plasmid DNA was extracted from bacterial cells using the QIAPrep Spin Mini-prep Kit (Qiagen, Crawley, UK) in accordance with the manufacturer's instructions. Briefly, a single transformed *E. coli* colony was inoculated into 3 mls of LB broth containing 50 mg/ml ampicillin and incubated overnight at 37°C in an orbital shaker. 2mls of the overnight culture were employed the following day for plasmid DNA extraction. Bacterial cells were resuspended and lysed under alkaline conditions. The lysate was then neutralised with acetic acid and bound to the silica membrane of the QIAprep spin column through centrifugation for 1 minute at 13,000rpm. Remaining impurities were washed away using an ethanol based buffer. The plasmid DNA was then eluted under low salt conditions into RNase-free water.

2.2.2.8 DNA quantification

The quality and quantity of plasmid DNA extracted was assessed by UV-Vis-Spectophotometry using 1 µl of mini-prep and a NanoDrop® ND-1000 UV-Vis Spectrophotometer (NanoDrop).

PCR products were also quantified by visualization using Ethidium bromide (Invitrogene) staining and agarose gel electrophoresis of a 5 µl aliquot of purified DNA with a DNA molecular marker.

2.2.2.9 EcoRI restriction digest

The introduced PCR product is flanked by EcoRI restriction sites in the pCR®2.1-TOPO vector. Therefore, digestion of the plasmids with EcoRI restriction endonuclease will release the PCR product and allow us to identify clones harbouring the expected insert. To this purpose, 5 µl of mini-prep were digested with 1µl *EcoRI* (Invitrogen, Paisley, UK), 2µl buffer containing 1 mM $MgCl_2$, and 12µl RNase free water. Digests were incubated for 1 hour at 37ºC and visualised using ethidium bromide staining and agarose gel electrophoresis. Positive clones were identified as those harbouring the correctly sized inserts.

2.2.2.10 DNA sequencing

All DNA was sequenced using Sanger methodology. Primers employed for Gag-PR sequencing are shown in table 2.2. Purified PCR products or Plasmids identified as containing the PCR insert were diluted to a concentration of around 20 ng/µl and sequenced using the BigDye Sequencing mix v3.1. Sequencing reaction contained 8 µl of PCR product or plasmid, 0.5 µM of the selected primer and nuclease-free water to a final volume of 20 µl. A total of 8 to 10 primers were required for full-length Gag-PR sequencing. Primers were employed in different combinations depending on the specific sample. Sequencing PCR conditions were as follows, 25 cycles of 96°C for 10 seconds, 50°C for 5 seconds and 60°C for 4 minutes and a hold at 4°C. Sequencing reactions were purified by precipitating the DNA with 52 µl of a mix containing 50 µl of 100% ethanol (EtOH) and 2 µl of 3M sodium acetate followed by a washing step with 150 µl of 70% EtOH. Purified sequencing reactions were then run on a 3730-Avant Genetic Analyzer (Applied Biosystems, UK) and the obtained

sequence analysed employing Sequence analysis version 5.0, Seqscape version 6.0 and Mega Molecular Evolutionary Genetic Analysis software version 4.0 programmes.

2.2.2.11 Site-directed mutagenesis

Site directed mutagenesis was carried out using Quickchange Multi/Site Directed Mutagenesis Kit (Stratagene, Cheshire, UK) to insert desired mutations. All primer combinations were designed specifically to incorporate the desired mutation and are shown in table 2.3. 50 ng plasmid and 125 ng of the primers containing the required mutation were used in the following PCR, 16 cycles of 95°C for 30 seconds, 90°C for 30 seconds, 55°C for 1 minute, 72°C for 2 minutes and a hold at 4°C. In order to degrade the parental DNA plasmid, the PCR product was incubated at 37ºC for at least one hour with the restriction enzyme *Dpn1*. 50 µl of XL1-blue supercompetent cells were then transformed with 2 µl of Dpn-digested DNA. Transformation was carried out by incubating cells and ligation reaction for 30 minutes on ice followed by heat shock at 42ºC for 30 seconds and cooled the reaction on ice for 2 minutes. 500 µl of NZY^+ broth were added to the cells and incubated in a 37ºC orbital shaker for 1 hour. Cells were then plated onto LB agar plates containing 50 mg/ml ampicillin and incubated overnight at 37ºC. The following days, a number of colonies were selected for screening for the presence of the correct mutation. Plasmidic DNA was extracted with QIAPrep Spin Miniprep Kit (Qiagen, Crawley, UK) as described in section 2.2.2.7.1 and the presence of the correct mutations was confirmed by sequencing the full Gag-PR region using Sanger sequencing as previously described in section 2.2.2.10.

2.2.3 General tissue culture techniques

All cells and pseudo-virus cultures were grown in humidified 37°C incubators with 5% CO_2 in varying volumes and passaged as required.

2.2.3.1 Cell thawing

HEK-293 cells were removed from liquid nitrogen and thawed rapidly at 37°C. Cells were added to 10 ml of pre-heated DMEM media (GIBCO, Invitrogene, Pasley, UK) supplemented with 100 U/ml penicillin, 100 µg/ml streptomycin (GIBCO, Invitrogene, Pasley, UK) and 10% FCS (Biosera, UK). The cells were subsequently pelleted at 325g for 5 minutes, washed once in 10 ml of DMEM media and re-suspended in 15 ml of DMEM media in 10 cm dishes. The following day the media was replaced with fresh media.

2.2.3.2 Cell passaging

HEK-293 cells were maintained in DMEM media (GIBCO, Invitrogene, Pasley, UK) supplemented with 10% FCS (Biosera, UK), 100 U/ml of penicillin and 100 µg/ml streptomycin (GIBO, Invitrogene, Pasley, UK). Cells were washed with phosphate buffered saline (PBS) [137 mM NaCl, 2.7 mM KCl, 4.3 mM $Na_2H(PO)_4$, 1.4 mM $KH2(PO)_4$], incubated with 2 ml of trypsin-EDTA (Gibco, Invitrogene, Pasley, UK) until the cells were detached from the dish. Cells were then pelleted at 325g for 5 minutes, the tripsine removed and the cells re-suspended in fresh DMEM media. Cells were split 1:4 to 1:8, depending on the cell density and rate of growth, two or three times a week and grown in 5% CO_2 at 37°C.

2.2.3.3 Cell freezing

HEK-293 cells were centrifugated at 325g for 5 minutes and re-suspended at 1x 10^7cells/ml in 40% DMEM media, 50% FCS and 10% dimethyl sulphoxide (DMSO, Sigma, UK). Cells were then aliquoted into cryovials (Nunc, USA) and gradually cooled to -80°C in an isopropanol-containing cryo-container (Nalgene, USA) before being transferred to liquid nitrogene.

2.2.4 Single cycle assay related techniques

2.2.4.1 Generation of resistance test vectors

RTVs were generated by cloning patient related Gag and/or PR sequences into the P8.9NSX HIV expression vector. The P8.9NSX was provided by Professor Pillay's group and was further modified to facilitate the cloning of patients' Gag, PR or Gag-PR sequences. Modification consisted of "*in vitro"* side-directed mutagenesis leading to introduction of two restriction sites and the blocking of other two restriction sites in order to generate unique sites at the beginning and end of Gag and protease genes allowing the cloning of these two genes either simultaneously or independently. After mutagenesis, the sequence of the P8.9NSX HIV backbone was confirmed by Sanger sequencing and the presence of the restriction sites verified by digestion with the corresponding restriction endonucleases. Primers employed for vector modification by SDM are shown in table 2.4.

The generation of RTVs comprised the following steps:

1. Introduction of appropriate restriction sites in patient's samples: the introduction of restriction sites in patient's samples was accomplished by amplification employing Invitrogene protocol and modified nested PCR primers (sequences are shown in table 2.5). Modified primers were designed with restriction sites in their 5' regions. Different restriction sites were including in the forward and reverse primers generating a PCR product whose termini now carry restriction sites that can be used for directional cloning.

2. Purification of PCR product containing restriction sites: Once the restriction site-containing PCR product was generated, this was purified to eliminate the excess of primers, dNTPs and DNA polymerase from the amplified product before digestion with restriction endonucleases. PCR products were purified with QIAQuick PCR purification Kit (Qiagen, Crawley, UK) as indicated in section 2.2.2.4.

3. Digestion with restriction endonucleases: the purified PCR fragment and the PGEM8.9 vector were subsequently digested with appropriate restriction enzymes and the digested products were separated by agarose gel electrophoresis and visualized using ethidium bromide staining. The right size digested PCR product and vector were then gel extracted using QIAQuick Gel Extraction Kit as indicated in section 2.2.2.4.

Three different digestion protocols were optimized depending on whether Gag, PR or Gag-PR was going to be cloned into the P8.9NSX vector:

- Gag Cloning: vector and patient sample were digested with NotI and BglII. Digestion reaction contained: 1 µg of DNA, 2 µl of 1:10 Bovine Serum Albumin (BSA), 5 units of NotI and BglII to a final volume of 20 µl of NE Buffer 3. Digestion proceeded at 37°C for 2 hours after which 2 more units of NotI were added and the reaction incubated for two further hours.

- PR Cloning: vector and patient sample were digested with BglII and BamHI. Digestion reaction contained: 1 µg of DNA, 2 µl of 1:10 BSA, 5 units of BglII and BamHI to a final volume of 20 µl of NEBuffer 3. Digestion proceeded for 2 hours at 37°C.

- Gag-PR cloning: vector and patient sample were digested with NotI and BamHI. Digestion reaction contained: 1 µg of DNA, 2 µl of 1:10 BSA, 5 units of NotI and BamHI to a final volume of 20 µl of NE Buffer 3. Digestion proceeded at 37°C for 2 hours after which 2 more units of NotI were added and the reaction incubated for two additional hours.

4. Cloning: the digested PCR product and PGEM8.9 vector were ligated together using Rapid Ligation Kit (Roche Diagnostics, USA) according to the manufacturer's instructions. Previous to ligation the digested PGEMP8.9 vector was treated with shrimp alkaline phosphatase (Roche Diagnostics, USA) to prevent self-ligation. A

typical ligation reaction contained vector: insert molar ratio 1:3 in a final volume of 21 µl and total DNA content of 200 ng.

5. Transformation of *E.coli* HB101 cells: Ligated products were purified employing QIAQUICK PCR purification Kit (Qiagen, Crawley, UK) to eliminate excess restriction endonucleases and transformed into *E.coli* HB101 cells. Transformation was carried out by incubating cells and ligation reaction for 30 minutes on ice followed by heat shock at 42ºC for 45 seconds and cooled on ice for 2 minutes. 250 µl of SOC medium (Invitrogen, Paisley, UK) was added to the cells and incubated in a 37ºC orbital shaker for 1 hour. Cells were then plated onto LB agar plates containing 50 mg/ml ampicillin and incubated overnight at 37ºC.

6. Miniprep: Plasmid DNA was extracted from bacterial cells using the QIAPrep Spin Mini-prep Kit (Qiagen, Crawley, UK) as indicated in section 2.2.2.7.1.

7. Sequencing: The miniprep containing the plasmid DNA was subsequently sequenced with HIV specific primers to verify the successful cloning of patient's sample into pGEMP8.9.

8. SDM: The plasmid containing ligated PGEMP8.9 and patient's Gag, PR or Gag-PR sequences could be employed as a template for SDM in case modified patient's sequences or modified wild type HIV sequences were required.

9. Transfer of Patient's sample from PGEM8.9 to the HIV expression vector P8.9NSX: patient's Gag, PR or Gag-PR sequences or modified sequences for these genes were transferred from PGEMP8.9 vector to the HIV expressing vector (P8.9NSX) by digestion of both plasmids with appropriate restriction endonuclease, ligated together using rapid Ligation Kit (Roche Diagnostics, US) and transformed into E. coli HB101 cells as described above. Upon sequencing of the HIV backbone of the P8.9NSX vector using HIV specific primers in order to confirm successful cloning of patient backbone into the P8.9NSX vector, the RTVs are ready to be employed in drug susceptibility and RC experiments.

2.2.4.2 Generation of pseudotyped viruses

Pseudotyped viruses were produced by transient transfection of HEK 293-T cells with three plasmids: RTV containing patient' related Gag and/or protease sequences or wild type HIV; pCSFLW containing the HIV packaging sequencing and the luciferase encoding gene and PMDG expressing vesicular estomatitis G protein. Briefly, HEK 293-T cells were seeded so that 10 cm dishes were just sub-confluent on the day of transfection. 18 µl of FuGENE-6 (Roche Diagnostics, US) was added to 200 µl of Opti-MEM medium (Invitrogene). 1.5 µg of pCSFLW, 1 µg of PMDG and 1 µg of P8.9NSX HIV Gag-pol expression vector were made up to 15 µl of TE buffer and added to the Fu-GENE-6 and Opti-MEM mixture. The

transfection mixture was incubated for 30 minutes at room temperature before being added dropwise to the sub-confluent HEK 293-T cells in 8 ml of fresh DMEM medium (Invitrogene) supplemented with 10% FCS (Biosera, UK), 100 U/ml penicillin and 100 µg/ml streptomycin (Invitrogene). HEK 293-T cells and transfection mixture were incubated overnight at 37°C and 5% CO_2. The following day, the cell culture medium was replaced for fresh medium. The pseudovirus containing supernatants were either directly employed in drug susceptibility assays or harvested at 48 and 72 hours, filtered with a 0.45 µm filter to eliminate cell debris and stored at -80°C in 1ml aliquots for subsequent applications.

2.2.4.3 Protease inhibitor susceptibility assay

Pseudovirus stocks used for PI susceptibility testing were obtained by co-trasfecting HEK 293 T cells with RTV, pCSFLW and PMDG plasmids as described above. The cells were tripsinized approximately 16 hours after transfection and distributed into 96-well plates containing serial dilutions spanning and empirical determined range for each PI (between 1000 nm -0.005 nm). Pseudoviral stocks generated in the presence of PIs were harvested at around 48 hours after transfection and employed to infect fresh HEK 293 T cell cultures in 96-well plates in the absence of drug. Replication was monitored by measuring luciferase expression in infected target cells at approximately 48 hours after infection. Luciferase expression was measured using Steady Glo and a Glomax Luminometer (both Promega). Data were analyzed by plotting the percent inhibition of luciferase activity *versus.* $\log_{10}$ drug concentration. The percent inhibition was derived as follows: [1-(luciferase activity in the presence of drug-background)/ (luciferase activity in the absence of drug-background)] x 100. Mean percent inhibition for each drug concentration was determined from independent

replicates and the deviation standard calculated. Inhibition curves, defined by the four-parametric sigmoidal function $f(x) = a- [b/ (1+(x/c)^d]$, were fit to the data by nonlinear least-squares and used to calculate the drug concentration required to inhibit virus replication by 50% (IC_{50}). The fold change (FC) in drug susceptibility is determined by comparing the IC_{50} for the tested virus to the IC_{50} of the WT reference virus (P8.9NSX) which contains the PR and RT sequences of the NL4-3 strain of HIV-1. All analysis was performed employing GraphPad PRISM version 5.

2.2.4.4 Replicative capacity assay

Pseudovirus stocks are prepared by co-transfecting HEK 293T cells with RTVs, pCSFLW and PMDG plasmids as described above. Cells and transfection mix are incubated for 48 hours in DMEM medium supplemented with 10% FCS, 100 U/ml penicillin and 100 µg/ml streptomycin. The medium will be replaced with fresh medium every 24 hours. Pseudo-virus supernatant will be filtered with a 0.45 µm filter to eliminate cell debris and 100 µl will be employed to infect fresh HEK 293T cells in 96 well plates. Thus, pseudo-virus containing supernatant will be titrated along the HEK 293- cells-containing 96 well plates. Replication will be monitored by measuring luciferase expression 48 hours after infection. Luciferase activity was determined with steady Glo and a Glomax Luminometer (both Promega) and expressed relative to the wild type reference virus (P8.9NSX). The relative light luciferase units (RLU) were plotted against both µl of virus supernatant and the ng of P24 produced by virus supernatants and the mean luciferase activity calculated by using at least four values within the linear range. Replicative capacity (RC) of the tested virus would be directly related

to the luciferase activity displayed and would be expressed as a percent of the WT reference virus (P8.9NSX) to which a 100% RC value would be assigned.

2.2.4.5 P24 ELISA

In order to correct for potential transfection efficiency, the relative luciferase units produced by RTVs were expressed normalized by ng of P24 protein. P24 protein in pseudovirus supernatant was measured by employing a twin-site sandwich Enzyme-linked immunosorbent assay (ELISA). The ELISA was performed with reagents supplied by Aalto Bioreagents LTd. Briefly, 96 well plates were coating with Anti-HIV-1-p24 Gag by adding 100 µl per well of affinity purified sheep anti-HIV-1-p24 Gag (D7320, Aalto bio Reagenst Ltd) reconstituted in water at 1 mg/ml. Plates were incubated overnight at room temperature. The following day the plate is washed twice with 200 µl of TBS buffer. P24 antigen is captured from the pseudo-virus containing supernatant which is previously inactivated by treated with Empigen zwitterionic detergent (Sigma Aldrich) at a final concentration of 1% per volume and incubated at 56ºC for 30 minutes. Serial dilutions of pseudo-virus supernatants are made in TBS/Empigen and 100 µl of the dilutions added to the Anti-HIV-p24-Gag-coated well and incubated for 3 hours at room temperature. Unbound p24 was washed away with 2 x 200µl of TBS. Bound p24 was detected by using alkalin- phosphatase-conjugated anti-HIV-1-p24 mouse monoclonal antibody (BC 1071-AP, Aalto Bio Reagenst) and the AMPAK ELISA amplification system. The HIV-1 p24 assay was calibrated using known amounts of a purified recombinant p24 protein (AG6054, Aalto Bio Reagenst). The calibration curved was obtained by plotting the optical density (OD) at 492 nm against the known amount of recombinant p24 protein. Comparison of OD displayed by pseudo-viral

supernatants with that of calibrators allowed us to determine the amount of p24 in pseudo-virus containing supernatants.

Table 2.1 Primers employed for the amplification of HIV-1 Gag and protease genes

Primer name[a]	Sequence 5'-3'	Position[b]	Description
GagFout	GTT GTG TGA CTC TGG TAA CTA GAG ATC CCT CAGA	570-603	Forward outer primer
GagBout	TCC TAA TTG AAC YTC CCA RAA GTC YTG AGT TC	2797-2828	Reverse outer primer
GagFin	TCT CTA GCA GTG GCG CCC GAA CAG	626-649	Forward inner primer
GagBin	GGC CAT TGT TTA ACC TTT GGD CCA TCC	2597-2623	Reverse inner primer
GagFin2	AAA TCT CTA GCA GTG GCG CCC GAACAG	623-649	Forward inner primer
GagBin2	TGG MCC AAA RGT TAA ACA RTG GC	2600-2622	Reverse inner primer

[a]GagFin and GagBin were employed as default inner primers in the nested PCR protocol. GagFin2 and GagRin2 were used when the default primers failed.

[b]Primer position is relative to HXB2 (GenBank accession number K03455).

Table 2.2 Primers employed for the sequencing of HIV-1 Gag and protease genes

Primer name	Sequence 5'-3'	Position[a]	Direction
G00*	GACTAGCGGAGGCTAGAAG	764-782	Forward
G50*	CACAGCAAGCAGCAGCTG	1133-1150	Reverse
G70*	ATGAGGAAGCTGCAGAATGGG	1406-1426	Forward
G01*	AGGGGTCGTTGCCAAAGA	2281-2264	Reverse
G05*	TGTTGGCTCTGGTCTGCTCT	2157-2138	Reverse
G35*	CATGCTGTCATCATTTCTTCTA	1838-1817	Reverse
G45*	TTGGACCAACAAGGTTTCTGTC	1761-1740	Reverse
Ana1	GGG CCA TCC ATT CCT GGC TT	2602-2586	Reverse
Ana2	CAG AGC CAA CAG CCC CAC CAG	2147-2167	Forward
Ana3	ATC KTT CYA GCT CCC TGC TTG	916-899	Reverse
Ana4	GCC ATA TCR CCT AGA ACY TT	1228-1244	Forward
Ana5	GGG ATT AAA YAA AAT AGT AAG	1593-1612	Forward
Ana6	TAG AAG RAA TGA TGA MAG	1820-1834	Forward
Ana7	ATA ATC CAC CTA TCC CAG	1547-1561	Forward
Ana8	GAC ACC AAR GAA GCY TTA	1078-1092	Forward
G85*	TGC ACT ATA GGG TAA TTT TG	1193-1173	Reverse
Ana9	GAT AGG GGG AAT TGG AGG TTT TAT CAA AGT	2390-2419	Forward
Ana10	ATG TTG ACA GGT GTA GGT CCT ACT AAT ACT GTC C	2503-2470	Reverse

Primers G00, G50, G70, G01, G35, G45 and G85 had been previously published (Sanders-Buel, 1995). [a]Primer position is relative to HXB2 (GenBank accession number K03455).

Table 2.3 Primers employed for Site-directed mutagenesis (SDM)

Primer name	Sequence 5'-3'	Description
Y132F-F	CAGGTCAGCCAAAATTTCCCTATAGTGCAGACC	Forward primer for introduction of Y132F
Y132F-R	GTTCTGCACTATAGGGAAATTTTGGCTGACCTG	Reverse primer for introduction of Y132F
T375A-F	CAAATCCAGCTGCCATAATGATACAGAAAGGC	Forward primer for introduction of T375A
T375A-R	GCCTTTCTGTATCATTATGGCAGCTGGATTTG	Reverse primer for introduction of T375A
ΔT375A-F	CCAAGTAACAAATCCAGCTACCATAATGATACAGAAAGGC	Forward primer for reversion of T375A
ΔT375A-R	GCCTTTCTGTATCATTATGGTAGCTGGGTTTGTTACTTGG	Reverse primer for reversion of T375A
V82I-F	CTTAGATCATTATATAATACAATAGCAACCCTCTATTGTGTG	Forward primer for introduction of V82I
V82I-R	CACACAATAGAGGGTTGCTATTGTATTATATAATGATCTAAG	Reverse primer for introduction of V82I
A115I-F	CAAAAGTAAGAAAAAAGTACAGCAAGCAGCAGCTGACAC	Forward primer for introduction of A115I
A115I-R	GTGTCAGCTGCTGCTTGCTGTACTTTTTTCTTACTTTTG	Reverse primer for introduction of A115I
A120S-F	CACAGCAAGCAGCAGTTGACACAGGACACAG	Forward primer for introduction of A120S

A120S-R	CTGTGTCCTGTGTCAACTGCTGCTTGCTGTG	Reverse primer for introduction of A120S
I437V-F	GCTAATTTTTTAGGGAAGGTCTGGCCTTCCCACAAGGG	Forward primer for introduction of I437V
I437V-R	CCCTTGTGGGAAGGCCAGACCTTCCCTAAAAAATTAGC	Reverse primer for introduction of I437V
Y441H-F	GAAGATCTGGCCTTCCCACAAGGGAAGGCCAG	Forward primer for introduction of Y441H
Y441H-R	CTGGCCTTCCCTTGTGGGAAGGCCAGATCTTC	Reverse primer for introduction of Y441H
G443E-F	GATCTGGCCTTCCTACAAGGGGAGGCCAGGGAATTTTTTTCAG	Forward primer for introduction of G443E
G443E-R	CTGAAAAAAATTCCCTGGCCTCCCCTTGTAGGAAGGCCAGATC	Reverse primer for introduction of G443E
ΔY132F-F	CAGGTCAGCCAAAATTACCCTATAGTGCAGAAC	Forward primer for reversion of Y132F
ΔY132F-R	GTTCTGCACTATAGGGTAATTTTGGCTGACCTG	Reverse primer for reversion of Y132F
ΔL449F-F	GGCCAGGGAATTTTCTTCAGAGCAGCC	Forward primer for reversion of L449F
ΔL449F-R	GGTCTGCTCTGAAGAAAATTCCCTGGCC	Reverse primer for reversion of L449F
ΔA431V-F	CTGAGAGACAGGCTAATTTTTTAGGGAAGATCTG	Forward primer for reversion of A431V
ΔA431V-R	CAGATCTTCCCTAAAAAATTAGCCTGTCTGTCTCAG	Reverse primer for reversion of A431V

Table 2.4 Primers employed for modification of P8.9NSX vector by SDM

Primer name	Sequence 5'-3'	Description
BglII-F	TCTCTTTGGCAGCGCTTCCCTCAGAT	Forward primer for introduction of BglII restriction site
BglII-R	CGCTGCCAAAGAGAGGTCTGAGGGAAG	Reverse primer for introduction of BglII restriction site
BamHI-F	CCAGGTATGGATCCCCCAAAAGTTAAACAATGGCC	Forward primer for introduction of BamHI restriction site
BamHI-R	GGCCATTGTTTAACTTTTGGGGGATCCATACCTGG	Reverse primer for introduction of BamHI restriction site
ΔBglII-F	CTAATTTTTTAGGGAGACCTGGCCTTCCCGAAGG	Forward primer for blocking BglII restriction site
ΔBglII-R	CCTTGTGGGAAGGCCAGGTCTTCCCTAAAAAATTAG	Reverse primer for blocking BglII restriction site
ΔBamHI-F	CATTCGATTAGTGAACGGGTCCTTGGCACTTATCTG	Forward primer for blocking BamHI restriction site
ΔBamHI-R	CAGATAAGTGCCAAGGACCCGTTCACTAATCGAATG	Reverse primer for blocking BamHI restriction site

Table 2.5 Primers employed for the introduction of appropriate restriction sites

Primer name	Sequence 5'-3'	Position[a]	Description
GagF-NotI	TCTCTA**GCGGGCCGC**GCAGTGGCGCCCGAACAG	626-649	Forward primer for introduction of NotI site at the beginning of *Gag* gene
GagB-BglII	GGCCAT**AGATCT**TGTTTAACYTTTGGDCCATCC	2597-2629	Reverse primer for introduction of BglII site at the end of *Gag* gene
PRF-BglII	CCC**AGATCT**CACCAGAAGAGAGCTTC	2159-2178	Forward primer for introduction of BglII at the beginning of *PR* gene.
PRB-BamHI	GGG**GGATCT**CCATCCATTCCTGGCTT	2585-2604	Reverse primer for introduction of BamHI at the end of *PR* gene.

[a]Primer position is relative to HXB2 (GenBank accession number K03455).

The above primers are employed for nested amplification of patient's samples in preparation for cloning into the P8.9NSX HIV-1 expression vector. Cloning of Gag gene will require amplification with GagF-NotI and GagB-BglII, cloning of PR gene will require amplification with PRF-BglII and PRB-BamHI and cloning of both Gag and PR simultaneously will require amplification with GagF-NotI and PRB-BamHI. Restriction sites are shown in bold.

3 Chapter three: development and optimization of an assay for the amplification and sequencing of full-length HIV-1 Gag and protease genes

3.1 Introduction

Since the discovery of HIV as the causative agent of acquired immunodeficiency syndrome (AIDS), research efforts have led to the development and clinical use of several drugs aimed at inhibiting the viral replication cycle at particular critical stages. Six classes of antiretroviral drugs are currently approved that target the fusion between viral and cellular membranes, the reverse transcription of viral RNA into cDNA, the integration of viral DNA into the host genome and the maturation of newly synthesized virions.

In 1996-1998, the combination of three antiretroviral drugs, which is known as highly active antiretroviral therapy (HAART), revolutionized the care of patients infected with HIV leading to a dramatic decrease in mortality and morbidity. However, inadequate treatment (e.g., lack of complete adherence to therapy or insufficient potency of some regimens) may produce incomplete viral suppression, which often results in the appearance of viral resistance and as a consequence in therapy failure.

The clinical relevance of HIV drug resistance is well established. The emergence of resistance to first-line HAART might not only facilitate failure to subsequent line of therapy due to cross resistance, but is also associated with an increased risk of death, particularly if all three major classes of drugs (nucleoside or nucleotide reverse transcriptase inhibitors (NRTIs), non-nucleoside reverse transcriptase inhibitors (NNRTIs) and protease inhibitors (PIs)) are involved (Hogg et al, 2006; Zaccarelli et al, 2005). In most cases, resistance arises from genetic mutations affecting the viral protein targeted by the antiretroviral drug. As a consequence, detection of resistance in clinical practice is usually achieved by sequencing the targeted gene, commonly reverse transcriptase (RT) or protease (PR), followed by the identification of resistance-associated mutations (RAMs).

The classic mechanism of resistance to PIs features the stepwise accumulation of substitutions in the viral protease (Croteau et al, 1997; Johnson et al, 2008b and Nijhuis et al, 1999). The first amino acids changes observed during PI exposure often involve the substrate-binding cleft of the viral enzyme. The mutations that encode these amino acid changes, termed primary mutations, are responsible for the resistance of the virus to PIs. However, these mutations also have a deleterious effect in the replication capacity of the virus as they code for a protease that displays a reduced cleaving activity. Long-term exposure to a non-suppressive PI-containing regimen may lead to the selection of compensatory mutations, whose principal role is to re-establish the original replication capacity of the wild type (WT) virus. Several compensatory mutations have been described, mainly within the protease itself (Eastman et al, 1998; Ho et al, 1994; Mammano et al, 2000; Molla et al, 1996; Nijhuis et al, 2001 and Quinones-Mateu and Arts, 2001;), but also in its natural substrate, the Gag gene (Clavel et al, 2000; Doyon et al, 1996; Miller et al, 2001;

Nijhuis et al, 2001 and Zhang et al, 1997). Most reports have identified the Gag C-terminal cleavage sites (CS) P7/P1 and P1/P6 as major mutation hot spots (Doyon et al, 1996; Malet et al, 2007; Nijhuis et al, 2007; Verheyen et al, 2006 and Zhang et al 1997). However, this observation may be biased as most studies have only targeted these cleavage sites. In fact, there have been anecdotal reports that describe mutations at other CS or even outside the CS areas in association with exposure to PIs (Gatanaga et al, 2002), which imply that such regions may also have a role in the development of PI resistance. Moreover, some groups have suggested that mutations in Gag outside CSs could be responsible for primary resistance to PIs (Gupta et al, 2010; Nijhuis et al, 2007 and Parry et al 2009). Taken together, these observations suggest that resistance to the PIs may be underestimated in routine clinical practice since only primary mutations in the protease are searched for.

The objective of the study presented in this chapter was to develop an assay for the co-amplification and sequencing of full-length HIV-1 Gag and protease genes in order to evaluate the contribution of Gag mutations to drug susceptibility and viral fitness, and investigate linkage between Gag and protease mutations in patients failing PI-based therapy. More specifically, this encompassed: 1) the design and selection of PCR and sequencing primers; 2) the optimization of the PCR reaction for full-length HIV-1 Gag and protease genes and 3) sequencing optimization for full-length HIV-1 Gag and protease genes.

3.2 Methods

3.2.1 Samples

Plasma samples derived from HIV-1 infected patients attending the HIV services at the Royal Free Hospital were used for this study. The HIV-1 subtype was determined by submitting polymerase (*pol)* sequences to the NCBI HIV-1 genotyping tool. The plasma HIV-1 RNA load ("viral load") was measured with the RealTime HIV-1 assay (Abbott Molecular, USA). A selection of samples representing a wide range of HIV-1 subtypes and circulating recombinant forms (CRFs) was employed for the study.

3.2.2 Primers

A total of 1,400 HIV-1 sequences from all available group M subtypes and CRFs were downloaded from the Los Alamos National Laboratory Database (www.hiv.lanl.gov) and aligned using Mega Molecular Evolutionary Genetic Analysis software version 4.0. Sequences that were relatively well conserved across different subtypes and close to the Gag and protease region of interest (Gag-Pr) were selected for primer design. PCR and sequencing primers were constructed employing Oligo software V7.0 and synthesized by Invitrogene (Invitrogene, UK).

3.2.3 RNA extraction

One milliliter of plasma was centrifuged at 25,000 g for one hour at 4°C to concentrate the virus. The supernatant was then removed and the pellet re-suspended to a final volume of 280 µl. The re-suspended pellet was used for nucleic acid extraction; two extraction methods were employed according to the manufacturer's instructions, the manual QIAamp Viral RNA MiniKit (Qiagen, Germany) and the semi-automated EasyMag (Nuclisens, France). Finally, the RNA was eluted into 55µl of buffer and stored at -80°C. The performance of the assay with both extraction methodologies was compared.

3.2.4 Amplification of Gag-protease region

A reverse-transcription-nested (rt-nested) PCR protocol was designed for the amplification of the HIV-1 Gag-Pr region. The procedure was divided in two stages:

Stage 1 encompassed two sequential reactions, an initial reverse transcription of the viral RNA into cDNA followed by a PCR reaction leading to the generation of a 2.2Kb DNA fragment; the Qiagen one-step RT-PCR kit (Qiagen, Germany) used for this purpose contains Omniscript® / Sensiscript® reverse transcritptase and HotStar® Taq DNA polymerase.

Stage 2 comprised a PCR reaction on the 2.2 Kb DNA fragment which led to the production of a 2.0 Kb DNA fragment. This stage was carried out using the HotStar® Taq DNA polymerase kit (Qiagen, Germany).

All experiments were initially performed following the manufacturer's instructions. However, subsequent modifications to different parameters affecting the development of reverse transcription and nested PCR reactions, such as annealing temperature, concentration of Magnesium [Mg^{2+}], concentration of desoxiribonucleotides [dNTPs], concentration of primers [primers], type of reverse transcriptase (AccuScript reverse transcriptase (Stratagene, Netherlands), SuperScript® III reverse transcriptase (Invitrogene, UK) and type of DNA polymerase (Pfu Ultra DNA polymerase (Stratagene, Netherlands), Platinum®TaqDNA Polymerase HF(Invitrogene, UK) were tested in order to optimize the yield.

The optimized reverse transcriptase-nested PCR protocol was validated with a variety of HIV-1 group M and CRFs.

3.2.5 Sequencing of the Gag-protease region

PCR products were purified using Microcon YM-100 centrifugal filters units (Millipore, UK) according to the manufacturer's instructions. Population sequencing was performed using the ABI PRISM BigDye Terminator v3.1 ready reaction cycle sequencing kit and reactions were run on an ABI3100 Genetic analyzer.

3.2.6 Cloning of PCR products

Clonal analysis was performed on samples were population sequencing was not conclusive. For this purpose, the 2.0 Kb PCR product encompassing the sequence of the HIV-1 Gag-Pr region was cloned into a plasmid vector (PCR2.1) using the TOPO TA Cloning Kit (Invitrogene, UK). Positive clones, identified as those harbouring inserts of the correct size after restriction enzyme digestion, were isolated and sequenced as previously described.

3.2.7 Sequence analysis

Sequences were analyzed using Sequence analysis version 5.0, Seqscape version 6.0 and Mega Molecular Evolutionary Genetic Analysis software version 4.0 programmes.

3.3 **Results**

3.3.1 Primers

Multiple sets of potential primers for amplification and sequencing were obtained by OLIGO software and their characteristics were examined in order to choose the most suitable pairs.

Selection criteria included:

- GC content >50%.
- No obvious tendency to form secondary structures.
- No complementary regions between primers of the set.
- Lack of homology with other sequences on either strand of the HIV-1 genome.
- Difference in melting temperature between forward and reverse primers < 5°C.

The region targeted by forward amplification primers was located within the 5'LTR region, which is present upstream of the Gag gene. As this sequence was relatively conserved across different HIV-1 strains only a single set of primers was designed for each of the reactions comprising the nested PCR. By contrast, no conserved region located downstream of Gag and protease genes was identified, so reverse amplification primers, consisted of a mixture of several primers whose sequences differ at certain positions (i.e., degenerated primers), were constructed.

We selected the following sets of primers for the amplification of HIV-1 Gag-Pr by nested PCR:

- A sole set of outer primers derived from an alignment of all group M HIV-1 subtypes and CRFs.

- Two sets of inner primers: a main set that was chosen based on an alignment constructed with the most prevalent subtypes found in the Royal Free Hospital population (i.e., subtypes B, C, A, D and CRF02) and an accessory set that was obtained from an alignment of all other group M HIV-1 subtypes and CRFs and was employed when the main set of primers did not render a PCR product.

Similarly, we designed several primers for sequencing purposes. A total of 18 primers were employed in different combinations, usually 8 to 10, in order to obtain a full-length Gag-Pr sequence. The location of the primers and the amplicon length for each set of amplification primers are indicated in Figure 3.1. Primer sequences for amplification and sequencing are given in chapter 2, tables 2.1 and 2.2, respectively.

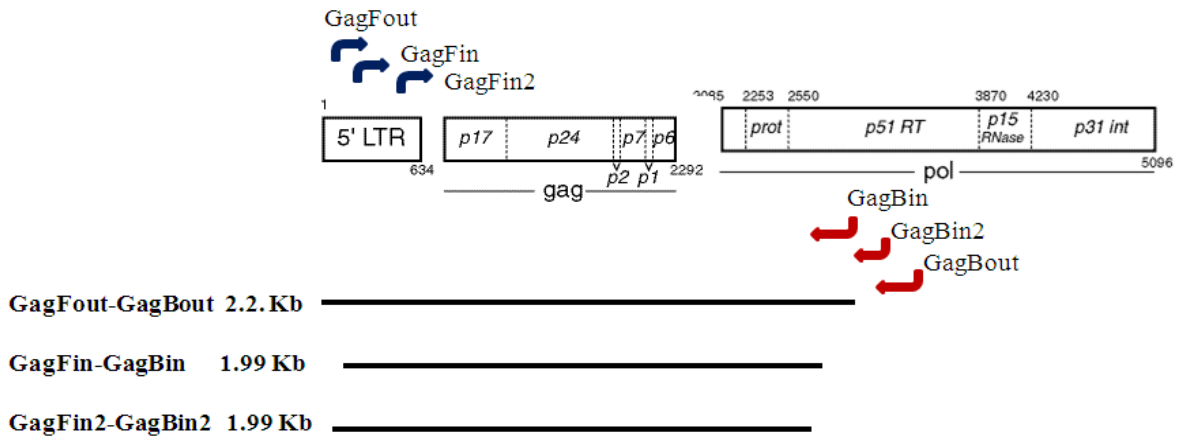


Figure 3.1 Location of amplification primers in HIV-1.

The length of the amplicons for different sets of primers is indicated.

3.3.2 Reverse transcription and amplification of Gag-protease region

The Qiagen one-step RT-PCR kit was used for reverse transcription of viral RNA and first PCR run. The kit consisted of a single enzyme mix, which contained Omniscript® and Sensiscript® reverse transcriptase and HotStar® Taq DNA polymerase.

The initial experiment was conducted under the conditions recommended by the manufacturer: 1X buffer containing 2.5 mM of Mg^{2+}, 400 µM of each dNTP, 0.6 µM of both forward and reverse primers and 2.5 units of RT-PCR enzyme mix. Reverse transcription occurred for 30 minutes at 50°C. By heating at 95°C for 15 minutes, the reverse transcriptase was inactivated and simultaneously the DNA polymerase was activated. Subsequent PCR cycling conditions included a denaturation step of 2 minutes at 95°C; 40 cycles of 15 seconds

at 95°C; 30 seconds at the annealing temperature calculated by OLIGO software for the set of primers (56°C) and 2 minutes and 30 seconds extension at 68°C and followed by a final extension step of 10 minutes at 68°C.

The experiment included HIV-1 positive (subtype B) and negative controls. The positive control was diluted at different concentrations prior to extraction.

The results of these preliminary experiments are shown in figure 3.2. Successful amplification was demonstrated by visualization of a 2.2 Kb PCR product on an agarose gel electrophoresis. A patent band of the expected size was present when viral loads were between 3,000 and 100,000 copies/ml. A very faint band could be appreciated at viral loads around 1,000 copies/ml. No amplification was detected when the viral load was 500 copies/ml. However, strong non-specific amplification was evident on the agarose gel at all tested viral loads.

As amplification of the target was achieved, the primer pair employed was considered appropriate. However, it was deemed that the sensitivity and specificity of the method should undergo further optimization. Parameters considered for optimization included PCR cycling conditions, particularly annealing temperature as well as $Mg^{2+,}$ dNTPs and primer concentrations.

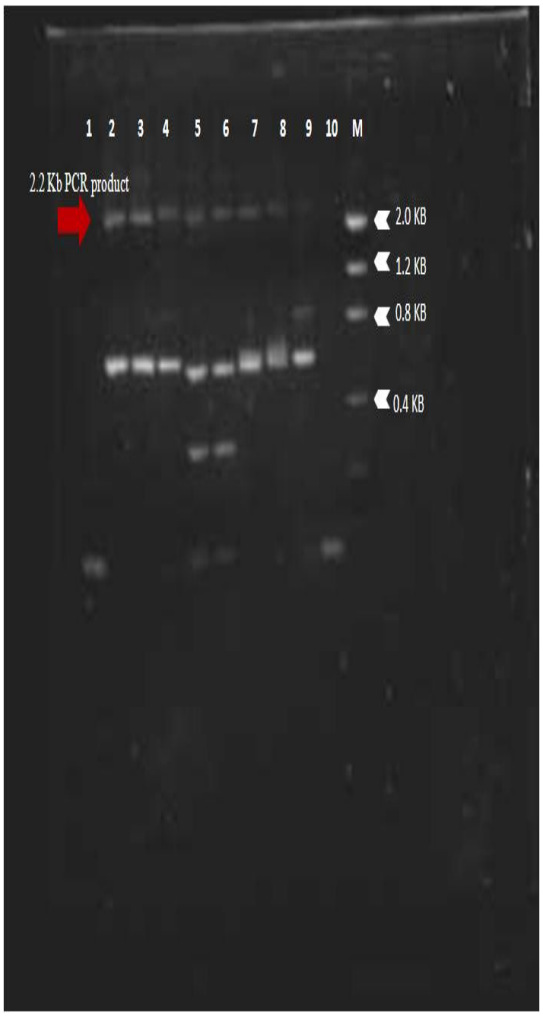


Figure 3.2 Agarose gel electrophoresis following RT and amplification of HIV-1 Gag-pr region. The band numbers from 1 to 10 on the gel correspond to different viral load of the dilutions tested : 1(negative control); 2 (100,000 cp/ml); 3 (50,000 cp/ml); 4 (25,000 cp/ml); 5 (12,500 cp/ml); 6 (10,000 cp/ml); 7 (5,000 cp/ml); 8 (3,000 cp/ml);9 (1,000 cp/ml); 10 (500 cp/ml) and M correspond to low DNA Mass ladder (Invitrogene, UK).

3.3.2.1 Optimization of annealing temperature

Optimization of the annealing temperature started by calculating the melting temperature (Tm) for the primer-template pairs with the OLIGO software v7.0. As recommended for most PCR applications, the annealing temperature (Ta) was preliminarily set 5°C below the calculated Tm, which was 57°C. The optimal Ta was finally determined by performing temperature gradient PCR studies within the range 52-62°C, employing Qiagen one Step RT-

PCR system. This was achieved by programming a 10°C gradient and setting the Ta to 57°C. Other PCR parameters such as [Mg^{2+}], [dNTPs], primers and target concentrations remained constant. Results are presented in Figure 3.3.

The amplification of the desired target (2.2 Kb product) was successful when the annealing temperature was between 55°C and 60°C, but strong non-specific reactivity was detected on the agarose gel. Below 55°C only non-specific amplification was observed. At annealing temperatures above 60°C there was no evidence of amplification on the agarose gel. The strongest amplification of the target was observed at 57 °C; therefore, this annealing temperature was maintained constant on subsequent experiments, while other parameters were modified for further optimization.

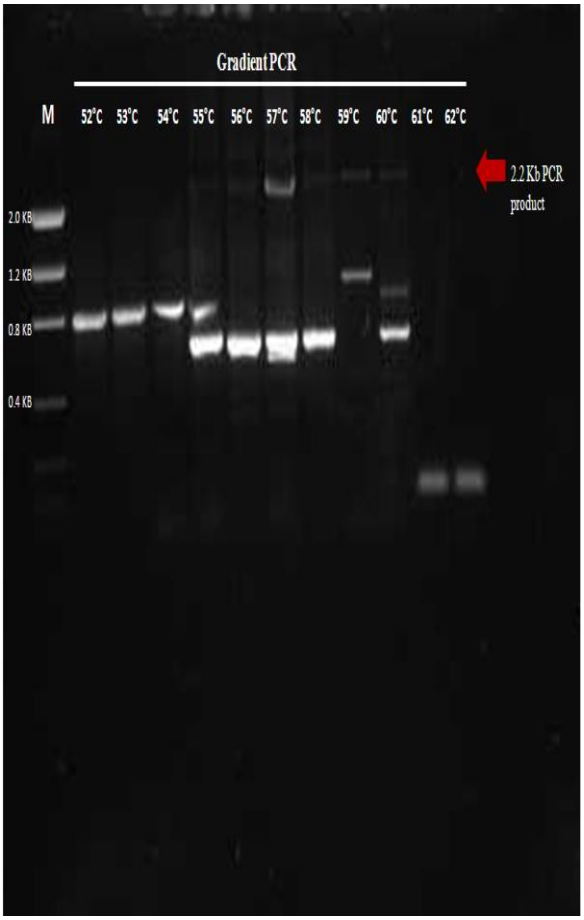


Figure 3.3 Agarose gel electrophoresis of the temperature gradient PCR experiments.

A subtype B HIV-1 positive control of 20,000 copies/ml was extracted manually with QiaAmp Viral RNA extraction kit and subjected to a temperature gradient PCR with Qiagen one step RT-PCR. Band M correspond to Low DNA Mass ladder (Invitrogene, UK) and the different annealing temperatures tested are shown.

3.3.2.2 Optimization of magnesium concentration

The optimal [Mg^{2+}] was determined empirically by setting a series of experiments between 0.5 and 5.0 mM of Mg^{2+}. Results are shown in figure 3.4. As demonstrated in the agarose gel electrophoresis, a minimum [Mg^{2+}] of 2.5 mM was required in order to achieve amplification of the target. Above this concentration, a faint band of the expected size was visualized on the gel and specificity and sensitivity of the PCR protocol did not significantly change within the range 2.5- 5 mM [Mg^{2+}]. Because an increase in the [Mg^{2+}] may have a detrimental effect on

the fidelity of the Taq DNA polymerase and because no difference in sensitivity and specificity of the PCR reaction was detected above 2.5mM of the cation, we considered this concentration appropriate for further experiments and no adjustment of [Mg^{2+}] and [dNTPs] with respect to those recommended by the manufacturer were judged necessary.

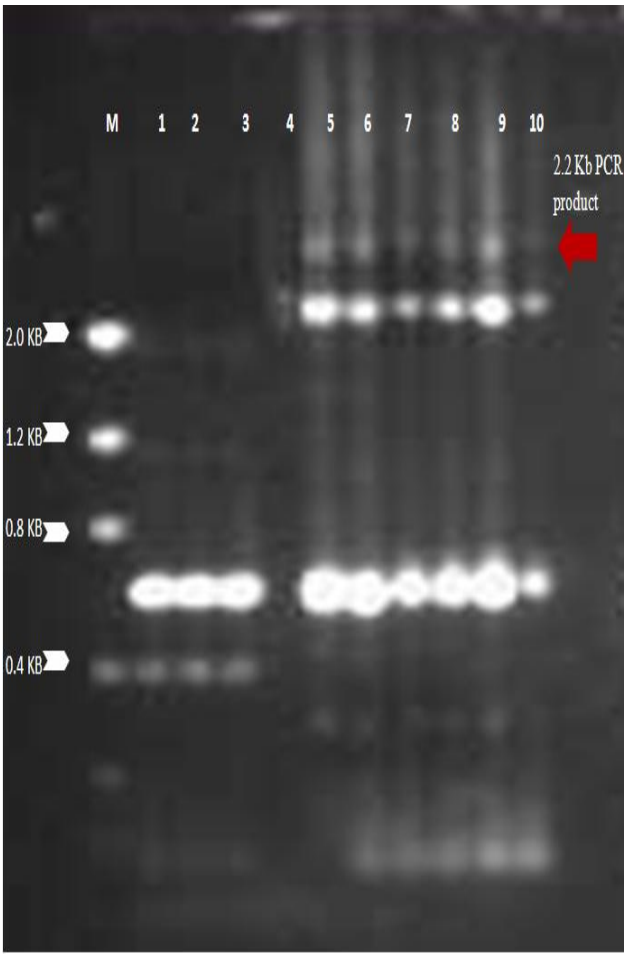


Figure 3.4 Agarose gel electrophoresis of [Mg^{2+}] titration experiments.

Optimization was carried out with a subtype B HIV-1 positive control of 20,000 copies/ml and Qiagen one step RT-PCR system. 400 µM of each dNTPs was included in all reactions. Band number from 1 to 10 correspond to 10 different PCR reactions with Mg^{2+} concentrations of (1) 0.5 mM; (2)1.0 mM; (3) 1.5 mM; (4) 2.0 mM; (5) 2.5 mM; (6) 3.0 mM; (7) 3.5 mM; (8) 4.0 mM; (9) 4.5 mM and (10) 5.0mM, respectively. Band M correspond to low DNA Mass ladder (Invitrogene, UK).

3.3.2.3 Optimization of primer concentrations

Most PCR applications give suitable results with primer concentrations within 0.1-1.0 µM. The use of the lowest concentrations of primers favours specific amplification and generally requires at least 30 cycles of amplification for a 1Kb segment. Longer templates, however, necessitate higher primer concentrations, which can result in mispriming and ultimately non-specific amplification. The Qiagen system recommends employing specific primers at a 0.6 µM concentration. Primer titration experiments were carried out in order to assess the optimal primer concentration. According to the agarose gel electrophoresis presented in figure 3.5, a 0.4 µM concentration of primers seems to be sufficient for amplification. At lower concentrations, no PCR product was detectable on the gel. By contrast, increasing primer concentrations led to substantial non-specific reactivity. Because keeping primer concentrations low is likely to have a positive effect on PCR specificity as well as fidelity, we decided to select the lowest concentration of primers at which amplification of the target was visible on agarose gel electrophoresis;(0.4 µM); this concentration was used in subsequent experiments.

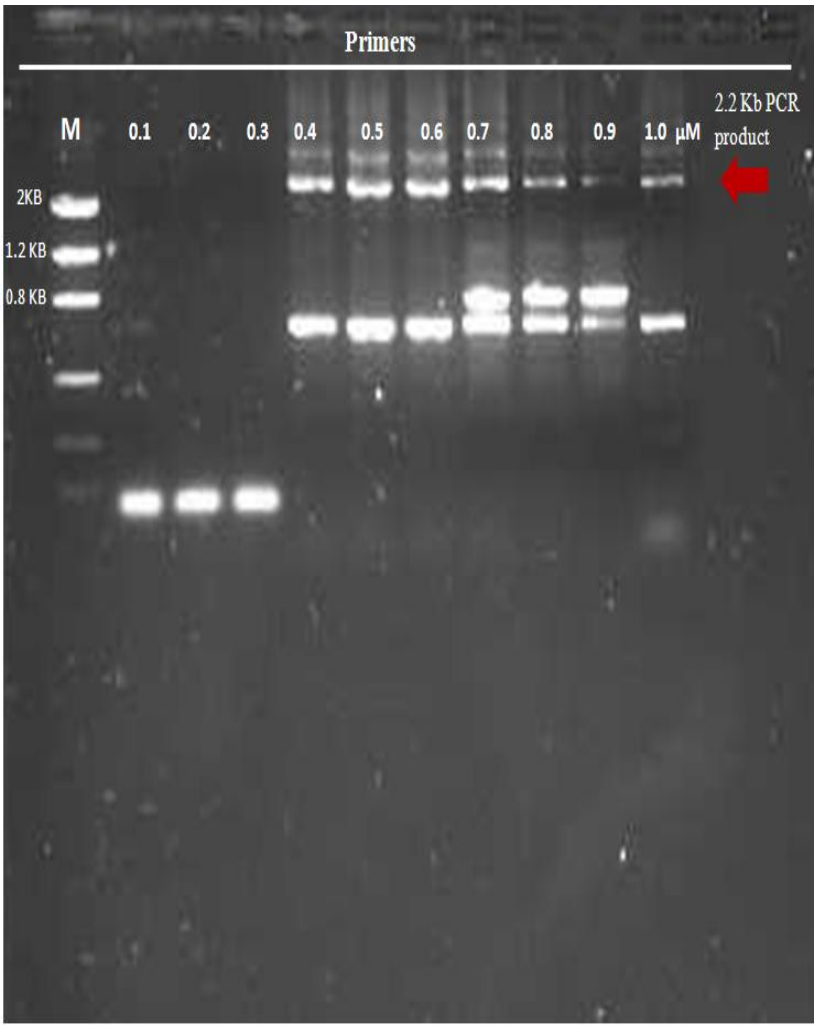


Figure 3.5 Agarose gel electrophoresis of primer titration experiments.

Amplification was performed after manual extraction with QiaAmp Viral RNA extraction Kit (Qiagen) of a HIV-1 subtype B positive control, employing Qiagen one step RT-PCR system. Primer concentrations between 0.1 µM and 1mM were tested and the band numbers from 1 to 10 on the gel correspond to the different primer concentrations: 1 (0.1µM); 2 (0.2µM); 3 (0.3µM); 4 (0.4µM); 5 (0.5µM); 6 (0.6µM); 7 (0.7µM); 8 (0.8µM); 9 (0.9µM) and 10 (1mM) and band M corresponds to low DNA Mass ladder (Invitrogene, UK).

3.3.2.4 Touchdown and nested PCRs

Previous experiments demonstrated that using the Qiagen one-step RT-PCR with the selected set of primers successfully generated a 2.2 Kb product comprising HIV-1 gag and protease genes. However, there were two main drawbacks to the procedure:

- Weak detection of target band on agarose gel electrophoresis, even after optimization of the RT-PCR conditions.
- Variable amount of non-specific reactivity.

Two strategies were envisaged to correct these defects:

1) Touchdown PCR

Touchdown PCR may be used to reduce non-specific amplification. A high annealing temperature is maintained during the initial PCR cycles favouring specific priming, but is decreased as the reaction progresses in order to facilitate amplification. This ensures initial selective generation of the target amplicon, so that it will be in a favorable position to outcompete any lagging PCR product during the remaining cycles. To this end, the first 3 PCR cycles were set at an annealing temperature of 60°C, which was the maximum annealing temperature at which we observed amplification in the temperature gradient PCR experiments, whereas the remaining 37 PCR cycles were maintained at the previously determined optimal Ta of 57°C.

2) Nested PCR

Nested PCR may be used to increase both specificity and sensitivity. Nested PCR is expected to reduce contamination in the PCR product resulting from amplification of unexpected primer binding sites. It is highly unlikely that any unwanted PCR products contain binding sites for both set of primers, ensuring that the product from this second PCR run has little contamination from unwanted PCR products of primer dimers, hairpirins and alternative primer targeted sequences. In addition, by conducting a second run of amplification, the sensitivity of the target amplification is also expected to increase. Therefore, a volume of two µl of the first PCR run product was subjected to an additional PCR run with the designed inner set of primers. Hotstar Taq DNA polymerase (Qiagen, Germany) was employed and we maintained the $[Mg^{2+}]$, [dNTPs] and primer concentrations determined for the first PCR run. The optimal annealing temperature for the inner PCR primers was calculated by oligosoftware and confirmed by setting temperature gradient experiments as previously described. Final cycling conditions consisted of a denaturation step of 2 minutes at 95°C; 3 cycles of 15 seconds at 95 °C, 30 seconds at 60°C and 2 minutes and 30 seconds at 68°C followed by 27 cycles of 15 seconds at 95°C, 30 seconds at the primer annealing temperature (55°C) and 2 minutes and 30 seconds at 68 °C and a final extension step of 10 minutes at 68°C.

Results can be seen in Figure 3.6. Touchdown and nested PCRs significantly reduced non-specific amplification and increased the sensitivity of the amplification reaction as demonstrated by the presence of distinct bands of the expected size (2.0 Kb) for all viral load levels tested (500-100,000 copies/ml).

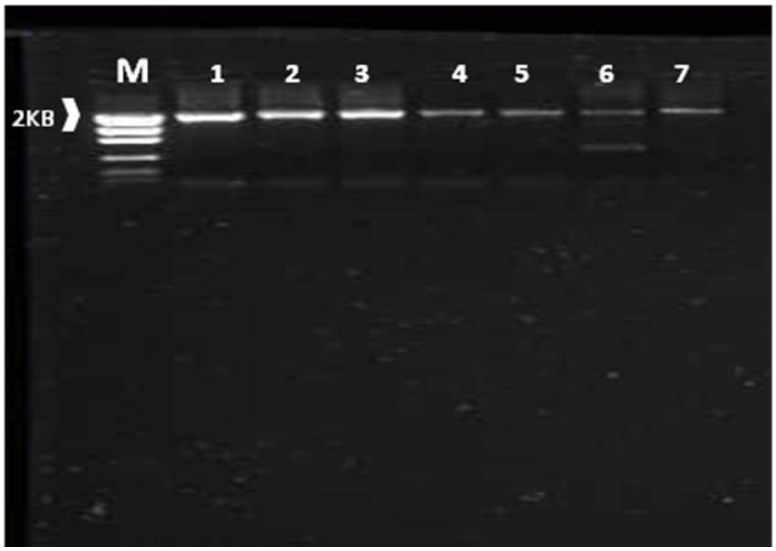


Figure 3.6 Agarose gel electrophoresis of nested and touchdown PCR.

Amplification was performed after manual extraction with QiaAmp Viral RNA extraction Kit (Qiagen) of a HIV-1 subtype B positive control. The RNA extract was subjected to reverse transcription and 1st run PCR and 2 µl of the 1st run PCR product was subjected to a second PCR. The eight bands presented on the gel correspond to the Mass ladder (M) and positive control at concentrations 100,000 cp/ml (1); 50,000 cp/ml (2); 25,000 (3); 10,000 cp/ml (4); 3,000 cp/ml (5); 1,000 cp/ml (6) and 500 cp/ml (7), respectively.

3.3.3 Evaluation of the RT-nested PCR protocol

RT-nested PCR conditions were as depicted in detail in section 2.2.2.2.1 of chapter 2 and briefly summarized below:

- Concentration of reagents, primers and enzyme for RT-PCR: 1 X Qiagen RT-PCR buffer, 400 µM dNTPs, 0.4 µM forward and reverse primers (GagFout, GagBout), 10µL of RNA extract and 2.5 Units of Qiagen RT/PCR enzyme mix.

- RT-PCR conditions: temperature and time:
    - Reverse transcription of viral RNA into cDNA: 30 minutes at 50°C.
    - Inactivation of reverse transcriptase and activation of DNA polymerase: 15 minutes at 95°C
    - 1st PCR run: 2 minutes at 95°C; 3x { 15 seconds at 95°C, 30 seconds at 60°C and 2:30 minutes at 68°C}; 37x {15 seconds at 95°C, 30 seconds at 57°C, 2:30 minutes at 68°C} and 10 minutes at 68°C.

- Concentration of reagents, primers and enzyme for 2nd PCR run : 2µ of 1st run PCR product, 1X Qiagen PCR buffer, 400 µM dNTPs, 0.4 µM forward and reverse primers (GagFin, GagBin) and 2 units of Qiagen Hotstar Taq polymerase.
- 2nd PCR run conditions: 2 minutes at 95°C; 3x {15 seconds at 95°C, 30 seconds at 60°C, 2:30 minutes at 68°C}; 27x {15 seconds at 95°C, 30 seconds at 55°C, 2:30 minutes at 68°C} and 10 minutes at 68°C.

The optimized RT-nested PCR protocol was tested against a panel of different HIV-1 group M subtypes and CRFs. The limit of detection for each subtype was determined by performing serial dilutions. The results of the evaluation are summarized in table 3.1. Amplification of HIV-1 Gag and protease genes was achieved in 21/28 (75%) samples. The majority of these samples (16/21, 76%) were amplifiable at viral load levels ≤ 1,000 copies/ml. However, 5 samples (1 HIV-1 subtype G; 2 CRF02; 1 CRF14 and 1 CRF06) required viral loads > 1,000

copies/ml in order to attain amplification. Samples that failed to amplify included 3 HIV-1 subtype G; 1 subtype J; 2 Complex mosaic HIV-1 sequences and 1 CRF13.

Amplification using an accessory set of inner primers (GagFin2 and GagBin2) was attempted in samples that did not amplify at all, or that did not amplify at viral loads below 1,000 copies/ml. All of the samples amplified (Figure 3.7) and serial dilutions of these samples demonstrated amplification at viral loads below 1,000 copies/ml for all them (Table 3.2).

Sample	Subtype	Viral load (copies/mL)	Last dilution positive by PCR (viral load copies/mL)	First dilution negative by PCR (viral load copies/mL)
B-1	B	123,237	1:200 (600)	1:400 (300)
B-2	B	200,237	1:400 (500)	1:800 (250)
B-3	B	1,075	1:2 (500)	1:4 (250)
C -1	C	130,307	1:200 (650)	1: 400 (300)
C-2	C	61,595	1:800 (750)	1:1600 (350)
C-3	C	16,000	1:200 (800)	1:400 (400)
D-1	D	46,168	1:200 (500)	1:400 (250)
D-2	D	10,015	1:20 (500)	1:40 (250)
D-3	D	867	1:1 (867)	1:2 (435)
A-1	A	9,000	1:16 (600)	1:32 (300)
A-2	A	26,579	1:40 (750)	1:80 (350)
A-3	A	138,783	1: 200 (700)	1:400 (350)
G-1	G	1,236	PCR FAILED	
G-2	G	714,852	1:100 (7000)	1:200 (3500)
G-3	G	7,550	PCR FAILED	
G-4	G	604,206	PCR FAILED	
F-1	F	24,737	1:100 (300)	1:200 (150)
J-1	J	218,794	PCR FAILED	
CRF02-1	CRF02	11,346	1:1 (10,000)	1:2 (5,000)
CRF02-2	CRF02	16,500	1:20 (800)	1:40 (400)
CRF02-3	CRF02	76,000	1:2 (35,000)	1:4 (19,000)
CRF01-1	CRF01	6,831	1:8 (850)	1:16 (425)
Cpx-1	Complex	22,000	1:20 (1,000)	1:40 (500)
Cpx-2	Complex	14,523	PCR FAILED	
Cpx-3	Complex	9,354	PCR FAILED	
CRF13-1	CRF13	824	PCR FAILED	
CRF14-1	CRF14	102,675	1:20 (5,000)	1:40 (2,500)
CRF06-1	CRF06	93,980	1:40 (2,000)	1:80 (1,000)

Table 3.1 Evaluation of the RT-nested PCR protocol (1).

The subtype and viral loads of 28 HIV-1 plasma samples employed in the evaluation are shown. Samples were manually extracted using QiaAmp Viral RNA extraction Kit (Qiagen). Undiluted samples were initially amplified with the optimized RT-nested PCR protocol. Positive samples were subsequently diluted before extraction and amplification to estimate the limit of detection of the assay.

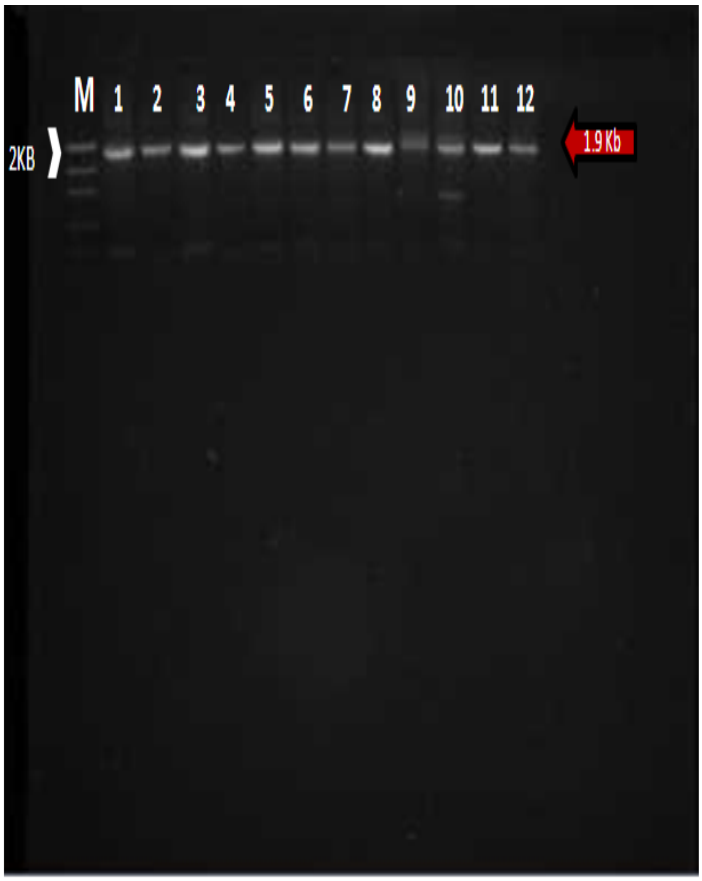


Figure 3.7 Agarose gel electrophoresis of RT-nested PCR.

Amplification of 12 plasma samples that failed to amplify with the initial set of selected primers was performed with the optimized RT-nested PCR protocol employing an accessory set of inner primers (GagFin2, GagBin2). Bands number from 1 to 12 represent different genotypes and viral loads: (1) G, 1,236 cp/ml; (2) G, 7,550 cp/ml; (3) G, 604,206 cp/ml; (4) J, 218,794 cp/ml; (5) Cpx, 14,523 cp/ml; (6) Cpx, 9,354 cp/ml; (7)CRF13, 824 cp/ml; (8) G, 1,236 cp/ml; (9) CRF02, 1,546 cp/ml; (10) CRF02, 76,000 cp/ml; (11) CRF14, 102,675 cp/ml; (12) CRF06, 93,980 cp/ml and band M refers to the low DNA mass ladder (Invitrogene).

Sample	Subtype	Viral load (copies/mL)	Last dilution positive by PCR (viral load copies/mL)	First dilution negative by PCR (viral load copies/mL)
G-1	G	1,236	1:2 (618)	1:4 (309)
G-2	G	714,852	1:800 (875)	1:1600 (440)
G-3	G	7,550	1:10 (750)	1:20 (375)
G-4	G	604,206	1: 800 (750)	1:1600 (375)
J-1	J	218,794	1:200 (1,090)	1:400 (545)
CRF02-1	CRF02	11,346	1:10 (1,134)	1;20 (567)
CRF02-3	CRF02	76,000	1:100 (760)	1:200 (380)
Cpx-2	Complex	14,523	1:20 (726)	1:40 (363)
Cpx-3	Complex	9,354	1:10 (935)	1:20 (467)
CRF13-1	CRF13	824	1:1 (824)	1:2 (412)
CRF14-1	CRF14	102,675	1:100 (1,026)	1:200 (513)
CRF06-1	CRF06	93,980	1:100 (940)	1:200 (470)

Table 3.2 Evaluation of RT-nested PCR protocol (2).

The subtypes and viral loads of the 12 samples amplified with a second set of inner PCR primers (GagFin2, GagBin2) are shown. Samples were manually extracted with QiaAmp viral RNA extraction kit. Undiluted samples were initially amplified and positive samples were subsequently diluted before extraction and amplification to estimate the limit of detection.

3.3.4 Evaluation of the amplification protocol with proofreading enzymes

Both the reverse transcriptase and Taq DNA polymerase employed for amplification of HIV-1 Gag and protease genes lack proofreading activity (3'-5'exonuclease activity) and as a consequence they may not be appropriate for applications where fidelity is paramount. We evaluated the performance of the amplification method with alternative polymerase enzymes. After extensive literature review, we selected AccuScript™ High Fidelity (HF) RT-PCR plus PfuUltra HF DNA Polymerase (Stratagene) and SuperScript III one-step RT-PCR plus Platinum®TaqDNA Polymerase HF (Invitrogene). The Stratagene kit separates reverse transcription and PCR reactions in two different steps: it employs proofreading enzymes for both reverse transcription (AccuScript) and PCR (pfuUltra DNA polymerase). RNA was reverse transcribed into cDNA using the following conditions: 1X AccuScript RT buffer; 10 mM DTT; 1mM dNTPs, 20 Units of RNAse inhibitors and 2 µM of outer reverse primer. Primers and template are incubated at 65°C for 1 hour time after which the reaction is cooled to room temperature, 2 Units of AccuScript RT enzyme are then added and reverse transcription occurs at 42°C for 1 hour. The resulting cDNA is then subjected to nested PCR using the same cycling conditions as for Qiagen. The optimal [Mg^{2+}], [dNTPs] and primer concentrations were determined by running titration experiments as previously described.

The final conditions were : 1st run PCR mix including: 1x Pfu Ultra HF buffer containing 2mM of Mg^{2+}; 200 µM dNTPs; 0.4 µM of forward and reverse primers and 2.5 units of Pfu Ultra DNA polymerase. 2 µl of the 1st run PCR was subjected to a second PCR with 1X Pfu

Ultra HF buffer; 200 µM dNTPs; 0.4 µM of forward and reverse primers, additional Mg^{2+} was required up to 4mM in the final mix and 2.5 Units of Pfu Ultra DNA polymerase.

The Invitrogene kit performs reverse transcriptase and PCR in a single step but in contrast to Qiagen, it employs a mix of a proofreading reverse transcriptase (Pfu) and non-proofreading DNA polymerase (Taq). Cycling conditions for rt-nested PCR were those previously established with Qiagen. As before, optimal [Mg^{2+}], [dNTPs] and primers for Invitrogene system were determined by performing titration experiments. The final conditions for amplification of Gag and protease with Invitrogene were: RT-PCR mix included 1X Invitrogene buffer containing 1.2 mM Mg^{2+} and 200 µM dNTPs; 0.4 µM of forward and reverse primer and 2 unit of RT-PCR enzyme mix. 2µl of the resulting PCR product was subjected to nested PCR with Platinum Taq DNA polymerase (Invitrogene, UK). Nested PCR mix contained 1x PCR buffer, 1.5 mM of Mg^{2+}, 200µM of dNTPs, 0.4µM of forward and reverse primers and 2.5 unit of Platinum Taq HF DNA polymerase.

For a more detailed explanation about reverse transcription and nested PCR protocols (Qiagen, Stratagene and Invitrogene), please refer to section 2.2.2.2 in chapter 2.

Ten plasma samples from patients infected with HIV-1 subtypes B (n=2), C (n=2), D (n=2), A (n=2), CRF02 (n=2) were diluted to achieve a final viral load concentration of 10,000 copies/ml and amplified according to the conditions specified in the three protocols (Qiagen, Invitrogene and Stratagen).

The agarose gel electrophoresis of the 10 plasma samples amplified with the three systems is presented in figure 3.8. The Qiagen and Invitrogene protocols amplified 10/10 (100%) samples, although bands were in general fainter when the Invitrogene kit was employed. By contrast 4/10 (40%) samples failed to amplify when the Stratagene kit was utilized.

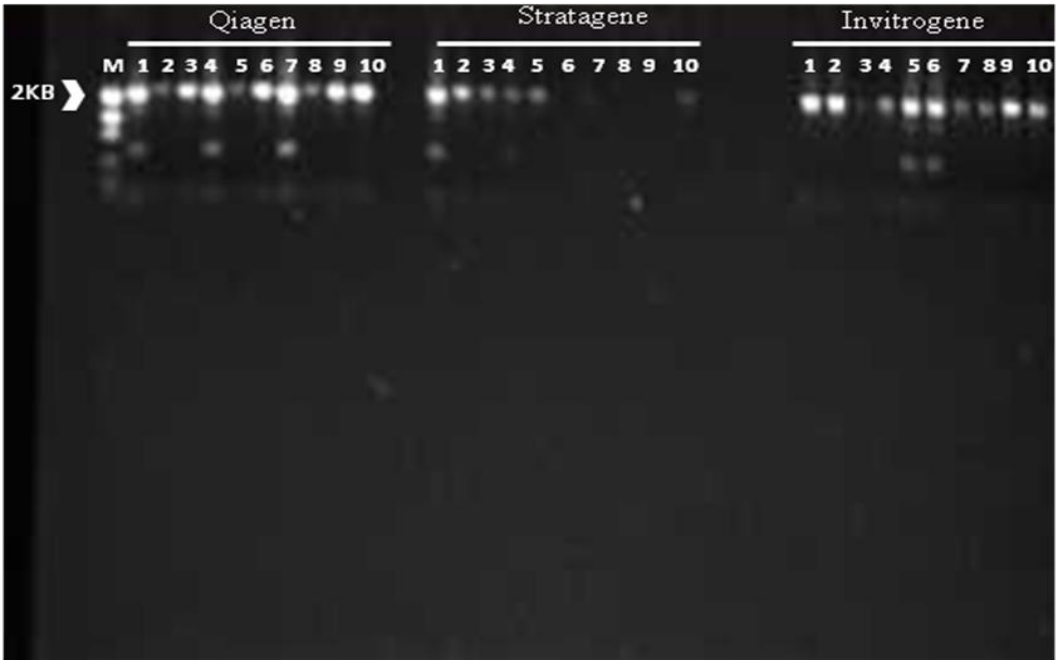


Figure 3.8 Agarose gel electrophoresis of nested PCR with Qiagen, Stratagene and Invitrogene kits.

Ten plasma samples (subtype B, n=2; C, n=2; D, n=2; A, n=2, CRF02, n=2) were tested with the three protocols, which differed in the reverse transcriptase and polymerase enzymes employed. Samples were diluted to a HIV-1 RNA load of 10,000 copies/ml and manually extracted with QiaAmp Viral RNA extraction Kit (Qiagen) before amplification. Band M correspond to low DNA mass ladder (Invitrogene) and bands from 1-10 represents subtypes B (1, 2); C (3, 4); D (5, 6); A (7, 8) and CRF02 (9, 10), respectively.

3.3.5 Manual *vs*. automated extraction

Plasma samples of subtypes A, B, C and CRF02 were diluted to achieve a viral load of 3,000, 1,000 and 500 copies/ml prior to extraction. Extraction was carried out using either a manual

(QiaAmp viral RNA extraction Kit (Qiagen,) or a semi-automated (Nuclisens EasyMag (France)) method. As shown in the agarose gel electrophoresis presented on figure 3.9, all samples extracted with QiaAmp viral RNA MiniKit amplified successfully. However, by using semi-automated extraction most samples with viral load of 1,000 and 500 copies/ml failed to amplify by PCR. The manual extraction method incorporates carrier-RNA that facilitates efficient RNA extraction. In addition, in the manual extraction, samples are previously concentrated by centrifugation at 25,000g at 4°C. Including the centrifugation step before semi-automated extraction may also increase the sensitivity of amplification as demonstrated by visualization of a band on agarose gel electrophoresis at a viral load of 1,000 copies/ml in all samples and in 2/4 (50%) of samples with 500 copies/ml.

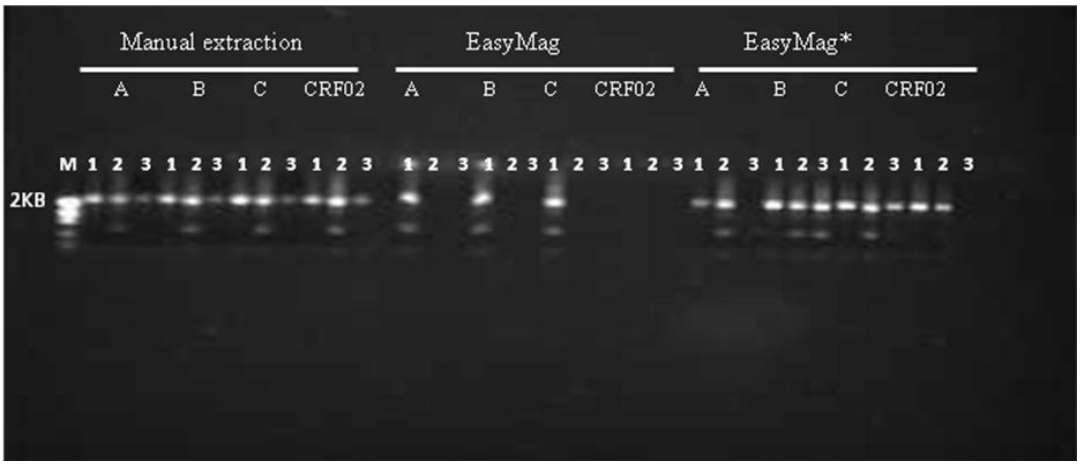


Figure 3.9 Agarose gel electrophoresis of nested PCR after manual, automated or modified automated extraction.

Four samples (subtype A, B, C, CRF02,) were diluted to a viral load of 3,000 1,000 and 500 copies/ml before being extracted manually (QiaAmp Viral RNA extraction kit, Qiagen) or with a semi-automated extractor (EasyMag, Nuclisens). EasyMag* included a high speed centrifugation for 1 hour at 4 °C prior to extraction. Band M corresponds to low DNA mass ladder (Invitrogene) and bands from 1 to 3 represents samples with viral load levels of 3,000 copies/ml; 1,000 copies/ml and 500 copies/ml, respectively.

3.3.6 Sequencing results

Twenty-eight plasma samples belonging to different group M subtypes and CRFs (B, n=3; C, n=3; A, n=3; D, n=3; G, n=4; F, n=1; J, n=1; CRF02, n=3; CRF01, n=1; Cpx, n=3; CRF13, n=1; CRF14, n=1 and CRF06, n=1) were amplified by nested PCR using the Qiagen rt-nested PCR protocol and subsequently sequenced. In addition, 5 samples (subtypes B, n=2; C, n=2 and D, n=1) were amplified in parallel with Stratagene and Invitrogene kits prior to sequencing in order to assess to what extent the use of a non-proofreading enzyme (Qiagen), a proofreading enzyme (Stratagene) or a mix of both affected the sequencing results. Different combinations of sequencing primers were used in order to achieve full-length Gag-PR sequencing. Population sequencing was achieved in 17/28 (61%) samples. The remaining 11 samples produced non-interpretable sequences by population sequencing in P7 (2/11); P6 and P7 (1/11); P6 and P17 (2/11); P6, P7 and P17 (6/11) and clonal analysis was required to obtain full-length Gag-PR sequencing. The primers employed for each sample are detailed in Figure 3.10.

The analysis of amino acid sequences showed no variation between the different DNA polymerases in the 17 samples analysed by population sequencing.

Clonal analysis of 11 samples demonstrated higher genetic variability in the regions where population sequencing was inconclusive compared with other regions of Gag and protease. Thus, mean inter-clone nucleotide variability was between 0.4-3.01 % in P17; 0.05-0.14% in P24; 0.71-8.51% in P7 and 0.08-5.96 in P6 (Table 3.4).

One of the samples requiring clonal analysis was amplified by the Qiagen, Invitrogene and Stratagene protocols. No significant difference was observed in the pattern of nucleotide variations regardless of the amplification protocol used. The most predominant nucleotide changes detected in Gag were A-to-G (23.8%) and G-to-A (16.25%) transitions, followed by T-to-C (15.14%) and C-to-T (13.89%) transitions. Transitions occurred around three times more frequently than transversions. Similarly, no significant difference was observed in the percentage of nucleotide variations among clones obtained after amplification with the three different protocols (Table 3.3).

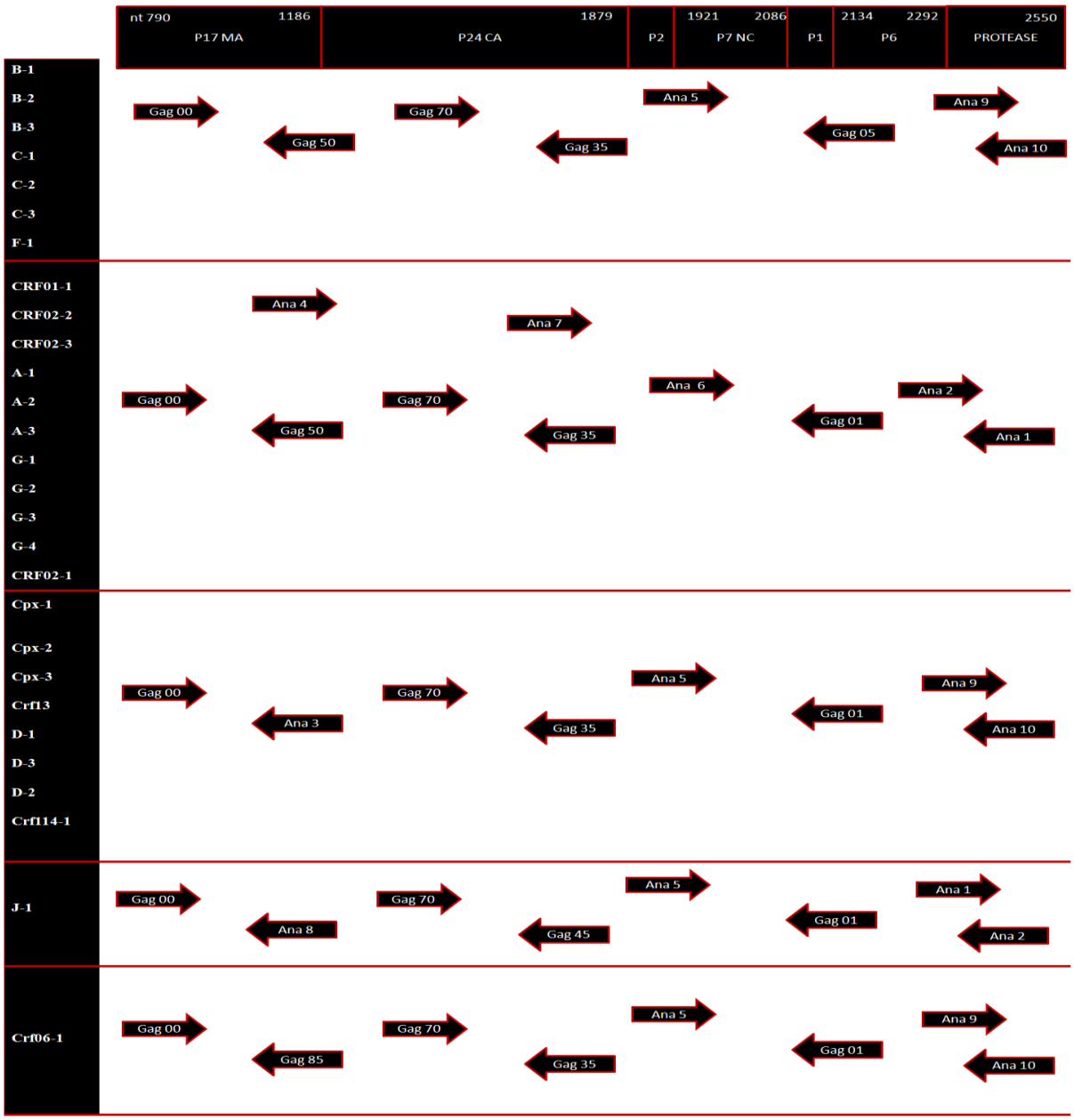


Figure 3.10 Primers employed for full-length Gag and protease sequencing.

A total of 28 plasma samples representing different subtypes and CRFs underwent population sequencing after manual extraction with QiaAmp Viral RNA extraction kit and amplification by nested PCR with the Qiagen system. Samples B-1, B-3, C-1 C-3 and D-2 were in addition amplified in parallel following Invitrogene and Stratagene protocols before sequencing. Samples C-3; CRF01-1; A-3; G-3; G-4; Cpx-1; CRF13-1; D-3; CRF14-1; J-1; CRF06-1 yielded non-interpretable sequences and required cloning prior to successful full-length sequencing.

	Mean % nucleotide variation					
Sample	**Gag**	**P17**	**P24**	**P7**	**P6**	**Region failing population sequencing**
C-3	2.1	3.01	0.05	0.70	3.30	P17, P6
CRF01-1	1.97	0.4	0.08	2.45	0.08	P7
A-3	2.03	0.47	0.12	2.50	0.08	P7
G-3	1.38	1.44	0.14	2.37	1.61	P17, P7, P6
G-4	1.87	0.34	0.13	1.18	5.96	P7,P6
Cpx-1	3.12	1.32	0.05	8.51	2.57	P17, P7, P6
CRF13-1	1.12	2.01	0.06	0.31	2.06	P17, P6
D-3	1.87	0.98	0.13	2.11	4.33	P17, P7, P6
CRF14-1	0.97	1.21	0.03	0.71	2.05	P17, P7, P6
CRF06-1	1.04	0.96	0.05	2.01	1.13	P17, P7, P6
J-1	1.63	2.55	0.05	1.94	1.98	P17, P7, P6

Table 3.3 Nucleotide sequence variation of Gag gene.

Variability across different Gag regions in 11 samples failing population sequencing was determined. Intra-clone distances were calculated using Mega software version 5.0.

	Nucleotide (%)	**Amino acid (%)**
Qiagen	3.29	5.21
Invitrogene	3.08	5.13
Stratagene	2.98	5.02

Table 3.4 Nucleotide and amino acid variability.

The nucleotide variability among 20 clones obtained from one plasma sample was compared employing non-proofreading (Qiagen), proofreading (Stratagene) and a mix of proofreading/non-proofreading (Invitrogene) enzymes. Inter-clone distances were calculated with Mega software version 5.0.

3.4 Discussion

The present chapter describes the design, optimisation and validation of an assay for the amplification and sequencing of HIV-1 Gag and protease. Two main characteristics were to be taken into account when designing the assay:

1) Heterogeneous target population:

Group M HIV-1 viruses, which are further classified into nine subtypes (A-D, F-H, J and K) and at least 51 circulating recombinant forms (CRFs), are responsible for most HIV infections globally. The distribution of these viruses varies by geographical areas. In Western Europe, including the UK, subtype B predominates. However, the prevalence of non-B subtypes in Europe has progressively increased owing to the influx of immigrants from Africa and Asia (Deroo et al, 2002; Machuca et al, 2001; Lospistao et al, 2005; Op de Coul et al, 2001; Snoeck et al, 2004; Thomson and Najera, 2001). Because of historical and current connections with many countries across all six continents, subtypes other than B, namely A, C, D, E, F, G and H, were reported early in the UK (Clewley et al, 1996; Devereux et al, 1999) and by 2001 it was estimated that up to 25% of HIV-1 infections were due to non-B subtypes and CRFs (Barlow et al, 2001). A study published in 2006 based on sequence analysis of the polymerase gene and representing approximately one-fifth of all UK HIV infections, showed that while B was the most common subtype, subtypes C and A were present at prevalence of 10% and 6%, respectively. In addition, all other subtypes as well as several CRFs and unclassified strains were also identified (Gifford et al, 2006).

Consistent with the increasingly diversity of the HIV-1 epidemic in the UK, a wide variety of subtypes are encountered among HIV-infected patients attending the Royal Free Hospital. Approximately half of the HIV-1 infected patients seen locally harbour subtypes other than B, most commonly subtype C followed by subtypes A, D and the recombinant form CRF02. However, all other subtypes as well as other recombinant forms and complex mosaic sequences are also detected (Booth et al, 2007). It should be emphasized that although subtypes other than B were historically linked to immigration (Barlow et al, 2001), more recently an increase has been observed in recent years in the number of UK indigenous population infected with non-B clade HIV-1(Fox et al, 2010). In line with these findings, non-B subtype infection among UK autochthonous population has also been observed within the Royal Free Hospital cohort, (Booth et al, 2007) possibly indicating a higher degree of mixture between UK native and non-native inhabitants. As a consequence of the diversity of HIV strains in our study population, we aimed at developing an assay able to detect a wide range of HIV-1 subtypes and CRFs.

2) Main purpose of the assay:

The primary intended use of the assay was the detection of amino acids changes in HIV-1 *Gag* and *PR* genes in patients failing a PI-based regimen, in order to identify markers of drug resistance associated with changes in these two genes. Genotypic resistance testing is the most convenient method to identify resistance to antiretrovirals as the cause of treatment failure and is recommended in patients experiencing failure of their current regimen and requiring a change in antiretroviral therapy (Hirsch et al, 2008).

Current guidelines for the treatment of HIV infection recommend durable and maximal plasma viral load suppression to <50 copies/ml as the desired outcome after starting antiretroviral treatment. They also indicate that therapy failure should be identified and managed promptly to achieve this outcome (Gazzard et al, 2008; Hammer et al, 2008 and US Department of Health and human services 2008). As a result, genotypic resistance testing is frequently performed early after the onset of treatment failure, on samples still displaying a low viral load. Therefore one key requirement for our assay was to have a high analytical sensitivity.

A number of factors were considered in order to optimize the assay performance characteristics:

1) Type of PCR

Nested PCR was chosen over conventional PCR in order to achieve high analytical sensitivity, which would allow detection of the virus at a low viral load. This PCR strategy significantly improved not only the analytical sensitivity but also the analytical specificity of the amplification protocol.

2) Primer design

Due to the complex mixture of HIV-1 subtypes and CRFs in our study population, good primer design was a critical step to achieve a successful outcome. However, the design of suitable primers for the amplification of the HIV-1 *Gag* gene from different subtypes was more challenging than it would be for other genomic regions, such as pol, due to the higher

genetic variability present within this region (Albert et al, 1994). To address this problem, we proceeded as follows: Firstly, we selected the most conserved sites adjacent to the target sequence to ensure optimal primer binding. Secondly, multiple sets of primers were constructed in order to cover major sequence variations among different strains. Lastly, ambiguity positions were introduced in some primers in order to account for minor genomic diversity at primer binding sites within strains.

3) Nucleic acid extraction methods

Efficient recovery of high-quality intact HIV-1 RNA is fundamental for the success of any RT-PCR-based procedure. A plethora of nucleic acid extraction methods, both manual and automated, are currently available. As it has been reported that HIV genotypic resistance testing achieves higher rates of success when manual extraction techniques are employed (Perandin et al, 2009), we initially selected a manual extraction kit, the QiaAmp Viral RNA extraction minikit (Qiagen). However, because manual extraction methods are labour intensive and more susceptible to variations in operator performance, we evaluated in parallel a semi-automated extraction platform, the Nuclisens EasyMag (BioMerieux), which is an easy-to-use bench top instrument based on silica extraction technology.

Consistent with previous studies, the success rate of amplification was higher when manual extraction was employed. However, the efficiency of the automated extraction procedure could be easily improved by including a high speed centrifugation step of the plasma sample prior to extraction. This additional step enabled us to amplify all samples with viral load around 1,000 copies/ml and 50% of samples with viral load around 500 copies/ml. Although,

the amplification success rate was slightly lower at viral load levels around 500 copies/ml when using the Nuclisens EasyMag platform compared to the QiaAmp Viral RNA minikit (100% vs. 50%), we considered it sufficient for the purpose of our assay. This performance is in line with that of commercial kits for HIV genotypic resistance testing of pol, which recommend a viral load of at least 1,000 copies/ml for reliable results. In addition, the benefit of performing resistance testing below a threshold of 1,000 copies/ml is still focus of controversy, and although multiple studies have demonstrated that resistance testing below this threshold is informative (Mackie et al, 2004 and Mackie et al, 2010), the clinical utility remains unclear.

In light of our results, we can conclude that although HIV-1 RNA extraction using the manual QiaAmp Viral RNA extraction minikit (Qiagen) increases diagnostic sensitivity of the amplification of HIV-1 Gag and protease from diverse subtypes in comparison to the semi-automated platform Nuclisens EasyMag (Biomerieux), the latter performs to an acceptable level when slight modifications are introduced and offers a more suitable methodology in the diagnostic setting where high throughput and reduced hands-on time are required.

4) Types of reverse transcriptase and DNA polymerase enzymes

The choice of enzymes involved in RT-PCR protocols represents another major parameter that influences a successful outcome.

There are currently a large number of reverse transcriptase and DNA polymerase enzymes to choose from, which differ among other characteristics in thermal stability, fidelity and processivity. Some manufacturers provide ready-to-use kits that incorporate both enzymes plus an optimized reaction buffer whereby reverse transcription and PCR reactions can be performed either in the same tube or independently in two tubes. Single-tube reverse transcription-PCR (RT-PCR) procedures are recommended in the diagnostic setting because they are simple to perform, allow high-throughput and reduce the risk of cross-contamination between samples.

Initially, we selected the Qiagen one-step RT-PCR kit (Qiagen), which contains a specially formulated enzyme blend for both reverse transcription (i.e., Omniscript / Sensiscript reverse transcriptases) and PCR (i.e., HotStarTaq DNA polymerase) reactions and a proprietary reaction mix containing a buffer with optimised concentrations of Mg^{2+} cations and dNTPs. Omniscript and Sensiscript are non-MMLV/AMV-derived reverse transcriptases (RTs). These enzymes may be superior to other commercially-available enzymes due to their higher affinity for RNA, which facilitates transcription through secondary structures where other RTs may be inhibited. In addition, the special composition of the buffer provided allows these RTs to operate at high temperatures (50°C), thus further improving reaction efficiency by disrupting secondary structures. As a result of their different abilities to copy small amounts of template, this enzyme mixture provides highly efficient and sensitive reverse transcription of any RNA quantity from 1 pg to 2 µg.

HotStarTaq DNA polymerase is an engineered version of the native Taq DNA polymerase, which was isolated from *Thermus aquaticus*. While maintaining the robustness and low cost of the original Taq DNA polymerase, this enzyme features a hot start procedure where the enzyme is activated by a 15-minute incubation at 95°C; this activation pattern ensures that polymerase activity does not start until the sample has reached a temperature where all DNA is denatured, thus avoiding extension from non-specifically annealed primers and primer-dimers that may have formed at lower temperatures. In addition, this enzyme, as all Taq DNA polymerases, has the capacity to incorporate adenosine overhangs at the 3' end of the PCR products; this characteristic is very useful for clonation studies when using TOPO or TA vectors, as the presence of thymidine (T) overhangs in these vectors enables ligation using topoisomerase or DNA ligase.

Using the Qiagen one-step RT-PCR kit (Qiagen) we achieved a high success rate for the amplification of gag and protease with a variety of group M HIV-1 subtypes. A success rate of 100% was achieved when samples with viral loads ranging from ≤ 1,000 to >100,000 copies/ml were tested following the manufacturer's recommendations. In addition, we carried out clonal analysis using TOPO/TA vectors on 11 samples that did not provide conclusive results by population sequencing. We successfully cloned all the 11 samples, using the TOPO TA cloning kit according to the manufacturer's recommendation. No optimization was required as it is often necessary when traditional cloning methods involving restriction digestion are applied. Cloning of PCR product with proof-reading enzymes such as those included in Invitrogene and Stratagene systems required additional steps to introduce the A overhangs and was in general less efficient.

Despite the advantages of Taq DNA polymerases, one of their main drawbacks is the lack of 3'→5' exonuclease activity (i.e., proof-reading activity), which removes a mispaired nucleotide from the 3' end of the growing strand, thus improving the fidelity of polymerization. The error rate of DNA polymerases is commonly expressed as the number of mutations per nucleotide per cycle; it depends not only on the intrinsic properties of the enzyme but also on the nature of the target sequence and the PCR conditions. Some studies have suggested that mutation rates may be artificially increased when employing Taq DNA polymerases (Bracho et al, 1998) due to their lack of proof-reading activity. In one report, the use of Taq DNA polymerase overestimated the proportion of minor hepatitis C virus quasispecies variants detected (Mullan et al, 2001), and it has been documented that quasispecies diversity is in general lower when proof-reading enzymes are used (Polyak et al, 2005). In another study the error rate of Taq DNA polymerase for the amplification of the HIV-1 Gag gene was estimated at about 1 in 83,000 nucleotides by cloning individual DNA molecules from the amplified population and determining the number of DNA sequences changes (Eckert and Kunkel, 1991).

To investigate how the fidelity of reverse transcription and PCR reactions could affect the detection of Gag and protease mutations, we performed further studies with alternative commercial kits. After extensive literature revision, we selected the AccuScript® HF RT-PCR system (Stratagene) and the SuperScript® one-step RT-PCR system with platinum Taq HF. AccuScript® HF RT-PCR system (Stratagene) contains a specially formulated enzyme blend for both reverse transcription (i.e., AccuScript reverse transcriptase) and PCR (i.e., *PfuUltra* HF DNA polymerase) reactions plus two proprietary reaction buffers containing optimized concentrations of Mg^{2+} cations and dNTPs; reverse transcription and PCR

reactions are performed independently in two separate tubes. AccuScript is a MMLV-derived reverse transcriptase, its main differential feature is the presence of 3'→5' exonuclease activity which improves reverse transcription fidelity by more than three fold when compared to other commercially-available RTs (43). *PfuUltra* HF DNA polymerase is an engineered version of the native *Pfu* DNA polymerase, which was isolated from *Pyrococcus furiosus*. Its accuracy has been documented to be 18 times higher than that of Taq DNA polymerases (Lundberg et al, 1991).

Through the high fidelity (HF) of the enzymes included in the AccuScript® HF RT-PCR system (Stratagene, Netherlands), we expected to improve the accuracy of detection of mutations in HIV gag and protease genes. However, the Stratagene system had a series of disadvantages. Firstly, a series of extra requirements, such as the incorporation of RNAse inhibitors in the RT-PCR mix, the setting of the reverse transcription and PCR reactions on ice or the initiation of the PCR reaction in a pre-heated thermocycler immediately after the addition of the enzyme to avoid primer degradation. These additional steps increase hands-on time and the length and cost of the procedure. In addition, we observed a very low success rate for the amplification of our target sequence when following the manufacturer's recommendations. To improve the performance of the assay, an increase in the Mg^{2+} cation concentration was required. It is known that while the lowest concentrations of Mg^{2+} favour specific priming and thus reduce non-specific amplification, the highest ones increase polymerization rates and as a result PCR sensitivity, but tend to facilitate non-specific primer binding, therefore diminishing PCR specificity. As a consequence, increasing Mg^{2+} concentration may have led to a decrease in the fidelity of the *Pfu* DNA polymerase. After optimization of the [Mg^{2+}], an amplification failure rate of 40% was observed compared to

0% with the Qiagen one-step RT-PCR kit. In addition, for those samples where amplification was successful, smaller amount of amplicons was produced by the Stratagene system than by the Qiagen kit, as demonstrated by the lower intensity of the PCR bands on agarose gel electrophoresis. This reduced yield can be due to primer degradation by the *Pfu* DNA polymerase as a result of its proof-reading activity and/or the stringent conditions of the PCR reaction (Takagi et al, 1997). In conclusion, although the Stratagene system may be recommendable for techniques that require HF DNA synthesis, such as clonal analysis, it is not convenient for population sequencing purposes in a diagnostic setting as it may decrease the success rate of RT-PCR reactions and increase the length and cost of the procedure.

Mixtures of proofreading and non-proofreading DNA polymerases have been reported to synthesize higher yields of PCR product (Barnes 1994) and fidelity comparisons with Pfu/Taq-containing polymerase blends have shown that the error rate of the mixtures appears to be intermediate between the error rate of Pfu and the non-proofreading enzyme (Cline et al, 1996) and is likely to depend on the ratio of non proof-reading to proof-reading enzyme. In order to obtain high reaction success rates and HF of template replication, we performed a one-step RT-PCR with SuperScript® III Platinum® one-step RT-PCR system with Platinum ® Taq DNA polymerase HF (Invitrogen, UK) followed by amplification with Platinum ® Taq DNA polymerase HF (Invitrogen, UK). SuperScript® III Platinum® one-step RT-PCR system with Platinum ® Taq DNA polymerase HF (Invitrogen, UK) combines SuperScript® III reverse transcriptase, Platinum® Taq DNA polymerase HF and a propietory reaction mix containing a buffer with optimised concentrations of Mg^{2+} cations and dNTPs. SuperScript® III is a version of MMLV RT that has been engineered to reduce RNAaseH activity providing more full-length cDNA than other RTs. Platinum ® Taq DNA polymerase HF (Invitrogen,

UK) is an enzyme mixture composed of *Pyrococcus* species GB-D polymerase, a recombinant enzyme that features 3'→5' exonuclease activity, and a recombinant non proof-reading enzyme, Platinum® Taq DNA polymerase. The reaction success rates obtained with the Invitrogene system were comparable to those obtained with Qiagen and the level of accuracy of mutation detection is likely to be between Qiagen and Stratagene.

We performed a comparative study between the three enzymatic systems. A total of six samples underwent reverse transcription followed by a nested PCR reaction and the PCR products were subsequently analysed by population sequencing. We did not find significant differences between the three systems in the rate of mutations encountered. With regard to the type of mutations, transitions were much more frequently found than transversions and were in order of decreasing frequency A-G (23.8%),G-A (16.5%), T-C (15.4%) and C-T (13.9%) . These results suggest that a proof-reading polymerase is not advantageous over a DNA polymerase lacking 3'→5' exonuclease activity for the purpose of population sequencing. The rate and type of mutations these enzymes produce is similar. Although the rate of mutations may be slightly higher with non-proofreading enzymes, it can be minimized by optimizing the PCR reaction conditions. In any case, the proportion of PCR product displaying enzyme-derived mutations represents only a minority of the quasispecies and as a consequence would go undetected by population sequencing (Alcorn and Faruki 2000). On the other hand reaction success rates were substantially higher when non-proofreading DNA polymerase was employed.

However, population sequence analysis was unsuccessful in 20% of the samples. These samples were further characterized by clonal analysis, an application that may be more affected by low fidelity of template replication. When we compared 20 individual clones derived from one sample, we could not found any significant differences in the rate of mutations observed. Therefore, we concluded that despite lacking proof-reading activity, Taq polymerase can be used for some down-stream applications such as population sequencing and in our study clonal analysis, offering the advantages of high success rates, low cost, robustness and ease of use in cloning applications. We need to emphasize that our clonal analysis studies were not aimed at the detection of minority variants and therefore we only analyzed a reduced number of clones. In applications where variants that represent a minority in the quasispecies are of interest, requiring the analysis of a high number of clones, the use of more accurate enzymes, such as Pfu, is recommended.

Another possible application of our clonal analysis was the study of linkage between mutations. Polymerising enzymes employed for this purpose must have the ability to complete strand synthesis and display HF of template replication. PCR-mediated recombination is a main concern when performing linkage analysis. Recombinants during PCR presumably arise due to the presence of incompletely extended primers annealing to a heterologous target (Meyerhans et al, 1989); this situation is avoided when complete rather than partial strand synthesis is achieved. Early studies demonstrated that the enzyme employed for DNA synthesis significantly affects the rate of artificial recombination (Fang et al, 1998). In general, proofreading enzymes facilitate complete strand synthesis and are the preferred enzymes in this setting.

The Invitrogen system may be best suited for this use as the reduced RNAase H activity if SuperScript® III reverse transcriptase maximises synthesis of complete cDNA strands and Platinum® Taq DNA polymerase HF contains a proofreading enzyme that ensures synthesis of HF complete DNA strands. However, adjustment of other parameters that may affect the rate of artificial recombination, such as the method of RNA isolation, reverse transcription time or number of cycles during PCR (Fang et al, 1998) may be required in order to optimise this application.

The greatest challenge we faced to develop a protocol for the amplification and sequencing of HIV-1 gag and protease genes was the high degree of genetic variability present within the HIV-1 gag gene, which is not only observed between different subtypes, but also within subtypes and intra-patient quasispecies (Brown and Monaghan, 1988; Louwagie et al, 1993; Markham et al, 1995; Mulder-Kampinga et al, 1995 and Yoshimura et al, 1996). Variability is not equally distributed across the gag gene. P17 and P7 were the most variable sites, whereas P24 is the most conserved region; Yoshimura and colleagues observed a variation of around 12% in P17 and P7 and 3% in P24 in a cohort of patients infected with the same HIV-1 subtype. Similarly, when sequences of the gag clones within HIV-1 infected patients were compared, the greatest genetic diversity was located in P17 and P7 regions, while P24 had the lowest sequence variability (Yoshimura et al, 1996).

Due to the diversity seen in the gag gene, sequencing of this region posed a big challenge. Mismatches between sequencing primers and complementary sequences in the gag gene led to foreshortened sequences or to complete failure of the sequencing reaction in numerous

occasions. To overcome this problem, we developed multiple overlapping sequencing primers. A total of 18 sequencing primers were required to obtain a full-length gag sequence and were employed in different combinations according to the specific sample. An additional challenge for gag sequencing is the intra-patient genetic variability, which according to previous reports is mainly concentrated in P17, P7 and P6 regions (Brown and Monaghan, 1988). In our study, we obtained unreadable P17, P7 and/or P6 sequences due to heterogeneous sequencing signals in 11/27 (41%) of the samples analyzed. In addition, quasispecies displaying insertion and deletions of different length were also found in one patient (1/27, 4%), which resulted in a shift of the chromatogram and yielded a non-interpretable population sequence. In order to circumvent these problems, we performed clonal analysis. Twenty clones per patient were analyzed to obtain full-length gag sequences of the independent dominant variants that constituted the patient's quasispecies. Although this approach was cumbersome, it allowed us to accomplish full-length gag sequencing with all of the samples.

In summary, we successfully developed an assay for the amplification and sequencing of HIV-1 gag and protease. The assay was validated by testing both manual and automated nucleic acid extraction techniques as well as different reverse transcriptase and DNA polymerase enzymes. The assay can be employed for different purposes, such as therapy monitoring or the study of linkage of mutations. The choice of manual or automated nucleic acid extraction methods, characteristics of RTs and DNA polymerases will depend on the goal of the study. The high variability present in the HIV-1 gag gene leads to a significant assay failure rate; the use of clonal analysis is required to characterize failed samples,

situation that makes the sequencing of Gag cumbersome in a high-throughput routine diagnostic setting.

4 Chapter four: cross-sectional comparison of prevalence and patterns of HIV-1 *Gag* mutations in PI-experienced and PI-naïve patients

4.1 Introduction

The HIV-1 protease (PR) plays a crucial role in the late phase of the HIV life cycle. It cleaves gag and Gag-Pro-Pol precursor polyproteins at particular sites thereby generating mature proteins, which are indispensable for the production of infectious virions. HIV-1 PR is a member of the aspartyl protease family. The functional enzyme exists as a symmetrical homodimer, each subunit comprising 99 amino acids and the two subunits interact non-covalently to form a long tunnel where the active site is located. The active site consists of two Asp-Thr-Gly sequences, each sequence derived from a single monomer, and the aspartic residues play an essential role in the catalytic process. Access of substrates to the active centre is regulated by two flexible flaps located at the top of the tunnel, which undergo significant conformational changes to allow the substrate to enter and leave the tunnel. .

The gag precursor polyprotein is cleaved to generate structural proteins. Cleavage occurs in a controlled manner at 5 unique cleavage sites (CSs), comprising P17/P24, P24/P2, P2/P7, P7/P1 and P1/P6 (Krausslich et al, 1989; Pettit et al, 1994 and Wiegers et al, 1998). An important factor governing the order and rate of cleavage is the amino acid sequence at the specific CS (Pettit et al, 2002). Each CS consists of 10 amino acids; their positions relative to the cleaved peptide bond designated from N to C terminus as follows: N-P5-P4-P3-P2-

P1/P1'-P2'-P3'-P4'-P5'-C, with cleavage occurring between P1and P1' (Pettit et al, 2002). The amino acid sequence of the different CSs within Gag differs strikingly from one site to another. In addition, some sites, most notably P2/P7, show a high degree of polymorphism between HIV-1 strains (Bally et al, 2000; Feher et al, 2002; Gallego et al, 2003; Malet et al, 2007 and De Oliveira et al, 2003). The difference in the amino acid sequence between these different CSs explains, at least in part, their differential rate of cleavage by the viral PR (Pettit et al, 2002 and Wiegers et al, 1998). Interestingly, in spite of the marked sequence diversity, these Gag sites show a strong similarity in their secondary structure (Bandaranayake et al, 2008; Prabu-Jeyabalan et al, 2002), which explains why they all constitute strong and specific substrates for the viral PR, albeit with different rates of cleavage.

The essential role that HIV-1 PR plays in the viral life cycle makes this enzyme an attractive target for antiretroviral drugs. There are currently nine protease inhibitors (PIs) licensed for the treatment of HIV-1 infection, namely Atazanavir (ATV), Darunavir (DRV), Fosamprenavir (FPV), Indinavir (IDV), Lopinavir (LPV), Nelfinavir (NFV), Ritonavir (RTV), Saquinavir (SQV) and Tipranavir (TPV). They act as competitive inhibitors that bind the active site of the viral PR blocking the entrance of the natural substrate. Unfortunately, resistance to all available PIs has been documented and represents a major obstacle for successful treatment of HIV-1 infected patients. The classic mechanism of PI resistance involves accumulation of substitutions in the viral PR (Croteau et al, 1997; Mammano et al, 2000; Nijhuis et al, 1999). In general, the first mutations to be selected (i.e, primary or major resistance mutations) are located within or close to the substrate-binding domain, and can differ from one PI to another. Subsequently, secondary mutations are selected, which involve amino acids located away from the substrate-binding cleft. These secondary mutations, also

termed minor resistance mutations, are less drug-specific than primary mutations, but still essential for high-level resistance to PI and many of them play a key compensatory role, restoring viral fitness in strains with primary mutations.

An alternative pathway of resistance to the PIs has been proposed, which is mediated by other genomic regions and focuses on cleavage sites of the viral PR (Nijhuis et al, 2007). The co-evolution of PR and C-terminal gag CSs in PI-resistant viruses was reported soon after the introduction of PIs in antiretroviral regimens (Doyon et al, 1996; Zhang et al, 1997). The occurrence of CS mutations (CSMs) was originally attributed to compensatory effects similar to those described with secondary PR mutations (Doyon et al, 1996; Mammano et al, 2000).However, growing evidence shows that CS mutations also confer PI resistance either in isolation or in combination with PR mutations (Dam et al, 2009; Maguire et al, 2002; Nijhuis et al, 2007; Prado et al, 2002; Zhang et al, 1997). Amino acids substitutions associated with PI resistance have been described at two gag CSs, namely P7/P1 and P1/P6. However, the role of amino acid changes at other gag CSs or outside CS regions of gag remains to be elucidated.

The aim of the present chapter was to obtain full-length HIV-1 gag and PR sequences from PI-experienced patients employing the assay described in the previous chapter. These sequences were compared to HIV-1 gag and PR sequences from PI-naive individuals in order to identify mutations associated with PI-exposure. Furthermore the sequences were analysed with the aim of determining significant associations between PR mutations and gag mutations that may identify novel pathways of PI resistance.

4.2 Methods

4.2.1 Sequences

The gag and PR genes from 191 PI-experienced and 200 PI-naïve patients were analyzed retrospectively.

Sequences from PI-experienced patients were obtained from different cohorts: patients recruited in the MaxCmin1, MaxCmin2 and COLATE trials, patients attending Cologne University Hospital and patients attending the Royal Free Hospital. The MaxCmin1 and MaxCmin2 were two open label, multicentre, phase IV trials that compared the safety and efficacy of ritonavir boosted Saquinavir (SQV/r), against ritonavir boosted Indinavir (IDV/r) and against ritonavir boosted Lopinavir (LPV/r). COLATE recruited patients failing a lamivudine-containing regimen, and assessed whether maintaining lamivudine in the subsequent regimen was of virological benefit. Patients from MaxCmin1, MaxCmin2 and COLATE whose regimens included one or more PIs and who showed major PR resistance mutations (n = 13) were included in the analysis. In order to maximize numbers, we also selected patients from Cologne University (n = 128) Hospital and Royal Free Hospital (n =50) who had been previously exposed to PIs, all of whom showed major PR resistance mutations.

As a comparator group, we constructed a database of PI-naive sequences which included sequences obtained from RFH patients who had been previously exposed to antiretrovirals

(ARVs) other than PIs (n = 52) and sequences from ARV-naïve subjects from the Los Alamos database (n =148).

Both PI-experienced (n= 191) and PI-naive (n = 200) sequences were retrieved from patients with long lasting (i.e.; 3-10 years of infection) subtype B HIV-1 infection.

4.2.2 RNA extraction

Prior to extraction, one millilitre of plasma was centrifuged at 25,000g for 1 hour at 4°C to concentrate the virus; the supernatant was then removed and the pellet re-suspended to a final volume of 280 µl. Samples with HIV-1 RNA load above 1,000 copies/ml underwent extraction employing the semi-automated extractor EasyMag (Nuclisens, France), whereas those with viral load below 1,000 copies/ml were extracted manually with the QIAamp Viral RNA MiniKit (Qiagen, Germany). Finally, the RNA was eluted into 55µl of elution buffer and stored at -80°C until required.

4.2.3 Amplification of the Gag-protease region

A 2 Kb PCR product comprising full-length HIV-1 Gag and PR was amplified by nested PCR employing the Qiagen Gag-PR amplification protocol described in section 2.2.2.2.1 in chapter 2.

4.2.4 Purification of PCR products

PCR products for population sequencing were purified using QIAQuick PCR purification Kit (Qiagen, Crawley, UK) according to the manufacturer's instructions**.**

4.2.5 Analysis of PCR products

Five µl of PCR product mixed with five µl of loading buffer were loaded into a 1% agarose gel containing 0.5 µg/ml of ethidium bromide. Three µl of mass ladder was also loaded to use as a reference, and the gel was run for one hour at 80V. The gel was examined with UV light and the quantity and size of DNA was evaluated by comparing the intensity and position of the bands to those of the DNA mass ladder.

4.2.6 Sequencing

Purified PCR products of the right size were diluted to a final concentration of approximately 10-20 ng/µl and population sequencing was performed using the ABI PRISM BigDye Terminator v3.1.ready reaction cycle sequencing Kit. Primers selected for PR and Gag sequences were those described in chapter 3 (Table 3.2).

4.2.7 Sequence analysis

Sequences were analyzed using Sequence analysis version 5.0, Seqscape version 6.0 and Mega Molecular Evolutionary Genetic Analysis software version 4.0 programmes.

4.2.8 Classification of protease mutations

PR resistance mutations were assessed according to the last available list of mutations of the International AIDS Society (IAS, 2011) and classified into primary and secondary mutations according to the Stanford HIV Drug Resistance Database (Figure 4.1).

4.2.9 Classification of Gag mutations

We analysed mutations in the entire Gag protein. Gag sequences were aligned with the reference sequence HXB2 and subtyped by submitting the sequence to two different subtyping tools (NCBI and Rega). Mutations were defined as any change relative to the reference sequence. Mutations were divided into those seen in CSs (CSMs) and those seen outside CSs (Non-CSMs). Each CS consisted of the five amino acids on both sides of the cleavage bond. P5 to P1 and P5' to P1' were designated for residues on the N and C terminal sides of the target, respectively (Figure 4.2).

4.2.10 Gag and protease analysis

PI-naive sequences were compared to the HXB2 reference sequence. The variability at each position in the Gag protein was given as a percentage, defined as the proportion of sequences with a non-wild type amino acid relative to the total number of sequences. We considered amino acid positions showing a variability ≤1 % as being conserved. The association between gag mutations and PI-exposure was analyzed by using the Fischer's exact test. Gag CS positions (n =50) and Non-CS positions (n=450) were analyzed separately. A p value of less than 0.05 was considered to indicate a statistically significant different between groups in a preliminary analysis. Subsequently, the Bonferroni correction was applied in order to account for multiple comparisons. Therefore a p value of less than 0.05/50 = 0.001 for CSMs and 0.05/450 = 0.0001 for Non-CSMs were considered as the threshold to show statistically significant differences. In addition, the association between presence of gag mutations and presence of PR mutations was also analyzed using the Fisher's exact test after, again after applying the Bonferroni's correction for multiple associations.

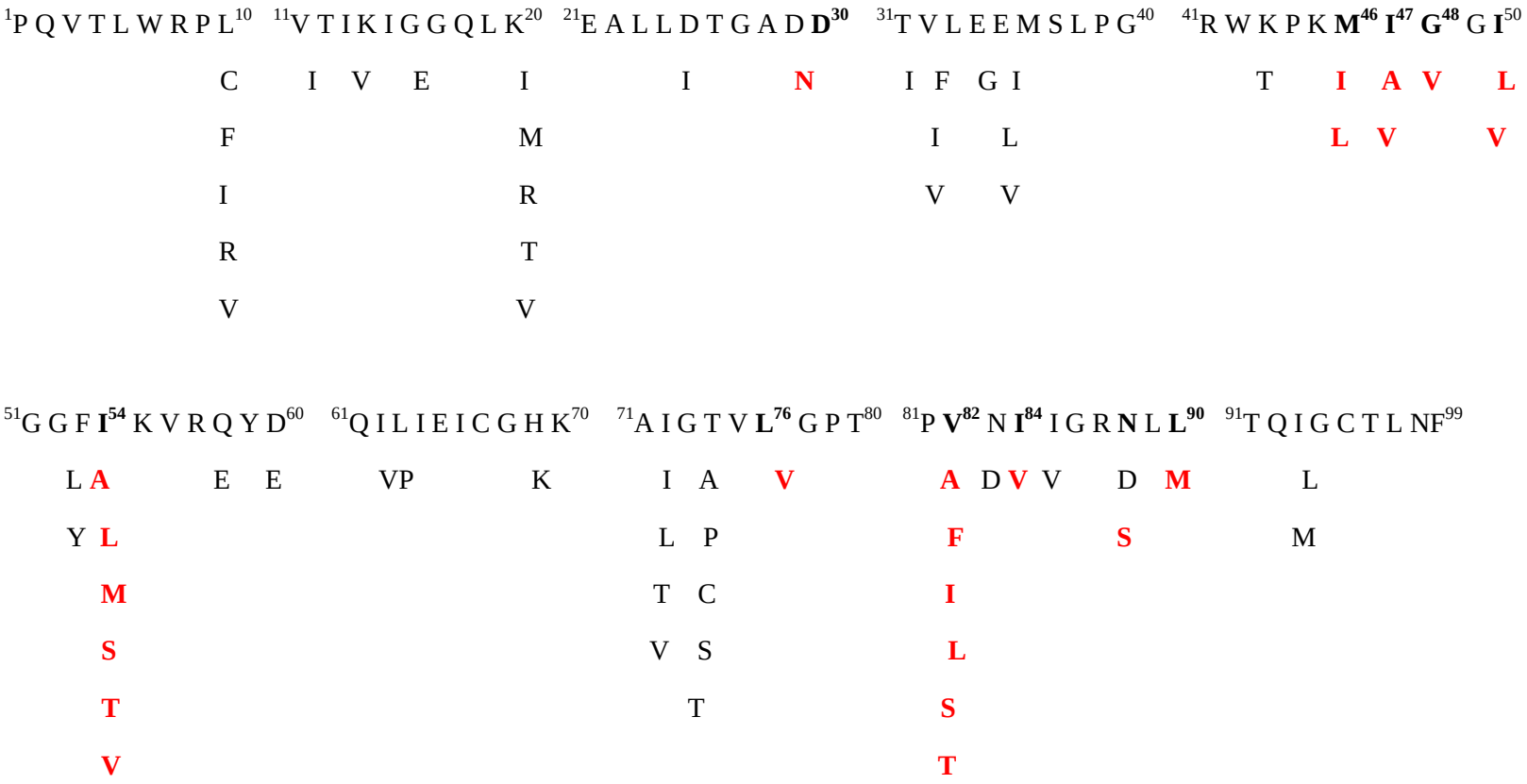


Figure 4.1 Protease inhibitor resistance mutations.

The first row of letters indicates the reference (HXB2) amino acid sequence. All amino acids are indicated by their one letter code. The position of the first and last amino acid of each series of 10 is indicated as well as positions where primary PR resistance mutations have been described which are also highlighted in bold. Letters below in bold red indicates primary resistance mutations and in black secondary mutations.

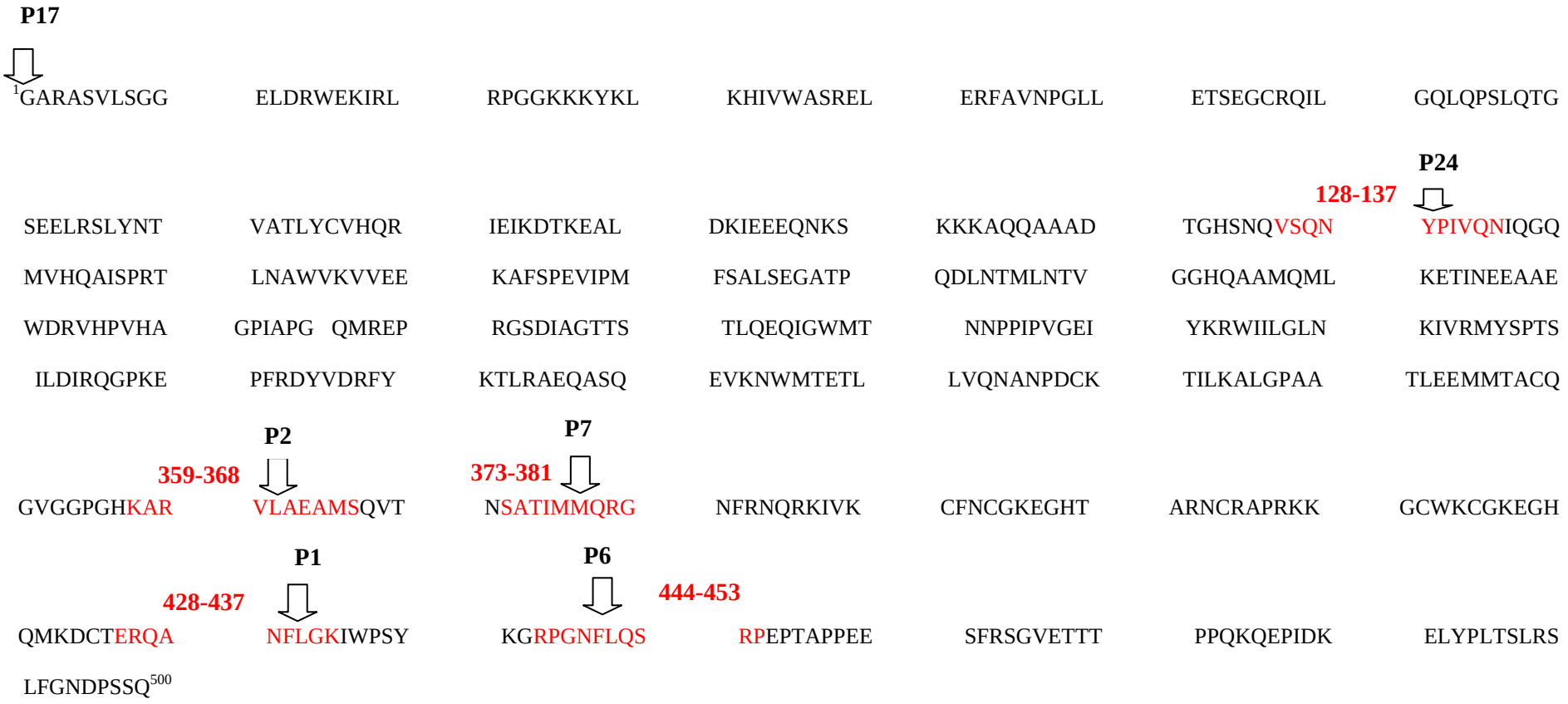


Figure 4.2 Gag HXB2 sequence.

The letters indicate the reference (HXB2) Gag amino acid sequence. All amino acids are indicated by their one letter code. The number position of the first and last amino acid of the Gag polyprotein is indicated. The beginning and end of each individual protein is indicated by arrows. CSs sequence and positions are indicated in red.

4.3 **Results**

4.3.1 Patient characteristics

PI experienced patients had been exposed to a median (range) of 2 PIs (1-4), 4 (2-5) NRTIs and 1 (0-2) NNRTIs. At the time of analysis, the most common PI individuals had previously been exposed to were: LPV/r (60/191, 31 %); SQV/r (42/191, 22%); IDV/r (30/191, 16%); APV/r (30/191; 16%); ATV/r (10/191, 5%); TPV/r (6/191, 3%) and DRV/r (4/191, 2%). All patients had major and minor protease resistance mutations. Similarly, 187/191 (98%) patients had NRTI-resistance-associated mutations and 125/191 (80%) had NNRTI-resistance associated mutations. The median (range) of major and minor PI resistance mutations at study entry were 3 (1-6) and 4 (0-8), respectively. The median (range) of NRTI and NNRTI-resistance associated mutations were 5 (0-11) and 1 (0-4), respectively.

PI-naïve patients were either completely treatment naïve (n = 148) or exposed to NRTIs and NNRTIs but not to PIs (n = 52). None of them had major PI resistance mutations and the median (range) number of minor PI resistance mutations was 1 (0 – 3). However, while none of the 148 treatment naïve sequences showed either NRTI nor NNRTI-resistance associated mutations, all of the 52 sequences ontained from patients previously exposed to ARVs other than PIs showed NRTI-resistance associated mutations, median 2, range (1-3) and 23/52 (44%) showed in addition NNRTI-resistance associated mutations, median 2 (0-2).

4.3.2 Gag variability at non-cleavage site positions

The 500 amino acid positions setting up the sequence of the Gag protein from 200 PI-naïve patients infected with subtype B HIV-1 were compared to the reference sequence HXB2. Non-cleavage site (Non-CS) amino acid positions (n=450) and CS amino positions (n=50) were examined separately. The latter comprised the 5 CSs (P17/P24; P24/P2; P2/P7; P7/P1 and P1/P6), each consisting of 10 amino acids. Figures 4.3a to 4.3f show the variability of Non-CS amino acid positions in PI-naïve individuals, divided into the different Gag domains Matrix (P17), Capsid (P24), Nucleocapsid (P7), P6, P2 and P1, and excluding amino acid positions within the CSs.

The analysis of the 450 Non-CS amino acid positions in the 200 PI-naïve individuals revealed the following:

- P17:

A total of 76/127 (60%) amino acid positions were conserved among subtype B HIV-1 strains. Some conserved residues were scattered, such as K27, Y29, L50, Q63, S77, I92, K103 and Q116, while others were contiguous forming conserved motives of between 2 and 10 amino acids. The remaining 51/127 (40%) positions showed a degree of variability ranging between 1% and 55% (Figure 4.3a).

- P24:

This region showed the highest degree of conservation. A total of 154/220 (70%) amino acid positions were conserved. Most conserved residues formed motives of between 2 and 11

amino acids, and only a few were individually scattered including 164F, 253N, 336A and 347M. The remaining 66/220 amino acids (30%) showed a degree of variability ranging between 1% and 82% (Figure 4.3b).

- P7:

A total of 24/45 (53%) amino acid positions were conserved. Most conserved residues formed motives of between 2 and 6 amino acids except for one that was individually scattered (C426). The remaining 21/45 amino acids (47%) showed a degree of variability ranging between 1% and 95% (Figure 4.3c).

- P6:

This region showed the highest degree of variability. A total of 22/47 (47%) amino acid positions were conserved, six of them individually scattered (P455, V467, P472, Q474, P485 and D496) and the remaining forming motives of between two and five amino acids. However, the majority of the amino acids (25/46, 54%) showed a degree of variability ranging between 1% and 75% (Figure 4.3c).

- Spacer peptides P1 and P2 :

The first spacer peptide encountered in the Gag polypeptide P2 was found to be highly polymorphic in the four Non-CS residues, with a degree of variability ranging between 8% and 22%. By contrast, the second spacer peptide P1 showed 3/6 (50%) conserved amino acids and 3/6 (50%) polymorphic residues (P439, S440, Y441H) with a degree of variability of 9%, 1% and 48%, respectively (Figure 4.3c).

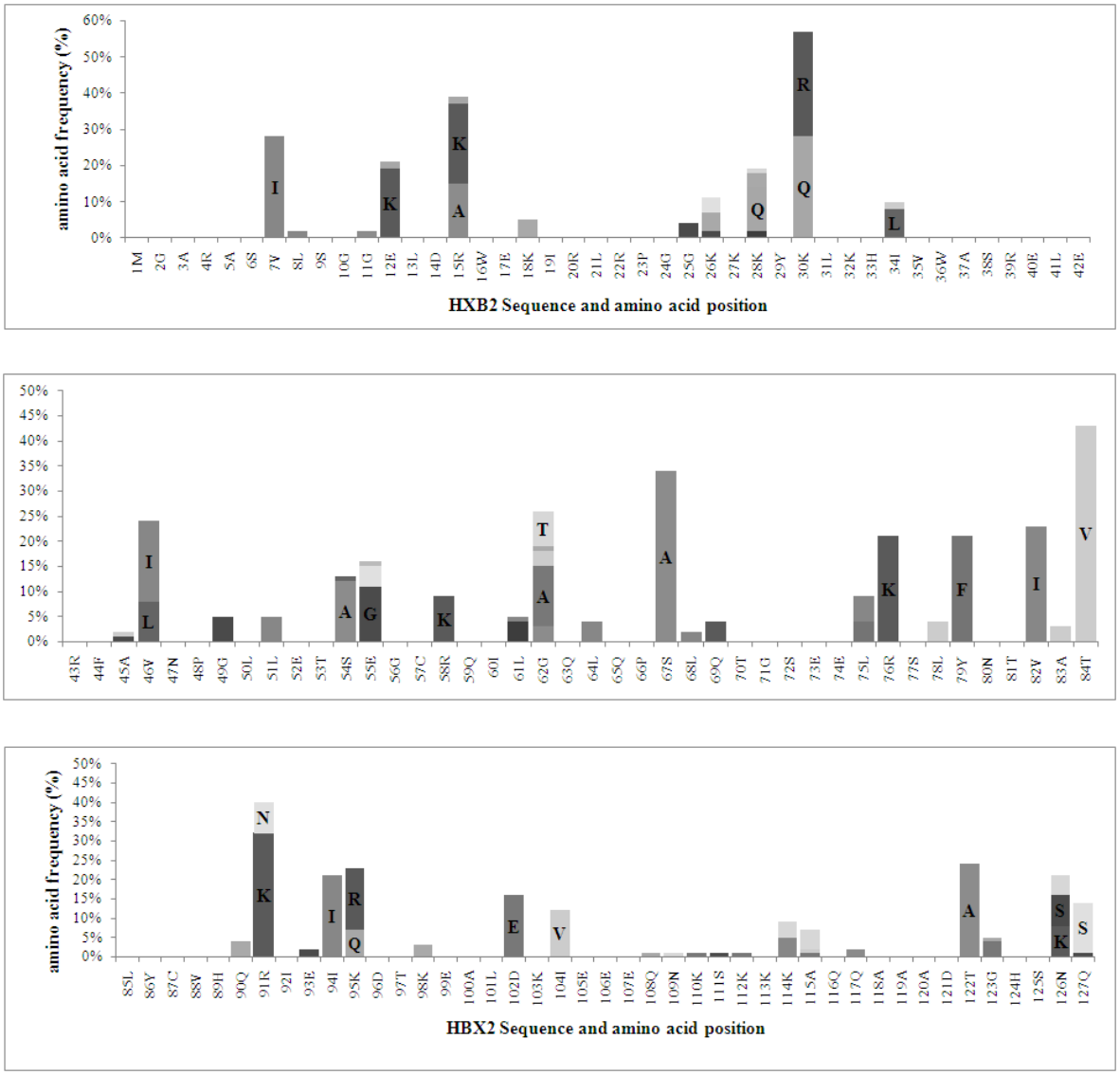
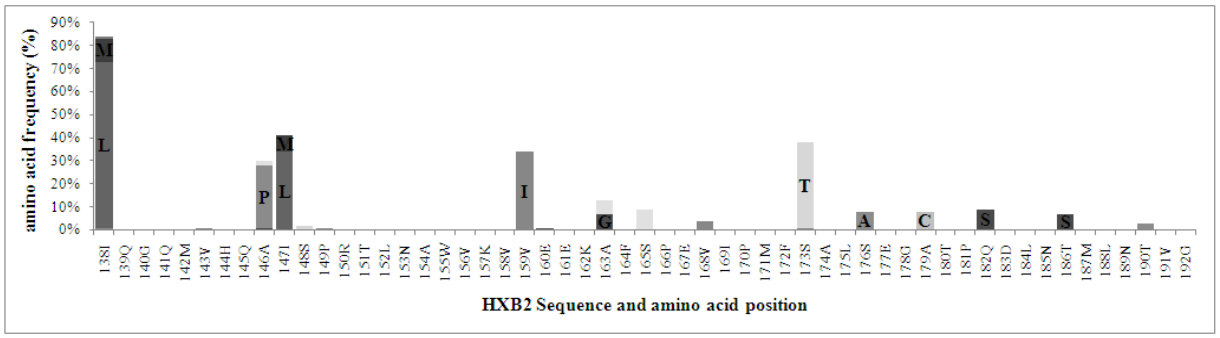
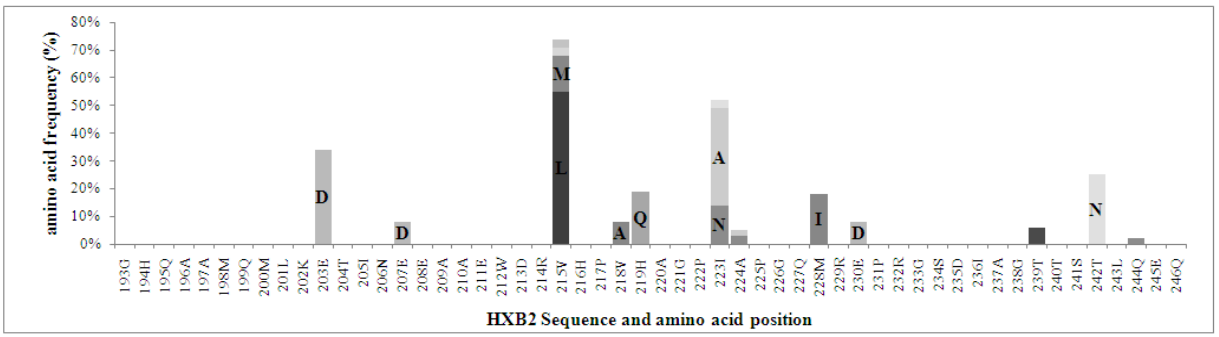
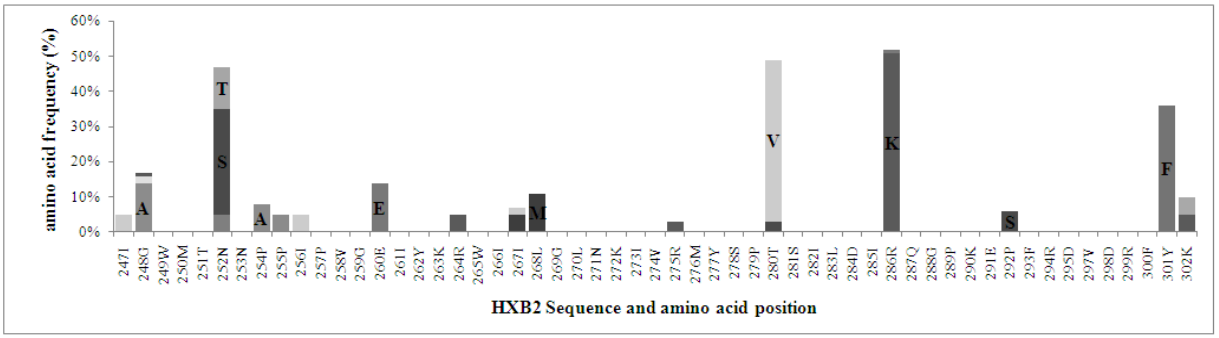


Figure 4.3a Polymorphisms of the HIV-1 P17 protein in subtype B strains from protease inhibitor-naïve patients.

Cleavage site positions (amino acid: 128-132) were excluded from this analysis. Sequences from 200 PI-naïve patients were compared to the HXB2 reference sequence (amino acid: 1-127**).** The letters specified on the bars indicate polymorphisms occurring at frequency ≥ 6%. HXB2 reference sequence and amino acid positions are shown.







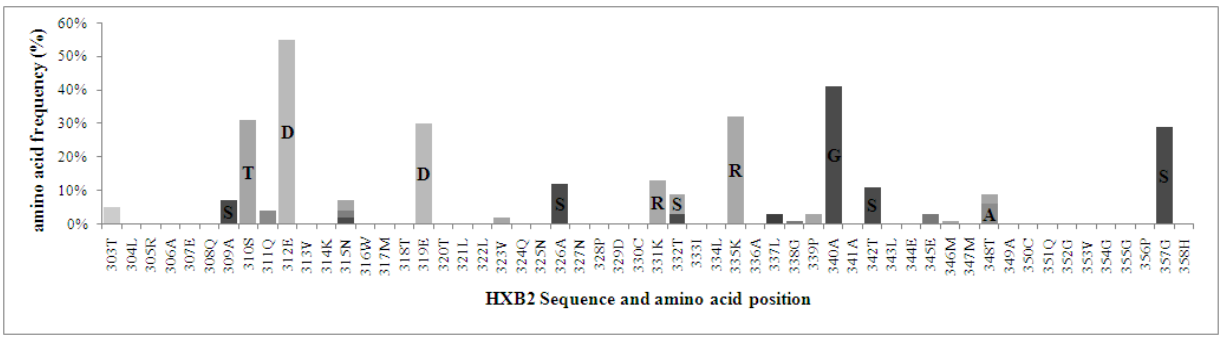


Figure 4.3b Polymorphisms of the HIV-1 P24 protein in subtype B strains from protease inhibitor-naïve patients.

Cleavage Site positions (amino acid: 133-138 and 359-362) were excluded. Sequences from 200 PI-naïve patients were compared to the HXB2 reference sequence (amino acid: 138-358). The letters indicate polymorphisms occurring at frequencies ≥ 6%. HXB2 reference sequence and amino acid positions are shown.

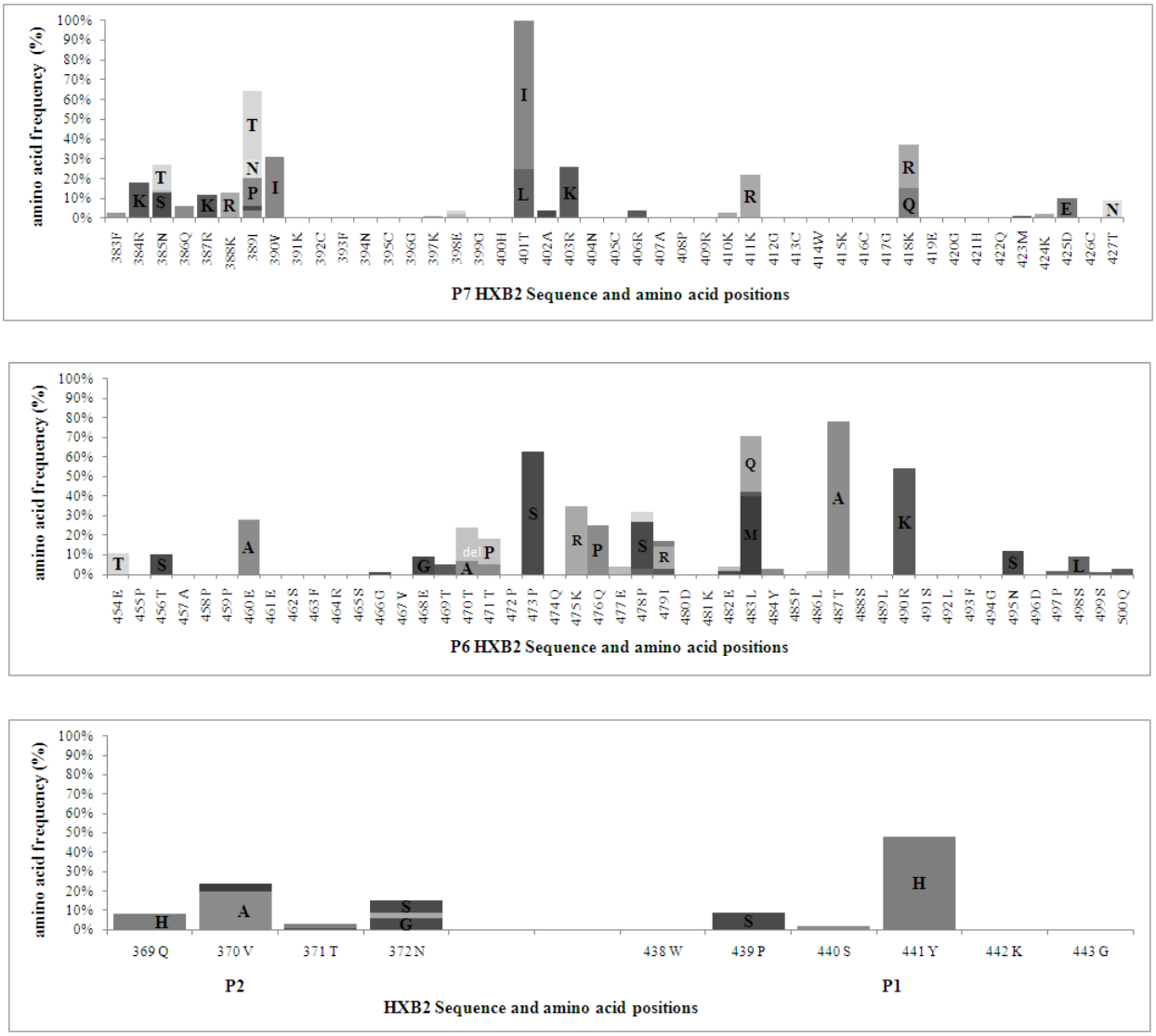


Figure 4.3c Polymorphism of the P7, P6, P1 and P2 proteins in subtype B strains from protease inhibitor-naïve patients.

Cleavage Site positions (amino acid: 377-381; 428-431; 448-453; 432-437; 444-447; 363-368 and 373-376) were excluded. Sequences from 200 PI-naïve patients were compared to the HXB2 reference sequence (amino acid 383-427; 454-500; 369-372 and 438-443 respectively).The letters indicate polymorphisms occurring at frequencies ≥ 6%. HXB2 reference sequence and amino acid positions are shown.

4.3.3 Gag variability at cleavage site positions

The CS sequences from the 200 PI-naïve patients were examined and compared to the reference sequence HXB2. Each cleavage site consists of 10 amino acids designated as P5 to P1 and P1' to P5' for residues on the N and C terminal sites of the CS, respectively. Variability of CS differed considerably. Whereas three CSs (P17/P24; P24/P2 and P7/P1) were relatively conserved, P1/P6 and P2/P7 showed moderate and high variability respectively. Specifically, the P17/P24 CS showed only three mutations: V128A in 8/200 (4%) patients, and V128S and Q130H in 2/200 (1%) patients each; the remaining eight positions were completely conserved. The P24/P2 CS showed only two mutations: V362I in 46/200 (23%) and R361K in 4/200 (2%) patients; the remaining eight positions were conserved. Concerning the P7/P1 CS, only two mutations were encountered among the 200 PI-naïve patients; these were E428D and K436R observed in 26/200 (13%) and 22/200 (11%) patients, respectively By contrast, at P2/P7 only two positions (Q379 and N382) were conserved, while all patients showed at least one mutation at the remaining eight positions and 144/200 (72%) showing more than one mutation (range 2-4mutations). Lastly, mutations at the P1/P6 CS were located at the C-terminal part, where four out of the five positions were variable; the most common mutations were L449P and L453P, found in 48/200 (24%) and 46/200 (23%) patients, respectively. Mutations at positions Q450 (Q450E, 22/200, 11%) and S451 (S451N 22/200, 11%; S451A 10/200, 5%) were less frequent. The five amino acids forming the N-terminal part of the P1/P6 CS were completely conserved in the PI-naïve population. Variability of the amino acids at the five Gag CSs is shown in figure 4.4.

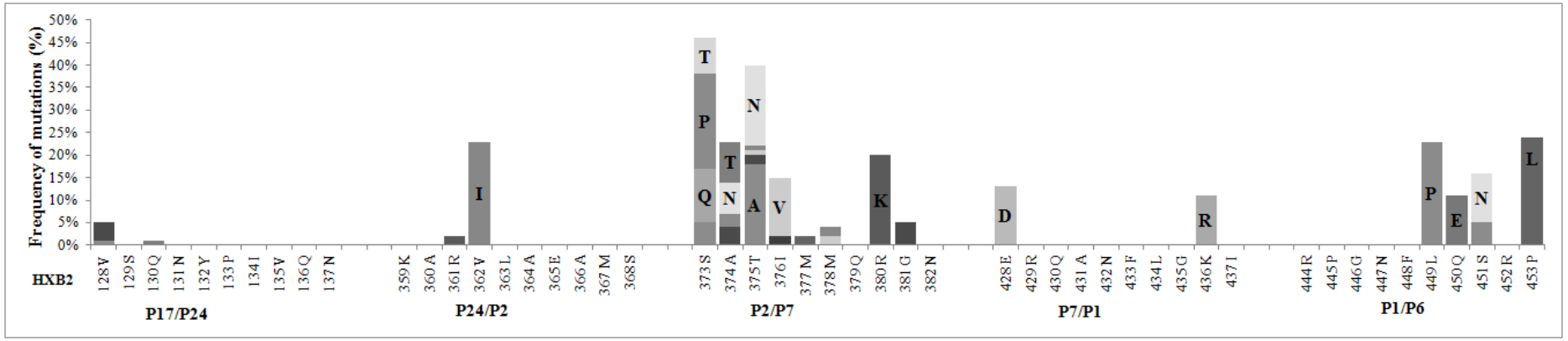


Figure 4.4 Polymorphisms of the HIV-1 Gag CSs (P17/P24, P24/P2, P2/P7, P7/P1 and P1/P6) in subtype B strains from protease inhibitor-naïve patients.

Sequences from 200 PI-naïve patients were compared to the HXB2 reference sequence. Letters indicate polymorphisms occurring at frequencies ≥ 6% HXB2 reference sequence and amino acid positions are shown.

4.3.4 Comparison of Gag non-cleavage site mutations according to history of exposure to protease inhibitors

The comparison of full-length Gag and PR sequences from 191 PI-experienced and200 PI-naïve individuals showed the following results:

- P17:

The 127 amino acids of the P17 protein were compared between sequences from PI-experience and PI-naïve patients. Overall 61 (48%) *vs.* 76 (60%) amino acids were conserved while 66 (52%) *vs.* 51 (40%) amino acids were variable, respectively. Among the 66 variable positions seen in PI-experienced patients, we detected a total of 150 different mutations. Although 39 mutations out of the 150 found were associated with PI exposure when a cut-off of 0.05 was considered, only 10 remained significantly associated when the Bonferroni's correction for multiple associations was applied ($p = 0.05/450$) (Table 4.1a).

- P24:

The 220 amino acids of the P24 protein were compared between sequences from PI-experienced and PI-naïve patients. As described for the PI-naïve individuals in section 4.3.1, the CA-P24 protein also showed the highest degree of conservation in PI-experienced patients. Overall 149 (68%) *vs.* 154 (70%) residues were conserved and 71 (32%) vs. 66 (30%) residues were variable in PI-experienced *vs.* PI-naïve individuals, respectively. Among the 71 variable positions seen in PI-experienced patients, we detected a total of 98 mutations. Although 17 out of the 98 mutations were associated with PI exposure when a cut- off of 0.05 was considered, only five remained significantly associated when the Bonferroni's correction for multiple associations was applied ($p = 0.05/450$) (Table 4.1b).

- P7:

The 45 amino acids of the P7 protein were compared between sequences from PI-experienced and PI-naïve patients. We found the same number of conserved and variable residues in the two groups; these were 24 conserved residues and 21 variable residues. Among the 21 variable amino acids in PI-experienced patients, we detected 42 mutations. Although seven out of the 42 mutations were significantly associated with PI-exposure when a cut-off of 0.05 was considered, none remain significantly associated when the Bonferroni's correction for multiple associations was applied (p= 0.05/450) (Table 4.1c).

- P6:

The 47 amino acids of the P6 protein was compared between sequences from PI-experienced and PI-naïve patients, overall 13 *vs.* 22 residues were conserved and 34 *vs.* 25 residues were variable in PI-experienced *vs.* PI-naïve individuals, respectively. We found 81 mutations distributed among the 34 variable residues seen in PI-experienced patients. Although, 19 out of the 34 mutations were associated with PI exposure when a cut-off of 0.05 was considered, only three remained significantly associated when the Bonferroni's correction for multiple associations was applied (p = 0.05/450) (Table 4.1d).

- Spacer peptides P2:

The four amino acids comprising the P2 peptide were compared between sequences from PI-naïve and PI-experienced subjects. As I described for PI-naïve individuals in section 4.3.1, all P2 residues (n =4) were variable in PI-experienced individuals. A total of 11 mutations were found distributed among the four variable residues in PI-experienced subjects. Although four out of the 11 mutations were significantly associated with PI exposure when a cut-off of 0.05

was considered, only one remained significantly associated when the Bonferroni's correction for multiple associations was applied (p = 0.05/450) (Table 4.1e).

- Spacer peptides P1:

Finally, the 6 amino acids comprising the P1 peptide were compared between sequences from PI-experienced and PI-naïve patients. Overall one *vs.* three residues was conserved and five *vs.* three were variable in PI-experience and PI-naïve patients, respectively. Among the five variable positions seen in PI-experienced subjects, we detected five mutations. Although one out of the five mutations was significantly associated with PI-exposure when a cut-off of 0.05 was considered, it did not remain significantly associated when the Bonferroni's correction for multiple associations was applied (p = 0.05/450) (Table 4.1e).

	MUTATION	PI-naïve (n= 200)		PI-experienced (n= 191)		
		n	%	n	%	P
	I34L	16	8	32	17	0.02
	Q59K	0	0	8	4	0.007
	L61I	**2**	**1**	**20**	**10**	**<0.0001**
	P66S	0	0	10	5	0.002
	T80A	0	0	10	5	0.002
	V82I	24	12	46	24	0.005
	T84V	60	30	86	45	0.009
	Y86W	0	0	6	3	0.003
	R91G	0	0	6	3	0.03
	I92V	0	0	6	3	0.03
	E93D	62	31	92	48	0.003
	I94V	**114**	**57**	**158**	**83**	**<0.0001**
	K103R	**0**	**0**	**16**	**8**	**<0.0001**
	E107I	0	0	8	4	0.007
	N109T	2	1	10	5	0.04
	N109S	0	0	6	3	0.03
	K113Q	**0**	**0**	**22**	**11**	**<0.0001**
	K113R	0	0	8	4	0.007
P17	**K114R**	**0**	**0**	**22**	**11**	**<0.0001**
	K114Q	0	0	10	5	0.04
	A115I	0	0	6	3	0.03
	Q117P	4	2	18	9	0.003
	Q117E	0	0	6	3	0.03
	A118T	0	0	8	4	0.007
	A119T	0	0	6	3	0.03
	A119E	0	0	10	5	0.04
	D121G	**0**	**0**	**16**	**8**	**<0.0001**
	D121A	**0**	**0**	**14**	**7**	**<0.0001**
	T122A	24	12	48	25	0.003
	T122K	0	0	6	3	0.03
	T122E	**0**	**0**	**18**	**9**	**<0.0001**
	G123K	0	0	6	3	0.03
	G123E	6	3	18	9	0.02
	H124K	0	0	8	4	0.007
	N126S	**16**	**8**	**130**	**68**	**<0.0001**
	N126G	0	0	12	6	0.0004
	N126K	0	0	6	3	0.03
	Q127P	0	0	8	4	0.007
	Q127K	**0**	**0**	**24**	**12**	**<0.0001**

Table 4.1a Non-cleavage site mutations in P17 significantly associated with PI-exposure.

Mutations showing significant different between 200 PI-naïve (none of them showing major PI mutations, 52 showing NRTI-resistance associated mutations and 23 showing NNRTI-resistance associated mutations) and 191 PI-experienced individuals (all of them showing major protease resistance associated mutations, 187/191 (98%) showing NRTI-resistance associated mutations and 125/191 (80%) showing NNRTI-resistance associated mutations) are presented. Total number f patients, percentage and p-values are shown. A p value of less than 0.05 was considered significant in a preliminary analysis. A Bonferroni's correction was subsequently applied and a p of less than 0.05/450 = 0.0001 was used instead. The mutations that remained significantly associated with PI exposure after the Bonferroni's correction was applied are shown in bold.

	MUTATION	PI-naïve (n= 200)		PI-experienced (n= 191)		
		n	%	n	%	P
P24	I138A	2	1	14	7	0.003
	V143T	2	1	10	5	0.04
	V143I	0	0	6	3	0.03
	A146S	2	1	10	5	0.04
	S173A	2	1	12	6	0.01
	Q182H	0	0	6	3	0.03
	T186M	**0**	**0**	**16**	**8**	**<0.0001**
	T190I	**0**	**0**	**16**	**8**	**<0.0001**
	M200I	0	0	12	6	0.0004
	A210S	**0**	**0**	**16**	**8**	**<0.0001**
	E211D	**0**	**0**	**16**	**8**	**<0.0001**
	V218A	0	0	6	3	0.03
	M228L	0	0	6	3	0.03
	L268M	10	5	22	11	0.04
	T280I	0	0	10	5	0.02
	S310T	**0**	**0**	**28**	**14**	**<0.0001**
	N315G	0	0	6	3	0.03

Table 4.1b Non-cleavage site mutations in P24 significantly associated with PI-exposure.

Mutations showing significant different between 200 PI-naïve (none of them showing major PI mutations, 52 showing NRTI-resistance associated mutations and 23 showing NNRTI-resistance associated mutations) and 191 PI-experienced individuals (all of them showing major protease resistance associated mutations, 187/191 (98%) showing NRTI-resistance associated mutations and 125/191 (80%) showing NNRTI-resistance associated mutations) are presented. A p value of less than 0.05 was considered significant in a preliminary analysis. A Bonferroni's correction was subsequently applied and a p of less than 0.05/450 = 0.0001 was used instead. The mutations that remained significantly associated with PI exposure after the Bonferroni's correction was applied are shown in bold.

	MUTATION	PI-naïve (n= 200)		PI-experienced (n= 191)		
		n	%	n	%	P
P7	I389V	0	0	6	3	0.03
	V390A	0	0	8	4	0.007
	R403K	52	26	74	39	0.02
	K415R	0	0	6	3	0.03
	M423I	0	0	6	3	0.03
	T427S	0	0	6	3	0.03
	T427I	0	0	10	5	0.002

Table 4.1c Non-cleavage site mutations in P7 significantly associated with PI-exposure.

Mutations showing significant different between 200 PI-naïve (none of them showing major PI mutations, 52 showing NRTI-resistance associated mutations and 23 showing NNRTI-resistance associated mutations) and 191 PI-experienced individuals (all of them showing major protease resistance associated mutations, 187/191 (98%) showing NRTI-resistance associated mutations and 125/191 (80%) showing NNRTI-resistance associated mutations) are presentedTotal number of patients, percentage and p-values are shown. A p value of less than 0.05 was considered significant in a preliminary analysis. A Bonferroni's correction was subsequently applied and a p of less than 0.05/450 = 0.0001 was used instead. The mutations that remained significantly associated with PI exposure after the Bonferroni's correction was applied are shown in bold.

	MUTATION	PI-naïve (n=200)		PI-experienced (n= 191)		
		n	%	n	%	P
P6	**F463L**	**0**	**0**	**22**	**11**	**<0.0001**
	F463V	0	0	6	3	0.03
	R464K	0	0	12	6	0.0004
	R464G	0	0	12	6	0.0004
	S465M	0	0	8	4	0.007
	G466R	0	0	6	3	0.03
	T469A	0	0	6	3	0.03
	T469I	**0**	**0**	**16**	**8**	**<0.0001**
	T470V	0	0	6	3	0.003
	T470A	14	7	34	18	0.003
	T471A	4	2	18	9	0.03
	T471S	0	0	10	5	0.002
	P472S	0	0	10	5	0.002
	Q474P	0	0	10	5	0.002
	E477G	0	0	10	5	0.002
	P478T	10	5	26	14	0.008
	P478Q	**0**	**0**	**22**	**11**	**<0.0001**
	I479V	0	0	8	4	0.007
	I479T	6	3	22	11	0.003

Table 4.1d Non-cleavage site mutations in P6 significantly associated with PI-exposure.

Mutations showing significant different between 200 PI-naïve (none of them showing major PI mutations, 52 showing NRTI-resistance associated mutations and 23 showing NNRTI-resistance associated mutations) and 191 PI-experienced individuals (all of them showing major protease resistance associated mutations, 187/191 (98%) showing NRTI-resistance associated mutations and 125/191 (80%) showing NNRTI-resistance associated mutations) are presented. A p value of less than 0.05 was considered significant in a preliminary analysis. A Bonferroni's correction was subsequently applied and a p of less than 0.05/450 = 0.0001 was used instead. The mutations that remained significantly associated with PI exposure after the Bonferroni's correction was applied are shown in bold.

	MUTATION	PI-naïve (n= 200)		PI-experienced (n= 191)		
		n	%	n	%	P
P1	K442R	0	0	8	4	0.007
P2	V370M	0	0	8	4	0.007
	T371A	0	0	8	4	0.007
	T371Q	**0**	**0**	**14**	**7**	**<0.0001**
	N372P	0	0	6	3	0.03

Table 4.1e Non-cleavage site mutations in P1 and P2 significantly associated with PI-exposure.

Mutations showing significant different between 200 PI-naïve (none of them showing major PI mutations, 52 showing NRTI-resistance associated mutations and 23 showing NNRTI-resistance associated mutations) and 191 PI-experienced individuals (all of them showing major protease resistance associated mutations, 187/191 (98%) showing NRTI-resistance associated mutations and 125/191 (80%) showing NNRTI-resistance associated mutations) are presented. A p value of less than 0.05 was considered significant in a preliminary analysis. A Bonferroni's correction was subsequently applied and a p of less than 0.05/450 = 0.0001 was used instead. The mutations that remained significantly associated with PI exposure after the Bonferroni's correction was applied are shown in bold.

4.3.5 Comparison of Gag cleavage site mutations according to history of exposure to protease inhibitors

At the five CSs, we detected 52 different mutations in PI-experienced individuals. The mutations were distributed as follow:

- P17/P24. Six mutations were present at three positions: V128, Q130 and Y132. Variants at these residues were detected in 34/191 (18%), 16/191 (8%) and 2/191 (1%) PI-experienced patients, respectively.

- P24/P2. Three mutations were present at three positions A360, V362 and S368. Variants at these residues were detected in 2/191 (1%), 34/191 (18%) and 4/191 (2%) PI-experienced patients, respectively.

- P2/P7. 20 mutations were present at eight positions: S373, A374, T375, I376, M377, M378, R380 and G381. Variants at these residues were detected at a frequency ranging between 1% and 36% in the PI-experienced population.

- P7/P1.Nine mutations were present at five positions: E428, R429, A431, K436, and I437. Variants at these residues were detected in 6/191 (3%), 2/191 (1%), 74/191 (39%), 14/191 (7%), 28/191 (15%) and 4/191 (2%) PI-experienced patients, respectively.

- P1/P6: 14 mutations were present at four positions: L449, S451, R452 and P453. Variants at these residues were detected at frequencies ranging between 1% and 21% in the PI-experienced population.

Overall, 17/52 CSMs were significantly associated with PI exposure when a cut-off of 0.05 was considered. The majority of them 14/17 remained significantly associated with PI exposure after Bonferroni's correction for multiple associations was applied (p= 0.05/50 = 0.01). The 14 mutations associated with PI-exposure were: two in P17/P24 (V128I and Y132F), none in P24/P2, four in P2/P7 (S373T, A374S, T375A and T375N), three in P7/P1 (A431V, K436R and I437V) and five in P1/P6 (L449F, S451T, S451R, R452S and P453T).

Details of CSMs in PI-experienced and PI-naïve individuals are presented in table 4.1f. Comparison of the number of mutations detected on PI-naïve vs. PI-experienced sequences is depicted in figure 4.5.

	Mutation	PI experienced patients (%)	PI-naïve patients (%)	p-value
P17/P24	**V128I**	**34 (18)**	**0 (0)**	**<0.0001**
Gag 128-137	**Y132F**	**16 (8)**	**0 (0)**	**< 0.0001**
VSQNY/PIVQN				
P2/P7	S373A	10 (5)	0 (0)	0.02
Gag 373-381	**S373T**	**16 (8)**	**0 (0)**	**<0.0001**
SATIM/MQRGN	**A374S**	**8 (4)**	**0 (0)**	**0.007**
	A374P	6 (3)	0 (0)	0.03
	T375A	36 (19)	15 (8)	0.002
	T375N	**36 (19)**	**10 (5)**	**<0.0001**
	G381S	10 (5)	2 (1)	0.02
P1/P7	**A431V**	**74 (39)**	**0 (0)**	**<0.0001**
Gag 428-437	K436R	14 (7)	2 (1)	0.003
ERQAN/FLGKI	**I437V**	**28 (15)**	**0 (0)**	**<0.0001**
P1/P6	**L449F**	**20 (10)**	**0 (0)**	**<0.0001**
Gag 444-453	**S451T**	**16 (8)**	**0 (0)**	**<0.0001**
RPGNF/LQSRP	S451R	8 (4)	0 (0)	0.007
	R452S	10 (5)	0 (0)	0.002
	P453T	10 (5)	0 (0)	0.002

Table 4.1f Gag CSMs significantly associated with PI-exposure.

Mutations showing significant different between 200 PI-naïve (none of them showing major PI mutations, 52 showing NRTI-resistance associated mutations and 23 showing NNRTI-resistance associated mutations) and 191 PI-experienced individuals (all of them showing major protease resistance associated mutations, 187/191 (98%) showing NRTI-resistance associated mutations and 125/191 (80%) showing NNRTI-resistance associated mutations) are presented. Total number of patients, percentage and p-values are shown. A p value of less than 0.05 was considered significant in a preliminary analysis. A Bonferroni's correction was subsequently applied and a p of less than 0.05/50 = 0.001 was used instead. The mutations that remained significantly associated with PI exposure after the Bonferroni's correction was applied are shown in bold.

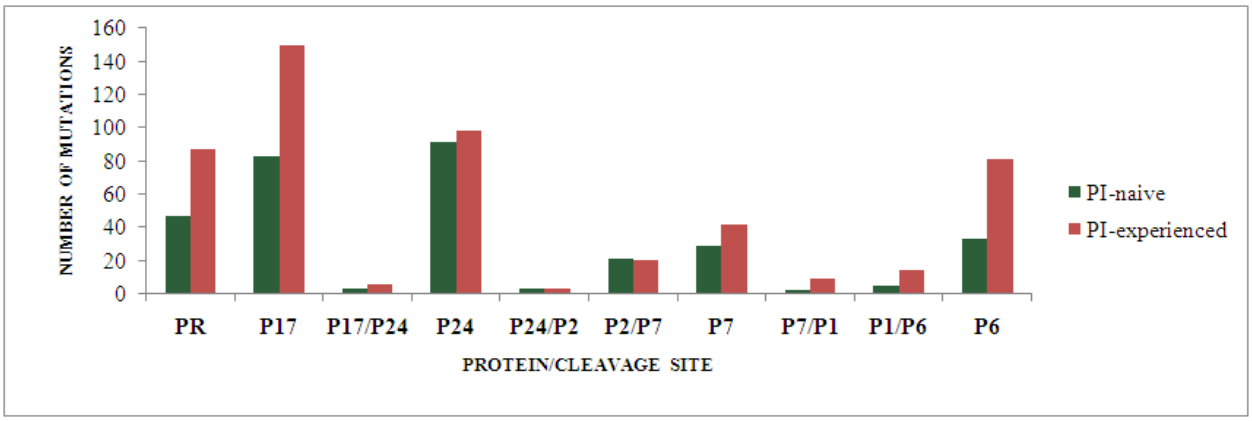


Figure 4.5 Number of different mutations detected in PR and Gag functional domains and CSs in PI-naïve and PI-experienced patients.

Sequences from 200 PI-naïve (none of them showing major PI mutations, 52 showing NRTI-resistance associated mutations and 23 showing NNRTI-resistance associated mutations) and 191 PI-experienced individuals (all of them showing major protease resistance associated mutations, 187/191 (98%) showing NRTI-resistance associated mutations and 125/191 (80%) showing NNRTI-resistance associated mutations) were compared.

4.3.6 PROTEASE CONSENSUS SEQUENCES

The reference sequence HXB2 for the 99 amino acids forming the HIV-1 protease and the two consensus sequences obtained from 200 PI-naive and 191 PI-experienced patients, respectively, are shown. Primary and secondary resistance positions are shown in red and blue, respectively. Dots denote identity with the reference sequence

Amino acid	1	2	3	4	5	6	7	8	9	10	11	12
HXB2	P	Q	V	T	L	W	Q	R	P	L	V	T
PI-naive	.	.	.	.	.	.	.	.	.	V	.	AINST
PI-experienced	.	.	.	.	.	.	.	.	.	IFV	I	AINST

Amino acid	13	14	15	16	17	18	19	20	21	22	23	24
HXB2	I	K	I	G	G	Q	L	K	E	A	L	L
PI-naive	AIV	.	IV	AEG	EG	KQ	ILT	IK	.	.	.	.
PI-experienced	AIV	.	IV	AEG	EG	KQ	ILT	IK	.	.	IL	IL

Amino acid	25	26	27	28	29	30	31	32	33	34	35	36
HXB2	D	T	G	D	D	D	T	V	L	E	E	M
PI-naive	.	.	.	.	.	.	.	.	FLV	.	ED	IM
PI-experienced	.	.	.	.	.	DN	.	IV	FLV	.	ED	IM

Amino acid	37	38	39	40	41	42	43	44	45	46	47	48
HXB2	S	L	P	G	R	W	K	P	K	M	I	G
PI-naive	ADES	.	KPQ	.	KR	.	KRT	.	KR	.	.	.
PI-experienced	ADES	.	KPQ	.	KR	.	KRT	.	KR	ILM	IV	AGMV

Amino acid	49	50	51	52	53	54	55	56	57	58	59	60
HXB2	G	I	G	G	F	I	K	V	R	Q	Y	D
PI-naive	.	.	.	.	.	.	.	.	KR	.	.	DE
PI-experienced	.	ILV	.	.	FL	ILMT	.	.	KR	EQ	.	DE

Amino acid	61	62	63	64	65	66	67	68	69	70	71	72
HXB2	Q	I	L	I	E	I	C	G	H	K	A	I
PI-naive	EQ	IV	AHLPVST	IMV	DE	.	CES	.	HQ	KR	.	ITV
PI-experienced	EQ	IV	AHLPVST	IMV	DE	.	CES	.	HQ	KR	AILV	ITV

Amino acid												
	61	62	63	64	65	66	67	68	69	70	71	72
HXB2	Q	I	L	I	E	I	C	G	H	K	A	I
PI-naive	EQ	IV	AHLPVST	IMV	DE	.	CES	.	HQ	KR	.	ITV
PI-experienced	EQ	IV	AHLPVST	IMV	DE	.	CES	.	HQ	KR	AILV	ITV

Amino acid	73	74	75	76	77	78	79	80	81	82	83	84
HXB2	G	T	V	L	V	G	P	T	P	V	N	I
PI-naive	.	S	.	.	IV	.	.	.	.	I	.	.
PI-experienced	CGST	PST	.	LV	IV	.	.	.	.	ACFSV	.	IV

Amino acid	85	86	87	88	89	90	91	92	93	94	95	96
HXB2	I	G	R	N	L	L	T	Q	I	G	C	T
PI-naive	.	.	.	.	LMV	.	.	.	IL	.	.	.
PI-experienced	.	.	.	DNS	LMV	LM	.	.	IL	.	.	.

Amino acid	97	98	99
HXB2	L	N	F
PI-naive	.	.	.
PI-experienced	.	.	.

4.3.7 GAG CONSENSUS SEQUENCES

The reference sequence HXB2 for the 500 amino acids forming the HIV-1 Gag polyprotein and the two consensus sequences obtained from 200 PI-naive and 191 PI-experienced patients, respectively, are shown below. CS positions are shown in red and the beginning of each domain is indicated. Dots denote identity with the reference sequence.

P17

Amino acid	1	2	3	4	5	6	7	8	9	10	11	12	13
HXB2	M	G	A	R	A	S	V	L	S	G	G	E	L
PI-naive	.	.	.	.	.	.	IV	IL	.	.	AG	EKQ	.
PI-experienced	.	.	.	.	.	.	.	.	.	.	.	.	.

Amino acid	14	15	16	17	18	19	20	21	22	23	24	25	26
HXB2	D	R	W	E	K	I	R	L	R	P	G	G	K
PI-naive	.	AKQR	.	.	KR	.	.	.	.	.	.	GS	KNRS
PI-experienced	.	AKR	.	.	KR	.	KQR	.	.	.	GM	.	KR

Amino acid	27	28	29	30	31	32	33	34	35	36	37	38	39	40
HXB2	K	K	Y	K	L	K	H	I	V	W	A	S	R	E
PI-naive	.	KMQR	.	KQR	.	.	.	ILV	.	.	.	.	.	.
PI-experienced	.	HKMQRT	.	KMQR	.	.	.	ILV	VL	.	.	GS	.	.

Amino acid	40	41	42	43	44	45	46	47	48	49	50	51	52
HXB2	E	L	E	R	F	A	V	N	P	G	L	L	E
PI-naive	.	.	.	.	.	ASV	ILV	.	.	GS	.	IL	.
PI-experienced	.	LM	.	.	FY	.	ILV	.	.	GS	.	.	.

Amino acid	53	54	55	56	57	58	59	60	61	62	63	64	65
HXB2	T	S	E	G	C	R	Q	I	L	G	Q	L	Q
PI-naive	.	ALS	DEGN	.	.	KR	.	.	ILM	AEGRTV	.	IL	.
PI-experienced	.	APST	ADEG	.	.	KQR	KQR	.	ILM	EGRSTV	HQ	.	HQ

Amino acid	66	67	68	69	70	71	72	73	74	75	76	77	78
HXB2	P	S	L	Q	T	G	S	E	E	L	R	S	L
PI-naive	.	AS	IL	KQ	.	.	.	.	.	FIL	KR	.	LV
PI-experienced	PS	AST	.	KQR	AT	.	ST	.	.	FLV	KR	.	.

Amino acid	79	80	81	82	83	84	85	86	87	88	89	90	91
HXB2	Y	N	T	V	A	T	L	Y	C	V	H	Q	R
PI-naive	FY	.	.	IV	AV	TV	.	.	.	.	.	QR	KNR
PI-experienced	FHY	ANT	.	ILV	AV	TV	.	FWY	CF	.	.	EQR	GKNQR

Amino acid	92	93	94	95	96	97	98	99	100	101	102	103	104
HXB2	I	E	I	K	D	T	K	E	A	L	D	K	I
PI-naive	.	DEG	IV	KQR	.	.	KR	.	.	.	DE	.	IV
PI-experienced	IV	DEGN	IV	KR	.	.	KNQ	.	.	.	DE	KQR	ILV

Amino acid	105	106	107	108	109	110	111	112	113	114	115	116	117
HXB2	E	E	E	Q	N	K	S	K	K	K	A	Q	Q
PI-naive	.	.	.	QR	NT	EK	CGRS	EK	.	IKT	AETV	.	PQ
PI-experienced	.	.	DEIQ	.	DKNST	EKQRT	CS	EKQR	EKQR	AKQR	AIKTV	PQ	EKNQPS

Amino acid	118	119	120	121	122	123	124	125	126	127	128	129
HXB2	A	A	A	D	T	G	H	S	N	Q	V	S
PI-naive	.	.	.	.	AT	EGR	.	.	KNSQ	NQS	ASV	.
PI-experienced	APTV	AEPTV	AENV	ADEGS	AEIKRTV	AEGNRV	ADHKNRS	DKNQS	EGNQRS	HKNPQST	AIV	.

P24

▼

Amino acid	130	131	132	133	134	135	136	137	138	139	140	141
HXB2	Q	N	Y	P	I	V	Q	N	I	Q	G	Q
PI-naive	QH	.	.	.	.	.	.	.	AILM	.	.	.
PI-experienced	HQRS	.	FY	.	.	.	.	.	AILMPV	PQ	.	.

Amino acid	142	143	144	145	146	147	148	149	150	151	152	153
HXB2	M	V	H	Q	A	I	S	P	R	T	L	N
PI-naive	.	IV	.	.	ANPS	ILM	ST	AP	.	.	.	.
PI-experienced	AMW	ITV	.	.	APS	ILMP	.	.	.	DT	.	.

Amino acid	154	155	156	157	158	159	160	161	162	163	164	165
HXB2	A	W	V	K	V	V	E	E	K	A	F	S
PI-naive	.	.	.	.	.	IV	EK	.	.	AGN	.	NS
PI-experienced	.	.	.	.	.	INV	.	.	.	AS	.	NS

Amino acid	166	167	168	169	170	171	172	173	174	175	176	177
HXB2	P	E	V	I	P	M	F	S	A	L	S	E
PI-naive	.	.	IV	.	.	.	.	AST	.	.	AS	.
PI-experienced	.	.	.	IN	.	.	.	AST	.	.	AS	DE

Amino acid	178	179	180	181	182	183	184	185	186	187	188	189
HXB2	G	A	T	P	Q	D	L	N	T	M	L	N
PI-naive	.	AC	.	.	QS	.	.	.	ST	.	.	.
PI-experienced	.	.	.	.	HQ	.	.	.	MTV	.	.	.

Amino acid	190	191	192	193	194	195	196	197	198	199	200	201
HXB2	T	V	G	G	H	Q	A	A	M	Q	M	L
PI-naive	AT	.	.	.	.	.	.	.	.	.	.	.
PI-experienced	IT	IV	.	.	.	.	.	.	.	.	IM	.

Amino acid	202	203	204	205	206	207	208	209	210	211	212	213
HXB2	K	E	T	I	N	E	E	A	A	E	W	D
PI-naive	.	DE	.	.	.	DE	.	.	.	.	.	.
PI-experienced	.	DE	AT	.	.	DE	.	.	ASV	DE	.	.

Amino acid	214	215	216	217	218	219	220	221	222	223	224	225
HXB2	R	V	H	P	V	H	A	G	P	I	A	P
PI-naive	.	ILMTV	.	.	AV	HQ	.	.	.	AINV	APV	.
PI-experienced	.	AILMTV	.	.	AV	HQ	.	.	.	AIV	APV	.

Amino acid	226	227	228	229	230	231	232	233	234	235	236	237
HXB2	G	Q	M	R	E	P	R	G	S	D	I	A
PI-naive	.	.	IM	.	DE	.	.	.	.	.	.	.
PI-experienced	.	.	IKLM	.	DE	.	.	.	.	.	.	AP

Amino acid	238	239	240	241	242	243	244	245	246	247	248	249
HXB2	G	T	T	S	T	L	Q	E	Q	I	G	W
PI-naive	.	ST	.	.	NT	.	AQ	.	.	IV	AGNT	.
PI-experienced	.	ST	PT	.	NST	LP	.	DE	.	IV	AGNT	.

Amino acid	250	251	252	253	254	255	256	257	258	259	260	261
HXB2	M	T	N	N	P	P	I	P	V	G	E	I
PI-naive	.	.	GNST	.	AP	AP	IV	.	.	.	DE	.
PI-experienced	.	.	GHNS	.	.	AP	ITV	.	.	.	DE	.

Amino acid	274	275	276	277	278	279	280	281	282	283	284	285
HXB2	V	R	M	Y	S	P	T	S	I	L	D	I
PI-naive	.	KR	.	.	.	.	STV	.	.	.	.	.
PI-experienced	.	KR	.	.	.	.	ACITV	.	.	.	.	.

Amino acid	286	287	288	289	290	291	292	293	294	295	296	297
HXB2	R	Q	G	P	K	E	P	F	R	D	Y	V
PI-naive	EKR	.	.	.	.	.	PS	.	.	.	.	.
PI-experienced	KR	.	.	.	.	.	PS	.	.	.	.	.

Amino acid	298	299	300	301	302	303	304	305	306	307	308	309
HXB2	D	R	F	Y	K	T	L	R	A	E	Q	A
PI-naive	.	.	.	FY	KRS	TV	.	.	.	.	.	AS
PI-experienced	.	.	.	FY	.	ATV	IL	.	.	.	.	.

Amino acid	310	311	312	313	314	315	316	317	318	319	320	321
HXB2	S	Q	E	V	K	N	W	M	T	E	T	L
PI-naive	ST	PQ	DE	.	.	HNST	.	.	.	DE	.	.
PI-experienced	ST	.	DE	.	.	GN	.	.	.	DE	ST	.

Amino acid	322	323	324	325	326	327	328	329	330	331	332	333
HXB2	L	V	Q	N	A	N	P	D	C	K	T	I
PI-naive	.	IV	.	.	AS	.	.	.	.	KR	NTS	.
PI-experienced	.	.	.	.	AS	.	.	.	.	.	ST	.

Amino acid	334	335	336	337	338	339	340	341	342	343	344	345
HXB2	L	K	A	L	G	P	A	A	T	L	E	E
PI-naive	.	KR	.	LM	EG	PQ	AG	.	ST	.	.	DE
PI-experienced	.	KR	.	.	.	PT	AG	AS	ST	LQ	EK	DE

Amino acid	346	347	348	349	350	351	352	353	354	355	356	357
HXB2	M	M	T	A	C	Q	G	V	G	G	P	G
PI-naive	IM	.	AIT	.	.	.	.	.	.	.	.	GS
PI-experienced	.	.	AST	.	.	.	.	LV	.	.	.	GS

P2

Amino acid	358	359	360	361	362	363	364	365	366	367	368	369
HXB2	H	K	A	R	V	L	A	E	A	M	S	Q
PI-naive	.	.	.	KR	IV	.	.	.	.	.	GS	HQ
PI-experienced	.	.	AS	.	IV	.	.	.	.	.	CS	HQ

P7

Amino acid	370	371	372	373	374	375	376	377	378	379	380	381
HXB2	V	T	N	S	A	T	I	M	M	Q	R	G
PI-naive	AIV	SNT	GNQS	APQST	ASNPT	AINSTV	IMV	LM	IMV	.	KR	GS
PI-experienced	AMV	ANQST	GNPQS	ACPST	ANPST	AINSTV	IMV	LM	IMV	.	KR	GS

Amino acid	382	383	384	385	386	387	388	389	390	391	392	393
HXB2	N	F	R	N	Q	R	K	I	V	K	C	F
PI-naive	NT	FY	KR	NST	PQ	KR	KR	INMPST	IV	.	.	.
PI-experienced	.	FY	NKRS	GKNRS	PQ	GKR	.	AINPRSTV	AIV	KR	.	.

Amino acid	394	395	396	397	398	399	400	401	402	403	404	405
HXB2	N	C	G	K	E	G	H	T	A	R	N	C
PI-naive	.	.	.	KR	DET	.	.	ILT	AS	KR	.	.
PI-experienced	.	.	.	KR	DEQT	.	.	ILTV	.	KR	.	.

Amino acid	406	407	408	409	410	411	412	413	414	415	416	417
HXB2	R	A	P	R	K	K	G	C	W	K	C	G
PI-naive	KR	.	.	.	KR	KR	.	.	.	.	.	.
PI-experienced	KR	.	.	.	KR	KR	.	.	.	KR	.	.

Amino acid	418	419	420	421	422	423	424	425	426	427	428	429
HXB2	K	E	G	H	Q	M	K	D	C	T	E	R
PI-naive	KQR	.	.	.	.	IM	KR	DE	.	NT	ED	.
PI-experienced	KQR	.	.	.	.	IM	KR	DE	.	AINST	EDG	GKR

P1

Amino acid	430	431	432	433	434	435	436	437	438	439	440	441
HXB2	Q	A	N	F	L	G	K	I	W	P	S	Y
PI-naive	.	.	.	.	.	.	KR	.	.	PS	PS	HY
PI-experienced	.	AV	.	.	.	.	GKR	ILV	.	PS	PS	HY

P6

Amino acid	442	443	444	445	446	447	448	449	450	451	452	453
HXB2	K	G	R	P	G	N	F	L	Q	S	R	P
PI-naive	.	.	.	.	.	.	.	LP	EQ	ANS	.	LP
PI-experienced	KR	EG	.	.	.	.	.	FHLPV	.	GNRST	GRS	ALPST

Amino acid	454	455	456	457	458	459	460	461	462	463	464	465
HXB2	E	P	T	A	P	P	E	E	S	F	R	S
PI-naive	ET	.	ST	.	.	.	AE	.	.	.	.	.
PI-experienced	AEP	.	ST	.	.	.	AE	EQ	ILS	FLV	EGKLR	CEFMS

Amino acid	466	467	468	469	470	471	472	473	474	475	476	477
HXB2	G	V	E	T	T	T	P	P	Q	K	Q	E
PI-naive	GR	.	EG	KT	AT	APT	.	PS	.	KR	PQ	DE
PI-experienced	GMQR	EV	EGK	AIKT	AITV	APQST	PS	PS	QP	.	.	DEG

Amino acid	478	479	480	481	482	483	484	485	486	487	488	489
HXB2	P	I	D	K	E	L	Y	P	L	T	S	L
PI-naive	LPST	IKRT	.	.	DEG	KLMQ	PY	.	LV	AT	.	.
PI-experienced	AKPQST	AEGIKLRTV	DEGNSV	EKNQR	DEG	KLMQP	APY	AP	.	ADSTV	AS	.

Amino acid	490	491	492	493	494	495	496	497	498	499	500
HXB2	R	S	L	F	G	N	D	P	S	S	Q
PI-naive	KR	.	.	.	.	NS	.	PL	LS	LS	KQ
PI-experienced	KR	.	.	.	.	NS	DS	PQ	LST	.	KQT

4.3.8 Association between protease and Gag mutations

We analyzed the 191 sequences obtained from PI-experienced patients and explore the relationship between major resistance mutations in PR (22 mutations at 14 positions listed along the 99 amino acids) and Gag mutations identified in this study and described in section 4.3.4 as being associated with PI-exposure, after Bonferroni's correction for multiple comparisons was applied (19 mutations at 18 non-CS positions and 14 mutations at 12 CS positions). Therefore, PR mutations including in this analysis were: D30N, V32I, M46I, M46L, I47V, G48A, G48M, G48V, I50L, I50V, I54L, I54M, I54T, L76V, V82A, V82C, V82F, V82S, I84V, N88D, N88S and L90M. Gag Non-CSMs associated with PI-exposure included in the analysis were: L61I, I94V, K103R, K113Q, K114R, D121G, D121A, T122E, N126S, Q127K, T186M, T190I, A210S, E211D, S310T, T371Q, F463L, T469I and P478Q. Finally, Gag CSMs associated with PI exposure included in the analyses were: V128I, Y132F, S373T, A374A, T375A, T375N, A431V, K436R, I437V, L449F, S451T, S451R, R452S and P453T. A total of 418 combinations of Non-CSMs and major PR resistance mutations were identified and analyzed with Fischer's exact test. Only two associations were selected based upon the Bonferroni's corrected p value cut-off ($< 0.05/418 = 1.2 \times 10^{-4}$). Similarly, we identified 308 combinations of major PR and CSMs and only one associations retained significance after applying the Bonferroni's correction for multiple associations ($p < 0.05/308 = 1.6 \times 10^{-4}$). The association between the CSM A431V and the PR mutations M46I/L as well as between CSM L449F and I84V that have been extensively described in the literature were close to the cut-off p value defined in the study ($p = 6 \times 10^{-4}$ and 2×10^{-4}, respectively) . Results are shown in table 4.2

	***P* values for indicated Gag residues[a]**					
Protease mutations	P17 N126S	P17/P24 Y132F	P2/P7 T375A	 G381S	P7/P1 A431V[b]	P1/P6 L449F[b]
D30N		0.008				
M46I/L	**$< 10^{-5}$**	0.03	0.004		6×10^{-4}	
I54V	**$< 10^{-5}$**				**$< 10^{-4}$**	
V82A			0.02			
I84V						**2×10^{-4}**
L90M					0.01	

Table 4.2 Association between major protease resistance mutations and Gag mutations.

The table shows the associations between Gag and major protease resistance mutations which were statistically significant when considered a p value = 0.05. [a]*p* values in bold correspond to the associations retained using the Bonferroni's correction method ($p= 1.2 \times 10^{-4}$ for NCSMs and 1.6×10^{-4} for CSMs, respectively). [b]The association between A431V and M46I/L and between L449F and I84V showed p values close to the Bonferroni's corrected p value defined in the study $p = 1.6 \times 10^{-4}$.

4.4 Discussion

In the present chapter, we performed a cross-sectional comparison of sequences from PI-naïve and PI-experienced individuals in order to assess differences in the prevalence and patterns of mutations in the Gag protein. We identified key substitutions in both cleavage and non-cleavage site residues of Gag that were significantly associated with PI-exposure and the presence of major resistance mutations in the *PR* gene.

Multiple pressures shape the evolution of *Gag* during long-term HIV-1 infection. Firstly, its key role in viral assembly and infectivity, which is mediated by the functional role of different gag cleavage products (Wang et al. 1993). Secondly, the immune system, as numerous CD8 epitopes are known to be located across the gag protein. Lastly, in those patients on suboptimal PI-based regimens who develop mutations on PR, Gag protein is likely to co-evolve in order to preserve the efficient cleavage of the polypeptide by the mutated PR.

During or soon after the release of the immature virion from the plasma membrane, the gag polyprotein precursor ($Pr55^{gag}$) is cleaved by the viral PR into four major Gag cleavage products, namely P17, P24, P7 and the C-terminal peptide P6. In addition, two small spacer peptides are generated - P2 which is closer to the N-terminus of Gag and P1. Polyprotein processing causes a dramatic transformation in viral structure. This process, known as maturation, gives rise to the condensed conical core, which is characteristic of fully infectious HIV viral particles. Each of the individual Gag proteins has multiple functions. The P17

protein plays a crucial role in targeting the Gag precursor to the site of assembly on the plasma membrane (Facke, et al. 1993; Spearman, et al. 1994; Wang et al, 1993; Yuan et al. 1993); it is also essential for the stable association of the envelope glycoprotein with the viral capsid (Dorfman et al, 1994; Wang et al. 1993; Yu, et al. 1992) and its association with the pre-integration complex suggests that it might be important in directing this complex to the nucleus (Burkrinsky et al, 1993a; Burkrinsky et al, 1993b). The P24 protein is the major structural component of the virion, forming the capsid that encases the ribonucleoprotein complex. It influences both viral assembly and replication activities and determines the internal organization of the assembled and budded viral material (Cairns and Craven, 2001). The P7 protein contains sequences that are essential for the efficient encapsidation of the viral genomic RNA into the assembled viral particles (Aldovini and Young, 1990; Dorfman et al, 1993; Gorelick, et al. 1988). The C-terminus proline-rich P6 peptide appears to be important in mediating viral budding (Gottlinger, et al. 1991). Little is known about the function of the two spacer peptides P1 and P2, which are present in all primate lentivirus. Although poor conservation in sequence and length of these two peptides has been reported among different viruses (Henderson et al, 1988), their consistent presence suggests that they may have an important function in the retrovirus life-cycle. In fact, in the case of P2, early studies showed that this spacer peptide is essential for virus replication (Henderson et al, 1988).

In agreement with the crucial roles that each functional Gag protein exerts in the life-cycle of HIV, we observed conserved amino acid motifs in each of them, including many known to be essential in the accomplishment of protein function during viral assembly. The first few amino acids located in the N-terminal of MA-P17 contain a myristilation signal that is essential for the formation of extracellular viral particles (Bryant et al. 1990; Gottlinger et al.

1989). In addition, a stretch of basic amino acids located between amino acid 17 and 31 is known to be responsible for the targeting of HIV-1 Gag to the plasma membrane during virus assembly by interacting with acidic phospholipids (Zhou et al, 1994). In an early mutagenesis analysis of the matrix protein performed by Freed and colleagues, mutations at some residues in this region, (i.e., R20, L21 and P23) were found to significantly decrease virus production (Freed et al, 1994). In the same study, several scattered residues in the C-terminus of matrix, such as A37, L50, E52, were also reported to decrease viral production and mutations at G56, C57 and I60 resulted in the complete absence of viral replication as well as mutations between positions L85 and H89.

A high degree of conservation was observed in P24. The major homology region (MHR), which is located between amino acids I285 and L304, is known to be conserved across all lentivirus and has an important role in viral replication (Gorelick et al, 1990). Mutations in this region often interfere with particle assembly (Borsetti et al, 1998; Dorfman et al, 1994; Ebbets-Reed et al, 1996; Gamble et al, 1997; Mammano et al, 1994).Mutations at residues Q155, E159 and Y164 substantially reduce gag proteins release. In addition, numerous mutations scattered throughout the P24 protein have been found to block viral assembly and release (Chazal et al, 1994; Mammano et al 1994; Zhao et al 1994).

The P7 protein contains the histidine box located between amino acids C392 and C405, which constitutes a metal-binding domain essential for effective RNA encapsidation, and mutations in this region yield viral particles with defective RNA encapsidation (Aldovini et al 1990; Gorelick et al 1988; Gorelick et al 1990). Three cysteines (C392, C395 and C405) and

one histidine (400H) are conserved. In addition, an aromatic amino acid at position Y393 and a basic amino acid at position K397 are also conserved.

Consistent with these findings, the residues mentioned above were also highly conserved across our study population.

P6 has been found to be the most variable protein in all primate lentiviruses including HIV (Accola et al, 2000). In agreement with this data, we observed some degree of variability in most of the amino acids of this protein in our population. However, even in this highly variable peptide we could recognise two highly conserved motifs. The first one was the PT/SAP motif located close to the amino terminus end of P6, from amino acids 455P to 459P; this has been reported to be essential for viral release (Gottlinger et al, 1993; Huang et al, 1995). It has been demonstrated that mutations in this domain cause a defect in the budding process, which results in a larger number of immature particles tethering to the plasma membrane (Gottlinger et al 1991; Huang et al, 1995). The second conserved domain in the P6 protein was located towards the C-terminus and included amino acid L489 and 491SLFG494. The conservation of this motif is less understood; the residues are known to be essential for the incorporation of the accessory protein Vpr into the assembling HIV-1 virion (Kondo et al, 1995; Kondo et al, 1996; Lu et al, 1995). However, since this motif is dispensable for the incorporation of Vpr or the equivalent Vpx in certain simian immunodeficiency virus (Accola et al, 1999), it can be proposed that the interaction of the motif with other crucial factors during viral assembly must be in fact responsible for the striking conservation across HIV-1 strains. Apart from this constraint, the rest of the P6

protein seems to exhibit high plasticity with the exception of a few scattered conserved residues.

With regard to the CSs, we observed a high degree of conservation in some of the amino acids that constitute the site of cleavage of HIV-1 PR, in particular those that directly constitute the cleavage bond (P1 and P1'). Sequential and ordered proteolytic processing of HIV-1 Gag is required to achieve fully infectious viral particles, and this process partially depends on the amino acid sequence within the processing site, thus explaining the necessity for some degree of conservation in the CS sequences. However, other determinants, such as the sequence of regions near the processing site that determines its accessibility are likely to contribute to effective processing. In addition, conservation across CSs is not equally distributed. Some of them, such as P2/P7, were highly variable, while others, such as P24/P2, were highly conserved. Amino acids at CSs are not only important for the concerted cleavage of the gag precursor polyprotein by the viral PR, but may in some cases be part of larger domains crucial for completion of the viral life-cycle. For example, amino acids located at the P24/P2 CS, and in particular those between amino acids L363 and S368, have been found to be part of a domain in P24 essential for viral replication. By contrast, those located in the P2/P7, and particularly those at the N terminus site between amino acids 373S and 378M, can be deleted without having a significant impact on viral replication, (Accola et al, 1998), which may at least partially explain the marked differences in variability in these two CSs.

When we compared the PR and gag sequences obtained from PI-naïve and PI-experienced patients, we observed that PI-experienced patients had a higher variability not only in PR but

also throughout the Gag polyprotein. Importantly, certain mutations were significantly more prevalent in PI-experienced patients and others were completely absent across the PI-naïve population, suggesting that selective pressure with PIs leads to genetic evolution not only in *PR* but also in *Gag*. A role for Gag mutations in mediating PI resistance was proposed early after the introduction of PIs in antiretroviral therapy. In 1996, Doyon and colleagues demonstrated that HIV-1 variants highly resistant to PIs *in vitro*, showed mutations in the PR and also in the two gag CSs P1/P6 and P7/P1. In addition, they observed that CSMs improved polyprotein processing in viruses with PR mutations, providing the first evidence of a possible mechanism by which mutations in Gag can compensate for impaired PR activity. Furthermore, they observed that PR-mutated viral clones that contained CSMs grew much better *in vitro* than clones in which such mutations were removed by site-directed mutagenesis, highlighting a potential compensatory role for the CSMs (Doyon et al, 1996). Since then, numerous studies have demonstrated a role for Gag CSMs in failure to PIs, in particular for those located at the C-terminal P7/P1/P6. Although most studies agree on attributing a compensatory role to CSMs, the mechanism by which these mutations exert their function seems to differ. Thus, Gatanaga and colleagues observed that mutations in Gag were indeed essential for the efficient replication of APV-resistant variants. However, while the mutated gag was required for viral replication, the rate of cleavage was comparable to that exhibited by wild-type gag, suggesting that an alternative mechanism must be responsible for the compensatory role of Gag mutations towards improved polyprotein processing (Gatanaga et al, 2002).

Most of the mutations at these two CSs (P7/P1 and P1/P6) that were associated with PI-exposure in our study, such as A431V, I437V, K436R or L449F, have been previously described. Thus, in line with previous studies (Verheyen et al, 2006), the A431V mutation

was found to be the most prevalent mutation in PI-experienced patients in our population. This mutation has been described before, both *in vitro* and *in vivo* (Mammano et al, 2000; Zhang et al, 1997; Maguire et al, 2002; Cote et al, 2001; Bally et al, 2000 and Dauber et al, 2002). The residue is in direct contact with the substrate-binding pocket and the mutation results in an enhanced cleavage of P7/P1 by wild-type PR. Increased processing of A431V-containing CS was also demonstrated in the background of primary PR resistance mutations at codons 46, 82, 84 and 90 (Feher et al, 2002). Similarly, the L449F mutation, which on its own does not confer PI resistance, was associated with reduced PI susceptibility in the background of the major PR mutation I50V (Maguire et al, 2002; Prado et al, 2002). In addition, mutations such as I437V, which was also found to be associated with PI-exposure in our study, have been associated with enhanced proteolytic processing and direct contribution to PI resistance. Of note, the mutation was selected in the absence of any substitution in the PR during *in vitro* passages employing novel PIs (Nijhuis et al, 2007).

In our study, we found mutations associated with PI-exposure not only at P7/P1/P6 CS but also at all other CS, except for P24/P2. This cleavage site, similarly to what we observed in PI-naïve individuals, was also highly conserved in PI-experienced patients and the rare mutations detected in this site did not show a significant different prevalence between PI-naïve and PI-experienced patients. The P24/P2 sequence has been reported to be one of the best HIV-1 processing sites (Richards et al, 1990 and Tozser et al, 1991), an observation that may account for the high degree of conservation observed in both PI-naïve and PI-experienced patients, as little improvement can be added by sequence modification. In addition, as mentioned above, this CS along with sequences located closely in P24 and P2 constitutes a domain that plays a pivotal role in viral replication (Accola et al, 1998).

We detected a number of mutations at P2/P7 and P17/P24 that were significantly associated with exposure to the PIs. The implication of mutations at these two CS in PI failure has been only briefly suggested in the literature. For example, the two mutations in P17/P24 associated with PI-exposure in our study, namely V128I and Y132F, have been previously reported. V128I was documented by Dierynck and colleagues as the only CS gag mutation emerging in patients whose viral load rebounded during DRV/r monotherapy (Dierynck et al, 2007). Similarly, the Y132F mutation was found to be linked to certain primary PR mutations and associated with improved rate of cleavage of P17/P24 (Myint et al, 2004; Ueda et al, 2005). Mutations at P2/P7 have also been briefly documented. Malet and colleagues reported an association between variants at codon 373 and impaired responses to ritonavir booster saquinavir (SQV/r)-based regimens. Nevertheless, a clear role for mutations at P17/P24 and P2/P7 in failure to PIs remains to be established, although it is worth mentioning that the number of studies looking at the P7/P1/P6 CS by far outnumbers that of studies looking at other CSs, explaining the substantial difference in information regarding mutations in PI-experienced patients at this CS and the others.

In the present chapter, we have seen that under PI-selective pressure mutations are likely to be selected not only in PR and its CSs but also in gag regions outside of the CSs. A role for Gag mutations outside cleavage site in PI failure has been previously suggested. Gatanaga and colleagues studied the effect of various substitutions on the development of HIV-1 resistance to APV: L75R in the P17 protein, H219Q in the P24 protein, V390D/A and R403K in the P7 protein and E468K in the P6 protein, together with the P1/P6 CSM L449F. They concluded that both the CSMs and the mutations located outside CSs were essential for the

efficient replication of APV-resistant HIV-1. They also observed that while some non-CSMs were selected before major APV resistance-associated mutations, others only emerged after the selection of major PR mutations; both groups were shown to be required for efficient viral replication. However, they failed to demonstrate any differences in the mutated PR cleavage patterns between the wild-type and the gag mutated virus, suggesting once again that gag evolution under PI-selective pressure may facilitate certain functions such as assembly, packaging and budding functions and contribute to resistance to PIs by means other than recovery of polyprotein cleavage function (Gatanaga et al, 2002). Similarly, Myint and colleagues demonstrated that non-CSMs are as important as CSMs for the recovery of fitness in PI-resistant viruses. In addition, they saw that non-CSMs differed in different viral clones, suggesting that while PR and CS mutations implicated in PI-failure may be consistent across different patients, the pattern of non-CSMs may vary widely (Myint et al, 2004).

In our study, we observed a trend towards a higher number of mutations in PI-experienced patients compared with PI-naïve patients at all gag domains. The effect was particularly noticeable in the P17 and the P6 proteins, suggesting a potentially more prominent role of the two proteins in PI failure compared with other gag regions. Recently, Parry and co-authors studied full-length *Gag* and *PR* genes from one multidrug resistant clinical isolate and showed that the P17 protein from the multidrug resistance virus was on its own able to rescue the replicative capacity (RC) of the mutated PR to the level observed in the WT virus. In addition, they also demonstrated that such mutated P17 protein could lead to a reduction in susceptibility to PIs in the absence of major PR resistance mutations, suggesting that major determinants of PI-resistance may be located in the HIV-1 P17 protein (Parry et al, 2009). They observed 12 mutations in the P17 protein from the clinical isolate when compared with

the HXB2 reference sequence; of these six (I34L, T84V, E93D, I94V, N124K and N126S) were found to be associated with PI selective pressure in our analysis. In particular two mutations (I94V and N126S) remained significant even after applying the Bonferroni's correction. Similarly, other studies have suggested that the P6 protein inhibits HIV-1 PR function (Paulus et al, 1999) and therefore it is speculated that variations in its sequence may change PR activity, thereby affecting viral fitness and PI susceptibility. Kaufman and co-workers explored this hypothesis by looking at mutations in the P7/P1/P6 CS and the P6 protein in patients who experienced virological failure while on a SQV/r-based therapy. They observed that mutations emerged in the P6 region and were in general associated with major PR mutations which suggest a compensatory role for mutations in this region. They also reported an association between insertions in the P6-region and failure (Kauffman et al, 2001). This observation has not been confirmed by other investigators, and in fact we did not find a higher prevalence of insertions in the P6 protein in PI-experienced patients compared with PI-naïve individuals in our study.

Several statistically significant associations between Gag and PR mutations were identified in the present analysis. The M46I/L mutation was associated with the CSM A431V as well as with the mutation N126S in P17. A431V was also associated with the I54V mutation in PR. The M46I/L and I54V mutations have been found to be associated in isolates obtained from patients receiving PI-based regimen (Wu, 2003) and may be selected in patients failing IDV/r (Condra et al, 1996). The relationship of the M46I/L and I54V cluster with the A431V mutation in Gag has been previously described (Bally et al, 2000; Koch et al 2001; Malet et al, 2007). In addition, we identified an association between the PR mutation I84V and the P1/P6 Gag CSM L449F. This association was also previously described by Verheyen and

colleagues (Verheyen et al; 2006). In addition, we observed an association between a non-CSM located in the P17 protein N126S and two major PR mutations M46I/L and I54V, which has not been previously reported. However, as it has been previously pointed, studies addressing the impact of mutations beyond the Gag CSs NC-P7/P1/P6 are scarce. In the study performed by Parry and colleagues, a multi-protease resistant patient showed the M46I, L33F, I54V and V82A mutations in PR and full length gag sequencing identified the N126S mutation in the P17 protein along with other seven mutations. However, the study was based on the analysis of a single patient and therefore no statistical association between mutations could be confirmed. As discussed above, the study showed the important role that the P17 protein plays in conferring resistance to PIs, but whether this can be attributed to specific residues or to the whole protein was not clarified in the study.

It should be emphasized that all of the sequences analyzed in the present chapter were subtype B HIV-1 and consequently those mutations identified as associated with PI exposure in our population could be naturally occurring polymorphism in other HIV-1 subtypes what may have implications for susceptibility to PIs of such non-B subtype viruses. Studies addressing the variability of gag in both B and non-B subtypes are scarce. De Oliveira and colleagues reported a greater variability of Gag for subtype C HIV-1 compared to B subtype particularly at certain CSs such as P2/P7 and P17/P24 (De Oliveira et al, 2003). Similarly, Jinnopat reported a higher Gag variability in subtype CRF01 HIV-1 (Jinnopat et al, 2009) and in particular the mutations L61I and P66S in P17 were frequently detected in drug-naïve CRF01 viruses. Both L61I and P66S mutations were associated with PI-exposure in our study and in the case of L61I mutation the association remained significant once the Bonferroni's correction for multiple associations was applied. In addition, Gupta and co-authors detected

the T84V mutation frequently in subtype A (Gupta et al, 2010), which have also been associated with PI-exposure in our study. Importantly, both Jinnopat and Gupta reported a decreased susceptibility of recombinant viruses containing CRF01 and A subtype Gag, respectively, compare to viruses containing subtype B Gag (Jinnopat et al, 2009 and Gupta et al, 2011). Overall, these findings suggest that indeed mutations selected under PI –selective pressure in subtype B HIV-1 can occur as natural polymorphism in non-B subtypes and may contribute to decrease susceptibility to PIs of non-B subtype HIV-1 viruses. However, further studies are required to confirm this hypothesis. In summary, in this chapter we have showed that several Gag mutations are more common in subtype B-infected PI-experienced patients than in those who have never been exposed to these drugs infected with the same subtype, and demonstrated that certain specific mutations were in fact statistically associated with PI-exposure.

In agreement with previous studies, we have found that mutations located at P2/P7/P1/P6 including T375A, A431V, I437V and L449F were associated with PI exposure and with the presence of specific PR mutations namely M46IL, I54V and I84V. In addition, by addressing full-length Gag sequencing, we have expanded previous reports and have identified a number of mutations strongly associated with PI selective pressure namely, V128I, Y132F, S373T, A374S, T375N and S451T in Gag CS and L61I, I94V, K103R, K113Q, K114R, D121AG, N126S, T186M, A210S, E211D, T371Q, F463L, T469I and P478Q outside gag CS.

As a consequence full-length PR and gag sequencing may be of importance for the full assessment of failure to PIs. A longitudinal analysis of *Gag* and *PR* genes in patients with

ongoing viraemia while on PI-based regimen may be helpful to clarify the role of gag mutations in failure to PI inhibitors. Clonal analysis and phenotypic analysis could be also important to confirm linkage and effect of specific gag mutations. These points will be addressed in subsequent chapters.

5 Chapter five: genetic evolution of HIV-1 *Gag* and *protease* in patients failing a ritonavir-boosted protease inhibitor-based antiretroviral regimen

5.1 Introduction

Ritonavir-boosted protease inhibitors (PI/r) are among the most effective antiretroviral drugs (ARVs) currently employed for the treatment of HIV infection. These compounds inhibit the proteolytic activity of the viral protease enzyme and as a result they exert a powerful inhibitory action on HIV replication both "*in vitro"* and "*in vivo"*. In the large majority of treated patients, combination regimens containing PI/r result in effective and sustained suppression of HIV replication, dramatically reducing HIV-related morbidity and mortality. However, as with all other ARVs, failure to fully suppress HIV replication leads to the development of PI drug-resistance.

HIV resistance to PIs is a stepwise process in which accumulation of amino acid substitutions in the viral protease causes a progressive increase in the level of resistance. (Condra et al, 1996; Molla et al, 1996). The first mutations to be selected are generally those affecting amino acids that are in or close to the substrate-binding site of the enzyme: these mutations, which are termed primary mutations, often differ from one PI to another. Subsequently, secondary mutations develop, which involve amino acids located away from the substrate-binding site: these mutations are generally less drug-specific than primary mutations. Primary and secondary mutations produce in conjunction an enlargement of the catalytic site of the

enzyme, which decreases the affinity of the viral protease for the inhibitor (Logsdon et al; 2004; Prabu-Jeyabalan et al; 2006a; Prabu-Jeyabalan et al, 2006b). In general, primary mutations are accountable for the resistant phenotype displayed by the mutated virus. By contrast, secondary mutations often do not have a substantial resistance effects, but they restore, at least partially, the fitness impairment caused by primary mutations. As a consequence, secondary mutations are still critical for the development of high-level PI resistance (Mammano et al, 2000; Martinez-Picado et al, 1999; Zennou et al, 1998).

As discussed in the previous chapter, a key feature of resistance to the PIs is that mutations implicated in drug resistance and viral fitness are located not only in the viral protease, but also in its natural substrate Gag. Similarly to secondary mutations in the HIV-1 protease, substitutions in Gag were initially described as compensatory mutations aimed at restoring the loss of viral fitness caused by primary mutations (Doyon et al, 1996; Zhang, et al. 1997). Following these observations, several studies pointed to the importance of Gag mutations in the evolution of HIV resistance to the PIs (Brumme et al, 2003; Banke et al, 2009; Brann et al, 2006; Cote et al, 2001; Maguire et al, 2002; Malet et al, 2007; Robinson et al, 2000). Since then, evidence has been accumulating that Gag mutations can also directly affect HIV susceptibility to the PIs independently of their effect on viral fitness (Zhang et al, 1997;; Maguire, et al. 2002; Prado, et al. 2002). Two recent studies in particular, clearly demonstrated that Gag mutations should be considered as authentic PI resistance mutations (Dam et al, 2009; Nijhuis et al, 2007). Nevertheless, it should be emphasized that all of the above studies have exclusively investigated mutations affecting cleavage site (CS) sequences in the P7-NC/P1/P6 region of Gag. A more recent paper however suggested that domains in Gag beyond its CSs can have both resistance and fitness effects in viruses bearing primary mutations in the protease (Parry et al, 2009).

Despite this body of knowledge, the impact of Gag mutations on clinical responses to PI-based regimens remains to be established. Several studies have demonstrated that patients experiencing viraemia during their first PI/r containing regimen infrequently show primary protease resistance mutations in routine testing (Kempf et al, 2004; Delaugerre et al, 2009; Gupta et al, 2008; Lathouwers et al, 2011). Suboptimal adherence and fast pharmacokinetics of PIs compared to other drug classes such as the non-nucleoside reverse transcriptase inhibitors (NNRTIs) (Bangsberg et al, 2004) have been suggested as determinants for the frequent lack of protease drug-resistance mutations at the time of PI/r failure. Rapid clearance in particular may explain both the more common occurrence of rebound low-level viraemia during PI/r therapy and the narrower window for effective drug-selective pressure (i.e., the optimal combination of sufficient levels of virus replication and sufficient drug levels) compared to the NNRTIs (Bangsberg et al, 2004; Geretti et al, 2008). However, emergence of protease mutations can be observed during prolonged failure of a PI/r based regimen, and is facilitated by pre-existing protease mutations (Bandaranayake et al, 2010). There are scarce longitudinal data addressing Gag evolution in this population.

In the previous chapter, we demonstrated that Gag mutations are observed in patients exposed to PIs and that many of these mutations are infrequent or absent in PI-naïve individuals. Importantly, mutations associated with PI-exposure were detected throughout the entire HIV-1 Gag protein and not exclusively in CSs. Based upon this observation; we proposed that full-length Gag sequencing may be required for full assessment of PI resistance.

The objective of the present chapter was to extend the above cross-sectional observations and longitudinally assess the emergence of mutations in the HIV-1 *Gag* and *PR* genes during failure of a PI/r-based regimen. Specifically, this involved the selection of patients with matched pre-treatment and failure plasma samples, the amplification and sequencing of full-length Gag and PR at the two time points and the comparison of the sequences obtained. In addition, we postulated that if Gag mutations compensate for the loss of viral replicative capacity caused by protease resistance mutations, Gag and protease mutations should be linked on the same viral genome. To assess this hypothesis, we performed clonal analysis in samples showing evolution of *Gag* and *PR* genes.

5. 2 Methods

5.2.1 Patient population

We examined two groups of patients. The first group (Group I) received PI/r-based therapy within the MaxCmin1, MaxCmin2 and COLATE trials, experienced virological failure (see below) and had a resistance test available at both study entry (baseline) and virological failure. Patient's treatment histories were examined in detail in order to exclude those that stopped therapy due to toxicity or poor adherence. The second group (Group II) consisted of PI/r-treated patients attending the HIV services of the Royal Free Hospital who presented with ongoing viraemia and evidence of accumulating protease resistance-associated mutations.

5.2.2 Definition of virological failure for trial population

Since the definition of virological failure differed between the three trials, we applied the MaxCmin2 definition on the combined trial population. This is:

- For patients entering the trial with a HIV-1 RNA load <200 copies/ml, a confirmed viral load ≥200 copies/ml at any time during the trial.
- For patients entering the trial with a viral load ≥200 copies/ml:
    - Any confirmed rise in viral load of $\geq 0.5\log_{10}$ copies/ml and/or
    - ≤ 0.5 log reduction in viral load at week 4

- ≤ 1.0 log reduction in viral load at week 12
- Viral load ≥ 200 copies/ml at week 24

5.2.3 RNA extraction

One milliliter of plasma was centrifuged at 25,000 g for 1 hour at 4ºC to concentrate the virus. The supernatant was then removed and the pellet re-suspended to a final volume of 280 µl. The re-suspended pellet was used for nucleic acid extraction. Samples with viral load above 1,000 copies/ml were subsequently processed employing the automated extractor EasyMag (Nuclisens, France) and those with viral load lower than 1,000 copies/ml were processed manually with the QIAamp Viral RNA Minikit (Qiagen, Germany). The RNA was eluted into 55 µl of elution buffer and stored at - 80ºC until required.

5.2.4 Amplification of Gag-protease region

The 2 Kb PCR product comprising the HIV-1 Gag and PR was amplified by nested PCR employing the invitrogene amplification protocol described in section 2.2.2.2.2 in chapter two.

5.2.5 Sequencing of Gag-protease region

Purified PCR products of the right size were diluted to a final concentration of approximately 10-20 ng/µl and population sequencing was performed using the ABI PRISM BigDye Terminator v3.1.ready reaction cycle sequencing Kit. Primers selected for PR and Gag sequences were those described in chapter 3 (Table 3.2). Sequences obtained were subsequently analyzed using Sequence analysis version 5.0 and Seqscape version 6.0.

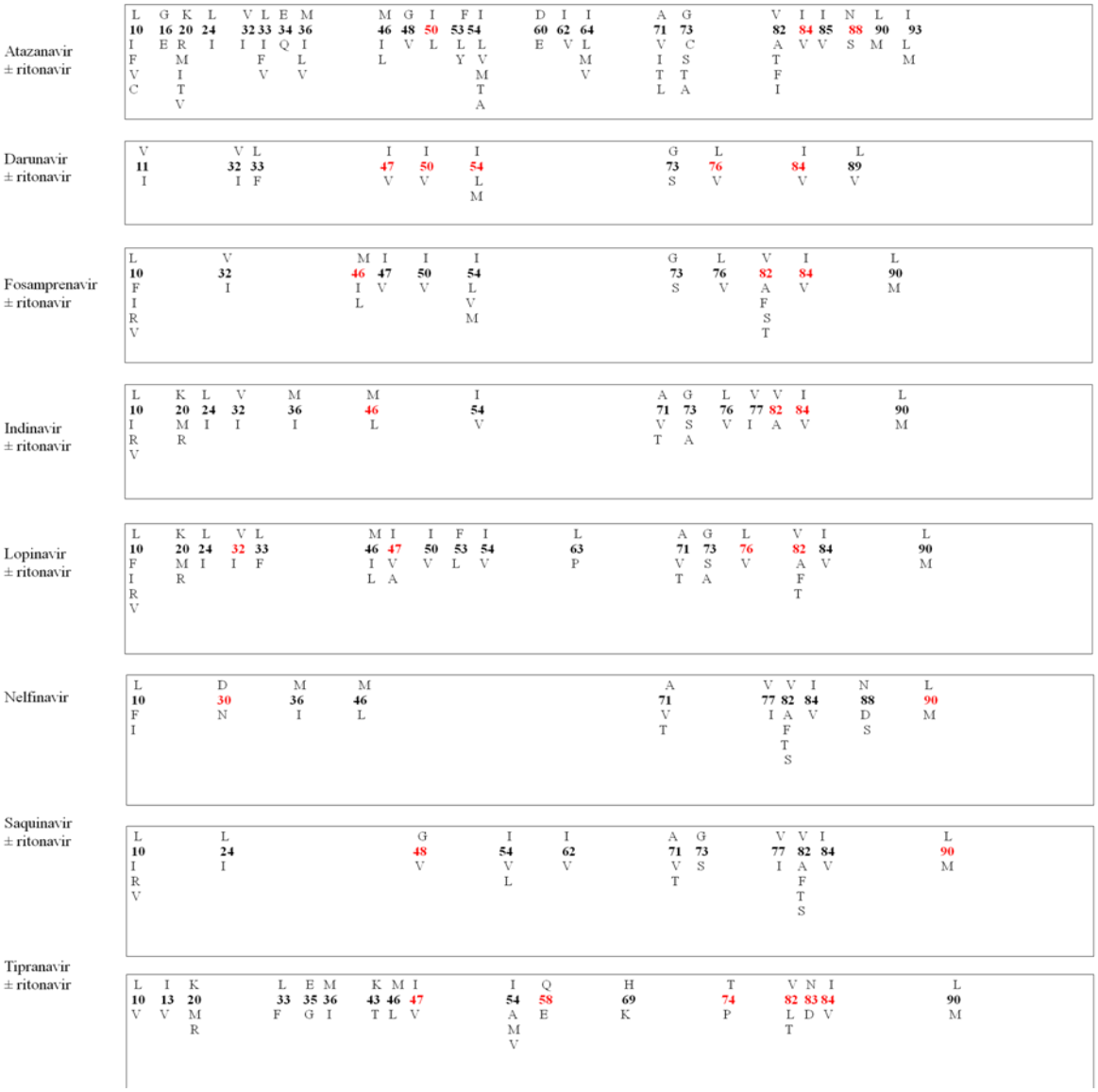
5.2.6 Classification of Gag and protease mutations

As explained in chapter 4 (sessions 4.2.9 and 4.2.10) PR and Gag sequences from patients were aligned against the HXB2 HIV-1 reference sequence. Mutations were considered as any change with respect to the reference sequence. PR resistance associated mutations were assessed according to the IAS-2011 list (Figure 5.1). Gag mutations were classified into cleavage site mutations (CSMs) and non-cleavage site mutations (non-CSMs) and stratified as those associated and not associated with PI-exposure according to the cross-sectional analysis performed in the previous chapter.

5.2.7 Cloning of Gag and protease genes

PCR products obtained from selected patients were purified and cloned into the commercial vector pCR 2.1-TOPO (Invitrogene). Positive clones, identified as those harbouring inserts of the correct size after a restriction digest were subsequently sequenced and analysed as described above.

Figure 5.1 Protease resistance associated mutations.



The first row of letters indicate the wild-type amino acid, numbers indicate codon position and letter below indicate amino acid substitution conferring resistance. Codons in red represent major mutation associated with resistance to the corresponding drug. All amino acids are shown by their one letter code.

5.3 **Results**

5.3.1 Study population

5.3.1.1 Group I

The first group of patients studied was selected from those experiencing virological failure in the MaxCmin1, MaxCmin2 and COLATE clinical trials. The three trials together included over 700 patients on ART treatment with regimens including SQV/r, IDV/r or LPV/r as either first-line or subsequent lines of therapy. After reviewing the patients' treatment histories to exclude those who had stopped therapy due to toxicity or poor adherence, we selected 28 patients as eligible for the study: 7 on IDV/r, 13 on SQV/r and 8 on LPV/r. Of these, five were receiving their first PI-based regimen, whereas the remaining 23 had been exposed to other PIs before entering the trial. Details of these patients are presented in table 5.1.

5.3.1.2 Group II

We selected this second group from a laboratory database of patients undergoing drug resistance testing at the Royal Free Hospital. The database contains the patients' *pol* gene sequences, current and past treatment regimens, plasma HIV-1 RNA load at the time of resistance testing, and HIV-1 subtype. Upon examination of the database, three patients with ongoing viraemia while on a PI/r-based regimen and whose routine genotype showed evidence of accumulating PI resistance-associated mutations were selected for the study. Details of these patients are shown in table 5.2.

Table 5.1 Group I: Patients from MaxCmin1, MaxCmin2 and COLATE who experienced virological failure.

Patient ID*	*Baseline HIV-1 RNA load (Copies/ml)	VF HIV-1 RNA load (Copies/ml)	Treatment status At **baseline
IDV/r, n=8			
Pt-1	6,025	32,359	PI-naïve
Pt-2	17,783	151,359	PI-experienced
Pt-3	2,399	10,715	PI-experienced
Pt-4	2,041	32,379	PI-experienced
Pt-5	9,772	346,737	PI-experienced
Pt-6	26,915	25,527	PI-experienced
Pt-7	4,786	4,169	PI-experienced
Pt-8	251,188	128,824	PI-naïve
SQV/r, n=13			
Pt-9	9,333	8,317	PI-experienced
Pt-10	549,540	40,738	PI-naïve
Pt-11	281,838	38,019	PI-experienced
Pt-12	3,467	6,761	PI-experienced
Pt-13	147,910	43,651	PI-experienced
Pt-14	21,380	589	PI-experienced
Pt-15	43,651	10,233	PI-naïve
Pt-16	117	301	PI-experienced
Pt-17	10,000	275	PI-experienced
Pt-18	1,047	363	PI-experienced
Pt-19	1,259	282	PI-experienced
Pt-20	1,995	229	PI-experienced
Pt-21	1,202	501	PI-experienced
LPV/r, n =7			
Pt-22	380,189	813	PI-experienced
Pt-23	1,549	1,122	PI-naïve
Pt-24	407,380	331	PI-experienced
Pt-25	15,488	1,778	PI-experienced
Pt-26	479	589	PI-experienced
Pt-27	5,248	380	PI-experienced
Pt-28	95,499	43,651	PI-experienced

PI regimen, HIV-1 RNA load at baseline and at VF, as well as treatment status at baseline is indicated.

*All patients were infected with subtype B HIV-1.

**Baseline is referred to the sample at study entry.

Table 5.2 Group II: Patients with ongoing viraemia while on PI/r-based therapy selected for studying the dynamics of emergence of PR and Gag mutations.

Pt	Date	HIV-1 RNA load (copies/ml)	HIV-1 subtype	Regimen	Protease resistance mutations
RFH-1	08-2002	100,000		D4T, 3TC, NVP	None
	11-2002	3,500		TDF, APV, LPV/r	M46I
	06-2004	1,300	C	TDF, APV, LPV/r	M46I, I84V
	11-2005	800		TDF, APV, LPV/r	M46I, I84V, L76V
	08-2007	74,642		TDF, APV, LPV/r	M46I, I84V, L76V, F53L
RFH-2	09-2006	428,688		NONE	None
	03-2007	1,368	CRF02	3TC, LPV/r, TDF	None
	07-2007	857		3TC, ABC, TDF, LPV/r	L76V
	10-2007	961		3TC, ABC, TDF, LPV/r	L76V, M46I
RFH-3	06-2004	64,545		ddI, TDF, EFV, LPV/r	M46I, I50V
	07-2005	49,861	CRF02	ddI, TDF, EFV, LPV/r	I54V, V82A
	11-2007	102,504		ZDV, 3TC, ABC, LPV/r, SQV/r	G48V, I54V, V82A

Date of resistance testing, HIV-1 RNA load and ARV regimen at the time of testing, as well as major PR resistance mutations detected in routine genotypic HIV resistance testing are shown.

Abbreviations: stavudine (D4T), Lamivudine (3TC), Tenofovir (TDF). Didadosine (ddI), zidovudine (ZDV), efavirenz (EFV) and nevirapine (NVP) and PIs: ritonavir boosted lopinavir (LPV/r), amprenavir (APV) and ritonavir boosted saquinavir (SQV/r).

5.3.2 Full-length Gag and protease sequencing results

5.3.2.1 Group I

- **Baseline**:

At baseline 13/28 (46%) patients had at least one primary PR resistance-associated mutation (median 3, range 1-5); 27/28 (96%) had secondary PI resistance-associated mutations (median 4, range 1-7); and 28/28 (100%) had protease polymorphisms which are not associated with PI resistance. All 13 patients with primary PR mutations had been exposed to PIs before commencing the trial. The primary PR mutations observed included: D30N (n = 1), V32I (n = 1), L33F (n = 1), M46I (n = 8), I47V (n =1), G48V (n = 1), I54V (n = 7), V82A (n = 7), I84V (n = 2), N88D (n = 1) and L90M (n = 5). The secondary PI mutations detected included: L10F (n = 1), L10I (n = 10), L10V (n = 4), I13V (n = 7), K20R (n = 4), K20T (n = 2), L24I (n = 2), M36I (n = 11), F53L (n = 1), D60E (n = 2), I62V (n = 9), L63P (n = 16), L63Q (n = 1), I64V (n = 4), H69K (n = 2), H69Q (n = 2), A71I (n =1), A71T (n = 3), A71V (n = 5), G73S (n = 1), G73T (n = 1), V77I (n =9), and I93L (n = 10).

Regarding *Gag*, at baseline, 28/28 (100%) patients had at least one mutation associated with PI-exposure in the P17 protein (median 4, range 1-9); 13/28 (46%) showed at least one mutation associated with PI-exposure in the P24 protein (median 1, range 1-2); 15/28 (54%) presented at least one mutation associated with PI-exposure in the P7 protein (median 1, range 1-3); 18/28 (64%) had at least one mutation associated with PI exposure in the P6 protein (median 3, range 1-4); 3/28 (11%) had at least one mutation associated with PI-exposure in the spacer peptide P2 (median 1, range 1-2); and 8/28 (28%) had the mutation

K442R associated with PI-exposure in the spacer peptide P1. In addition, at baseline 17/28 (60%) showed at least one mutation associated with PI-exposure in one of the CSs, distributed as follow: 3/28 (11%) had ≥1 mutations in MA-P17/CA-P24; 15/28 (53%) had ≥1 mutations in P2/P7; 3/28 (11%) had 1 mutation in NC-P7/P1 and 1/28 (3%) had 1 mutation in P1/P6. Most patients (24/28, 86%) presented polymorphisms in the CSs not associated with PI exposure and located mainly in P2/P7 (22/28, 78%) but also in P17/P24 (2/28, 7%), P24/P2 (4/28, 14%) and P1/P6 (2/28, 7%). Protease, Gag non-CS and Gag CS mutations at baseline for the 28 patients are shown in tables 5.3, 5.4a to 5.4f and 5.5, respectively.

Patient	Protease mutations at baseline		
	MAJOR	**MINOR**	**OTHER**
Pt-1*	NONE	I13V, V77I	S37N, I64V, L90LF
Pt-2	NONE	I13V, M36I, I62V, L63P	V3I, I15V, S37DN, K43R
Pt-3	M46I, L90M	L63P, A71AT, V77I	V3I, T12ATM S37N, V82L
Pt-4	NONE	L10LV, L63P, I93L	V3I, T12A, K14KR, S37N, R41K
Pt-5	NONE	M36I, H69HQ, I72IV	V3I, S37D, Q61H
Pt-6	V32I, M46I, I47V, V82A	L10IL, K20R, M36I, I62IV, L63P, I93L	E35D, S37D, K55R, Q58E
Pt-7	L33F, I54V, I84V, L90M	L10I, I13V, I62V, L63P, A71V, G73SG	V3I, L19I, E21DE, E35D, S37D, D60E, I72L
Pt-8*	NONE	I13V, K20R, M36I, H69K, I93L	I15V, S37N, R41K, L89M
Pt-9	L90M	I13V, K20T, M36I, D60E, I93L	V3I, I15V, E35D, S37N, P39Q, R41K, R57K, Q61D, I61IV
Pt-10*	NONE	L10LV, M36I, D60E, I62IV	V3I, I15V, E35D, S37N, P39Q, R41K, R57K, Q61D
Pt-11	M46IM, I84IV, L90LM	L10FIL, K20KT, I62IV, L63P, A71AV, V77IV, I93L	V3I, K14KR, I15IL, G16AG, S37N, I85IV
Pt-12	NONE	M36I, I93L	E35D, S37N, R41K, K45R, R57K, Q61S
Pt-13	NONE	L63P, V77IV	L19Q, D25DN, S37CS
Pt-14	NONE	I62IV, L63P	V3I, S37N, K45R, K70R, I72E
Pt-15*	NONE	NONE	V3I, S37N, L63A, E65DE
Pt-16	NONE	V77I	S37N
Pt-17	V82A	M36I, L63P, H69K, I93L	V3I, T12S, I15V, L19T, S37N, R41K, Q61E, L89M
Pt-18	I54V, V82A	L10I, L24I, I62IV, L63P, A71AITV, V77I, I93L	S37T, R41K
Pt-19	M46I, I54V, L90M	L10IV, L63P, G73T	I15V, S37N, K55R, V82C
Pt-20	D30N, M46I, N88D	L10IV, L63PQ, I64V, A71AV, V77IV	S37N, E65D, V75IV,
Pt-21	M46IM, G48V, I54IV, V82AV, L90LM	I64IV, V77IV	S37N, T74AT
Pt-22	I54IV, V82A	L10I, K20R, M36I, L63LP	V3I, I15V, E35D, S37N, R41K, R57K, Q61N, I72IT, T74AT
Pt-23*	NONE	L10I, I13V, L63P, H69HQ, A71AT	V3I, L19IL
Pt-24	NONE	L63P, V77I	R41K, I64L
Pt-25	M46I, I54V, V82A	L10I, I13V, K20R, M36I, I62V, I64V	V3I, S37N, R41K, L63T, E65D
Pt-26	NONE	L63P, A71AV, I93L	V3I, T12I, E35D, R41K, I62M
Pt-27	M46I, I54V, V82A	L10I, L24I, F53FL, I62V, I64V	V3I, I15IV, E35D, M36V, S37N, K43T, K55R, L63A
Pt-28	NONE	M36I, I93L	E35D, S37N

Table 5.3 Protease mutations observed at baseline.

*These patients were treatment-naïve at baseline.

Pt	P17 mutations at baseline	
	PI –exposure associated	Other mutations
Pt-1*	V82I, E93D, **I94V, N126SG**	S54P, G62VL, R76K,R91Q, A118TAins
Pt-2	V82I, T84V, A119T, **N126S**	E55EG, G62KREG, Q69QR, R76K,K114I, AT117ins
Pt-3	T84V, E93D, I94V,Q123K, **N126S**	A83AV, D102E, H124N
Pt-4	I34L, **I94V**, N109NT, T122KT, **N126S**	K28Q, K30Q, E55D, T70TS, G71RG, S72PS, E74QE, R76K, Y79YF, S11G, H124N, S125N
Pt-5	E93D, **I94V**, K114R, **N126S**	K28M, S38G, S54A, S67A, R76K, Y79F, K95R
Pt-6	T84V, **I94V**, A119T, **N126S**	R15S, K26S, K30Q, S54A, Q69K, R76K, Y79F,K113N, A120K, H124N
Pt-7	G123K, **N126S**	K20Q, K28R, R76K, R91N, E93K, I104V,K110E, Q117K, H124N
Pt-8*	L61I, **I94V**, H124K	R15T, R20K, K28NT, K30M, G49D, R58K, T81A, Q90E, E93A, A115T
Pt-9	Q59K, **L61I**, V82I, T84V, I92V, K113R, Q117P, A119EA, T122K	G62, L75F, A83V, C87F, R91K, I104L, K112Q, A115T, GHSN123ins, Q127G.
Pt-10*	I34L, **L61IL**, V82IM, T84V, **K103R**, K113KR, **N126S**	R43QR, F44L, K58KR, G62A, Q63R, R76K, R91K, H124N
Pt-11	T84V, **I94V**, A119T, T122A, Q127P	K30Q, R76K, T81A, D102E, A118V, H124N, S125N
Pt-12	T84V, **K113R**, A119T,	K26N, K30R, E55G, G62E, R76K, Y79F, R91K, K98R, A120T, D121N, TGH122del
Pt-13	V82I, **N126S**	Q90E, R91K, D102E, I104V, A119E, AA119ins, G123R, H124N
Pt-14	V82IV, E93D, I94V, G123E, **N126S**	G62AV, R76K, Y79YF, I104V, H124N
Pt-15*	I94IV, **N126S**	K28T, K30R, G62M, R76K, E93N, S111C, H124N, S125N,
Pt-16	E93D, **I94V**, D121G, **N126S**	D102E, AT117ins, H124N
Pt-17	D93E	K58R, M61I, K62KNRS, A67S, Q69KN, H79F, K123del
Pt-18	I34L, V82I, **K113Q**,	K30R, R76K, D102E, A115V, H124N
Pt-19	T84V, **I94V**, N109S, K114Q, **N126S**	K30R, Q59R, Q65H, R76K, E93N, S111C, K112R, A115T, H124N
Pt-20	T84V, E93D, **N126S**	L61M, H124N, N125S
Pt-21	**N126S**	K28T, R76K, H124S, S125C, Q127S
Pt-22	Q59K, I92V, **I94V,**	R30K, V35I, P54T, K58R, R62G, A83V, C87F, K91Q, E102DE, KSQQK109ins, K110T, G111S, K114N
Pt-23*	I34L, **I94V, N126S**	K28Q, K30Q, E55D, T70TS, G71GR, S72PS, E74QE, R76K, Y79YF, N109NT, S111SG, H124N, Q125N
Pt-24	T84V, E93D, **I94V, N126S**	V46IV, G62S, P66S, Q69K, R76K, H124N, Q127S
Pt-25	I34L, **L61I, I94V**, K114Q, D121A, **N126S**	K30R, V46L, N47D, E52D, S54A, Q69K, R76K, Y79F, Q90R, R91N, D102E, H124N
Pt-26	**I94IV**, Q117P, T122A, **N126S**	G62V, R76K, T81A, D96G, S111A, H124S
Pt-27	T84V, **I94V**, N109S, **N126S**	K30R, Q59R, Q65H, R76K, E93N
Pt-28	T84V, Q127P	K30R, R76K

Table 5.4a P17 mutations observed at baseline.

Mutations are classified as those associated with PI-exposure (as determined in the analysis presented in chapter 4) and other mutations. In bold are shown mutations which remained associated after Bonferroni's correction.

* These patients were treatment-naïve at baseline.

Pt	P24 mutations at baseline	
	PI –exposure associated	Other mutations
Pt-1*	None	I138L, V215L, I223A, G248N, N252S, L268M, G357S
Pt-2	None	I138M, V215LV, R286K
Pt-3	M200I, **E211D**	I138L, A146P, I147L, S173T, V215L, N252S, E312D
Pt-4	None	N137T, I138IL, Q139P, Q141K, A146P, S176A, V215L, I223IV, T242TS, E245D, G248N, N252G, T280S, E312D, G357S
Pt-5	None	I138L, I223V, P292S
Pt-6	L268M	I138L, I147L, V215L, N252H, A340G
Pt-7	L268M	I147L, V215L, H219Q, I223V, M228I, G248A, R257KR, A326S, A340G
Pt-8*	None	A147P, V157I, L213V, M226I, V254I, K284R
Pt-9	T280I, N315G	I138L, A146P, E203D, T204A, V215L, H219Q, G221R, A224P, M228I, G248Q, M252G, I256V, Y301F, S310T,E319P, A340G
Pt-10*	None	I138L, I147L, V215L, H219Q, M228ML, T242N, G248A, T280V, E312D
Pt-11	S173A	I138L, S173T, T239S, N252S, I256IV, E312D, G357S
Pt-12	None	I138L, I147L, V159I, E203D, V215L, A224P, G248Q, N252S, I256V, T280V, Y301F, S310T, E319D, A340G
Pt-13	None	I138M, A146P, V159I, M200L, V215L, G248A, M250T, G357S
Pt-14	A146S, Q182H	I138L V191I, V215M, I223V, N252S, N253T, R286K, A326S
Pt-15*	None	I138L, V215L, M228L, G248A, N253T, P255S, A340G
Pt-16	None	V215L, N252H
Pt-17	None	P144A, A147P, V157I, V166I, H217Q, V211IV, M226IV, V254I, T301GVCF, S240T
Pt-18	L268M, T280I	V147l, V159I, V215L, H219Q, I223V, T242N, G248A, N252G, P255A, A340G, G357S
Pt-19	None	I138L, V215L, A340G, E345D
Pt-20	T280I	I138L, I147L, V215L, N252H, E312D, G357S
Pt-21	L268M	V215L, H219Q, I223V, N252S, P292S, A340G
Pt-22	N315G	P142S, I143L, A214V, L220P, S316T, S322A
Pt-23*	None	I138L, V215L, I223IV, E312D, G357S
Pt-24	T280I, N315G	I138L, V159I, V215L, I223V, T242S, G248A, N252S, I256V, A340G
Pt-25	T280I	I138L, V147L, V159I, V215L, I223V, E312D
Pt-26	None	V215L, Q246P, E312D, A326S, A340G
Pt-27	None	I138IL, V215L, I223A, T239S, N252S
Pt-28	L268M, T280I	I138L, I147L, S173A, V215M, R264K, R286K

Table 5.4b P24 mutations observed at baseline.

Mutations are classified as those associated with PI-exposure (according to chapter 4) and other mutations. In bold are shown mutations which remained associated after Bonferroni's correction.

*These patients were treatment-naive at baseline.

Pt	P7 mutations at baseline	
	PI –exposure associated	Other mutations
Pt-1*	None	V390C, A402S
Pt-2	R403K, T427S	T401I, N404H, R406KR, K410R
Pt-3	None	N385S, I389T, T401I
Pt-4	I389V, R403K, T427S	R387G, E398A, T401I
Pt-5	None	N385K, T401I
Pt-6	I389V, R403K	T401I
Pt-7	I389V, R403K	T401I
Pt-8*	None	R384K, N385G
Pt-9	R403K	N385Q, Q386H, R387K, K388R, E419K
Pt-10*	R403K	N385H, K397R, T401I
Pt-11	I389V, R403K	R384K, R387K, K388R, K397R, T401I
Pt-12	V390A, R403K	R384K, T401I
Pt-13	I389V, V390A, R403K	K411R
Pt-14	None	T401I, K411R, D425E
Pt-15*	None	T401I, K411R
Pt-16	T427S	T401I
Pt-17	None	I387V, L398I, R400K, K415R, V420M
Pt-18	R403K	R384K, R387K, K388R, E398V, T401I
Pt-19	None	R384K, R387KR, T401I
Pt-20	I389V	T401I, K418R
Pt-21	None	R384K, I389P, N394I, T401I, N404I, K424E
Pt-22	None	R383K, K413T, H416Q
Pt-23*	None	T401I, K411R, K418R
Pt-24	R403K	T401I, K418R
Pt-25	None	R387K, K388R, T401I
Pt-26	None	T401I
Pt-27	I389V	T401I
Pt-28	I398V, R403K	N385S, T401I, K418R

Table 5.4c P7 mutations observed at baseline.

Mutations are classified as those associated with PI-exposure (according to chapter 4) and other mutations.

*These patients were treatment-naïve at baseline.

Pt	P6 mutations at baseline	
	PI-exposure associated	Other mutations
Pt-1*	None	T456S, E460A, S465F, V467E, P473S, E477DE, D480E, T487A, R490K
Pt-2	R464KR, T471A,	S465F, I479R, T487A, R490K
Pt-3	None	V467E, T470A, I479V, T487A, R490K
Pt-4	None	454EPTAins, S462C, S465F, V467E, T487AY, R490K, P497Q
Pt-5	T470A, T471A	E460A, S465F, V467E, 483LMPTins, R490K, S498L
Pt-6	**F463L**, E477G, P478T	S465L, V467E, P473S, T487A
Pt-7	S465M, T469A	460EPTAPPEins, V467E, P473S, T487A, R490K
Pt-8*	None	E460A, E482K, L483P
Pt-9	T469A, Q474P, P478T, I479T	454EPTAPPAEins, E460A, P473A, K481R, P485A, R490K, N495S, S498L
Pt-10*	F463V, T470A	E454A, T456S, S465F, V467E, E482D, R490K
Pt-11	T470A, I479V	APS459ins, S465F, V467E
Pt-12	**F463L**, E477G, P478T	S465L, V467E, P473S, T487A
Pt-13	S465M, T470A, I479T	V467E, T487A, 496-500del
Pt-14	**F463L**, R464K, T469I, Q474P	E460A, 465PPAESFins,V467F, P473A, P478A, I479R, K481R, P485A, R490K, N495S, S498L
Pt-15*	None	E454A, 460EPTAPPEins, V467E, R490K
Pt-16	T470A	E454A, S462IV, S465F, V467E, P478PL, E482D, T487A, R490K
Pt-17	None	E457D, T487A, R490K
Pt-18	None	P459PS, S465F, V467E, E482D, R490K
Pt-19	None	E460A, E482K, L483P
Pt-20	G466R	R464L, D480V, K481N, S488A, D496S
Pt-21	T471S, E477G, P478T	S465F, V467E, T487A, R490K
Pt-22	**P478Q**, I479T	E460A, S465F, V467E, P473S, E482D, S488A
Pt-23*	None	459APSins, S465F, V467E, I479R
Pt-24	None	T456S, S465F, V467S, E482D, R490K
Pt-25	S465M, **P478Q**	V367E, E482D, T487A, R490K
Pt-26	**F463L**,S465M, T471A, E477G	V467E, P478A, S488A, R490K
Pt-27	T470V, T471A, I479V	E460A, S465F, V467E, P478S, D480E, R490K, N495S
Pt-28	**F463L**, G466R	454EPTAins, S462C, S465F, V467E, T487AY, R490K, P497Q

Table 5.4d P6 mutations observed at baseline.

Mutations are classified as those associated with PI-exposure (according to chapter 4) and other mutations. In bold are shown mutations which remained associated after Bonferroni's correction.

*These patients were treatment-naïve at baseline.

Pt	P2 (SP1) mutations at baseline	
	PI-exposure associated	Other mutations
Pt-1*	None	T371N
Pt-2	None	V370A
Pt-3	None	None
Pt-4	None	None
Pt-5	None	T371N
Pt-6	None	Q369H
Pt-7	None	None
Pt-8*	None	T371N
Pt-9	None	T371N
Pt-10*	None	V370A, N372S
Pt-11	None	None
Pt-12	None	None
Pt-13	None	V370A
Pt-14	None	None
Pt-15*	None	None
Pt-16	None	None
Pt-17	None	V370A
Pt-18	T371A	Q369H, V370A
Pt-19	None	None
Pt-20	None	V370A
Pt-21	None	V370A, N372K
Pt-22	None	Q369H, T371V
Pt-23*	None	None
Pt-24	None	Q369H
Pt-25	V370M	T371del
Pt-26	None	T371S
Pt-27	None	V370A
Pt-28	V370M, T371A	None

Table 5.4e P2 mutations observed at baseline.

Mutations are classified as those associated with PI-exposure (according to chapter 4) and other mutations.

*These patients were treatment-naïve at baseline.

Pt	P1 (SP2) mutations at baseline	
	PI exposure associated	Other mutations
Pt-1*	None	Y441H
Pt-2	K442R	None
Pt-3	None	Y441H
Pt-4	K442R	Y441H, G443D
Pt-5	K442R	None
Pt-6	None	Y441H
Pt-7	K442R	Y441H
Pt-8*	None	None
Pt-9	None	None
Pt-10*	None	Y441S
Pt-11	None	None
Pt-12	None	None
Pt-13	K442R	Y441H
Pt-14	None	None
Pt-15*	None	Y441S
Pt-16	None	Y441H
Pt-17	None	R439K
Pt-18	None	Y441H
Pt-19	None	None
Pt-20	None	R439K
Pt-21	K442R	None
Pt-22	K442R	None
Pt-23*	None	None
Pt-24	None	Y441H
Pt-25	K442R	Y441H
Pt-26	None	Y441H, G443E
Pt-27	None	Y441H
Pt-28	None	None

Table 5.4f P1 mutations observed at baseline.

Mutations are classified as those associated with PI-exposure (according to chapter 4) and other mutations.

*These patients were treatment-naïve at baseline

Patient	Gag CS Mutations									
	P333I-exposure associated					Other CS mutations				
	P17/P24	P24/P2	P2/P7	P7/P1	P1/P6	P17/P24	P24/P2	P2/P7	P7/P1	P1/P6
Pt-1*	None	None	**A374S**	None	None	None	A360S	T375Y	None	None
Pt-2	None	None	None	None	None	None	None	S373A, R380K	None	None
Pt-3	None	None	None	None	None	None	None	S373Q	None	None
Pt-4	None	None	**A374**S, T375A	None	None	None	None	None	None	None
Pt-5	None	None	T375A	None	None	None	None	None	None	None
Pt-6	**V128I, Y132F**	None	None	None	None	None	None	S373A, A374N	None	None
Pt-7	None	None	**S373T, T375N**	None	None	None	None	A374del	None	None
Pt-8*	None	None	G381S	None	None	Q130H	None	I376V, R380K	None	None
Pt-9	None	None	G381S	A431V	None	None	None	A374T, M377L, R380K	None	None
Pt-10*	None	None	None	None	None	None	V362I	S373A, A374N, I376V	None	None
Pt-11	None	None	**T375N**	A431V	None	None	None	S373A, A374T	None	None
Pt-12	None	None	**S373T**, G381S	None	None	None	None	T375del, I376V, M377L	None	None
Pt-13	None	None	A374P	None	None	None	None	S373P, R380K	None	None
Pt-14	None	None	T375A	None	None	Q130H	None	None	None	P453L
Pt-15*	None	None	None	None	None	None	None	S373P, R380K	None	None
Pt-16	None	None	T375A	None	None	None	None	AT373-374ins	None	None
Pt-17	None	None	None	None	None	None	None	None	None	L445F, N448S
Pt-18	None	None	None	A431V	None	None	None	TSA374ins, R380K	None	None
Pt-19	**Y132F**	None	**T375N**	None	None	None	None	S373A	None	None
Pt-20	**Y132F**	None	**A374S**	None	None	None	V362I	S373A, M378I	None	None
Pt-21	None	None	None	None	None	None	None	None	None	None
Pt-22	None	None	G381S	None	None	None	A366V	M378V	None	None
Pt-23*	None	None	None	None	None	None	None	A374N, T375V, R380K	None	None
Pt-24	None	None	None	None	S451T	None	None	R380K	None	None
Pt-25	None	None	**A374S**, G381S	None	None	None	None	S373P, I376V, R380K	None	None
Pt-26	None	None	None	None	None	None	None	I376V, R380K	None	None
Pt-27	None	None	None	None	None	None	None	None	None	None
Pt-28	None	None	None	None	None	None	None	R380K	None	None

Table 5.5 Cleavage site mutations observed at baseline.

Mutations are classified as those associated with PI-exposure (according to chapter 4) and other mutations. In bold are shown mutations which remained associated after Bonferroni's correction. *These patients were treatment-naïve at baseline.

- **Virological failure:**

At the time of virological failure, 6/28 (21%) and 13/28 (46%) patients showed emergence of primary and secondary PR resistance-associated mutations, respectively. Regarding *Gag* 9/28 patients (32%) had P17 emergent mutations, 8/28 (28%) had P24 emergent mutations, 7/28 (25%) had P7 emergent mutations, 5/28 (18%) had P1 emergent mutations, 4/28 (14%) had P2 emergent mutations and 1/28 (3%) had P6 emergent mutations. In addition, 13/28 (46%) patients had treatment emergent Gag CSMs including four patients with one mutation at P2/P7, five patients with one mutation in P7/P1 and four patients with one mutation in P1/P6.

We next examined the Gag emergent mutations and classified them as associated or not associated with PI exposure. Among the 18 patients with G*ag* mutations emerging outside CSs, seven (39%) showed one mutation associated with PI exposure, three (17%) showed two mutations, one (5%) showed three mutations, and two (11%) showed four mutations. Overall, 24 mutations associated with PI exposure emerged at the time of VF and of them the majority were located in P17 (15/24, 62%); the remaining were found in P24 (5/20, 25%), P7 (2/24, 8%) and P1 (2/24, 8%). In addition a total of 24 polymorphisms not associated with PI exposure emerged in 15/28 (53%) patients and where located in P17 (2/24, 8%), P24 (4/24, 17%), P7 (7/24, 29%), P6 (2/24, 8%), P2 (5/24, 21%) and P1 (4/24, 17%). Among the 13 patients with Gag CSMs, ten had one emergent mutation associated with PI exposure located in P2/P7, n = 3; P7/P1, n =4 and P1/P6, n =3. Moreover, three patients showed polymorphism not associated with PI exposure emerging in P2/P7, n =1; P7/P1, n =1 and P1/P6, n= 1.

Details of mutations emerging at the time of virological failure are shown in table 5.6. The comparison of the number of mutations observed at baseline and at virological failure is displayed in figure 5.2.

Treatment Emergence mutations							
Patient	**Protease mutations**			**Gag Mutations**			
	Primary	**Secondary**	**Polymorphisms**	**CS**		**non-CS**	
				PI-associated	**others**	**PI-associated**	**others**
Pt-1*	None	None	None	P7/P1: I437V	None	P17: N109T; P7: I389V	P7: T401I; P2: V370A
Pt-2	None	None	T12ST	P2/P7: T375A	None	None	P2: T371N
Pt-3	None	I13V, M36I, I62V	I15V, K43R	None	None	P17: Q59K	P24: E167EK
Pt-4	None	None	None	None	None	None	None
Pt-5	None	None	I51IV	None	P2/P7: R380K	P17: L61I, P66S, H124K; P7: I389V	P24: A163AV
Pt-6	None	None	None	None	None	None	P24: Q139QR
Pt-7	None	None	None	P2/P7: G381S	None	None	None
Pt-8	None	None	K70KR	None	None	None	None
Pt-9	I84V	A71V	None	None	P1/P6: P453L	None	None
Pt-10*	None	None	T4NT	P1/P6: P453T	None	P17: Q59K, P66S, E93D; P1: K442R	None
Pt-11	None	G73S	None	None	None	P24: V218A	None
Pt-12	None	None	V3I	None	None	None	P7: R387K, K388R
Pt-13	None	None	None	P1/P6: L449F	None	None	P7: N404D
Pt-14	M46IM, I54V, V82A, L90M	L10I, I13IV, A71V, G73S, V77IV	L19IL, Q61KQR	P7/P1: A431V	None	P17: Q117P	P1: Y441H
Pt-15*	None	I13ILPT	P9PS	P1/P6: P453T	None	P24: E211D	P6: E454A, T476P
Pt-16	None	L10V, I62V, L63P	K14R, I15V, E35D,	None	None	P1: K442R	None
Pt-17	None	None	K45KR	None	None	None	None
Pt-18	I84V	F53FL	None	None	None	None	None
Pt-19	None	F53L	None	None	None	None	None
Pt-20	I84V	T74P	None	P7/P1: I437V	None	P17: V82I, A115I	P17: A120V; P1: Y441H, G443E
Pt-21	I84IV	M36I, L10I	None	None	None	None	None
Pt-22	None	None	K43KR	P7/P1: K436R	None	P17: D121G; P24: S310T	P1: R439G, P440Q
Pt-23*	None	None	V3I, E35DE	None	None	P17: K103R, T122K; P24: N315G	P17: E105K, P7: G420R
Pt-24	None	L24IL, I93L	None	None	None	None	None
Pt-25	None	None	I19IL	None	None	P17: I92V	P7: Q422L; P2: V370del, T371M
Pt-26	None	K20R	None	None	P7/P1: E428K	P24: T280I	P24: R286K, P7: K418R
Pt-27	None	None	E34EK	P2/P7: T375A	None	None	P2: N372A
Pt-28	None	L10V, I62V, L63P	K14R, I15V, E35D,	None	None	None	None

Table 5.6 Mutations emerging at virological failure.

*These patients were treatment-naïve before commencing the PI-base regimen.

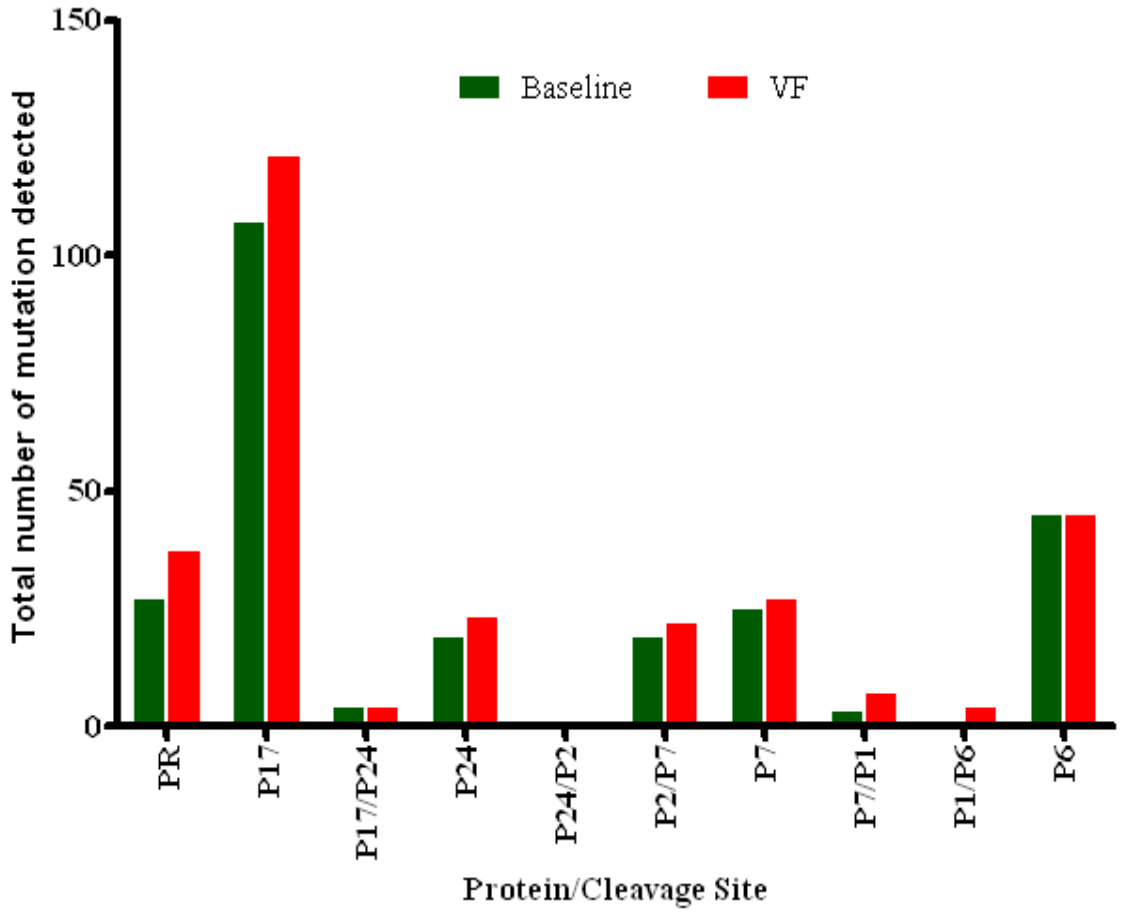


Figure 5.2 Total number of mutations observed at baseline and at the time of virological failure in 28 patients failing a PI/r based regimen with IDV/r, SQV/r or LPV/r.

Only major PI resistance-associated mutations and Gag CS and non-CS mutations associated with PI-exposure are shown.

5.3.2.2 Group II

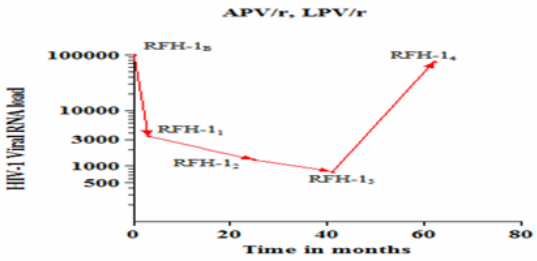
We selected three patients with ongoing viraemia while on a PI/r-based regimen.

- **Patient 1 (RFH-1)**

The first patient first attended the Royal Free Hospital in August 2002. At the time he was receiving an NNRTI-based regimen consisting of stavudine (D4T), Lamivudine (3TC) and nevirapine (NVP) and had never been exposed to PIs. The plasma HIV-1 RNA load was 100,000 copies/ml. Routine HIV-1 genotypic resistance testing revealed subtype C infection, the presence of two major NRTI resistance-associated mutations and three secondary protease mutations which are commonly seen in drug naïve individuals as natural polymorphism, and no major protease-resistance mutation. Analysis of the *Gag* gene at this time point showed three CSMs and numerous non-CSMs. Two of the three CS and most of the non-CS mutations were polymorphisms; one CS and four non-CS mutations (P17: I34L, V82I, E93D, N126S) were associated with PI exposure. The patient was switched to a double PI/r-based regimen containing tenofovir (TDF), ritonavir boosted amprenavir (APV/r) and ritonavir boosted lopinavir (LPV/r) and in November 2002 his viral load was 3,500 copies/ml. Analysis of full-length PR and Gag genes showed the emergence of a major resistance-associated mutation in the protease (M46I) and no changes in *Gag*. In June 2004, while still on the above regimen, the viral load remained detectable (1,300 copies/ml). At this time point sequencing showed the emergence of the major protease mutations L76V and I84V and of 21 non-CS mutations distributed as follow: 3 mutations in P17, 6 mutations in P24, 5 mutations in P7, and 7 mutations in P6. Of these 21 non-CS mutations, 9 were associated with PI-exposure and were located in MA-P17 (n=3), CA-P24 (n=1) and P6 (n=5). The remaining 12 were polymorphisms found in P24 (n=5), P7 (n=5) and P6 (n=2). The patient continued the same

regimen and in November 2005 showed a viral load of 800 copies/ml. While no major protease mutation emerged at this time, new mutations appeared in *Gag*: the P1/P6 CS mutation L449F, one PI-associated mutation in P17, and two polymorphisms in P24. In August 2007, while still on the same regimen, the viral load increased to 74,642 copies/ml. Full-length PR and Gag sequencing showed the emergence of the secondary protease mutation F53L, 5 mutations in P17 (3 of which associated with PI exposure: Q59K, K103R, and Q117P), the P17/P24 CS Y132F mutation (associated with PI exposure), 4 mutations in P6 (2 of which associated with PI exposure; R469K, E477G), and 3 polymorphisms in P7.

To summarize, overall, 37 mutations emerged in Gag over time in this patient, including two CS and 35 non-CS mutations. The two CS mutations and most mutations emerging at P17 and P6 (7/9 and 7/11, respectively) were indicative of PI selective pressure according to the analysis showed in chapter 4. The HIV-1 RNA viral load and PR and Gag evolution over time are shown in figure 5.3



Baseline time point						
Sample ID	**Date**	**Protease Mutations**			**Gag mutations**	
		Primary	**Secondary**	**Others**	**Gag CSs**	**Gag non-CSs**
RFH-1B	Aug-2002	None	M36L, L63P , I93L	I15V, L19I, N37S	**P17/P24:** none **P24/P2**: V362I **P2/P7:** A374S, R380K **P7/P1:** none **P1/P6:** none	**P17:** K20Q, K28R, I34L, R76K, V82I, E93D, H124N, N126S **P24:** I138L, I147L, T186S, V215L, M228I, L267V, K302R, Q386P, K418R **P7:** Q386P, K418R **P6:** T456S, E460A, T469K, T470A, E477D, T487A, R490K, P497L, S498L **P1:** None **P2:** T371N
Treatment emergent mutations						
Sample ID	**Date**	**Protease Mutations**			**Gag mutations**	
		Primary	**Secondary**	**Others**	**Gag CSs**	**Gag non-CSs**
RFH-11	Nov-2002	M46I	No new mutations	No new mutations	None	None
RFH-12	Jun-2004	L76V, I84V	No new mutations	No new mutations	None	**P17:** I94V, K114R, D121G **P24:** S148T, Q182H, M228I, E230D, E260D, R286K **P7:** R384K, N389I, V390I, K411R, T427N **P6:** E460A, F463L, T471A, P472S, Q474P, I479T, L486V
RFH-13	Nov-2005	No new mutations	No new mutations	No new mutations	**P1/P6:** L449F	**P17:** A118T **P24:** A146P, S176A
RFH-14	Aug-2007	No new mutations	F53L	No new mutations	**P17/P24:** Y132F	**P17:** R15K, S38G, Q59K, K103R, Q117P **P7:** N385G, K410R, K418Q **P6:** R464K, E477G, N495S

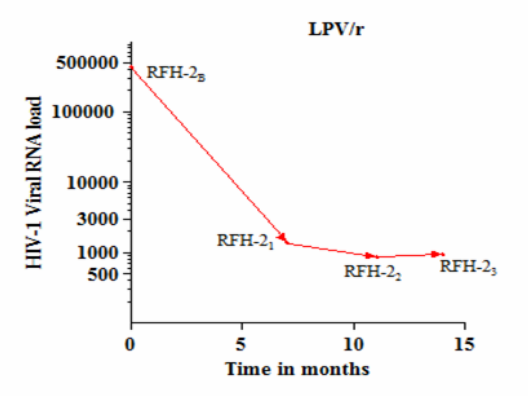
Figure 5.3 Dynamic of emergence of PR and Gag mutations during PI/r-based therapy with ongoing viraemia - patient RFH-1.

PI treatment and viral load at each time point are indicated. The table below shows the PR and Gag mutations emerging at each time point. Major PR mutations and Gag mutations associated with PI exposure are shown in red.

- **Patient 2 (RFH-2)**

The patient first attended the Royal Free Hospital in September 2006 when he was ART-naïve and his viral load was 4278,688 copies/ml. Full length PR and Gag sequencing demonstrated infection by a CRF02_AG recombinant HIV-1 virus, no major protease resistance-associated mutations and several Gag CS and non-CS mutations. The latter were mainly polymorphisms except for four mutations in P17, one in P7 and two in P6. The patient was started on

Lamivudine (3TC), Tenofovir (TDF) and LPV/r and in March 2007 his viral load remained detectable (1,368 copies/ml). Sequencing of PR and Gag genes showed no emergent mutations in the protease, but 6 P17 mutations (3 of which associated with PI exposure), and two polymorphisms in P7 emerged at that time. The patient continued on this regimen and in July 2007 the viral load was 857copies/ml. At this time, sequencing showed the emergence of the major protease mutation L76V and four polymorphisms in Gag (P24, n=3; P7, n=1). Therapy was intensified with the NRTI Abacavir (ABC) but in October 2007 low-level viraemia persisted. Sequencing showed the emergence of the major protease mutation M46I and one gag CS mutation (P7/P1 A431V). The HIV-1 RNA viral load and PR and Gag evolution over time are shown in figure 5.4.



Baseline time point						
Sample ID	**Date**	**Protease Mutations**			**Gag mutations**	
		Primary	**Secondary**	**Others**	**Gag CSs**	**Gag non-CSs**
RFH-2_B	Sep-2006	None	K20R, M36I, L63P	I13V, R41K, I62V, I72V	**P17/P24:** V128A **P24/P2:** none **P2/P7:** S373P, T375N, R380K **P7/P1:** none **P1/P6:** none	**P17:** R76K, Y79F, V82I, T84V, I94V, A119T, H124N **P24:** I138L, A146P, I267V, A309S **P7:** N385K, I389V, R406K **P6:** G466R, T471A, E477D, P478S, L483Q, R490K, N495S, Q500K **P1:** Y441H **P2:** T371N
Treatment emergent mutations						
Sample ID	**Date**	**Protease Mutations**			**Gag mutations**	
		Primary	**Secondary**	**Others**	**Gag CSs**	**Gag non-CSs**
RFH-2_1	Mar-2007	No new mutations	No new mutations	No new mutations	No new mutations	**P17:** R15K, K28Q, K30R, K114R, D121G, G123K, N126S **P7:** N382T, V390I
RFH-2_2	Jul-2007	L76V	No new mutations	No new mutations	No new mutations	**P24:** I147L, S148T, V159I **P7:** R384K
RFH-2_3	Oct-2007	M46I	No new mutations	No new mutations	**P7/P1:** A431V	No new mutations

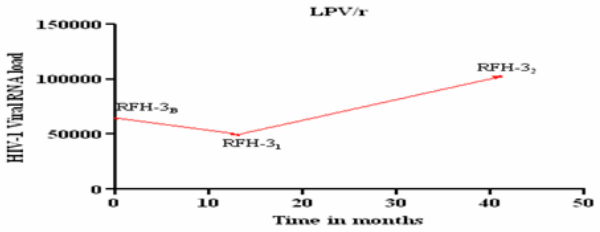
Figure 5.4 Dynamic of emergence of PR and Gag mutations during PI/r-based therapy with ongoing viraemia-patient RFH-2.

PI treatment and viral load at each time point are indicated. The table below shows the PR and Gag mutations emerging at each time point. Major PR mutations and Gag mutations associated with PI exposure are shown in red.

- **Patient 3 (RFH-3)**

The patient first attended the Royal Free Hospital in June 2004 when he was already highly ART-experienced and failing an IDV/r-based regimen consisting of didadosine (ddI), tenofovir (TDF), efavirenz (EFV) and ritonavir boosted indinavir (IDV/r). PR and Gag Sequencing showed infection by a CRF02_AG recombinant HIV-1 strain, two major protease mutations (M46I, I50V) and several gag CS (2 associated with PI-exposure) and non-CS mutations (12 associated with PI-exposure). A ritonavir boosted lopinavir (LPV/r)-containing

regimen was started (ddI, TDF, EFV and LPV/r) and in July 2005 the patient had a viral load of 49,861 copies/ml. PR and Gag sequencing did not show the two previous major protease mutations but all previously seen Gag mutations were still detectable. Two new major protease mutations emerged at this time (I54V, V82A) as well as two polymorphic CS mutations (P2/P7), and 8 non-CS mutations: three in MA-P17 (all associated with PI-exposure), three in P24 (none associated with PI-exposure) and two in P6 (one associated with PI-exposure). The patient was switched to a new regimen consisting of zidovudine (ZDV), lamivudine (3TC), abacavir (ABC), LPV/r and saquinavir (SQV) and in November 2007 his viral load was 102,504 copies/ml and the major protease mutation G48V as well as 2 mutations in P17 associated with PI-exposure, 6 mutations in P24 of which only one was associated with PI-exposure, 5 mutations in P7 of which none was associated with PI-exposure, 3 mutations in P6, two of which were associated with PI-exposure and one P1/P6 CS mutation associated with PI-exposure had also emerged. The HIV-1 RNA viral load and PR and Gag evolution over time are shown in figure 5.5.



Baseline time point						
Sample ID	**Date**	**Protease Mutations**			**Gag mutations**	
		Primary	**Secondary**	**Others**	**Gag CSs**	**Gag non-CSs**
RFH-3B	Jun-2004	M46I, I50V	K20I, M36I,	V11I, I13V, K14R, I15V, G16A, R41K, H69K, L89V	**P17/P24:** None **P24/P2:** None **P2/P7:** A374S, T375V, R380K **P7/P1:** A431V **P1/P6:** P453L	**P17:** R15K, N109T, D121G, T122E, G123K, N126S, G127P **P24:** I138L, A146P, I267V, L268M **P7:** I389A, K397R, K418R **P6:** E454T, T456S, E460A, F463L, R464K, T469I, T470A, Q474P **P1:** Y441H **P2:** T371N, N372S
Mutations emerging (disappearing) at each time point						
Sample ID	**Date**	**Protease Mutations**			**Gag mutations**	
		Primary*	**Secondary**	**Others**	**Gag CSs**	**Gag non-CSs**
RFH-3I	Jul-2005	I54V, V82A (M46I, I50V)	No new mutations	No new mutations	**P2/P7:** I376V, M377I	**P17:** A115I, A120S, H124K **P24:** I147L, V159I, V191I **P6:** T471A, S473P
RFH-3L	Nov-2007	G48V	No new mutations	No new mutations	**P1/P6:** S451T	**P17:** I94V, K114R **P24:** S148T, Q182H, M228I, E230D, E260D, R286K P6: P472S, I479T, L486V

Figure 5.5 Dynamic of emergence of PR and Gag during PI/r-based therapy with ongoing viraemia-patient RFH-3.

PI treatment and viral load at each time point are indicated. The table below shows the PR and Gag mutations that emerged or were lost at each time point. Major PR mutations and Gag mutations associated with PI exposure are shown in red.

* Primary protease mutations shown in braquets were lost with respect to the baseline sample.

5.3.3 Clonal analysis

We performed clonal analysis using plasma samples from the last available time point for the three patients within the group II (RFH-1, RFH-2 and RFH-3) and for patient Pt-20 from group I. A total of 10 clones per patient were analysed.

5.3.3.1 Full-length protease and Gag clones from patient RFH-1

At the last time point, population sequencing in this patient had shown three major, four minor and three polymorphic protease mutations, respectively. Regarding Gag, population sequencing revealed five Gag CS mutations and 58 non-CS mutations. Out of the five CSMs detected, three were associated with PI-exposure according to the analysis presented in the previous chapter located in P17/P24, P2/P7 and P1/P6, respectively. With regard to non-CSMs, 20 out of the 58 mutations detected were associated with PI exposure, 11 in P17, 8 in P6 and 1 in P24. The 10 clones analysed all showed the four major protease resistance mutations and Gag mutations associated with PI-exposure were observed in between 54% and 100% of the clones. In particular, all those Gag mutations strongly associated with PI-exposure (the association remained after a most conservative cut-off was applied, Bonferroni's correction) including CSMs: Y132F and L449F; non-CSMs in P17: I94V, K103R, K114R, D121G, N126S and the non-CSMs in P6: F463L were observed in over 90% of the clones. Results are shown in Figure 5.6.

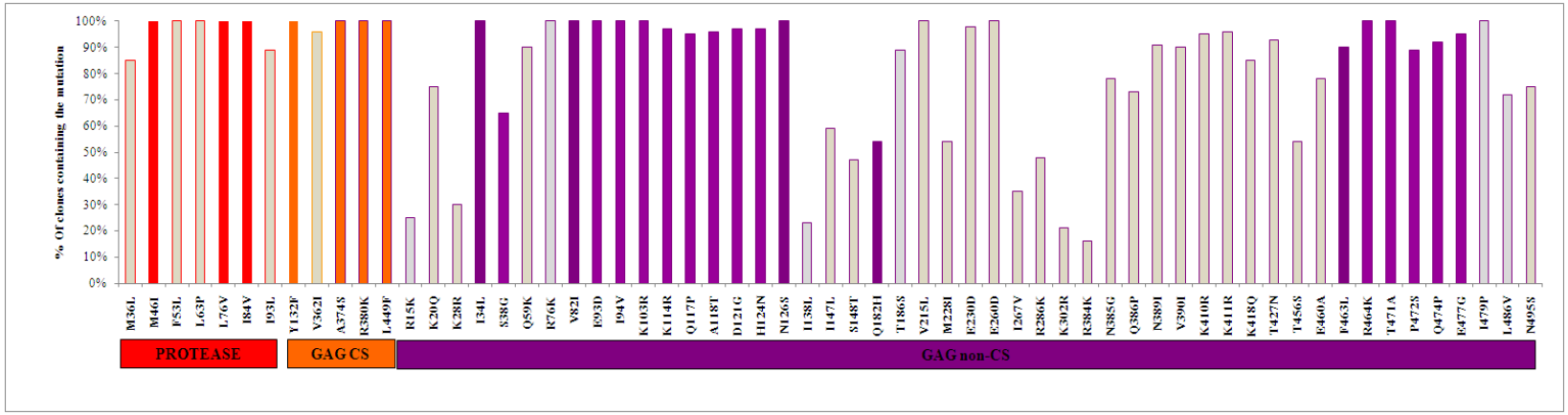


Figure 5.6 Clonal analysis from patient RFH-1.

Full-length protease and gag genes from patient RFH-1 were cloned and sequenced. The proportions (%) of clones (n = 10) with mutations are shown. Sequences were analyzed for the presence of mutations relative to HXB2 wild type. Protease mutations were classified according to the IAS 2011-list of mutations. Bars in red denote major protease mutations, bars in orange represents Gag CS mutations associated with PI exposure and bars in purple shows gag non-CS mutations associated with PI exposure. Bars in grey indicate secondary protease mutations and gag mutations not associated with PI exposure, respectively.

5.3.3.2 Full-length protease and Gag clones from patient RFH-2

At the last time point, population sequencing for this patient revealed the presence of two major, three minor and four polymorphic protease mutations, respectively. In addition, 5 CS and 37 non-CS mutations were also detected. Regarding Gag mutations, two out of the five CSMs were associated with PI-exposure located at P2/P7 and P7/P1, respectively and 11 out of the 37 non-CSMs were also associated with PI-exposure, located at P17 (n = 8), P6 (n = 2) and P7 (n = 1). Clonal analysis demonstrated a very homogeneous population and the major PR mutations as well as the Gag CS and non-CS mutations associated with PI-selective pressure were present in all the ten clones. Figure 5.7 shows the protease and Gag mutations present in the 10 clones for patient RFH-2.

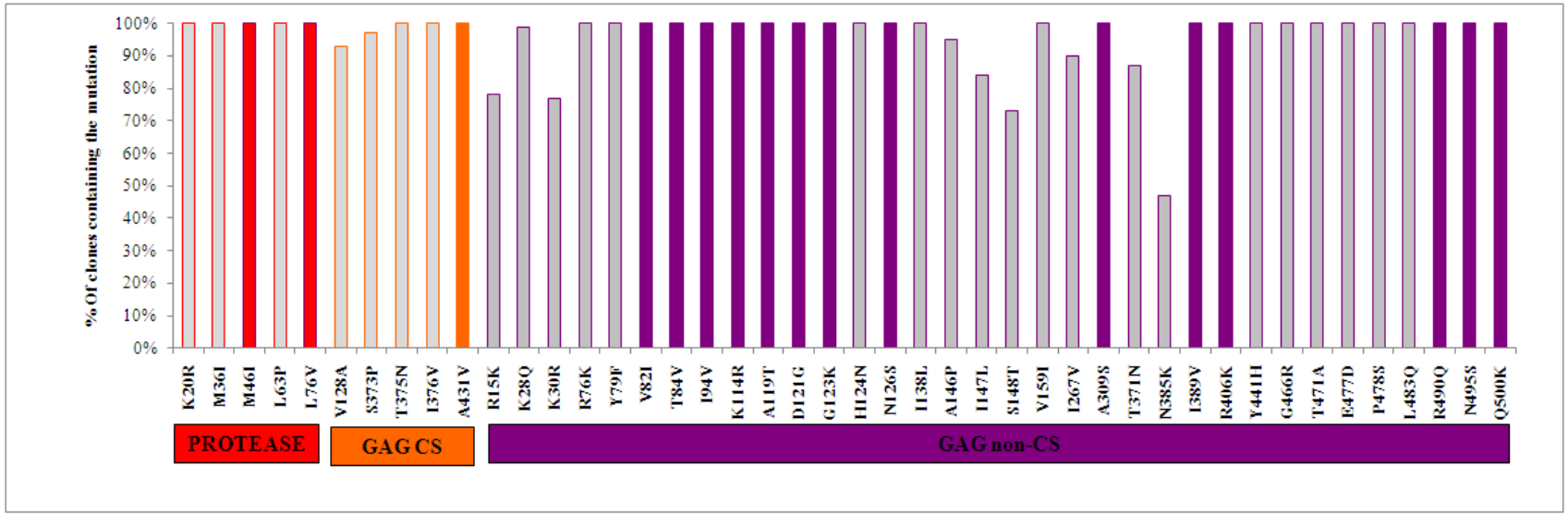


Figure 5.7 Clonal analysis from patient RFH-2.

Full-length protease and gag genes from patient RFH-2 were cloned and sequenced. The proportions (%) of clones (n = 10) with mutations are shown. Sequences were analyzed for the presence of mutations relative to HXB2 wild type. Protease mutations were classified according to the IAS 2011-list of mutations. Bars in red denote major protease mutations, bars in orange represents Gag CS mutations associated with PI exposure and bars in purple shows gag non-CS mutations associated with PI exposure. Bars in grey indicate secondary protease mutations and Gag mutations not associated with PI exposure, respectively.

5.3.3.3 Full-length protease and Gag clones from patient RFH-3

At the last time point, population sequencing for this patient showed the presence of three major, two minor and eight polymorphic protease mutations, respectively. Moreover, seven Gag CS and 46 non-CS mutations were also detected. With regard to Gag mutations, three of the seven CSMs were associated with PI exposure and were located at P2/P7, P7/P1 and P1/P6, respectively. Similarly, 20 of the 46 non-CSMs were associated with PI-selective pressure and were located at P17 (n = 11), P24 (n = 2) and P6 (n = 7). Clonal analysis demonstrated a homogeneous population and major protease mutations and Gag mutations associated with PI-exposure were present in between 99% and100% of the clones analyzed.

Results are presented in figure 5.8.

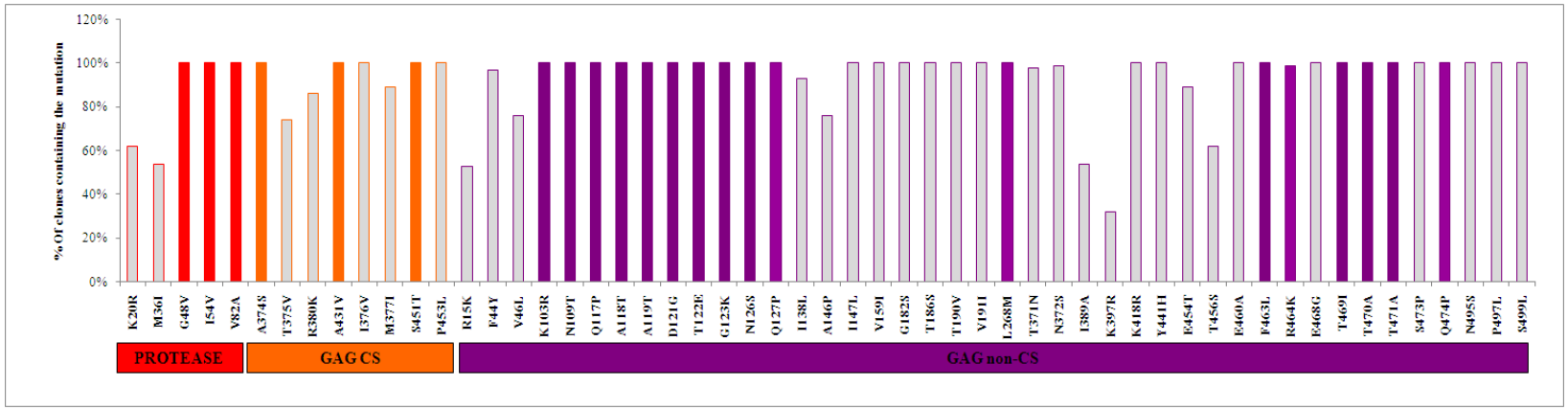


Figure 5.8 Clonal analysis from patient RFH-3.

Full-length protease and gag genes from patient RFH-3 were cloned and sequenced. The proportions (%) of clones (n = 10) with mutations are shown. Sequences were analyzed for the presence of mutations relative to HXB2 wild type. Protease mutations were classified according to the IAS 2011-list of mutations. Bars in red denote major protease mutations, bars in orange represents Gag CS mutations associated with PI exposure and bars in purple shows gag non-CS mutations associated with PI exposure. Bars in grey indicate secondary protease mutations and gag mutations not associated with PI exposure, respectively.

5.3.3.4 Full-length protease and Gag clones from patient Pt-20

At the time of VF, population sequencing had shown four major, six minor and three polymorphic protease mutations. In addition, six Gag CS and 26 Gag non-CS mutations were also detected. With regard to Gag, three of the six CSMs were associated with PI exposure and were located at P17/P24, P2/P7 and P7/P1, respectively. Furthermore, eight out of the 26 non-CSMs were also associated with PI-selective pressure and were located at P17 (n = 5), P24 (n =1), P7 (n =1) and P6 (n = 1). Clonal analysis showed a homogeneous population and all major protease mutations and Gag mutations associated with PI exposure were detected in all of the ten clones analysed.

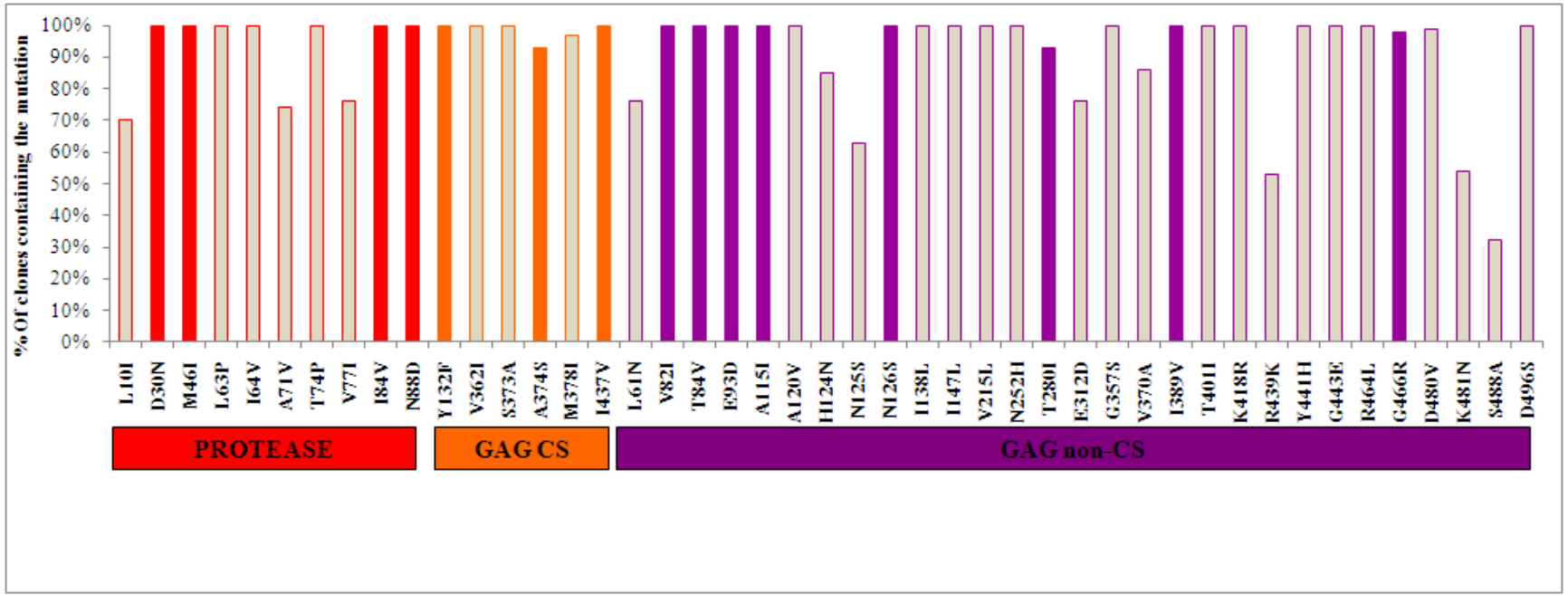


Figure 5.9 Clonal analysis from patient Pt-20.

Full-length protease and gag genes from patient Pt-20 were cloned and sequenced. The proportions (%) of clones (n = 10) with mutations are shown. Sequences were analyzed for the presence of mutations relative to HXB2 wild type. Protease mutations were classified according to the IAS 2011-list of mutations. Bars in red denote major protease mutations, bars in orange represents Gag CS mutations associated with PI exposure and bars in purple shows gag non-CS mutations associated with PI exposure. Bars in grey indicate secondary protease mutations and Gag mutations not associated with Pi exposure, respectively.

5.4 Discussion

Numerous studies have shown that patients failing ritonavir-boosted protease inhibitors often lack major resistance-associated mutations in the protease gene that may account for the clinical resistance observed (Kempf et al, 2004; Delaugerre et al, 2009; Gupta et al, 2008; Lathouwers et al, 2011). Growing evidence however indicates that complete assessment of PI resistance is a more complex process than revealed by the detection of mutations in the protease gene. Analysis of the protease substrate Gag may be required to obtain a more complete picture. In this respect, the involvement of certain gag regions located at the protease cleavage sites P7/P1/P6 has been clearly established (Brumme et al, 2003; Banke et al, 2009; Brann et al, 2006; Cote et al, 2001; Maguire et al, 2002; Malet et al, 2007 and Robinson et al; 2000). However, the role of changes occurring elsewhere in the full-length gag protein has only rarely been addressed. The few studies available suggest that determinant of PI resistance may be located beyond gag CSs (Parry et al, 2009).

In the current chapter we performed a longitudinal analysis of full-length protease and gag genes in patient experiencing virological failure while on PI/r-based regimens. We selected two different groups of patients to perform the analysis. The first group consisted of 28 patients infected with subtype B HIV-1 who had been enrolled in clinical trials and started on IDV/r, SQV/r or LPV/r. All of these patients experienced virological failure during up to 24 weeks of follow-up. The majority of the patients were treatment-experienced before commencing on one of these three regimens and in fact all but five had been exposed to PIs previously. Consistent

with this, we observed a high prevalence of primary and secondary protease mutations at baseline. In addition, we also detected a high frequency of Gag mutations at both CSs and non-CSs which, based upon the analysis shown in the previous chapter, were associated with PI exposure. A substantial number of these baseline mutations, markedly those detected at the CSs, occurred very uncommonly in PI-naïve patients suggesting that long term PI-selective pressure could account for the significant variability observed in Gag in these baseline samples. Consistent with this hypothesis, the five trial patients who were PI-naïve at baseline showed no major protease mutation and the mutations detected in Gag were predominantly polymorphisms not associated with PI-experience. This was the case for the CSs as well as most non-CS regions except for the matrix protein (P17). P17, showed a high degree of variability even in PI-naïve patients.

At the time of virological failure, our population showed a high prevalence of emergent primary, secondary or polymorphic protease mutations (17% of patients and 46% of patients, respectively) consistent with true virological failure and pointing at emergence of resistance variants as the cause of detectable viraemia. In addition, 54% of patients showed emergent polymorphic protease mutations, which are not known to confer PI resistance. Interestingly, emergence of mutations during virological failure was not restricted to the protease gene and in fact occurred more frequently in the Gag gene both within CSs (46% of patients) and in regions outside the CSs (64% of patients). While the vast majority of emergent mutations in Gag CSs and in P17 were seen predominantly or exclusively in the context of PI-selective pressure (77% and 94%, respectively), a substantial number of mutations emerging at other non-CSs such P24,

P7 and P2 were polymorphisms which are also seen relatively commonly in PI-naïve patients. This observation suggests that PIs exert a greater selective pressure on CSs and the non-CS, P17 protein. As a result, one may postulate that mutations in these regions play a key role in PI failure.

The high number of polymorphisms that were seen both at baseline and emerging under PI-selective pressure in Gag may also be explained by pressure exerted by the immune system. Indeed, early studies showed that cellular immune pressure represent a dominant selective force in viral evolution, accounting for up to half of the intra-host amino-acid sequence diversity selected over the course of infection in some cases (Allen et al, 2005; O'Connor et al, 2004 and Jones et al, 2004). Importantly, certain Gag mutations observed frequently in PI-naïve and PI-experienced patients in our population, such as K28QR, I34L, V82I, T84V, Y79F located in P17 and I138L, A146P, G357S and V362I located in P24, are present within well defined CD8+ T-cell epitopes (Frahm et al, 2008) and early studies performed by Philips and colleagues demonstrated that some of these variants found in HIV seropositive haemophilic donors, such as I34L, led to loss of CD8+ T cell recognition (Phillips et al, 1991). Similarly, Yokomaku and co-authors showed that certain Gag variants located in immunodominant CD8+ epitopes in P17, such as K28Q, I34L or Y79F, failed to be killed by CD8+ T cell (Yokomaku et al, 2004) due to an impaired antigen processing and presentation. Overall, these studies suggest that evolution of Gag is greatly influenced by the host immune response and specifically by adaptive CD8+ T-cell selective pressure.

With regard to Gag CSMs emerging at VF, we observed that mutations emerged in P7/P1, P1/P6 and P2/P7. Thus, at the time of VF we observed mutations emerging at P7/P1 in five patients. The mutations detected were: I437V in a patient failing IDV/r who did not show major protease mutations either at baseline or at VF and in a patient failing SQV/r who by contrast, presented major PR mutations at baseline and emerging at VF; the A431V mutation emerged in a patient failing SQV/r who showed major PR mutations at baseline but not emerging at VF; the K436R mutation appeared in a patient failing LPV/r who showed major PR mutations at baseline, but no new major PR mutations emerged at VF and finally, the E428K emerged in a patient failing LPV/r in the absence of major PR mutation either at baseline or at VF. All the mutations described above, except for E428K, have been identified in our study (see chapter four) and others as associated with PI selective pressure (Nijhuis et al, 2007; Verheyen et al, 2006; Bally et el, 2000; Koch et al, 2000; Malet et al, 2007 and Maguire et al, 2002) . The I437V mutation was described during *in vitro* selection experiment with an experimental PI (RO033-4649) and was shown to confer 5 to 8-fold resistance to multiple PIs in the absence of protease mutations (Nijhuis et al, 2007). The mutation was shown to enhance P7/P1 processing of the substrate by wild type protease. Similarly, The A431V mutation is one of the best-characterized CS mutations observed in the setting of PI failure. A number of studies have probed the association of A431V with protease resistance-associated mutations and demonstrated that the presence of this mutation increase the rate of Gag cleavage by a mutated protease indicating a plausible compensatory role for A431V (Verheyen et al, 2006; Bally et al, 2000; Koch et al 2001; Malet et al, 2007). In line with these studies, in the cross-sectional analysis presented in chapter 4 we found a high prevalence of A431V in PI-experienced patients and its absence in PI-naïve individuals. Furthermore, there was a statistically significant association with several major

protease mutations including M46IL and I54V, both of which were present at baseline in the patient who acquired A431V. Finally, the K436R was identified in our cross-sectional analysis as associated with PI exposure in line with previous reports (Verheyen et al, 2006). Overall, our results suggest that mutations emerging at this CS were indeed the result of PI-selective. The fact that patients in the three different regimens developed mutations associated with PI selective pressure at NC-P7/P1 suggests that this CS may broadly influence PI susceptibility and mutations selected at this CS may contribute to PI resistance by a common mechanism to the drug class. In the majority of cases major PR mutations were present either at baseline or emerging at VF along with CSMs suggesting that emergent of mutations at this site, in general, follow the appearance of PR resistance-associated mutations (RAMs) and may act as compensatory mutations in order to restore the impaired replication usually observed in viruses harbouring major PR mutations. However, CSMs were also observed in a few cases emerging at VF in the absence of major PR mutations either at baseline or at VF indicating that mutations at this CS may occasionally precede the emergence of PR mutations and they may be the first signal of virological failure to PIs.

Four patients developed mutations in P1/P6 including P453L in one patient failing SQV/r who showed the L90M and I84V major PR mutations at baseline and emerging at VF, respectively; the P453T mutations emerged in two patients failing SQV/r, in both cases in the absence of major PR mutations either at baseline or at the time of VF. Finally, the L449F mutation appeared in a patient failing SQV/r, once again in the absence of major PR mutations at baseline or at VF. The P453L mutation has been previously reported in the literature as a naturally occurring

polymorphism but has been also associated with major protease resistance mutations including I84V and V82A (Maguire et al, 2002). Similarly in the cross-sectional analysis previously presented we observed this mutation to occur frequently in both PI-naïve and PI-experienced patients and can be speculated that although natural polymorphism may favor the selection of protease resistance mutations by facilitating viral replication. Interestingly, the I84V mutation emerged concomitant with the P453L in our patient, confirming the association between both mutations that Maguire and co-authors (Maguire et al, 2002) reported. The L449F mutation has been previously documented and demonstrated to have *in vitro* effects on both viral fitness and phenotypic resistance to APV (Maguire et al, 2002). The P453T mutation has not been previously described. However, this mutation was found in our study (chapter four) as associated with PI exposure ($p < 0.002$). Overall, in our study mutations at this site occur in the absence of major PR mutations suggesting that they may precede the appearance of the latter and facilitate their selection. This finding contrast with the results observed by Maguire and colleagues who found that the L449F mutation was selected after the emergence of the major PR mutation I50V. However, agreed with those recently reported by Ghosn and colleagues who described that the L449F mutation detected at baseline in a patient on LPV/r led to the selection of the major PR mutation I50V at the time of VF (Ghosn et al, 2011). Our study and those conducted by Maguire and Ghosn differed on the PI failing which may account for the different encounters. In the study conducted by Maguire, the failing PI was APV, which is a second generation PI that was designed with greater resemblance to the natural substrate Gag and which interacts much tighter with the viral PR than earlier generation PIs and consequently displays a higher genetic barrier to resistance. Given these differences, it is not surprising that specific mutations may have a different impact on resistance to different PIs and that resistance pathways differed between both

of them. In fact, the order of accumulation of Gag and PR mutations is likely to depend on the failing PI and likely on the backbone viral sequence. This phenomenon has been clearly established for mutations in the PR. Bandaranayake and colleagues showed that PR sequence polymorphisms were capable of altering protease activity and inhibitor binding and consequently were able to alter the pathway of inhibitor resistance (Bandaranayake et al, 2010). Another interesting finding in our study is that all patients developing mutations at this site at the time of VF were failing SQV/r. This data may be indicative of a greater impact of P1/P6 mutations on this drug compared to other PIs. However, the number of patients studied is low as to establish a definitive linked.

Four additional patients developed mutation in P2/P7 at VF including T375A in a patient failing IDV/r who did not show major PR mutations at baseline or at VF and in another failing LPV/r who presented three major PR mutations at baseline, although did not develop any new major PR mutation at VF; the R380K mutation was seen emerging in a patient failing IDV/r who did not show major PR mutations at baseline or at VF and finally the G381S emerged in a patient failing IDV/r who showed four major PR mutation at baseline and none emerging at VF. Data on the role of mutations at this CS in PI resistance are scarce, since most studies have exclusively addressed the function of mutations at P7/P1/P6. A recent study conducted by Ghosn and colleagues showed that baseline mutations at this CS were predictive of virological failure of LPV/r monotherapy in patients enrolled on the monark trial (comparison of first line LPV/r and LPV/r + ZDV/3TC). However, they did not see emergence of mutations at this CS or others at the time of VF (Ghosn et al, 2011). The R380K mutation has been found in our study and others

to occur frequently as a polymorphism (Malet et al, 2007). The G381S mutation has been first time described in our study as associated with PI exposure ($p < 0.02$) (chapter four). Finally, the T375A mutation has been found to be associated with PI-experience both in our study ($p < 0.002$) (chapter 4) and by other investigators (Malet et al, 2007).

We did not observe emergence of mutations in the two remaining CSs P17/P24 and CA-P24/P2. We and others have indeed observed that these two CSs are highly conserved in both PI-naïve and PI-experienced patients (Malet et al, 2007). We also found in the previous chapter that only a limited number of mutations such as V128I and Y132F were detected in P17/P24 in PI-experienced patients and these were indeed significantly associated with exposure to PIs ($p < 0.0001$). TheY132F mutation was observed in 3 of our 28 patients at baseline, all of whom had been exposed to other PIs before entering the trial. One of the patients presented in addition the V128I mutation. All patients with mutations at this CS showed in adition to three to four major protease mutations, suggesting that the patients were heavily treated and therefore that evolution in this CS does occur, but probably requires prolonged PI-selective pressure. Overall, our findings suggest that while evolution of P17/P24 under PI pressure is possible, greater effects occur at CSs located in the C-terminal site - P7/P1, P1/P6 and P2/P7. In detail treatment history and length of exposure to PIs was unfortunately unavailable to confirm this judgment.

We also observed frequent emergence of mutations in non-CS regions of Gag at the time of virological failure. The effect was most prominent in the P17 protein, but was also noticed in all

other domains of Gag. Interestingly, while most mutations emerging in P17 were associated with PI-exposure (94%), several emerging mutations in other non-CSs were polymorphisms. These findings point once again to a prominent role of P17 in PI failure.

The role of the P17 protein in PI failure was discussed in the previous chapter. As mention there, the role of non-CS regions in PI resistance has been addressed in only a limited number of studies. One of them demonstrated that P17 plays a pivotal role in the rescue of the replicative capacity of multi-protease resistance virus. The same study also revealed that mutations in P17 are sufficient to confer resistance to all PIs (Parry et al, 2009).

In the previous chapter, we described a high prevalence of mutations associated with PI-exposure not only in P17 but also in P6. Although, no PI-associated mutations emerged in P6 at the time of virological failure, at baseline patients who had been previously exposed to PIs often presented PI-associated mutations at this site. By contrast, only one of the five PI-naïve patients showed PI-associated mutations in P6 before starting PI treatment. Therefore, we can speculate a situation similar to the one mentioned above for the P17/P24 site, in that P6 evolution can occur under PI-selective pressure, but effects are mainly observed in the P17 protein. As a consequence, the short period of time between baseline and VF for this patient may account for lack of emergence of mutations at the P6 protein. Once again, a formal analysis on the association of presence of specific mutations in particular regions of Gag with duration of PI treatment would be required to support this verdict. However, this information was unavailable for all of the patients.

It is worth mentioning that we also detected Gag mutations associated with PI-exposure at baseline even in the five patients who had never been exposed to PIs. The effect was more noticeable in MA-P17 but also to a lesser extend in other regions such as P7, P6 and the P2/P7 CS. In fact, all patients showed mutations associated with PI selective pressure at MA-P17, one patient showed in addition one mutation associated with PI selective pressure at P7 and two at P6 and two further patients presented a P2/P7 CSM at baseline. Conversely, mutations described as polymorphisms in our study and others such as R380K (Malet et al, 2007) and P453L (Maguire et al, 2002 and Verheyen et al, 2006) were seen emerging at VF in the absence or concomitantly with major PR mutations, respectively.

The detection of mutations associated with PI-exposure among PI-naive patients is not unusual as naturally occurring polymorphisms can confer an advantage to strains growing under PI-selective pressure and be enriched during therapy and play a role in virological responses. For example the secondary protease mutation L63P, which is frequently observed in untreated patients, does not confer resistance by itself but provides a significant replication benefit for certain viral mutants, particularly under drug pressure (Martinez-Picado, 1999; Sune et al, 2004), explaining its higher prevalence in PI-experienced patients. Similar considerations may apply to certain Gag mutations, explaining their presence in drug-naïve individuals and their increase in prevalence in the setting of PI-exposure.

In order to gather further insights into the evolution of Gag under prolonged PI selective pressure, we studied in detail three patients who had long-term viraemia while on a PI/r-based regimen. We observed that CS mutations frequently occurred at the same codons and often showed similar patterns in these patients. For example, the A431V CS mutation in P7/P1 was seen in two of the three patients. By contrast, mutations at non-CSs differed considerably among patients. As also seen with trial patients, the P17 protein showed a significant number of PI-associated mutations emerging over time under PI-selective pressure. In addition, we observed a high number of non-polymorphic changes emerging at other sites of Gag, predominantly in P6s. These results provide further support to the hypothesis that prolonged viral replication under selective drug pressure may drive evolution outside the previously mentioned hotspots such as P17 and CS P2/P7/P1/P6. Furthermore, the wide variety of non-CS mutations encountered in different patients indicates that while CS mutations emerge in consistent patterns, the evolutionary pathways in non-CS are complex. This was also earlier suggested by Myint and colleagues, who found that while protease and CS mutations were consistent between different viral clones, non-CS mutations differed considerably (Myint et al, 2004).

Finally, we performed clonal analysis to confirm co-occurrence of protease and Gag mutations in the same genome, which would strengthen a role for Gag mutations in PI failure. As the main aim of this clonal analysis was the identification of possible linkage between protease and Gag mutations, we were only interested in the dominant quasispecies and as a consequent we only examined 10 clones in each patient. We observed very homogenous population in all the three patients analysed and in most cases protease mutations and gag mutations associated with PI-

exposure co-existed on the same viral genome. This observation supports a contribution of Gag mutations as compensatory mutations, true resistance mutations, or both. We frequently observed the association between the A431V and the I437V CS mutations in Gag with major protease mutations, particularly M46I and I84V respectively. We and others previously observed a significant statistical association between occurrence of A431V and detection of M46I and other protease mutations (Verheyen et al, 2006). Here we have shown that the statistical relationship correlates with a genetic linkage between these two mutations.

In order to account for random mutations that may result as a consequence of PCR errors, we performed the clonal analysis from the PCR product obtained from five independent PCR reactions. The clones we obtained were rather homogenous in each of the patients examined. This is likely to reflect the fact that the dominant replicating quasispecies is also rather homogenous during prolonged replication under the same treatment regimen. A number of technical considerations should be made however. Firstly, two of the plasma sample selected for clonal analysis showed a low HIV-1 RNA load (961 and 229 copies/ml respectively). In addition, due to insufficient sample, two other samples were diluted to a final viral load of 750 and 1000 copies/ml, respectively. The relatively low input of HIV-1 RNA may have reduced the variability of RNA templates used for cloning. However, it should be pointed out that the scope of our clonal analysis was not to address viral genetic diversity within each patient, but rather to assess the linkage of protease and gag mutations in the dominant quasispecies.

A further technical consideration is that it may be argued that nested PCR amplification may introduce errors through the incorporation of incorrect nucleotides, which can be subsequently cloned and lead to misleading results. However, all the mutations seen in the clones were previously detected by population (Sanger) sequencing, which can only identify variants that represent over 20% of the population (Alcorn and Faruki, 2000). If PCR errors are introduced in the early cycles the spurious quasispecies can reach such a threshold of representation. However, this scenario is unlikely when employing optimized PCR reactions and particularly when proof-reading enzymes are used.

Another consideration is the possibility of recombination occurring *in vitro* during the PCR reaction, leading to a false interpretation of linkage of mutations. Recombinants presumably arise during PCR reactions due to the presence of incompletely extended primers annealing to a heterologous target. Consequently to prevent artificial recombination it is necessary to achieve complete strand synthesis. One effective strategy is performing limiting dilutions of the cDNA so that a single template is employed in the PCR and sequencing reactions. In addition, this also minimizes errors in the PCR reaction, as even if an error is introduced during the early PCR stages, the error will not be present in more than 25% of the bases in the mixture and will not result in an erroneous base call during sequencing (Learn et al; 1996). Therefore, limiting dilution is considered the reference technique for the identification of linkage of mutations. However, it requires a high volume of sample which was not available in our study. Nevertheless, we introduced a number of steps in order to minimize the risk of artificial recombination. Firstly, we employed a proof-reading enzyme with reduced RNAse H activity

during PCR and we also employed long PCR extension times of up to 3 minutes. Both actions are known to promote complete strand synthesis and as a result minimize artificial recombination. In addition, we used a low input of HIV-1 RNA, which limits the number of initial templates in the PCR reaction, in turn reducing the possibility of artificial recombination. We are therefore confident that the risk of artificial recombination was small to negligible in our study, and that the mutations shown to co-exist were indeed linked on the same viral genome. We, however, also confirmed the relationship between different mutations by performing *in vitro* replicative capacity and drug susceptibility experiments, as shown in the subsequent chapter.

To summarize, the results obtained in this chapter showed that under PI selective pressure evolution of Gag occurs primarily, although not exclusively, at CSs (P2/P7, P7/P1 and P1/P6) and at the P17 protein outside CSs. The fact that the majority of emergent mutations at Gag have been shown in our studies and others to be associated with PI exposure support a role for Gag mutations in failure of PI-based regimens. We should emphasize, however, that a high proportion of patients also showed emergent protease mutations which although are not known to confer PI resistance, we cannot exclude that such protease mutations may have led to PI failure in these particular patients. In the subsequent chapter we will focus on studying the contribution of the emergent Gag mutations on drug susceptibility and replicative capacity in order to clarify the potential role of Gag mutations in failure of PIs.

6 Chapter six: effect of Gag mutations on replicative capacity and susceptibility to protease inhibitors

6.1 Introduction

The results obtained in the previous chapters suggest that full-length Gag sequencing may be required for a complete assessment of resistance to protease inhibitors (PIs). Mutations in the Gag polyprotein associated with exposure to PIs and rare or absent in PI-naïve individuals have been identified. Importantly, mutations were not restricted to cleavage sites (CSs) but were detected throughout the whole Gag protein. Furthermore, by performing longitudinal studies in patients failing ritonavir boosted protease inhibitors (PI/r), we described emergence during treatment failure of mutations both within CSs and outside. Interestingly emergence of Gag mutations occurred in same cases in the absence of major protease resistance mutations. In the present chapter I describe the "*in vitro"* phenotypic characterization of some of the virus strains identified in patients in terms of drug susceptibility and replicative capacity (RC). For this purpose, I employed a single cycle recombinant assay.

There are two main methods for characterizing the viral phenotype using recombinant vectors, namely the single cycle system and the multiple cycle system. Both assays employ recombinant viruses obtained by cloning patient-derived sequences into a defective molecular clone. The resulting recombinant virus is then incubated in the presence of increasing concentration of drug, and the drug concentration required to reduce replication of the test virus by 50% relative to the control virus (IC_{50}) is calculated. Results are expressed as n-fold change (FC) in the IC_{50}

compared to the wild-type (WT) reference virus. A FC greater than one indicates reduced susceptibility of the virus tested compared to the wild type. By contrast, hypersusceptibility is usually defined as a virus that has a FC ≤ 0.4 compared to the wild type virus (Clark et al, 2006). The main difference between multiple and single cycle systems is that the first generates an infectious molecular clone capable of multiple rounds of replication. By contrast, the single cycle assay employs a replication-deficient vector that undergoes only a single round of infection. While the multiple cycle assay are proposed to more closely mimic the conditions that the virus experiences "*in vivo"*, the single cycle assay offers the advantages of high sensitivity and greater reproducibility. Furthermore, a single cycle system may be especially appropriate when studying the effect of a specific mutation, as the format limits the opportunity for "*in vitro"* selection of genetically diverse virus subpopulations, which may not accurately reflect the effect of the mutation of interest. In the studies presented in this chapter, a single cycle assay was employed, which is based on the system previously described by Petropoulos (Petropoulos et al, 2000), to characterize drug susceptibility and RC of virus strains and mutations of interest identified in the previous chapters.

6.2 Methods

6.2.1 Patient samples

Sample for drug susceptibility testing was available for the three patients selected from the HIV services at Royal Free Hospital, these are patients identified in chapter five as RFH-1, RFH-2 and RFH-3 as well as for 9 out of the 28 patients selected form the MaxCmin1, MxCmin2 and COLATE clinical trials, these ones include patients Pt-1, Pt-2, Pt-3, Pt-4, Pt-5, Pt-8, Pt-20, Pt-10 and Pt-26 described in chapter five. We selected patients showing evolution on Gag and in particular at sites identified throughout this study as associated with PI selective pressure in order to further explore the role of this specific sites in resistance to PIs. We were interested in addressing the role of both CS and non-CS mutations and as a consequence we selected patients showing evolution at both CSs and beyond. In particular, regarding CS mutations we wished to address the effects of mutations firstly identified in our study, but we also wanted to evaluate how our phenotypic results compare to those previously published. Therefore, we selected some patients with novel mutations (Y132F and T375A) and others with mutations previously described (I437V and A431V). With regard to non-CS mutaions, we were particularly interested in addressing the effects of P17 mutations specially those associated with PI selective pressure in my study, as the results showed in chapter four and five suggest a most prominent role of P17 in PI failure. We, therefore, selected four patients for drug susceptibility and replicative capacity.

- The two first patients were selected from the MaxCmin1, MaxCmin2 clinical trials. One of the patients was infected with a subtype B HIV-1virus and was failing a ritonavir boosted indinavir (IDV/r)-containing regimen with a viral load of 151,359 copies/ml. The other patient was

infected with a subtype B HIV-1 virus and was failing a ritonavir boosted saquinavir (SQV/r)-containing regimen with a viral load of 229 copies/ml. These two patients were identified in chapter five as Pt-2 and Pt-20, respectively.

- The third patient was selected from the HIV clinic at the Royal Free Hospital. The patient was infected with a subtype C HIV-1 virus and had longstanding ongoing viraemia while on a ritonavir boosted amprenavir (APV/r) plus ritonavir boosted lopinavir (LPV/r) containing regimen. The patient was identified in chapter five as RFH-1.

- The fourth patient was infected with CRF02 HIV-1virus and had long lasting ongoing viraemia while on ritonavir boosted lopinavir (LPV/r)-containing therapy. The patient was identified in chapter five as RFH-2.

We selected patients RFH-1 and RFH-2 as both showed siginificant evolution in Gag both at CSs and beyond. Therefore, the study of these two patients allowed us to explore the effect of such eveolution on PI susceptibility and differentlially assess the role of both Gag CS and non-CS mutations. In addition, while patient RFH-1 showed the Y132F CS mutation emerging, the patient RFH-2 demonstrated the A431V appearing. The Y132F mutation has been for the first time described in the present study as associated with PI-selective pressure. By contrast, A431V is well stablished in the literature as implicated on resistance to LPV. Consequently, by selecting these two patients we were able to further explore the effect of the novel Y132F mutation and also to evaluate how our result compare with those previously reported in the case of the A431V mutation.

In addition, we studied patients Pt-2 and Pt20. Patient Pt-2 has the P2/P7 CS mutation T375A emerging at the time of VF. One again, the T375A mutation has only been briefly reported in the literature as associated with PI-selective pressure (Malet et al, 2007). However, its effect on PI susceptibility has never been evaluated. In chapter five, we have also demonstrated an association of this mutation with PI exposure. As a result, we selected this patient in order to describe the effect of this newly described mutation in PI susceptibility and viral replication capacity. Finally, patient pt-20 has the P1/P7 CS mutation I437V together with mutations outside CS, in particular 3 P17 mutations and 2 P1 mutations. The CS mutation I437V has been extensively described in the literature as implicated in PI resistance and in fact it was the first Gag CS mutation found to confer resistance to PIs independently of protease resistance mutations (Nijhuis et al, 2007). Most mutations in P17 emerging in this patient at VF has been found to be associated with PI selective pressure in my study (chapter 4). In addition, the results showed in both chapter four and five suggest that Gag evolution under PI selective pressure is most prominent in the P17 subdomain of Gag and consequently pointed to a main role of P17 in failure to PIs. Therefore, the selection of this patient allowed us not only to compare our results with those previously reported for the case of the I437V mutation, but also to explore the hypothesis of role of P17 in resistance to PIs.

Overall, we considered these four patients representative of the studied population and sufficient to explore our hypothesis.

6.2.2 Side-directed mutants

In addition to using wild-type viruses of interest, we produced site-directed mutants clinical by introducing or reverting specific Gag mutations. Furthermore, we independently introduced the

mutations Y132F and T375A into a wild-type backbone. SDM was in all cases performed using the Quick Change Site-Directed Mutagenesis Kit (Stratagene) as described in section 2.2.2.11 and employing primers described in table 2.3 in chapter 2. The resulting viruses were subsequently studied for the effect of the mutations on drug susceptibility and RC.

6.2.3 Generation of resistance test vectors

Appropriate restriction sites were introduced in patients' samples by PCR employing modified nested PCR primers containing the corresponding restriction site in the 5' end (primer sequences are detailed in table 2.5 in chapter 2). Patient-derived Gag, Protease or whole Gag-Protease regions were subsequently cloned into the Gag-pol expression vector P8.9NSX by employing suitable restriction sites.

6.2.4 Drug susceptibility testing

PI susceptibility testing was performed as indicated in section 2.2.4 and briefly summarized here. Pseudovirus stocks used for PI susceptibility testing were produced by co-transfecting confluent human embryonic kidney 293 (HEK293T) cells with a resistance test vector DNA plasmid containing patient-derived HIV sequences; PMDG encoding the vesicular stomatitis virus G protein; and pCSFLW, encoding the firefly luciferase and the HIV packaging sequence. Cells were harvested 16 hours after transfection and seeded in the presence of different PI concentrations. Pseudovirus stocks produced in the presence of PIs were harvested approximately 24 hours later and used to infect fresh target HEK293T cells. Replication was

monitored by measuring luciferase production in infected target cells 48 hours after infection and compared to a control in the absence of drug. The IC_{50} was calculated by plotting the percentage luciferase inhibition *vs.* log_{10} drug concentration and using GraphPad Prism v5.0 to fit the inhibition curve by nonlinear least-squares analysis. Results were expressed as FC in the IC_{50} compared to the wild-type subtype B HIV-1 reference (P8.9NSX). Experiments were done in duplicate and the calculated IC_{50} represented the mean of at least two independent determinations.

In this study, a technical cut-off has been obtained by repeat testing of the wild type virus used as a reference throughout the experiments (P8.9NSX). In addition, the results were compared with the BCO and/or the two CCO proposed by Virco where available. As described in the introduction, hypersusceptibility was considered a FC of ≤ 0.4 (Clark et al, 2006).

6.2.5 Replicative capacity testing

Replicative capacity was measured as described in section 2.2.4.4. Briefly, as described above, pseudovirus stocks were produced by transfecting HEK293T cells with the three plasmids (RTVs, PMDG, and pCSFLW) and the RC determined by titration of serial dilutions on HEK293T cells and quantification of luciferase activity 48 hours after infection. Luciferase activity was determined with SteadyGlo and a Glomax luminometer (both Promega) and a mean was calculated by using at least four values within the linear range. In order to control for transfection efficiency, the expression of luciferase activity was normalized for the amount of

P24 protein in pseudovirus supernatants and expressed relative to the P8.9NSX reference virus (relative light units, RLU).

6.2.6 Statistical analysis

Replication capacity and fold-changes in IC_{50} were compared using one-way ANOVA test (or unpaired t-test if only two constructs were compared). A $p < 0.05$ was considered to be statistically significant. When groups differed significantly, a Bonferroni's multiple comparison post-test was performed to make two by two comparisons.

6.3 Results

6.3.1 Patients

We studied clinical samples from four patients. The first two patients (Pt-2 and Pt-20) were already PI-experienced before starting IDV/r and SQV/r, respectively. However, patient Pt-2 showed no major protease resistance mutations at both baseline and the time of virological failure, while showing the emergence of Gag mutations at virological failure. By contrast, patient Pt-20 had pre-existing major protease resistance mutations at baseline and showed emergence of both additional major protease resistance mutations and Gag mutations at virological failure. The selection of these two patients allowed us to study the effect of gag mutations on drug susceptibility and RC both in the presence and absence of major PI resistance mutations.

The other two patients (RFH-1 and RFH-2), were PI-naïve at baseline, although patient RFH-1 had been exposed to other antiretrovirals. Neither showed major protease mutations at baseline. Patient RFH-1 was studied over a five-year period during which he showed persistent viraemia while on a regimen composed of LPV/r, APV/r and TDF. Over this period, longitudinal samples showed the emergence of a total of 4 major protease mutations, 35 Gag non-CS mutations and 2 Gag CS mutations. Patient RFH-2 was studied over a one-year period of ongoing viraemia while receiving a LPV/r-based regimen. Over this time, longitudinal samples showed the emergence of 2 major protease mutations, 13 gag non-CS mutations and 1 CS mutation. The selection of these two patients allowed us to assess the long-term evolution in the Gag gene and its impact on drug susceptibility and RC.

Further details on these four patients can be found in chapter five. To facilitate the interpretation of the phenotypic data, a summary of protease and Gag mutations at baseline and at virological failure is presented in figures 6.1 to 6.4.

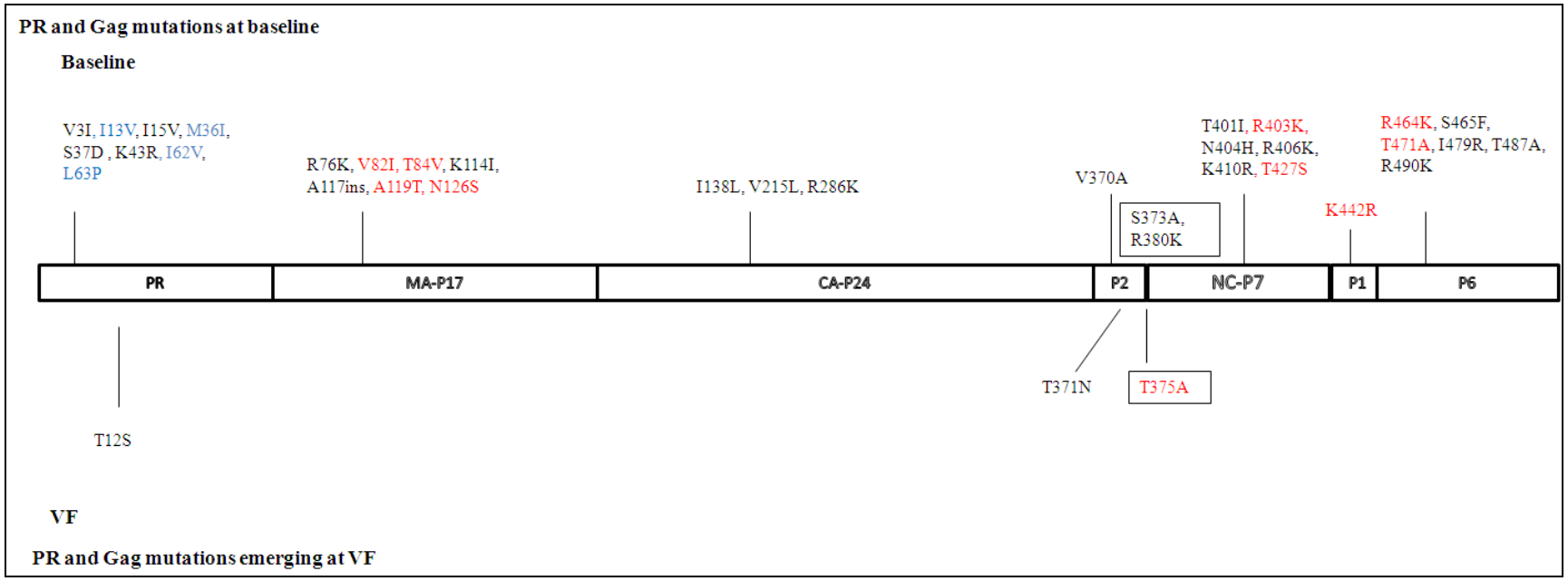


Figure 6.1 Schematic representation of HIV Protease and Gag protein for patient Pt-2.

Amino acid changes found in plasma virus of patient Pt-2 are illustrated. Protease and functional Gag matrix (P17), capsid (P24), P2, Nucleocapsid (P7), P1 and P6 are shown with protease cleavage sites indicated. Changes are numbered according to the HXB2 consensus sequence. Changes detected just before starting IDV/r therapy (baseline) are shown above and changes emerging at the time of IDV/r failure (VF) are presented below. Mutations in blue in the protease represent minor mutations and those in black other polymorphisms. In Gag, mutations in red are those associated with PI-exposure and those in black other polymorphisms. In addition, Gag cleavage sites mutations are presented in solid boxes.

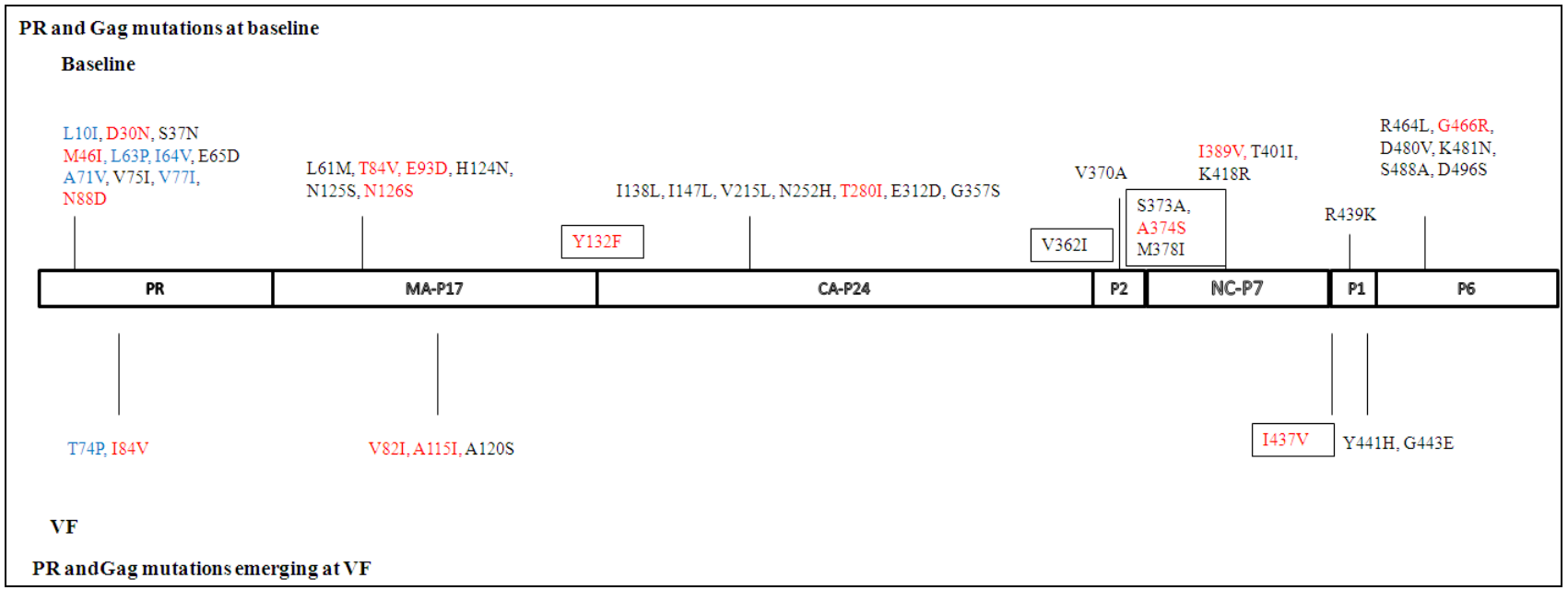


Figure 6.2 Schematic representation of HIV Protease and Gag protein for patient Pt-20.

Amino acid changes found in plasma virus of patient Pt-20 are indicated. Protease and functional Gag matrix (P17), capsid (P24), P2, Nucleocapsid (P7), P1 and P6 are shown with protease cleavage sites indicated. Changes are numbered according to the HXB2 consensus sequence. Changes detected just before starting IDV/r therapy (baseline) are shown above and changes emerging at the time of SQV/r failure (VF) are presented below. Mutations in red in the protease represent major resistance mutations, in blue minor mutations and in black other polymorphisms. In Gag, mutations in red are those associated with PI-exposure and those in black other polymorphisms. In addition, Gag cleavage sites mutations are presented in solid boxes.

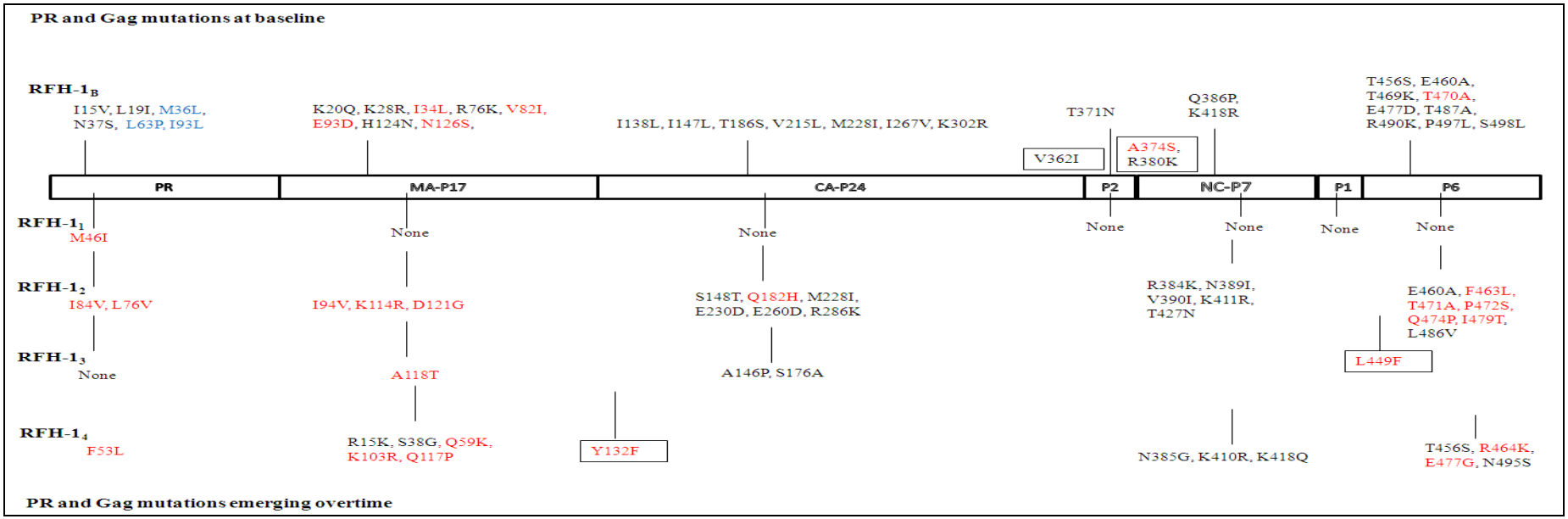


Figure 6.3 Schematic representation of HIV Protease and Gag protein for patient RFH-1.

Amino acid changes found in plasma virus of patient RFH-1 are indicated. Protease and functional Gag matrix (P17), capsid (P24), P2, Nucleocapsid (P7), P1 and P6 are shown with protease cleavage sites indicated. Changes are numbered according to the HXB2 consensus sequence. Changes detected at baseline (RFH-1_B) are shown above and changes emerging at different time points (RFH-1_1 to RFH-1_4) are presented below. Mutations in red in the protease represent major resistance mutations, in blue minor mutations and in black other polymorphisms. In Gag, mutations in red are those associated with PI-exposure and those in black other polymorphisms. In addition, Gag cleavage sites mutations are presented in solid boxes.

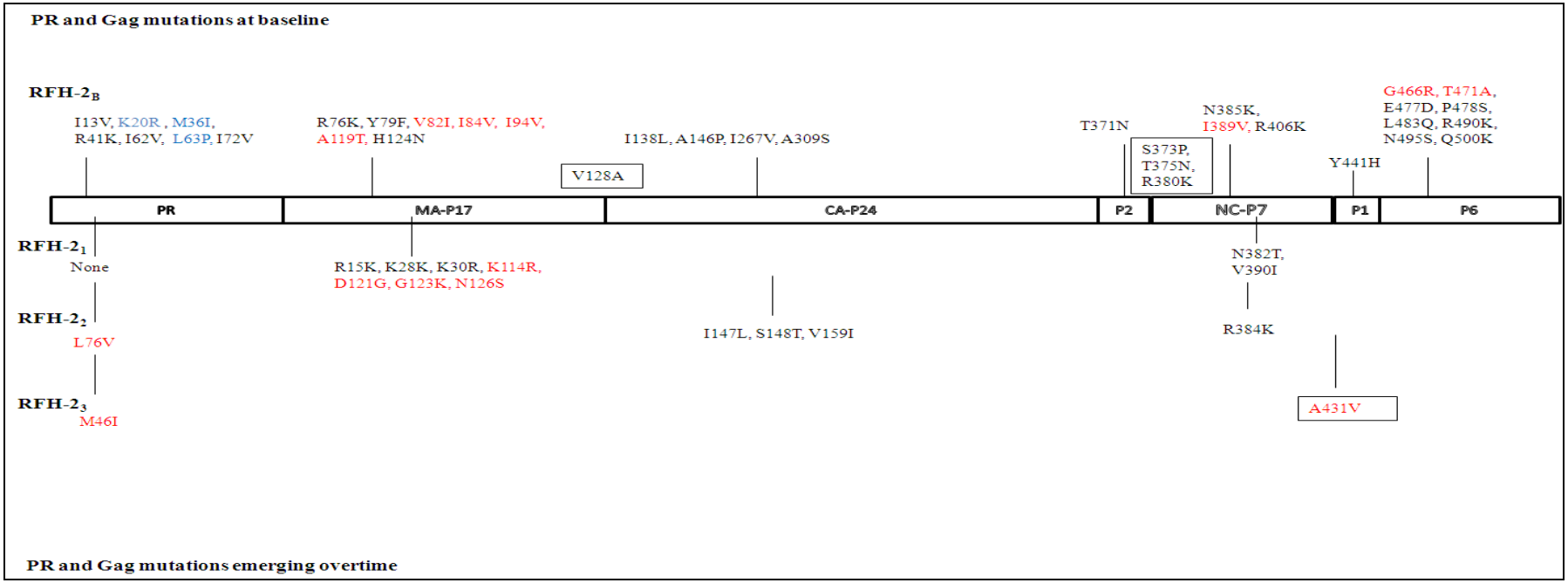


Figure 6.4 Schematic representation of HIV Protease and Gag protein for patient RFH-2.

Amino acid changes found in plasma virus of patient RFH-2 are illustrated. Protease and functional Gag matrix (P17), capsid (P24), P2, Nucleocapsid (P7), P1 and P6 are shown with protease cleavage sites indicated. Changes are numbered according to the HXB2 consensus sequence. Changes detected at baseline ($RFH\text{-}2_B$) are shown above and changes emerging at different time points ($RFH\text{-}2_1$ to $RFH\text{-}2_3$) are presented below. Mutations in red in the protease represent major resistance mutations, in blue minor mutations and in black other polymorphisms. In Gag, mutations in red are those associated with PI-exposure and those in black other polymorphisms. In addition, Gag cleavage sites mutations are presented in solid boxes.

6.3.2 Resistance test vectors

A total of 20 RTVs were produced:

1) **8.9Pt2Baseline**: containing baseline PR and Gag genes from patient Pt-2.

2) **8.9Pt2VF**: containing virological failure PR and Gag genes from patient Pt-2.

3) **8.9Pt2VFΔT375A**: containing virological failure PR and a modified virological failure Gag gene. Modification of Gag consisted of reversion to WT of the T375A mutation by SDM, previous to cloning into P8.9NSX.

4) **8.9Pt20Baseline**: containing the baseline PR and Gag genes from patient Pt-20.

5) **8.9Pt20 Baseline +Gag**: containing the baseline PR and a modified baseline Gag gene from patient Pt-20. Modification of baseline Gag consisted of introduction of the Gag mutations emerging at virological failure by SDM, previous to cloning into P8.9NSX.

6) **8.9Pt20Baseline+Gag+PR**: containing modified baseline PR and Gag genes from patient Pt-20. Modification of PR and Gag consisted of introduction of PR and Gag mutations emerging at virological failure by SDM, previous to cloning into P8.9NSX.

7) **8.9Pt20Baseline+P1**: containing baseline PR and a modified baseline Gag genes from patient Pt-20. Modification of Gag consisted of introduction of the P1 mutations emerging at virological failure, previous to cloning into P8.9NSX.

8) **8.9Pt20Baseline+P17**: containing baseline PR and a modified baseline Gag gene from patient Pt-20. Modification of Gag consisted of introduction of the P17 mutations emerging at virological failure by SDM, previous to cloning into P8.9NSX.

9) **8.9Pt20Baseline+I437V**: containing the baseline PR and a modified Gag gene from patient Pt-20. Modification of Gag consisted of introduction of I437V mutation by SDM previous to cloning into P8.9NSX.

10) **8.9RFH-1PR$_B$Gag$_B$**: containing baseline PR and Gag genes from patient RFH-1.

11) **8.9RFH-1PR$_4$Gag$_4$**: containing the last time point PR and Gag from patient RFH-1.

12) **8.9RFH-1PR$_B$Gag$_4$**: containing baseline PR and last time point Gag from patient RFH-1.

13) **8.9RFH-1PR$_4$Gag$_B$**: containing last time point PR and baseline Gag from patient RFH-1.

14) **8.9RFH-1PR$_B$Gag$_4$ΔY132F**: containing baseline PR and a modified last time point Gag from patient RFH-1. Modification of Gag consisted of reversion to WT of the Y132F mutation by SDM, previous to cloning into P8.9NSX.

15) **8.9RFH-1PR$_B$GAG$_4$ΔCS**: containing baseline PR and a modified last time point Gag from patient RFH-1. Modification of Gag consisted of reversion to WT of the two CS mutations emerging over time (Y132F and L449F) by SDM, previous to cloning into P8.9NSX.

16) **8.9RFH-2PR$_B$Gag$_B$**: containing baseline PR and Gag from patient RFH-2.

17) **8.9RFH-2PR$_1$Gag$_1$**: containing PR and Gag from the first time point after baseline from patient RFH-2.

18) **8.9RFH-2PR$_2$Gag$_2$**: containing PR and Gag from the second time point after baseline for patient RFH-2.

19) **8.9RFH-2PR$_3$Gag$_3$**: containing PR and Gag from the third time point after baseline from patient RFH-2.

20) **8.9RFH-2PR3Gag3ΔA431V**: containing 3rd time point PR and a modified third time point Gag. Modification of Gag consisted of reversion to WT of the A431V mutation by SDM, previous to cloning into P8.9NSX.

6.3.3 Determination of IC_{50} for the wild-type reference virus (P8.9NSX)

The PI susceptibility profile for the wild-type reference virus P8.9NSX was determined. The mean IC_{50} for each drug was obtained by repeat testing (n =10). Differences between the IC_{50} from replicate assays were consistently below 2 fold and as a result we established a 2-fold change in IC_{50} as the technical cut-off for this study. The mean IC_{50}, standard deviation and 95% confidence interval calculated from 10 separate determinations in the presence of 6 different PIs are shown in table 6.1.

Virus and statistics	PI IC_{50}(nM)*					
P8.9NSX	**APV**	**ATV**	**DRV**	**IDV**	**LPV**	**SQV**
Mean IC50	1.86	6.17	1.57	6.18	1.26	3.35
SD	0.14	0.24	0.09	0.25	0.09	0.23
95% CI	(1.6-2.3)	(5.6-6.8)	(1.4-1.8)	(5.6-6.7)	(1.2-1.4)	(2.9-3.8)

Table 6.1 IC_{50} for the WT reference virus P8.9NSX.

The mean IC_{50}, standard deviation (SD) and 95% confidence interval (95% CI) derived from 10 independent determinations are shown for the wild type virus (P8.9NSX) employed as a reference in the present study.

* 10% FCS was employed throughout the experiments.

Abbreviations: amprenavir (APV), atazanavir (ATV), darunavir (DRV), indinavir (IDV), lopinavir (LPV) and saquinavir (SQV).

6.3.4 Effect of Gag mutations on susceptibility to protease inhibitors

We tested the susceptibility to ATV, DRV, IDV, LPV and SQV of all the resistance test vectors produced from patient Pt-2 (8.9Pt-2Baseline, 8.9Pt-2VF and 8.9Pt-2VFΔT375A) with the following results:

At baseline (8.9Pt2Baseline) the FC was 1.5 for ATV, 2 for DRV, 3 for IDV and SQV and 6 for LPV. At the time of virological failure (8.9Pt2VF), the FC significantly increased for all PI tested from 1.5 to 3-fold for ATV, from 2-fold to 9-fold for DRV, from 3 fold to 8 fold for IDV, from 6 fold to 13 fold for LPV and from 3 fold to 15 fold for SQV. Reversion to WT of the treatment emergent T375A mutation (8.9Pt20VFΔT375A) led to a decreased in the level of

resistance to all PIs respect to the virological failure sample and was comparable to the FC observed at baseline.

Figure 6.5 shows the drug susceptibility profile for the three RTVs from patient Pt-2.

DRUG	WT (mean IC_{50} nM)	8.9Pt-2Baseline		8.9Pt-2VF		8.9Pt-2ΔT375A	
		Mean IC50 nM	Mean FC ± SD	Mean IC50 nM	Mean FC ± SD	Mean IC50 nM	Mean FC ± SD
ATV	6	9	1.5 ± 0.15	18	**3 ± 0.08**	10	1.6 ± 0.11
DRV	1.6	3.4	2± 0.10	15	**9 ± 0.5**	3	2 ± 0.06
IDV	6	18	3± 0.15	46	**8 ± 0.29**	15	2.5 ± 0.14
LPV	1.2	7.2	6± 0.361	16	**13 ± 0.34**	7	6 ± 0.17
SQV	3	9	3± 0.32	46	**15 ± 0.32**	10	3 ± 0.2

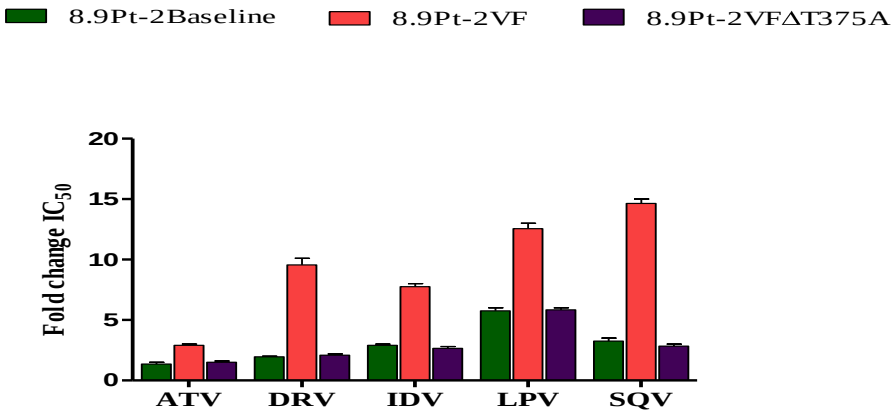


Figure 6.5 PI susceptibility profile for patient Pt-2.

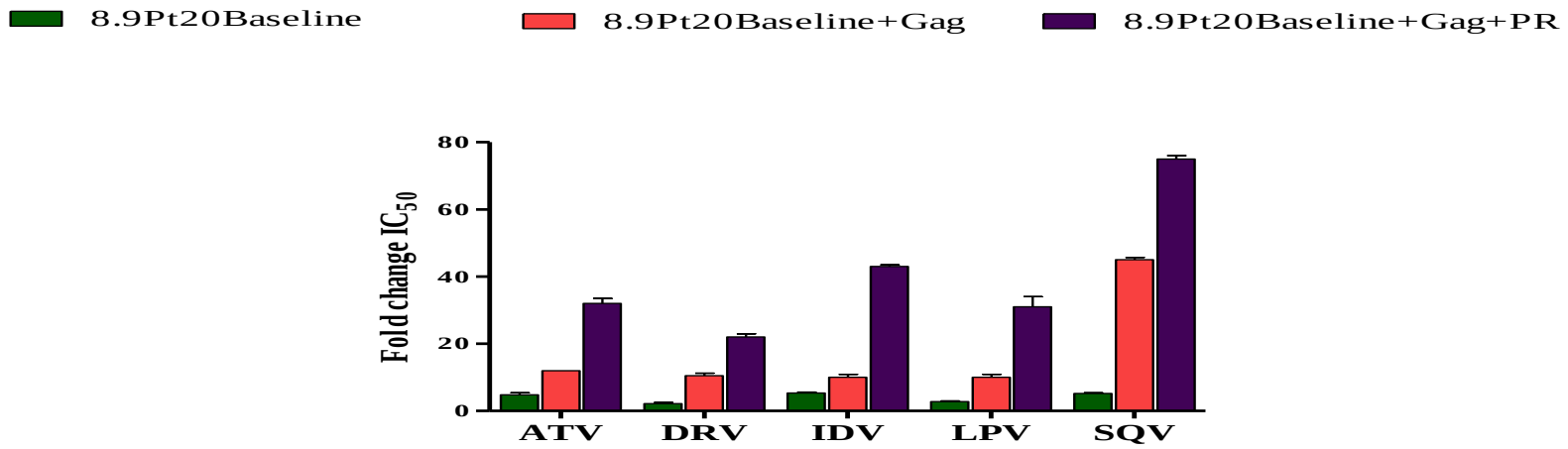
The PI susceptibility profiles at baseline (8.9Pt2Baseline), at VF (8.9Pt2VF) and after reversion to wild-type of the treatment emergent Gag cleavage site mutation T375A (8.9Pt2VFΔT375A) are shown. Susceptibility is shown as change in IC_{50} compared to that of the wild-type virus; thus a FC of 1 indicates same susceptibility as the wild-type P8.9NSX. Error bars represent the standard error of the mean of three separate experiments. Numbers are the mean of three experiments ± SD. Bold numbers indicate that the different FC respect to the baseline was found statistically significant using one way ANOVA followed by Bonferroni's multiple comparison post-test.

We tested the susceptibility to ATV, DRV, IDV, LPV and SQV of all the resistance test vectors

Obtained from patient 20, these were: 8.9Pt20Baseline, 8.9Pt20Baseline+Gag, 8.9Pt20Baseline+Gag+PR, 8.9Baseline+P1, 8.9Baseline+P17 and 8.9Baseline+I437V.

At baseline (8.9Pt20Baseline), according to the Virco biological and clinical cut-offs, the baseline sample showed complete susceptibility to DRV (FC: 2.0) and LPV (FC: 2.8) and reduced activity for ATV (FC: 4.8), IDV (FC: 5.3) and SQV (FC: 5.2). Introduction of the treatment emergent Gag mutations into the baseline Gag by SDM (8.9Pt20Baseline+Gag) led to an significant increase in the level of resistance for all the PIs tested, from 4.8-fold to 12 fold for ATV, from 2 fold to 10.5 fold for DRV, from 5.3 fold to 10 fold for IDV, from 2.8 to 10 fold for LPV and from 5.2 fold to 15 fold for SQV. Introduction of both treatment emergent Gag and protease mutations into the baseline (8/9Pt20Baseline+Gag+PR) led to a more substantial increase in the level of resistance for all the PIs tested, from 4.8 fold to 32 fold for ATV, from 2 fold to 22 fold for DRV, from 5.3 fold to 43 fold for IDV, from 2.8 fold to 31 fold for DRV and from 5.2 fold to 75 fold for ATV.

Figure 6.6 shows the drug susceptibility profile of 8.9Pt20Baseline, 8.9Pt20Baseline+Gag, 8.9Pt20Baseline+Gag+PR.



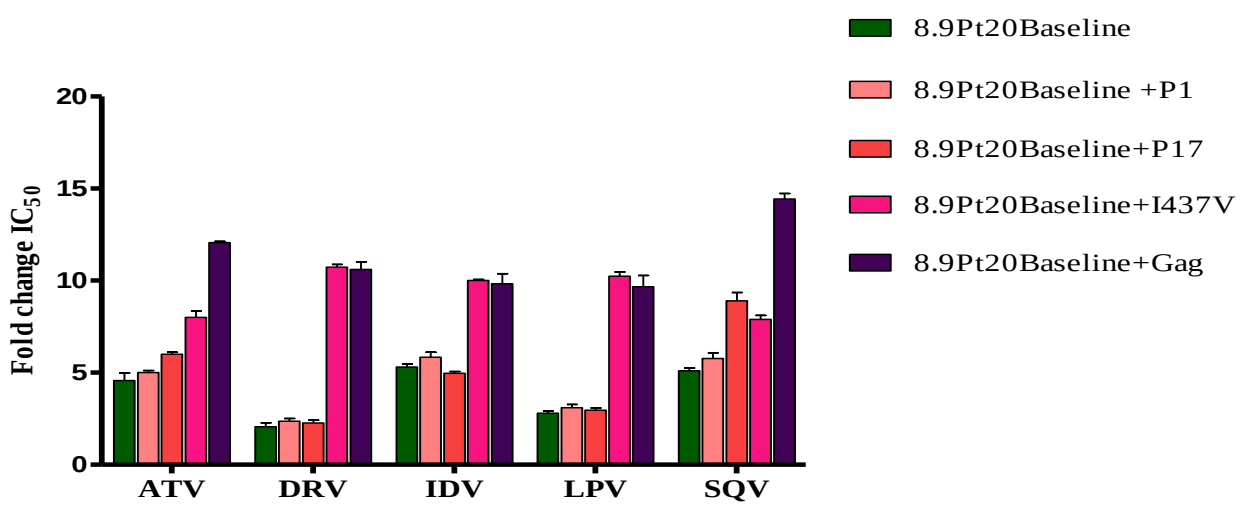
DRUG	WT	8.9Pt20Baseline		8.9Pt20Baseline +Gag		8.9Pt20Baseline+Gag+PR	
	Mean IC_{50} nM	Mean IC_{50} nm	Mean FC ± SD	Mean IC_{50} nm	Mean FC ± SD	Mean IC_{50} nm	Mean FC ± SD
ATV	6	28.8	4.8 ± 0.85	72	**12 ± 0.9**	192	**32 ± 2**
DRV	1.6	3.2	2 ± 0.5	16.8	**10.5 ± 1.0**	35.2	**22 ± 1.4**
IDV	6	32	5.3 ± 0.4	60	**10 ± 1.3**	258	**43± 0.7**
LPV	1.2	3.4	2.8 ± 0.28	12	**10 ± 1.3**	37.2	**31 ± 4.2**
SQV	3	15.6	5.2 ± 0.28	45	**15 ± 1.0**	225	**75 ± 1.4**

Figure 6.6 PI susceptibility profile of RTVs: 8.9Pt20Baseline, 8.9Pt20Baseline+Gag and 8.9Pt20Baseline+Gag+PR. The figure shows the PI susceptibility of the baseline sample, after introduction of emergent Gag mutations (P17-MA: V82I, A115I and A120S; P1: Y441H and G443E) and after introduction of emergent Gag and PR (T74P and I84V) mutations. Susceptibility is shown as the change in IC_{50} compared to that of the wild-type P8.9NSX; thus a FC of 1 indicates same susceptibility as the wild-type. Error bars represent the standard error of the mean from three separate experiments. Bold numbers indicate that the different FC respect to the baseline was found statistically significant using one way ANOVA followed by Bonferroni's multiple comparison post-test.

The susceptibility testing of the resistance test vectors containing modified Gag genes consisting of step-wise introduction of Gag mutations (8.9Pt20Baseline+P1, 8.9Pt20Baseline+MA and 8.9Pt20Baseline+I437V) showed the following:

Introduction of the two treatment emergent P1 mutations (Y441, G443E) did not change significantly the susceptibility profile of any of the PI tested compared to the baseline sample. However, introduction of the 3 treatment emergent P17 mutations (V82I, A115I and A120S) led to a significant increase of resistance to SQV from 5.2 fold to 9 fold compared with the baseline and from 2.8 fold to 6.0 fold for ATV sample but had no discernible effect on susceptibility to the other PIs. Finally, introduction of the I437V mutations led to a significant increase of resistance to all PIs. From 4.8 fold to 8 fold for ATV, from 2 to 10.6 fold for DRV, from 5.3 fold to 10 fold for IDV, from 2.8 fold to 10 fold for LPV and from 5.2 to 8 fold to SQV.

Figure 6.7 shows the comparison of the drug susceptibility profile of 8.9Pt20Baseline, 8.9Pt20Baseline+P1, 8.9Pt20Baseline+MA, 8.9Pt20Baseline+I437V and 8.9Pt20Baseline+Gag.



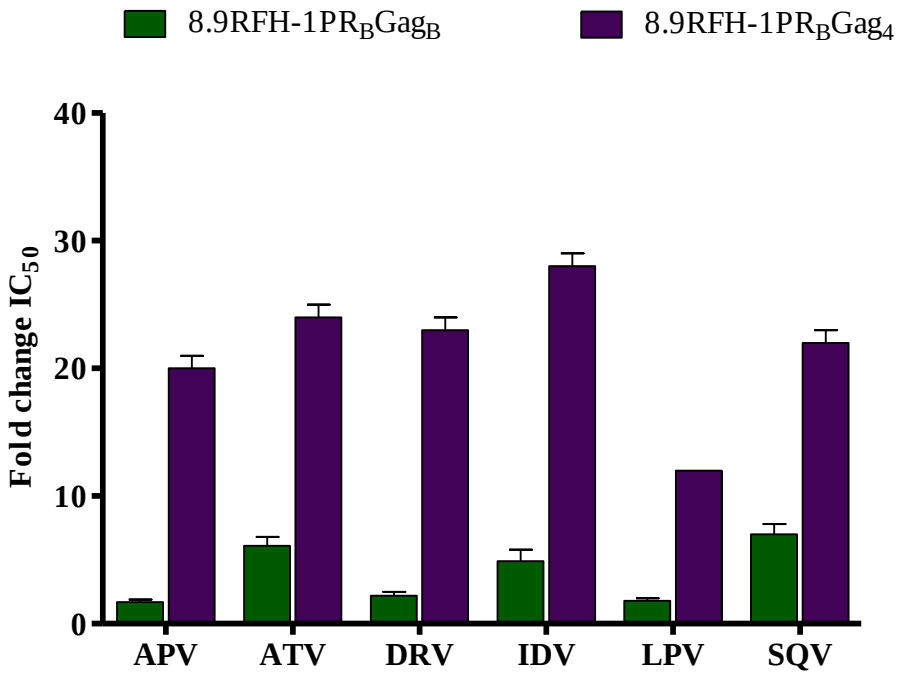
DRUG	WT (mean IC_{50} nM)	P8.9Baseline		P8.9Baseline+P1		P8.9Baseline+P17		P8.9Baseline+I437V		P8.9Baseline+Gag	
		Mean IC50 nM	Mean FC ± SD	Mean IC50 nM	Mean FC ± SD	Mean IC50 nM	Mean FC ± SD	Mean IC50 nM	Mean FC ± SD	Mean IC50 nM	Mean FC ± SD
ATV	6	28.8	4.8 ± 0.7	30	5 ± 0.2	36	**6 ± 0.2**	48	**8 ± 0.6**	72	**12 ± 0.1**
DRV	1.6	3.2	2 ± 0.4	3.7	2.3 ± 0.2	3.4	2.1 ± 0.2	17	**10.6 ± 0.2**	16.8	**10.5 ± 0.7**
IDV	6	32	5.3 ± 0.3	34.2	5.7 ± 0.5	29.7	4.9 ± 0.1	60	**10 ± 0.1**	60	**10 ± 0.9**
LPV	1.2	3.4	2.8 ± 0.2	3.5	2.9 ± 0.3	3.6	3 ± 0.2	12	**10 ± 0.4**	12	**10 ± 1.1**
SQV	3	15.6	5.2 ± 0.3	18.1	6 ± 0.5	27	**9 ± 0.8**	24	**8 ± 0.4**	45	**15 ± 0.5**

Figure 6.7 Individual effects of treatment emergent Gag mutations on PI susceptibility. Susceptibility profile of the baseline sample (8.9Pt20Baseline), after introduction of P1 mutations (8.9Pt20Baseline+P1), after introduction of matrix mutations (8.9Pt20 Baseline+P17), after introduction of the I437V mutation (8.9Pt20Baseline+I437V) and after introduction of all Gag mutations simultaneously (8.9Pt20Baseline+Gag). Susceptibility is shown as the change in IC_{50} compared to that of the wild-type P8.9NSX; thus a FC of 1 indicates the same susceptibility as the wild-type. Error bars represent the standard error of the mean from three separate experiments. Bold numbers indicate that the different FC respect to the baseline was found statistically significant using one way ANOVA followed by Bonferroni's multiple comparison post-test.

Next, it was compared the susceptibility to APV, ATV, DRV, IDV, LPV and SQV of patient RFH-1's resistance test vectors including 8.9RFH-1PR_BGag_B, 8.9RFH-1PR_BGag_4, 8.9RFH-1PR_BGag_4ΔY132F and 8.9RFH-1PR_BGag_4ΔCS. The drug susceptibility results were as follow:

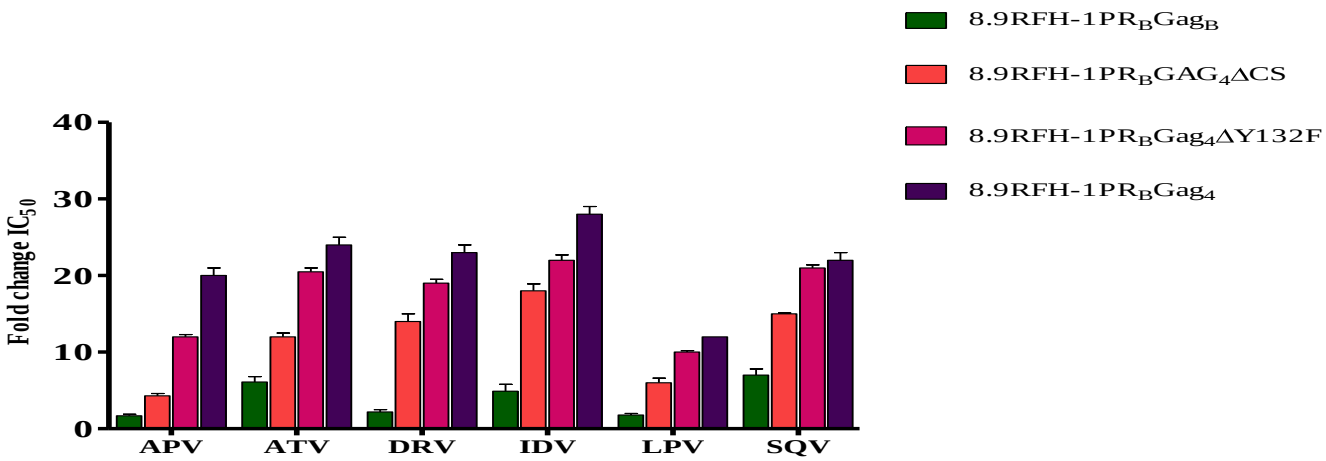
At baseline (8.9RFH-1PR_BGag_B) it was found around 2-fold resistance to APV, DRV and LPV, a 3-fold resistance to ATV, a 5-fold resistance to IDV and a 7-fold resistance to SQV. When the Gag baseline was replaced by the last time point Gag in the baseline sample (8.9RFH-1PR_BGag_4), a significant increase in resistance relative to the baseline sample was found. Thus, the FC for APV increased from 2.0 to 20, from 3 to 24 for ATV, from 2.2 to 23 for DRV, from 5 to 28 fold for IDV, from 2 to 12 for LPV and from 7 to 22 for SQV. The reversion to WT of the Y132F mutation in Gag_4 (8.9RFH-1PR_BGag_4ΔY132F) showed a decrease in the level of resistance for all PIs except SQV/r relative to the resistance test vectors containing intact Gag_4 but the level of resistance still remained significantly above that seen with baseline Gag (Gag_B). Thus, the FC for 8.9RFH-1PR_BGag_4ΔY132F changed to 12-fold for APV, 19-fold for ATV and DRV, 22-fold for IDV, 10-fold for LPV and 21-fold for SQV. Finally, both Y132F and L449F mutations were reverted to wild-type in Gag_4 (9RFH-1PR_BGag_4ΔCS) to evaluate the effect that treatment emergent Gag CS mutations had on PI susceptibility. A further decrease in the level of resistance relative to that seen when only Y132F was reverted to wild-type was found. However, the FC still remained significantly higher than those seen with Gag_B for all PIs except for APV. Thus, FC with RFH-1PR_BGag_4ΔCS was 4.3-fold for APV and IDV, 10-fold for ATV and DRV, 6-fold for LPV and 14-fold for SQV.

The contribution of Gag_4 to PI susceptibility is depicted in figure 6.8 and figure 6.9.



DRUG	WT (mean IC_{50} nM)	8.9RFH-1PR_BGag_B		8.9RFH-1PR_BGag_4	
		Mean IC50 nM	Mean FC ± SD	Mean IC50 nM	Mean FC ± SD
APV	1.8	3.1	2 ± 0.3	36	20 ± 1
ATV	6	18	3 ± 1.0	144	24 ± 1.4
DRV	1.6	3.5	2.2 ± 0.4	37	23 ± 1.3
IDV	6	29.4	5 ± 1.3	168	28 ± 0
LPV	1.2	2.2	2 ± 0.3	15	12 ± 1.3
SQV	3	13.5	7 ± 1.1	66	22 ± 1.3

Figure 6.8 Contribution of emergent Gag mutations to PI resistance. The PI susceptibility of the RTVs containing baseline protease and last time point Gag (8.9RFH-1PR_BGag_4) is shown. Susceptibility is shown as change in IC_{50} compared to that of the WT virus; thus a FC of 1 indicates same susceptibility as the WT P8.9NSX. Error bars represent the standard error of the mean of two separate experiments.



DRUG	WT (mean IC_{50} nM)	8.9RFH-1PR_BGag_B		8.9RFH-1$PR_BGag_4\Delta CS$		8.9RFH-1$PR_BGag_4\Delta Y132F$		8.9RFH-1PR_BGag_4	
		Mean IC50 nM	Mean FC ± SD	Mean IC50 nM	Mean FC ± SD	Mean IC50 nM	Mean FC ± SD	Mean IC50 nM	Mean FC ± SD
APV	1.8	3.1	2 ± 0.3	15	4.3 ± 0.4	37	**12 ± 0.4**	36	**20 ± 1.4**
ATV	6	18	3 ± 1	60	**10 ± 0.7**	72	**19 ± 0.7**	144	**24 ± 1.6**
DRV	1.6	3.5	3 ± 0.4	16	**10 ± 1.4**	22	**19 ± 0.7**	37	**23 ± 1.0**
IDV	6	29.4	5 ± 1.3	48	**8 ± 1.2**	60	**22 ± 1.0**	168	**28 ± 1.2**
LPV	1.2	2.2	2 ± 0.3	7	**6 ± 0.8**	16	**10 ± 0.3**	15	**12 ± 0**
SQV	3	13.5	7 ± 1.1	42	**14 ± 0.1**	51	**21 ±0.6**	66	**22 ± 1.0**

Figure 6.9 Contribution of Gag CSMs to PI resistance. The PI susceptibility of the RTVs containing baseline protease and last time point Gag after sequential reversion to WT of the emergent CSMs (8.9RFH-1$PR_BGag_4\Delta Y132F$ and 8.9RFH-1$PR_BGag_4\Delta CS$) compared to the baseline and last time point (8.9RFH-1PR_BGag_B and 8.9RFH-1PR_4Gag_4) is shown. Susceptibility is shown as change in IC_{50} compared to that of the WT virus; thus a FC of 1 indicates same susceptibility as the WT P8.9NSX. Error bars represent the standard error of two different experiments. Bold numbers indicate that the different FC respect to the baseline was found statistically significant using one way ANOVA followed by Bonferroni's multiple comparison post-test.

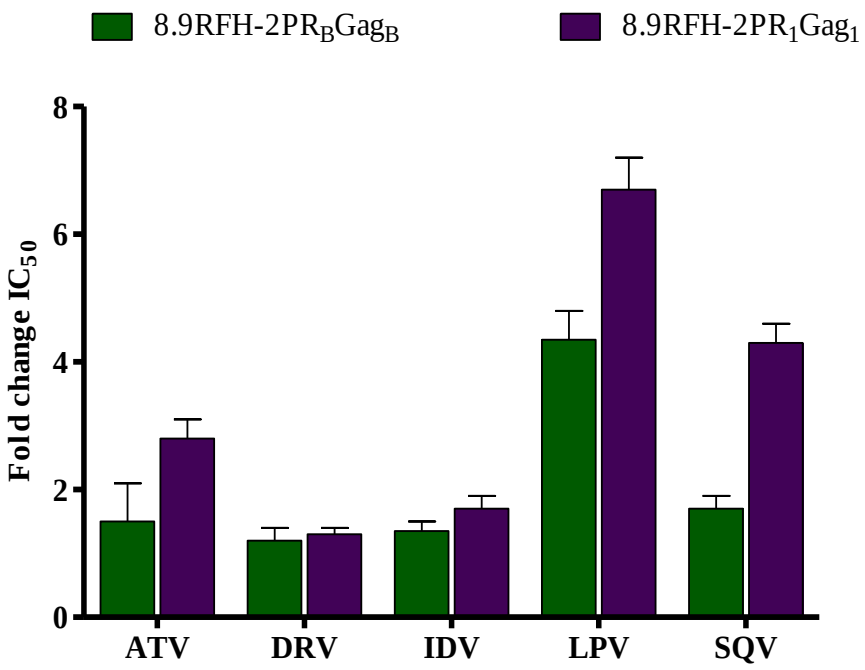
Finally, it was evaluated the susceptibility to ATV, DRV, IDV, LPV and SQV of the resistance test vectors obtained by cloning the PR and Gag sequences from patient RFH-2. These are 8.9RFH-2$PR_B$$Gag_B$, 8.9RFH-2$PR_1$$Gag_1$, 8.9RFH-2$PR_3$$Gag_3$ and 8.9RFH-2$PR_3$$Gag_3$ΔA431V.

The results of the PI susceptibility for the construct tested were as follow:

At baseline (8.9RFH-2$PR_B$$Gag_B$) a 4-fold resistance to LPV and below 2 fold change for all other PIs was observed. At the first time point after the baseline (8.9RFH-2$PR_1$$Gag_1$), it was found that the level of resistance to ATV increased from 1 to 2.8 fold, to LPV from 4 to 6.7 fold and for SQV from 1.6 to 4 fold and no change was observed for DRV and IDV. The FC difference was only significant for SQV.

When the last time point (8.9RFH-2$PR_3$$Gag_3$) was analysed, it was seen that the level of resistance further increased to a 12 fold change for LPV, 4 FC for DRV and 2 FC for IDV but hypersusceptibility for ATV and SQV (FC: 0.28 and 0.32, respectively) was detected. Applying the one-way ANOVA test followed by Bonferroni's multiple comparison tests showed that in this case, the difference FC respect to the baseline was significant for LPV and DRV but not for other PIs. Reversion to WT of the A431V mutation in the above construct (8.9RFH-2$PR_3$$Gag_3$ΔA431V) led to a modest decrease in the LPV FC from 12 to 10 fold and no change in the other PIs. The difference fold change respect to the baseline still remained significance for LPV.

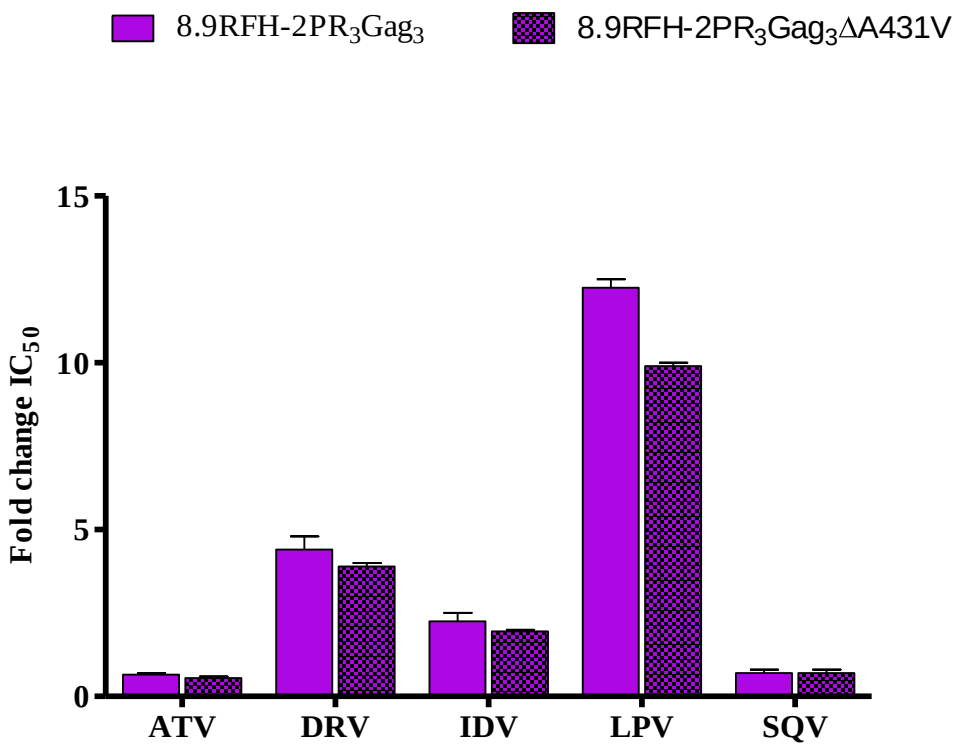
PI susceptibility profiles for the RTVs obtained from patient RFH-2 are shown in figures 6.10 and 6.11.



DRUG	WT (mean IC_{50} nM)	8.9RFH-2$PR_B$$Gag_B$		8.9RFH-2$PR_1$$Gag_1$	
		Mean IC50 nM	Mean FC ± SD	Mean IC50 nM	Mean FC ± SD
ATV	6	7	1 ± 0.8	17	2.8 ± 0.4
DRV	1.6	2	1.2 ± 0.3	2.1	1.3 ± 0.1
IDV	6	8	1.3 ± 0.2	10	1.7 ± 0.3
LPV	1.2	5	4 ± 0.6	8	6.7 ± 0.7
SQV	3	5	1.7 ± 0.3	13	**4.3 ± 0.4**

Figure 6.10 PI susceptibility profiles at baseline and 1st follow-up time for patient RFH-2.

The PI susceptibility of the RTVs containing baseline sample and 1st time point after the baseline is shown. Susceptibility is shown as change in IC_{50} compared to that of the wild-type virus; thus a FC of 1 indicates same susceptibility as the wild-type P8.9NSX. Error bars represent the standard error of two different experiments. Bold numbers indicate that the different FC respect to the baseline was found statistically significant using one way ANOVA followed by Bonferroni's multiple comparison post-test.



DRUG	WT (mean IC_{50} nM)	RFH-2_3		RFH-2_3ΔA431V	
		Mean IC50 nM	Mean FC ± SD	Mean IC50 nM	Mean FC ± SD
ATV	6	3	0.4 ± 0.1	3	0.28 ± 0.1
DRV	1.6	6	4 ± 0.5	6	4 ± 0.1
IDV	6	14	2 ± 0.1	12	2 ± 0.1
LPV	1.2	15	12.5 ± 0.7	12	10 ± 0.1
SQV	3	2	0.2 ± 0.6	2	0.32 ± 0.1

Figure 6.11 Contribution of emergent Gag CSMs to PI resistance. The PI susceptibility of the RTVs containing the last time point Gag after reversion to WT of the emergent A431V mutation (8.9RFH-2PR_3Gag_3ΔA431V) is shown. Susceptibility is shown as change in IC_{50} compared to that of the WT virus; thus a FC of 1 indicates same susceptibility as the WT P8.9NSX. Error bars represent the standard error of the mean of two separate experiments. Bold numbers indicate that the different FC respect to the baseline was found statistically significant using one way ANOVA followed by Bonferroni's multiple comparison post-test. Virco Biological cut off (BCO), lower and upper clinical cut-offs (CCO1 and CCO2) are shown.

6.3.5 Effect of Gag mutations on replicative capacity

The replicative capacity (RC) of the 20 resistance test vectors generated was determined and compared to that of the wild type reference 8.9PNSX virus. The results were as follows:

- Comparison of the RC of the resistance test vectors obtained from patient Pt-2 demonstrated that the baseline virus 8.9Pt-2Baseline showed 80% ± 1.4 RC relative to wild-type virus. At the time of virological failure, the virus replicated more efficiently than the WT virus (RC 146% ± 5.7) ($P < 0.0004$). However, if the T375A mutation was reverted to WT in the latter virus (8.9Pt2VFΔT375A), the RC returned to levels observed at baseline (from 146% ± 5.7 to 73% ± 7.8). Results are shown in the figure below (Figure 6. 12).

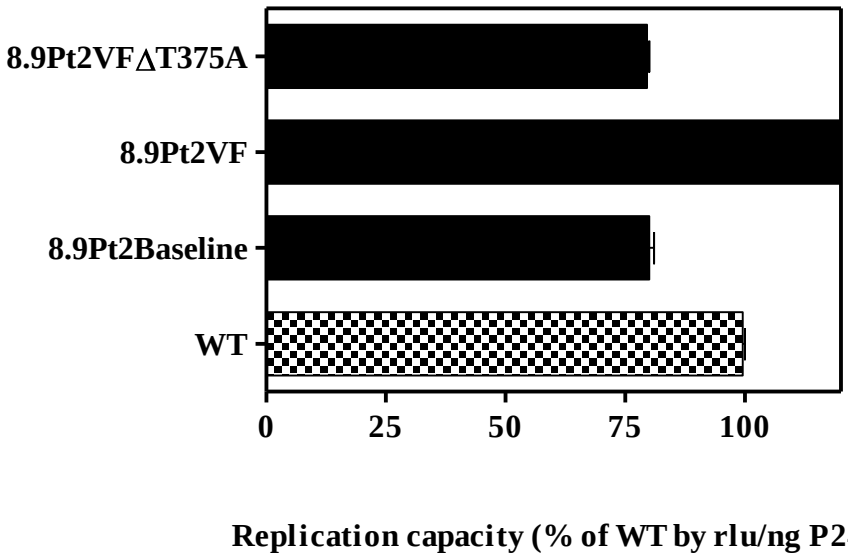


Figure 6.12 Replication capacity of RTVs from patient Pt-2.Recombinant resistance test vectors were prepared containing patient's derived PR and Gag genes: 8.9Pt2Baseline contained baseline PR and Gag, 8.9Pt2VF contained virological failure PR and Gag and 8.9Pt2VFΔT375A contained virological failure PR and a modified virological failure Gag in which the T375A mutation have been reverted to WT by SDM. RTVs were produced encoding luciferase and titrated. Luciferase activity was determined with SteadyGlo and a Glomax luminometer (both Promega); a mean was obtained using at least four values within the linear range and is expressed as the luciferase activity of the sample relative to the 8.9PNSX reference virus (relative light units, rlu). Virus titres were previously normalized for the amount of P24 protein produced in supernatants to correct for transfection efficiency. Results are shown as a percentage RC relative to the wild-type control virus.

- Comparison of the RC of the RTVs obtained from patient Pt-20 showed that the baseline virus (8.9Pt20Baseline) replicated around 50% less efficiently than the WT reference virus RC: (56% ± 2.82, P < 0.0001). Introduction of the treatment emergent PR mutations (8.9Pt20Baseline+PR) led to a significant decreased in viral RC from 56% ±2.82 to 32.5% ± 3.54). However, introduction of treatment emergent Gag mutations (8.9Pt20Baseline+Gag) rescued the RC from 32.5 % ± 3.54 to 69% ± 1.4 (P < 0.0001) relative to the wild-type reference virus. Step-wise introduction of the Gag emergent mutations demonstrated that introduction of the MA-P17 mutations led to a recovery in RC comparable to that obtained when all Gag mutations were introduced (69% ± 1.4 *vs.* 68% ± 0). However, when either P1 mutations or the I437V mutation were introduced, no significance changed on RC was found: 32.5 ± 3.54 *vs.* 33.5 ± 2.12 and 32.5 ± 3.54 *vs.* 38 ± 2.83, respectively. Results are shown in figure 6.13.

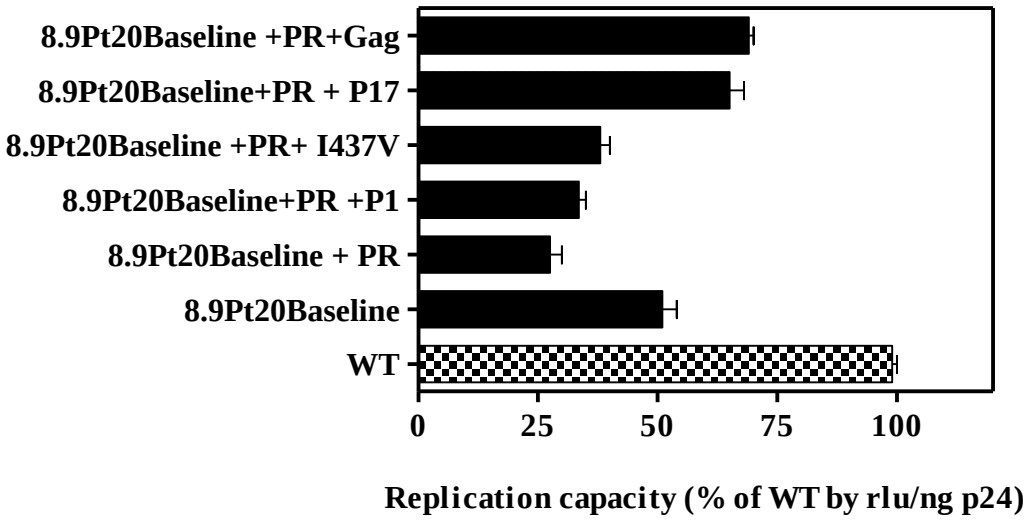


Figure 6.13 Replication capacity of RTVs from patient Pt-20.

Recombinant resistance test vectors were prepared containing patient's derived protease and Gag genes: 8.9Pt20Baseline contained baseline PR and Gag, 8.9Pt20Baseline+PR contained baseline Gag and mutated PR in which all treatment emergent PR mutations (I84V + T74P) were introduced by SDM. 8.9Pt20Baseline+PR+P1 contained the mutated PR and mutated Gag in which all treatment emergent P1 mutations were introduced by SDM. 8.9Pt20Baseline+PR+ I437V contained the mutated PR and a mutated Gag in which the I437V mutation was introduced by SDM. 8.9Baseline+PR+P17 contained baseline PR and a mutated Gag in which treatment emergent P17 mutations were introduced by SDM and finally 8.9Pt20Baseline+PR+Gag contained the mutated PR and a mutated Gag in which all treatment emergent Gag mutations were introduced simultaneously by SDM. RTVs were produced encoding luciferase and titrated. Luciferase activity was determined with SteadyGlo and a Glomax luminometer (both Promega); a mean was obtained using at least four values within the linear range and is expressed as the luciferase activity of the sample relative to the 8.9PNSX reference virus (relative light units, rlu). Virus titres were previously normalized for the amount of P24 protein produced in supernatants to correct for transfection efficiency. Results are shown as a percentage RC relative to the wild-type control virus.

- Comparison of the RC of the RTVs obtained from patient RFH-1 showed that the baseline virus (8.9RFH-1$P_{RB}Ga_{gB}$) had a RC comparable to that of the WT virus (98.5 ± 2.12 *vs.* 99 ± 0.71). At the last time point tested (8.9RFH-1$P_{R4}Ga_{g4}$), the virus showed an increase in RC (131.5% ± 9.2, $P < 0.0001$) respect to the wild-type virus. In order to assess the role that Gag evolution played in this increased in RC, the Gag from the two time points were exchanged. It was found that the virus containing baseline PR and last time point Gag (8.9RFH-1$P_{RB}Ga_{g4}$) had a RC relative to WT virus of 159.5% ± 10.6. This difference respect to the baseline virus reached statistical significance ($P < 0.0001$). By contrast, the virus containing last time point PR and baseline Gag (8.9RFH-1PR_4Gag_B) decreased significantly its RC respect to the baseline virus 98.5 ± 2.12 vs. 10% ± 2.83 ($P< 0.0001$). Furthermore, the effect that treatment emergent Gag CS mutations had on RC was assessed by step-wise reversion to WT of the two emergent mutations Y132F and L449F in the last time point Gag. It was observed that reversion of Y132F (8.9RFH-1$PR_BGa_{g4}\Delta$Y132F) led to a non-significance decrease in RC from 159% ±1 0.6 to 150% ± 2.8 and further reversion of L449F (8.99RFH-1$PR_BGa_{g4}\Delta$CS) led to a significant decrease from 159% ± 10.6 to 129 ± 3.0% ($P < 0.0001$). Results are shown in figure 6.14.

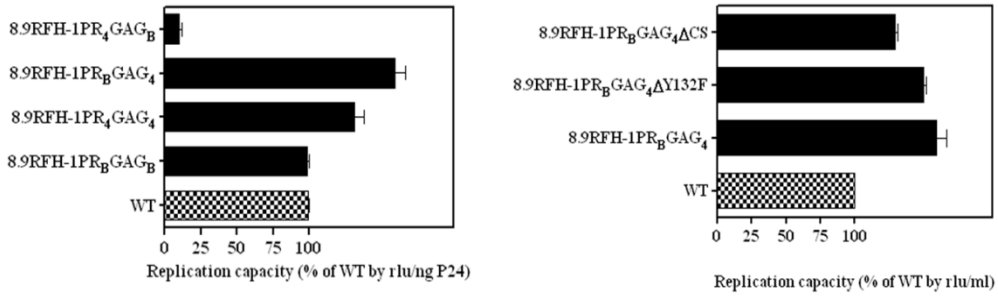


Figure 6.14 Replication capacity of RTVs from patient RFH-1.

Recombinant resistance test vectors were prepared containing patient's derived PR and Gag genes: 8.9RFH-1PR_BGag_B contained baseline PR and Gag, 8.9RFH-1PR_4Gag_4 contained last time point PR and Gag, 8.9RFH-1PR_BGag_4 contained baseline protease and last time point Gag, 8.9RFH-1PR_4Gag_B contained last time point PR and baseline Gag, 8.9RFH-1PR_BGag_4ΔY132F contained baseline PR and a modified last time point Gag in which the Y132F mutation have been reverted to WT by SDM and 8.9RFH-1PR_BGag_4ΔCS contained baseline PR and a modified last time point Gag in which both Y132F and L449F mutations have been reverted to WT by SDM. Luciferase activity was determined with SteadyGlo and a Glomax luminometer (both Promega); a mean was obtained using at least four values within the linear range and is expressed as the luciferase activity of the sample relative to the 8.9PNSX reference virus (relative light units, rlu). Virus titres were previously normalized for the amount of P24 protein produced in supernatants to correct for transfection efficiency. Results are shown as a percentage RC relative to the wild-type control virus.

- Comparison of the RC of the RTVs obtained from patient RFH-2 showed that the baseline virus (8.9RFH-2PR_BGag_B) had a 97% ± 0.71 RC respect to the WT virus. At the first time point after the baseline (8.9RFH-2PR_1Gag_1) the RC increased to130% ± 1.41. The difference in RC between baseline and first time point reached statistical significance (P, 0.0001). At the second time point after the baseline (8.9RFH-2PR_2Gag_2) the replicative capacity was significantly reduced from 97% ± 0.71 to 24 ± 4.24 (P < 0.0001). However, the RC was rescued at the third time point after the baseline (8.9RFH-2PR_3Gag_3) from 24 ± 4.24 to 74.5 ± 6.36, P < 0.0001. Reversion to WT of the treatment emergent Gag CS mutation A431V in the third time point (8.9RFH-2PR_3GAG_3ΔA431V) led to a significant reduction in RC from 74.5% ± 6.36 to 55% ± 4.24, P < 0.0001. Results are depicted in the figure below (Figure 6.15).

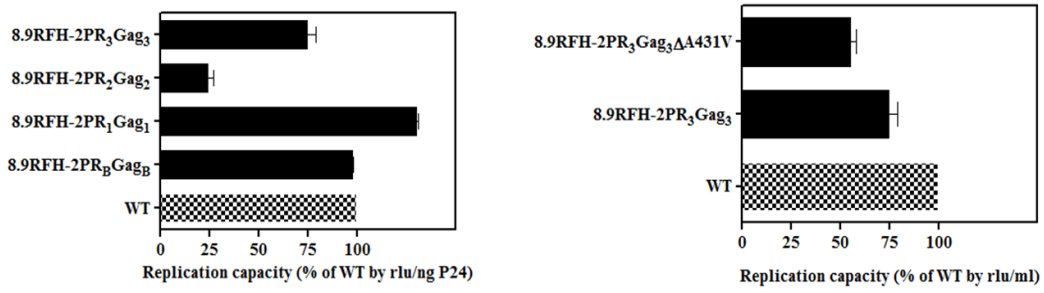


Figure 6.15 Replication capacity of RTVs from patient RFH-2.

Recombinant resistance test vectors were prepared containing patient's derived PR and Gag genes: 8.9RFH-2PR_BGag_B contained baseline PR and Gag, 8.9RFH-2PR_1Gag_1 contained first time point after baseline PR and Gag, 8.9RFH-2PR_2Gag_2 contained second time point after baseline PR and Gag, 8.9RFH-2PR_3Gag_3 contained third time point after baseline PR and Gag, 8.9RFH-2PR_3Gag_3ΔA431V contained third time point PR and a modified third time point Gag in which the A431V mutation have been reverted to WT by SDM. Luciferase activity was determined with SteadyGlo and a Glomax luminometer (both Promega); a mean was obtained using at least four values within the linear range and is expressed as the luciferase activity of the sample relative to the 8.9PNSX reference virus (relative light units, rlu). Virus titres were previously normalized for the amount of P24 protein produced in supernatants to correct for transfection efficiency. Results are shown as a percentage RC relative to the wild-type control virus.

6.3.6 Characterization of Y132F and T375A mutations

Y132F and T375A are CS mutations that were found to emerge during PI failure in our study, and which appeared to have effects on PI susceptibility and RC. In order to confirm the effect of these mutations we introduced each mutation in the P8.9NSX WT backbone by SDM and studied the phenotype of the resulting mutated viruses (P8.9NSX-T375A and P8.9NSX-Y132F).

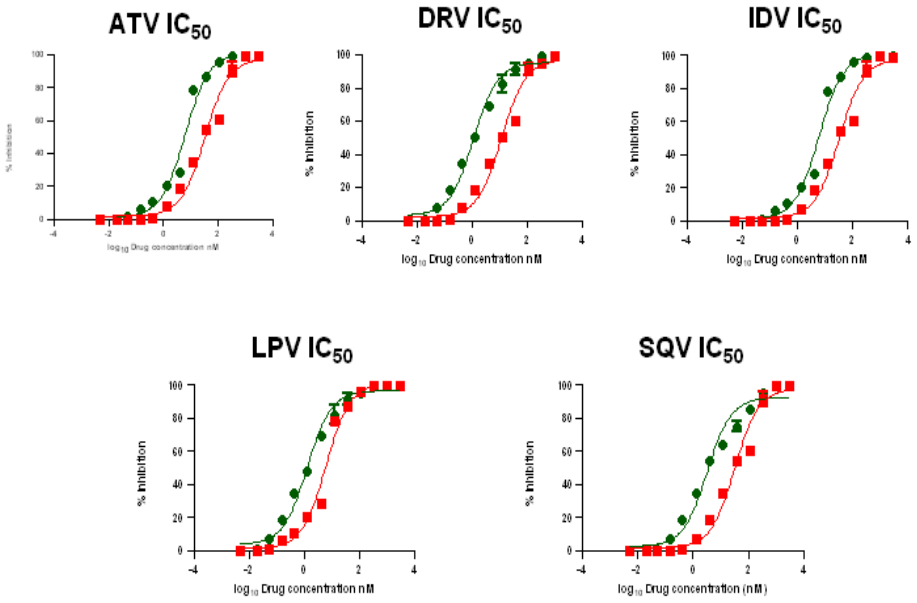
The results were as follows:

Introduction of the P2/NC-P7 CS mutation T375A into the wild-type backbone by SDM led to a significant increase in resistance to all PIs. The FC in IC_{50} of the mutated virus compared to the wild-type virus was: 5 for ATV, DRV and LPV and 10 for DRV and SQV. The mutated T375A virus showed an RC of 161.5% ± 4.95, which was significantly higher compared to the WT virus (P = 0.003).

Introduction of the MA-P17/CA-P24 CS mutation Y132F into the wild-type backbone by SDM also produced around a 10-fold increase in resistance to all PIs. However, it negatively affected RC as this declined to 61.5% ± 3.5, P = 0.009.

Drug susceptibility and RC profiles for the P8.9NSX-T375A are shown in figures 6.16 and 6.17, respectively. Drug susceptibility and RC profiles for P8.9NSX-Y32F are shown in figures 6.18 and 6.19, respectively.

P8.9NSX *vs.* P8.9NSX-T375A



DRUG	WTP8.9NSX (mean IC_{50} nM)	WTP8.9NSX-T375A	
		Mean IC50 nM	Mean FC ± SD
ATV	6.1	33	5.4 ± 0.3
DRV	1.14	11.6	10 ± 1.4
IDV	6.1	32.6	5 ± 0.7
LPV	1.2	6.2	5.2 ± 0.6
SQV	3.3	33	10 ± 0.7

Figure 6.16 Phenotypic PI susceptibility profile of SDM containing T375A.

Phenotypic drug susceptibility testing was performed with the wild-type reference virus P8.9NSX and a mutated wild-type virus (P8.9NSX-T375A) which contained the same HIV backbone as the reference virus except for the Gag CS mutation T375A which had been introduced by SDM. Green curves represent the reference virus and red curves represent the mutated virus. Inhibition curves shifted to the right (higher drug concentration) indicates reduced susceptibility. The fold change was calculated by comparing the IC_{50} of the reference virus to the IC_{50} for the mutated virus.

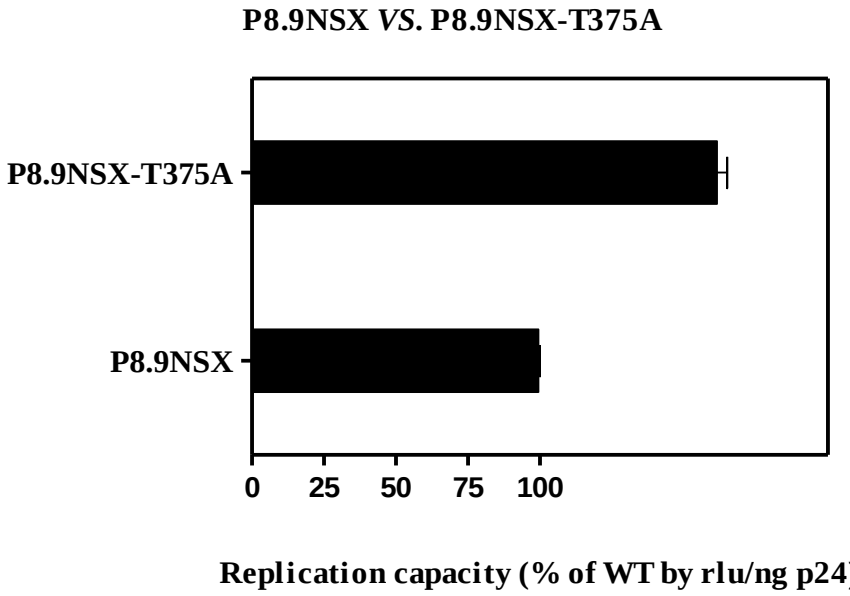
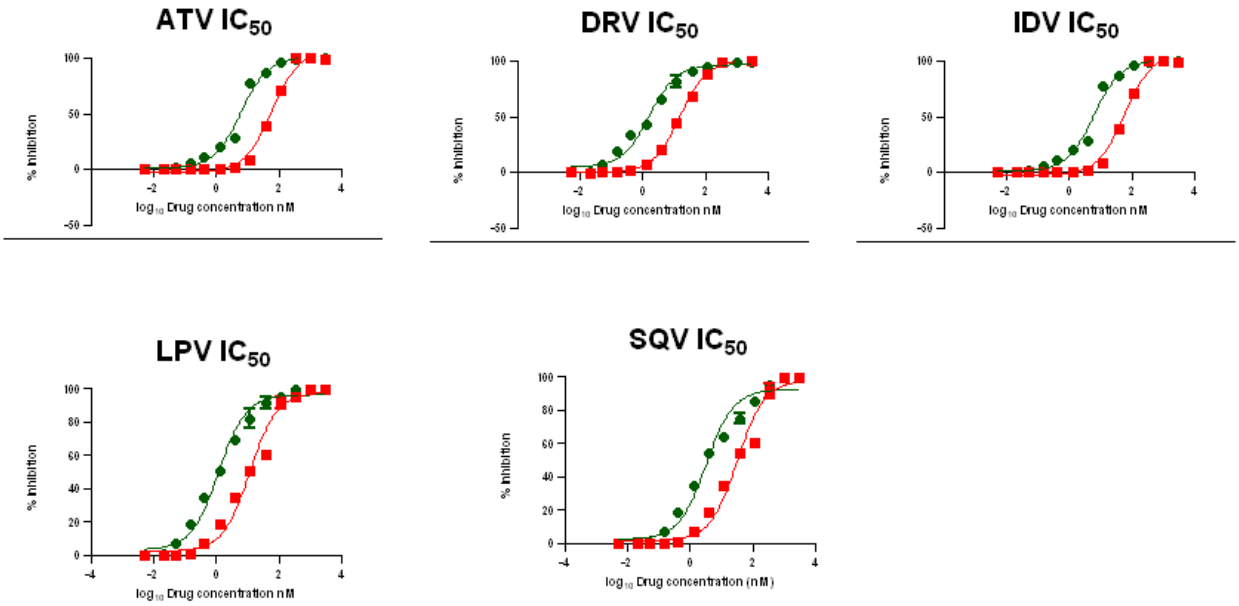


Figure 6.17 Replication capacity of SDM containing T375A.

Recombinant resistance test vectors were prepared containing either the WT reference virus (P8.9NSX) or a mutated virus (P8.9NSX-T375A) in which the T375A Gag CS mutation was introduced by SDM. Luciferase activity was determined with SteadyGlo and a Glomax luminometer (both Promega); a mean was obtained using at least four values within the linear range and is expressed as the luciferase activity of the sample relative to the 8.9PNSX reference virus (relative light units, rlu). Virus titres were previously normalized for the amount of P24 protein produced in supernatants to correct for transfection efficiency. Results are shown as a percentage RC relative to the wild-type virus control.

P8.9NSX *VS.* P8.9NSX-Y132F



DRUG	WTP8.9NSX (mean IC_{50} nM)	WTP8.9NSX-Y132F Mean IC50 nM	WTP8.9NSX-Y132F Mean FC ± SD
ATV	6.1	61	10 ± 0.7
DRV	1.63	16	9.8 ± 1.8
IDV	6.2	60.6	9.8 ± 0.2
LPV	1.2	12	10 ± 0.8
SQV	3.3	33	10 ± 1.6

Figure 6.18 Phenotypic PI susceptibility profile of SDM containing Y132F.

Phenotypic drug susceptibility testing was performed with the WT reference virus P8.9NSX and a mutated WT virus (P8.9NSX-Y132F) which contained the same HIV backbone as the reference virus except for the Gag CS mutation Y132F which had been introduced by SDM. Green curves represent the reference virus and red curves represent the mutated virus. Inhibition curves shifted to the right (higher drug concentration) indicates reduced susceptibility. The fold change was calculated by comparing the IC_{50} of the reference virus to the IC_{50} for the mutated virus..

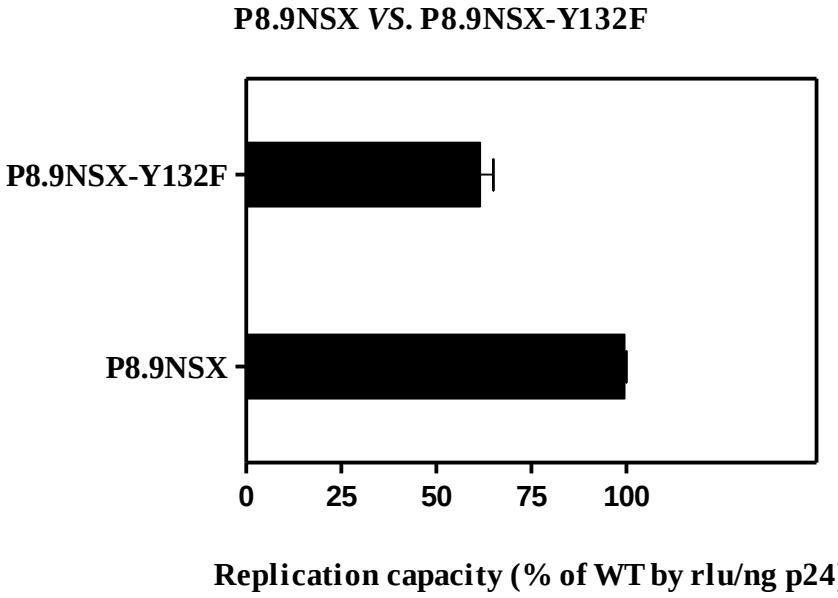


Figure 6.19 Replication capacity of SDM containing Y132F.

Recombinant resistance test vectors were prepared containing either the WT reference virus (P8.9NSX) or a mutated virus (P8.9NSX-Y132F) in which the Y132F Gag CS mutation was introduced by SDM. Luciferase activity was determined with SteadyGlo and a Glomax luminometer (both Promega); a mean was obtained using at least four values within the linear range and is expressed as the luciferase activity of the sample relative to the 8.9PNSX reference virus (relative light units, rlu). Virus titers were previously normalized for the amount of P24 protein produced in supernatants to correct for transfection efficiency. Results are shown as a percentage RC relative to the wild-type virus control.

6.4 Discussion

In the present chapter we have evaluated the impact of HIV-1 Gag mutations on PI susceptibility and viral replication capacity.

Overall our results showed that cleavage site mutations affect viral replicative capacity and resistance to protease inhibitors. While all CS mutations evaluated in this study increased to some extend the level of resistance to one or more PIs, the impact on RC differed among mutations and among viral constructs. These results suggest that CS mutations may be selected as truly resistance mutation independent of their role as compensatory mutations. In addition, we observed that non-CS mutations also increased the levels of PI resistance and in fact confer PI resistance on their own. By contrast with CS mutations, non-CS mutations had in all cases a positive effect on viral RC.

We employed a single cycle assay for the assessment of drug susceptibility and RC. The system incorporates a self-inactivating vector capable of a single round of infection and was initially developed by Petropoulos and colleagues (Petropoulos et al, 2000). Versions of this assay such as Phenosense are currently licensed for diagnostic purposes. The ability of single cycle assays to predict drug susceptibility "*in vivo"* has been debated. It has been stated that multiple cycle assays may be more appropriate as they more closely reproduced "*in vivo"* conditions. However, a few studies have shown comparable drug susceptibility results with both formats (Maguire et al, 2002). In addition, we should emphasize that our primarily objective was to address the effect of specific Gag mutations detected at the time of virological failure. In this respect, single cycle assays may be more suitable as the restriction to a single round makes virtually impossible the

selection of virus subpopulations that may not accurately reflect the initial viral population. The ability of our system to generate reproducible measurements of PI susceptibility was evaluated by repeatedly testing the drug susceptible reference strain P8.9NSX which derived from the NL4-3 molecular clone. The assay rendered highly reproducible measurements of PI susceptibility and the IC_{50} were within the range observed in published studies (Petropoulos et al, 2000).

We were interested in evaluating the contribution to PI resistance and RC of the emergent Gag mutations both at CSs and beyond. The I437V CS mutation located in P7/P1 emerged in patient Pt-20 together with one major (I84V) and one minor (T74P) protease resistance mutations as well as three non-CS mutations located in P17 and two in P1. Maximum levels of PI resistance for the recombinant construct containing patient derived Gag and PR sequences were observed when all PR, Gag CS and Gag non-CS mutations were incorporated. The stepwise introduction of Gag mutations demonstrated that the maximum effect on PI resistance was attributed to the I437V CS mutation. Thus the introduction of this mutation in the baseline sample led to a different in FC of between 2.8 and 8.6 depending on the PI and it was in all cases statistically significance. While P1 have no effect on resistance to PIs, a 3.8 FC difference for SQV and 1.2 for ATV were found after introduction of the P17 mutations in the baseline sample. These differences were statistically significant when applying the one-way ANOVA with the Bonferroni's correction for multiple comparisons. The fact that two of the three emergent P17 mutations were found to be associated with PI selective pressure in our study (chapter 4), that the patient was failing a SQV-based regimen and that the introduction of the mutations led to an increase resistance mainly to this drug suggest that the selection of the mutations was driven but

SQV-selective pressure and contribute to therapy failure. With regard to the RC experiments, we saw that emerging Gag mutations improved the replication of the PR-mutated-containing virus and in this case most of the replicative benefit was conferred by the P17 emerging mutations while I437V and P1 mutations had no discernible effect. This result suggest that the CS mutation, I437V, was selected as a truly resistance mutation rather than as a compensatory mutation. These findings are consistent with those obtained by other authors. For instance, Nijhuis and colleagues (Nijhuis et al, 2007) reported the selection of I437V mutation in P7/P1 without any preceding PR mutation during *in vitro* passages with an experimental PI (RO033-4649); after introduction of the I437V mutation in a reference strain they observed between 2-5 fold increase in IC_{50} to all PIs in a multiple cycle drug resistance assay and 2-3 fold increase in a single cycle assay. They demonstrated that incorporation of the mutation enhanced Gag polyprotein processing by the WT protease and propose this as a possible mechanism of PI resistance. Similarly, Dam and co-investigators (Dam et al, 2009) evaluated the effect of Gag on phenotypic PI resistance and observed the I437V in two patients infected with viruses containing several major protease resistance mutations. They observed that incorporation of the I437V mutation in the recombinant construct led to a 1.6 to 5 fold increase in resistance to all PIs in a single cycle assay while very little effect on resistance was attributed to other Gag regions (Dam et al 2009). Also consistent with our results, they reported that while reversion to WT of the I437V mutation clearly impacted the level of PI resistance, it did had very little effect on RC, strengthening the hypothesis that I473V has a role as a resistance mutation. By contrast, the RC benefit mainly lied on the non-CS mutations and in particular on those located in the P17 region. In agreement with our results, Parry and co-authors reported that the P17 protein was sufficient

to recover the otherwise compromised RC of a multi-PI resistant virus to levels observed in the wild-type virus (Parry et al, 2009).

We also observed the emergence of the T375A mutation at P2/P7 in patient Pt-2. Interestingly this patient had no major protease mutations either at baseline or emerging at VF. In fact the emergence of T375A along with T371N located in the spacer peptide P2 were the only changes observed at the time of virological failure in this patient. When we compared the level of PI resistance of the construct containing patient's PR and Gag at baseline and at virological failure, we observed a significance difference in FC of between 1.5 and 12 depending on the PI. In order to assess the contribution of T375A to this increased level of resistance, we reverted the mutation to WT in the virological failure sample and observed that the level of resistance returned to the levels found at baseline. We also evaluated the RC of the different viral constructs and observed that in this case the T375A mutation not only increased PI resistance but also the RC of the virus. Thus, at baseline the patient virus showed 80% RC compared to the wild-type reference virus and this was increased to 146% at the time of VF. Importantly, the increased in RC correlated with a one log increased in the patient's viral load. While we and others have reported a significantly higher prevalence of T375A mutation in PI-experienced than in PI naïve subjects (Malet et al, 2007) a role for this mutation in resistance to PIs had never been evaluated. Our results suggest that the selection of T375A mutation at the time of VF was not a random event, but was indeed driven by PI selective pressure. Due to the novelty of the T375A mutation, we employed SDM to introduce the mutation in the wild-type reference strain in order to address the independent impact of the mutation on PI resistance and viral RC. We found that the T375A mutation led to a 5 to 10 fold increase in the IC_{50} for all PIs and increased significantly the RC of

the wild-type reference strain (100% to 146%). These results further support a role of T375A mutation in PI resistance. The mechanism by which the T375A mutation exerts its effect on PI susceptibility and RC has not been addressed in this study. However, a role similar to the one documented for the well characterized A431V mutation can be postulated. A431V is located in the P7/P1 CS and it has been demonstrated to be selected in the presence of major protease resistance mutations. The A431V mutation increases the rate of processing of the CS compensating for the catalytic deficiency displayed by the mutated protease (Zhang et al, 1997 and Zennou et al, 1998). We propose that the T375A mutation may also lead to better substrates for protease thereby facilitating Gag processing. The processing of Gag is a coordinated process in which every site is cleaved in a specific order and a specific rate. T375A is located in P2/P7 CS, which in this study, and others, have found to be highly polymorphic (Malet et al, 2007). The extreme variation in this CS compared to others is probably related to the fact that this is the first site to be cleaved and its cleavage takes place rapidly. As a consequence, it is anticipated that variations in the rate of cleavage of this site may not be as crucial as those affecting other intermediate and rate limiting steps. Therefore, it is possible that the presence of the T375A mutation leads to an increase rate of cleavage of P2/P7 and as consequence to an overall improvement of Gag processing and viral replication. Maximum benefit of the mutation will be however achieved in the context of impaired protease activity, explaining its preferential selection in PI-exposed subjects. A similar mechanism has been also demonstrated for compensatory mutations in the PR, such as L63P which is a common polymorphism in pre-treatment isolates but is even more frequently selected during PI failure and it is known to improve the catalytic activity of the viral PR (Martinez-Picado et al, 1999).

Over a one year period on unsuccessful LPV/r therapy, patient RFH-2 developed the two major protease resistance mutations M46I and L76V, the CS mutation A431V in P7/P1 and several Gag non-CS mutations mainly located in the P17 protein. The emergence of P17 mutations preceded the appearance of both protease resistance mutations and the A431V mutation and was accompanied by a significance increase in viral replication capacity from 97% to 140% respect to the WT reference virus. In addition, an increased in the fold change for ATV (from 1 to 2.8 FC), LPV (from 4 to 6.7 FC) and SQV (from1.7 to 4.3 FC) was also seen, which only reached statistical significance for SQV. The appearance of the L76V resistance mutation led to a significant decrease in viral replication (from 140% to around 24% respect to the wild type reference virus) and finally the RC was rescued to 74% following the emergence of the M46I mutation in the PR and the A431V at P7/P1 CS. At this time the level of resistance to LPV also increased from 4 to 12 fold, while hypersusceptibility, defined as a fold change $\leq$0.4 (Clark et al, 2006), to ATV and SQV was also seen. Reversion to wild type of the A431V mutation led to a significance decrease in RC from 74% to 55% and slight decrease in the IC_{50} for LPV from 12 to 10 fold change, which was not statistically significance. The LPV/r resistance pathway consisting of emergence of L76V and M46I mutations in the protease was first time observed in patients failing LPV/r monotherapy in the MONARK trial (Delaugerre et al, 2009) and it was later confirmed by Nijhuis and colleagues. Consistent with our results, Nijhuis reported the concomitant occurrence of the two protease mutations with the A431V CS mutation and this mutation conferred a replicative benefit to the otherwise severely impaired replication of the L76V-containing virus. However, they observed that the triple mutant containing M46I and L76V in the PR and A431V in Gag conferred 12 fold LPV resistance, while the single mutant

A431V conferred 3 fold resistance to LPV. In addition, they also found that the A431V was the most frequent single mutant and proposed that the A431V may precede and facilitate the selection of the L76V protease resistance mutation. By contrast, our result showed that the A431V was selected after the emergence of the L76V protease resistance mutation and increased the RC of the virus while having no significance effect on resistance. Differences in viral subtype and the format of the assay employed for drug susceptibility testing may be accountable for these discrepancies. Nijhuis and colleagues employed a multiple cycle assay and studied a subtype B HIV-1 virus, while we employed a single cycle assay and evaluated a CRF02 HIV-1 virus. Different viral subtypes may differ in the pathway and order of accumulation of mutations (Wainberg et al, 2011). Also, subtle differences in fold change may be observed in different phenotypic assays.

Patient RFH-3 was followed over a five year period on unsuccessful LPV/r + APV/r therapy. During this time, the patient developed four major protease mutations (M46I, I84V, L76V and F53L), 35 non-CS mutations and 2 CS mutations (P17/P24: Y132F and P1/P6: L449F). Interestingly, we observed that despite the presence of 4 major protease resistance mutations, which are known to have a negative effect on RC (Martinez-Picado et al, 1999; Nijhuis et al, 1999), the patient's last time point virus replicated significantly more efficiently than the baseline virus which did not contain major protease resistance mutations (97% *vs*. 160%). In order to assess the contribution of Gag evolution to the efficient replication, we exchanged the Gag from baseline and last time point. We observed that the resistance test vector containing the last time point PR combined with the baseline Gag displayed a significantly impaired RC of 4% compared to the 160% observed when both last time PR and Gag were incorporated. This result

illustrated that Gag evolution was required for effective replication of the virus containing multiple major protease resistance mutations. We next stepwise reverted to wild type the two emergent CS mutations so that we could determine their effect on viral replication. We found that reversion of both mutations led to a non-significance decrease in RC from 159% to 150% when Y132 was reverted to wild-type and the RC was further decreased if also L449F mutation was reverted to wild type (159% to 129%) and in this case the difference in replication reached statistical significance. This result indicates that although emergent Gag CS mutations slightly contributed to the efficient replication of the multi-PI-resistant virus, most of the replicative benefit exhibited by the virus is likely to lie on the non-CS mutations. The importance of non-CS Gag mutations for the full recovery of viral fitness of multi-PI resistant viruses has been previously reported. We have already mentioned that Parry and co-authors described the importance of the P17 Gag domain for the efficient replication of multi-PI resistant viruses (Parry et al 2009). In the same line, Gatanaga and colleagues studied the effect of various non-CS substitutions on the development of HIV-1 resistance to APV and concluded that both the CS and non-CS mutations were essential for the efficient replication of APV-resistant HIV-1 (Gatanaga et al, 2002).

We also evaluated the PI susceptibility of the construct containing baseline Gag and PR and baseline PR with last time point Gag with and without emergent CS mutations in order to assess the effect of both CS and non-CS emergent mutations on PI susceptibility. Interestingly, we observed a significant increase in the level of resistance to all PIs when the baseline Gag was replaced with last time point Gag. The difference in FC was between 9 and 33 fold depending on the PI. If applying the Virco clinical cut-off to these data, this result means that Gag alone was

sufficient to elevate the level of resistance well above the lower clinical cut-off for all PIs and even above the upper clinical cut-off for APV. A further consideration of clinical relevance is the importance of addressing treatment failure early to avoid continuous viral replication and virus evolution which importantly, can occur not only in the PR but also in its substrate, the protein Gag and therefore be missed by routine genotypic testing. In this line, studies have reported that patients are most likely to fail a PI regimen if they were previously PI experienced even if the level of cross-resistance predicted by genotypic analysis of the PR is low (Dronda et al, 2001). Our results suggest that continuous evolution of the *Gag* gene under unsuccessful PI therapy may be at least in part be accountable for this finding. Although the stepwise reversion to wild type of the two emergent CS mutations (Y132F and L449F) led to a slight decreased in the level of PI resistance, the changes in IC_{50} were still significantly above those observed in the resistance test vector containing baseline Gag indicating that determinant of PI resistance were in this patient located in Gag both at CS and beyond. Importantly, at CSs we found the mutation L449F at P1/P6 and its reversion to wild type decreased the FC for all PI tested. Consistent with this finding Maguire and colleagues reported a 5-fold increase in APV resistance caused by this mutation (Maguire et al, 2002). Interestingly, we observed the mutation Y132F at P17/P24. The involvement of mutations at this CS on PI resistance has not been previously reported as we have already mentioned that studies have primarily focused on addressing the effect of mutations in P7/P1/P6. This focus reflects the fact that cleavage at these sites is the rate-limiting step in Gag processing and therefore it was anticipated that its cleavage would be severely affected by the loss of catalytic activity displayed by viral PR containing major PR resistance mutations. As a result, these CS were expected to rapidly evolve in order to maintain the efficient processing (Croteau et al, 1997; Zenou et al, 1998 and Martinez-Picado et al, 1999). However, it should be

noted that Maguire and colleagues showed that mutations in protease reduced cleaving efficiency for all known wild type CS substrates suggesting potential selective pressure on CS others than P7/P1/P6. In fact, in our study we have seen that mutations associated with PI exposure were seen in all CS except for P24/P2. In fact, we have found that the Y132F mutation was strongly associated with PI selective pressure (chapter four) and we and others have seen the Y132F mutation emerging in patients who developed PI resistance (Mammano et al, 1994). The fact that the reversion to wild type of the treatment emergent CS mutation Y132F led to an increase on PI susceptibility further support a role for the mutation in PI resistance.

Due to the scarce information available on the phenotypic impact of Y132F, we employed SDM to introduce the mutation in a drug susceptible reference strain. It was observed that the Y132F mutation led to a 10 fold reduction on PI susceptibility while also having a detrimental effect on RC. The negative impact on RC is consistent with the observation that Y132F occurs exclusively in PI–experienced patients (chapter four). The mechanism by which Y132F reduce PI susceptibility have not been addressed. It can be postulated that as other CS mutations may render the CS better substrate and facilitate its cleavage. To support this hypothesis, early studies demonstrated that the Y132F mutation caused an increase in the rate of P17/P24 processing (Tritch et al, 1991). It may be envisaged that in the context of a wild-type protease displaying unimpaired catalytic activity Y132F may lead to premature cleavage and production of aberrant viral particles.

It should be noticed that the effect of Y132F mutation on PI susceptibility was higher in the wild-type subtype B reference strain than it was on the CRF02 subtype patient's virus

highlighting the importance of including the whole virus backbone in the phenotypic assay. Thus, it has been documented that polymorphism and secondary protease resistance mutations may have different PI resistance effects on different viral contexts (Wainberg et al, 2011 and Martinez-Cajas et al, 2012).

Overall, our results showed that both CS and non-CS mutations are involved in modulating both PI susceptibility and RC. While most studies have agreed on the importance of Gag non-CS substitutions for the full recovery of virus RC of multi-PI resistant viruses (Gatanaga et al, 2002; Myint et al, 2004; Parry et al, 2009), the role of non-CS mutations in PI resistance have been debated. In line with our findings Parry and co-investigators reported that Gag alone was able to confer 10-15 fold resistance to all PIs. By contrast Dam and colleagues documented that no Gag region other than P7/P1/P6 was accountable for PI resistance. Several factors may account for the discrepancy such as the format of the phenotypic assay employed for drug susceptibility testing. In the present study and the one performed by Parry, a single cycle incorporating luciferase as reported gene was used. Dam and co-authors employed also a single cycle but galactosidase activity was in this case the indicator what may be less sensitive for the detection of subtle fold changes. In addition, as previously mentioned the backbone and subtype of the virus tested my also affect the results.

In summary, we have shown that Gag CS and non-CS mutations modulate PI susceptibility and RC. Gag alone was sufficient to reduced PI susceptibility and in some cases the fold change in PI resistance was above the upper clinical cut-off reported by some commercial assays. In

conclusion, our results indicated that exclusion of the Gag gene in phenotypic resistance testing may overestimate PI susceptibility.

7 Chapter seven: general discussion

The development of protease inhibitors (PIs) in the early 1990s was a landmark breakthrough in the treatment of HIV-1 infection. PIs made possible the dual-class triple combination therapy that became known as highly active antiretroviral therapy (HAART) and that has since constituted the standard of care of HIV-1 infected patients worldwide. However, first-generation PIs, such as saquinavir (SQV), ritonavir (RTV) indinavir (IDV) and nelfinavir (NFV) showed significant pharmacokinetic limitations. Poor oral absorption, high serum-protein binding, and rapid liver enzyme metabolism resulted in PIs needing to be ingested often and in large quantities to maintain effective antiviral concentrations in blood, with significant food intake requirements. Consequently, first generation PI-based HAART often led to suboptimal adherence and limited long-term viral inhibition. Moreover, when therapy failure occurred, frequently multiple protease resistance mutations were present which were capable of producing significant cross-resistance, since all first- generation PIs occupied the same cavity within the HIV-1 protease (PR) enzyme. A major advance in the use of PIs came with the discovery that ritonavir (RTV) was a potent inhibitor of the hepatic cytochrome P450 isoenzyme CYP3A4 which, in turn, is the main enzyme responsible for PI metabolism. Concomitant administration of low-dose RTV with a PI (PI/r) led to "boosting" of the most important pharmacokinetic parameters of most PIs, which ultimately simplified otherwise complex regimens by reducing the frequency and number of pills to be administered and often obviating to complex food intake requirements. In addition, by elevating drug plasma levels, boosted PIs became les prone to select for drug resistance and more effective against PI-resistant HIV-1 variants. Subsequently, novel PIs were licensed for the treatment of HIV-1 infection which displayed improved

pharmacokinetic profiles and genetic barrier to resistance. Lopinavir (LPV) was the first and currently remains the only PI co-formulated with a low dose RTV. LPV/RTV was also the first boosted PI to be compared head-to-head with a non-nucleoside/nucleotide reverse transcriptase inhibitor (NNRTI) for the initial treatment of HIV-1 infection and is still currently one of the most frequently prescribed PI for HIV-1 treatment, especially in the developing world. Another approach to overcoming PI resistance consisted of the development of PIs with greater resemblance to the PR substrate, the protein Gag. This resulted mutational profiles different from those observed for previously developed PIs. Amprenavir (APV) and Darunavir (DRV) were the the first and last second-generation PI respectively designed with this approach. DRV shows a particularly high binding affinity for HIV-1 PR what further increases the genetic barrier compared to all previous PIs. DRV was initially introduced for the treatment of patient infected with multi-PI resistant viruses. However, its antiviral potency, good adverse event profile and high genetic barrier to resistance led to the drug being evaluated and approved for the treatment of antiretroviral naïve patients. Furthermore, boosted PIs, including atazanavir (ATV), LPV and DRV, have also been tested in clinical trials as monotherapy with results ranging from suboptimal (ATV) to good (DRV) relative to standard triple therapy. As a result of all the progress made with PI-based therapy, since the introduction of the first PI nearly 20 year ago, these drugs continue today to be a cornerstone in the treatment of HIV-1 infection.

Despite the extraordinary improvement in PI therapy, the life-long nature of HIV treatment, together with the enormous genetic plasticity of the virus, make the development of PI resistance and the subsequent loss of efficacy unavoidable. While failure of first generation PI-based

treatment strategies was generally characterised by the presence of multiple PR resistance mutations, treatment failure of combination therapy including newer PIs, such as LPV/RTV or DRV/RTV in subjects previously antiretroviral-naïve rarely showed the emergence of resistance mutations (Kempf et al, 2004 and Gupta et al, 2008). In contrast, it has been observed that their genetic barrier to resistance can be lowered by mutations selected during previous PI-based regimens (Mo et al, 2005). The main determinants for failure of PI/r in the absence of detectable PR resistance mutations remain poorly understood. Incomplete adherence or altered absorption and metabolism of PIs may be possible explanations in some patients. However, unexplored pathways of resistance involving the Gag protein have also been considered. Along this line, numerous studies have demonstrated an association between the selection of protease mutations and the emergence of mutations in Gag, predominantly in the cleavage sites (CS) P7/P1 and P1/P6 (Maguire et al, 2002; Prado et al, 2002; Nijhuis et al, 2007 and Dam et al, 2009). A study performed by Maguire and colleagues (Maguire et al, 2002) demonstrated that mutations in P1/P6 CS (L449F and P453L), which individually did not confer PI resistance, reduced APV sensitivity in the context of the major protease mutation I50V, thus providing evidence that PR and Gag mutations can interact to increase PI resistance. Nijhuis and co-authors showed that variants in the Gag CS P7/P1 (A431V, K436E and/or I437V/T) were selected by PIs *in vitro*, in the absence of any substitution in the viral PR. The introduction of these Gag mutations in a reference strain led to low-level resistance to all PIs (Nijhuis et al, 2007). A more recent analysis of clinical isolates of patients on PI therapy carried out by Dam and colleagues (Dam et al, 2009) reported that mutations in P7/P1 (A431V and I437V) strongly and directly contributed to PI resistance in addition to compensating for the loss fitness caused by major PI resistance

mutations. Overall, these findings showed that selection of P7/P1 cleavage site mutations (CSMs) may represent an alternative pathway of PI resistance.

The role of mutations at other gag CS or even beyond CS regions has been poorly evaluated. A few studies suggest that determinants of PI resistance may be present at CS other than NC-P7/P1 and P1/P6 as well as at non-cleavage sites (non-CS). Malet and colleagues showed that mutations at amino acid 373 in the P2/P7 CS were predictive of impaired virological responses to SQV/r (Malet et al, 2007). Recently, Ghosn and colleagues evaluated the impact of amino acid variability in the five Gag CSs on failure to LPV/r monotherapy within the MONARK trial. The study compared the efficacy of LPV/r monotherapy to that of triple therapy with LPV/r, zidovudine (ZDV) and lamivudine (3TC) for treatment of antiretroviral-naïve HIV-1 infected patients. They showed that having more than two mutations in P2/P7, especially if involving position 374, at baseline was predictive of virological failure of LPV/r monotherapy. Similarly, Parry and colleagues showed that determinants of PI resistance are likely to be located outside PR and its CSs, and that Gag conferred low-level resistance to all PIs in the context of a wild-type PR (Parry et al, 2009).

Despite this body of knowledge, the clinical management of PI failure remains based on the sequencing of the PR gene, in search for recognised protease resistance mutations. Similarly, the phenotypic evaluation of PI resistance involves the use of recombinant viruses containing patient's derived PR and, in some cases, NC-P7/P1/P6 CS sequences. However, the contribution of full-length gag to PI resistance remains unaddressed.

In the course of my PhD studies, I developed and optimised an assay for amplification of full-length HIV-1 PR and gag genes. This was made challenging by the considerable variability occurring in gag. Using this assay, I was able to amplify and sequence full-length PR and gag from a variety of HIV-1 subtypes and circulating recombinant forms (CRFs) at viral loads of around 500 to 1,000 copies/ml. Consequently, the assay sensitivity was regarded satisfactory for monitoring patients experiencing PI failure. However, due to the intra-patient variability of the gag gene, non-interpretable population sequencing was not unusual and clonal analysis was required to circumvent the problem in some samples. This represents an important limitation to the implementation of full-length gag sequencing in a high-throughput routine diagnostic setting.

By applying the assay to the study of full-length Gag and PR genes from PI-naïve and PI-experienced patients, we obtained sequences from 200 PI-naïve .and 191 PI-experienced patient samples. Comparison of the two groups demonstrated that PI-experienced subjects showed greater variability than PI-naïve individuals not only in PR, but also throughout the Gag protein. Importantly, there were significant differences in the prevalence of certain Gag mutations in the two groups, suggesting that HIV-1 genetic evolution under PI-selective pressure is not restricted to the PR gene but occurs in its natural substrate, the Gag protein. Consistent with previous reports, we found that mutations in P7/P1 and P1/P6 CSs, such as A431V, K436R, I437V and L449F, were associated with PI selective pressure and occurred concomitantly with specific major PR resistance mutations, namely M46IL, I54V and I84V (Mammano et al, 2000; Zhang et al, 1997; Maguire et al, 2002; Cote et al, 2001; Bally et al, 2000 and Dauber et al, 2002,

Verheyen et al, 2006, Feher et al, 2002, Prado et al, 2002 and Nijhuis et al, 2007). However, by analysing full-length Gag, we significantly expanded previous knowledge and have identified a number of novel Gag mutations strongly associated with PI-experience. In the present study, mutations associated with PI-selective pressure were found in all CS except for CA-P24/P2. Mutations strongly associated with PI-experience included V128I and Y132F in P17/P24, and S373T, A374S and T375N in P7/P2.

Interesting, in our study, we found that mutations associated with PI exposure were also present outside the CSs. Thus, we observed a trend towards a higher number of mutations in PI-experienced patients compared with PI-naïve patients in all Gag domains. The effect was particularly noticeable in the MA-P17 and the P6 proteins, suggesting the two may play a prominent role in PI failure compared with other Gag regions. Consistent with this hypothesis, Parry and colleagues (Parry et al, 2009) reported that the P17 protein from a multidrug PI-resistant virus was on its own able to rescue the otherwise impaired replicative capacity (RC) of the mutated PR to the level observed in the wild-type virus. In addition, they demonstrated that the mutated P17 reduced PI susceptibility in the absence of major PR resistance mutations, suggesting that major determinants of PI-resistance may be located in this protein. Interestingly, the clinical isolate they studied showed 12 mutations in MA-P17 (compared with the HXB2 reference sequence) including six (I34L, T84V, E93D, I94V, N124K and N126S) that were found to be associated with PI selective pressure in our analysis. Two of these mutations in particular (I94V and N126S) showed a strong association with PI experience remaining significantly associated even after applying the Bonferroni's correction.

In order to gather more insight into the role of Gag mutations in patient experiencing PI failure we performed longitudinal analyses of patients that were failing a PI-based treatment regimen. A group of 28 patients receiving IDV/r, SQV/r or LPV/r were followed for up to 24 weeks and their Gag and protease sequences obtained pre-treatment (baseline) and at the time of virological failure were compared. In addition, I was able to follow three patients for up to 5 years and obtain Gag and protease sequences at multiple time points (3-5) during treatment failure of .LPV/r for two patients and LPV/r plus APV/r for one patient. As expected, we observed a high prevalence of treatment-emergent major and minor protease resistance mutations. Interestingly, we observed an even higher prevalence of emergent Gag mutations both in and outside its CSs. In the first group of patients, we detected treatment- emergent mutations in three CSs: P2/P7, P7/P1 and P1/P6. In total we observed nine different mutations emerging in these three CSs during failure. The majority of these mutations (6/9) had been found to be associated with PI selective pressure in both our study and studies by others (Verheyen et al, 2006; Malet et al, 2006). The role of most mutations occurring in P7/P1 (A431V and I437V), and in P1/P6 (L449F) has been described in the literature (Mammano et al, 2000; Zhang et al, 1997; Maguire et al, 2002; Cote et al, 2001; Bally et al, 2000 and Dauber et al, 2002, 2002 and Nijhuis et al, 2007). Emergence of CS mutations was observed both in the absence and presence of major protease mutations. Furthermore, the mutations did not appear to depend generally on the specific PI in the failing regimen, although P1/P6 mutations were more common in patients failing SQV/r and mutations in P2/P7 were seen mainly in patients failing IDV/r. Nevertheless, the number of patients studied longitudinally was limited and did not allow a formal analysis in the association between specific regimen and the presence of certain Gag mutations.

Emergent mutations outside the CSs, and predominantly in P17, were also detected, that had been found to be associated with PI exposure, thus strengthening the hypothesis that in addition to gag CSs, P17 plays a significant role in PI failure.

Although the cross-sectional analysis found several mutations associated with PI exposure in P17/P24 and P6, in our longitudinal analysis we did not observe emergence of mutations in these regions. However, we did observed the P17/P24 mutations Y132F and V128I in the baseline samples of three patients who had been previously exposed to PIs and showed several major protease resistance mutations. In addition, 18 patients had one or more mutations associated with PI selective pressure in P6, all of whom had been previously exposed to PIs. These findings suggest that that PI initially exert selective pressure mainly on the CSs P2/P7, P7/P1 and P1/P6 (which are located in the gag terminal site) and on the P17 protein outside the CSs. Long-term selective pressure in contrast may trigger evolution in other CSs (such as P1/C-P24) and in P6 outside the CSs. This concept was further supported by the observations made in the three patients who were followed by a longer period of time, and in whom we saw emergent mutations at these sites. Thus, the Y132F mutation in P17/P24 emerged in one patient after five years of ongoing viral replication while on APV/r and LPV/r, and P6 mutations also emerged in two patients overtime. Unfortunately, for most patients, we did not have access to data on the length of PI exposure at baseline. Therefore we could not formally analyse the association between duration of PI exposure and presence of specific Gag mutations.

An important observation is that while emergent Gag CS mutations were common among patients and consistent with previous reports, the emergent non-CSMs were highly heterogenous and did not show consistent patterns. This indicates that evolution outside CS is complex and likely to be driven by different pressures including virological factors (genetic make-up, level and duration of virus replication, emerging mutations in protease and CSs), pharmacology factors (drug type and drug exposure), and host-related factors (immune response).

In the final chapter we applied a single cycle assay for the phenotypic measurement of PI susceptibility and viral RC. We studied samples collected from four patients at the time of virological failure and assessed the impact of several CSMs (Y132F, T375A, A431V, I437V and L449F) on drug susceptibility and RC. Overall, we observed that CSMs contributed to PI resistance. Thus, the reversion to wild-type of Y132F, T375A or L449F led in all cases to a 3-5 fold reduction in PI susceptibility. Similarly, the introduction by side-directed mutagenesis of the I437V mutation in the patient's baseline sample produced a 2-5 fold increase in PI resistance. In addition, most CSMs led to an improvement in viral RC, with the exception of I437V which did not have a discernible effect. We confirmed the independent effect of mutations Y132F and T375A on PI susceptibility and RC by introducing the mutations in a wild type backbone. We found that Y132F conferred around 10 fold resistance to all PIs while decreasing RC, while T375A produced 5-10 fold resistance to all PIs but increased RC. In addition, in line with previous studies (Gatanaga et al, 2002; Myint et al, 2004; Dam et al, 2009 and Parry et al, 2009), we found that non-CS mutations, in addition to contributing to PI resistance, were also required for efficient replication of multi-PI resistance viruses. Importantly, we found that emergent Gag

mutations were not generally associated with a specific PI regimen but rather the same mutations were seen emerging under different PI regimens. This finding, combined with the observation that the Gag mutations conferred a degree of cross-resistance to most PIs in the phenotypic assay, suggests a common mechanism for conferring resistance.

PIs, with the exception of IDV, are more than 90% protein bound *in vivo* (Bilello et al, 1996). Our phenotypic drug susceptibilities studies have been performed in HEK T293 cells grown and maintained in cell culture media supplemented with 10% FCS. As the calculation of the IC_{50} relies on the presence of unbound drug concentrations, it will be highly dependent on the composition of the incubation media. As a result, caution should be applied when comparing the results obtained in the present studies with those obtained under different tissue culture conditions and also before extrapolating the IC_{50} in this study to *in vivo* conditions.

Overall, our results clearly indicate that continuous viral replication under PI selective pressure leads to evolution of the viral PR as well as its substrate, the Gag protein. Consequently, analyzing PR alone after PI failure may underestimate the level of PI resistance. Given our result and in general the expanding body of evidence indicating that Gag mutations contribute to PI resistance in treated patients, full-length Gag sequences should be incorporated in phenotypic assays to determine PI susceptibility of clinical isolates.

Expansion of full-length Gag sequencing in patients failing PI-based regimens would be also of interest to populate the relevant databases employed for the clinical assessment of HIV drug resistance and to assist in the identification of specific Gag determinants of PI susceptibility. Only then, the importance of Gag genotypic determination to guide patient care could be evaluated.

In addition, the findings of my PhD studies could also have important implications for drug development and as a result I propose that when novel PIs are being developed, it should be checked whether Gag evolution provides an alternative mechanism of escape for the virus.

It should be emphasised that all patients studied were failing PIs as part of triple combination therapies. However, due to financial constraints, the interest on PI-monotherapy regimens is increasing in many countries around the world. Nevertheles, it shold be emphasized that the analysis of patients failing LPV/r monotherapy in the MONARK trial demonstrated a lower genetic barrier to resistance compared with triple therapy. Therefore, it can be proposed that the effect of Gag mutations may be more prominent in the context of PI/r monotherapy. Consistent with this concept, the analysis of baseline Gag sequences from patients assigned to receive LPV/r monotherapy in MONARK showed that patients with at least two mutations in the gag CS P2/P7 were more likely to experience virological failure of LPV/r than those who did not show mutations at this CS (Ghosn et al, 2011). Consequently, it would be of interest to address the evolution of the Gag in patients starting PI/r monotherapy and its effect of drug susceptibilty. Currently, PI/r monotherapy is only established as a treatment switch in patients with stable

undetectable viraemia while on HAART. Consquently, performing such analysis may be hampered by the potential lack of access to the baseline samples for the patients.

It should be also noticed that most of my PhD work has been performed with subtype B HIV-1 viruses. By the end of 2007, nearly three million people were on HAART worldwide as a consequence of the therapy roll-out in developing countries. NNRTI-based regimens are recommended by the World Health Organisation (WHO) as first-line therapy in low-income countries. For second line therapy, the WHO advises initiation of a PI-based regimen, currently LPV/r or ATV/r, plus two NRTIs. As the number of patients having access to antiretroviral therapy increase in developing countries so does the number of patients needing to move to second line PI-based therapy. However, some of the Gag mutations that we have described as associated with PI exposure and that indeed we have found to decrease PI susceptibility "*in vitro"* may occur as natural polymorphisms in non-B subtypes thus potentially affecting responses to second-line therapy. This is especially important when considering that patients in these regions commonly start second-line therapy having already accumulated significant resistance to the NRTI component of the regimen. A few studies have shown that Gag variability is greater in non-B subtypes than in B subtype HIV-1 (Jinnopat et al, 2009 and Ghosn et al, 2011). For instance, Jinnopat reported a higher Gag variability in subtype CRF01 (Jinnopat et al, 2009), and in particular the mutations L61I and P66S in P17 were frequently detected in drug-naïve patients. Both L61I and P66S were associated with PI-exposure in our study and in the case of L61I the association remained significant after the Bonferroni's correction for multiple associations was applied. In addition, Gupta and co-authors frequently detected the T84V

mutation in subtype A (Gupta et al, 2010), which have also been associated with PI-exposure in our study. Importantly, Gupta and colleagues reported that in the context of wild-type subtype B PR, replacing the subtype B Gag with subtype A Gag decreased PI susceptibility. Similarly, Ghosn and colleagues showed that the risk of virological failure of LPV/r monotherapy was significantly higher in patients with polymorphisms at P2/P7 CS and polymorphism at this CS were in turn more common in non-B subtypes compared to subtype B HIV-1 viruses. Considering that resistance testing is not routinely perfomed in developing countries and the important implications that Gag mutations in non-B subtypes and their correnponding effects on PI susceptibility may have for the public health approach to antiretroviral treatment, in future work, we would like to expand the full-length Gag sequencing to patients infected with non-B subtypes who are due to start a PI-based regimen and to evaluate the prevalence of subtype-B mutations associated with PI exposure in non-B subtypes at baseline as well as the impact of such polymorphisms on virological responses to PIs.

8 Chapter eight: references

1. Accola MA, Bukovsky AA, Jones MS and Göttlinger HG. 1999. A conserved dileucine-containing motif in p6gag governs the particle association of Vpx and Vpr of simian immunodeficiency viruses SIV (mac) and SIV (agm). *J. Virol*. 73: 9992-9.

2. Accola MA, Höglund S and Göttlinger HG. 1998. A putative alpha-helical structure which overlaps the capsid-p2 boundary in the human immunodeficiency type 1 Gag precursor is crucial for viral particle assembly. *J. Virol*. 72: 2072-8.

3. Accola MA, Strack B and Gottlinger HG. 2000. Efficient particle production by minimal gag constucts which retain the carboxyterminal domain of human immunodeficiency virus type 1 capsid-p2 and a late assembly domain. *J. Virol*. 74: 5395-402.

4. Albert J, Walhberg J, Leitner T, Escanilla D and Uhlen M. 1994 Analysis of a rape case by direct sequencing of the human immunodeficiency virus type 1 pol and gag genes. *J. Virol*. 68: 5918-24.

5. Alcorn TM and Faruki H. 2000. HIV resistance testing: methods, utility and limitations. *Mol. Diagn*. 5: 159-68.

6. Aldovini A and Young RA. 1990. Mutations of RNA and protein sequences involved in human immunodeficiency virus type 1 packaging result in production of noninfectious viruses. *J. Virol*. 64: 1920-6.

7. Allen TM, Altfeld M, Geer SC, Kalife ET, Moore C, O'Sullivan KM, Desouza I, Feeney ME, Eldridge RL, Maier EL, Kaufman DE, Lahaie MP, Reyor L, Tanzi G, Johnston MN, Brander C, Draenert R, Rockstroh JK, Jessen H, Rossenberg ES, Mallal SA and Walker BD. 2005. Selective escape from CD8þ T-cell responses represents a major driving force of human immunodeficiency virus type 1 (HIV-1) sequence diversity and reveals constraints on HIV-1 evolution. *J. Virol*. 79: 13239-49.

8. Andries K, Azijn H, Thielemans T, Ludovici D, Kukla M, Heeres J, Janssen P, De Corte B, Vingerhoets J, Pauwels R and de Béthune MP. 2004. TMC125, a novel next generation nonnucleoside reverse transcriptase inhibitor active against nonnucleoside reverse transcriptase inhibitor-resistant human immunodeficiency virus type 1. *Antimicrob. Agents and chemother*. 48:4680-6.

9. Arhel NJ, Souquere-Besse S, Munier S, Souque P, Guadagnini S, Rutherford S, Prévost MC, Allen TD and Charneau P. 2007. HIV-1 DNA flap formation promotes uncoating of the pre-integration complexat the nuclear pore. *EMBO J*. 26: 3025-37.

10. Arion D, Kaushik N, McCormick S, Borkow G, and Parniak MA. 1998. Phenotypic mechanism of HIV-1 resistance to 3'-azido-3'-deoxythymidine (AZT): increases polymerization

processivity and enhanced sensitivity to pyrophosphate of the mutant viral reverse transcriptase. *Biochemistry*. 37: 15908-17.

11. Auewarakul P, Wacharapornin P, Srivhatrapimuk S and Puthavathana P. 2005. Uncoating of HIV-1 requires cellular activation.*Virology*. 337: 93-101.

12. Bacheler L, Jeffrey S, HannaG, D'Aquila R, Wallace L, Logue K, Cordova B, Hertogs K, Larder B, Buckery R, Baker D, Gallagher K, Sacarnati H, Tritch R and Rizzo C. 2001. Genotypic correlates of phenotypic resistance to efavirenz in virus isolates form patients failing non-nucleoside reverse transcriptase inhibitor therapy. *J. Virol*. 75: 4999-5008.

13. Bally F, Martinez R, Peters S, Sudre P and Telenti A. 2000. Polymorphism of HIV type 1 gag P7/P1 and P1/P6 cleavage sites: clinical significance and implications for resistance to protease inhibitors. *AIDS Res. Hum. Retroviruses*. 16: 1209-13.

14. Bandaranayake RM, Kolli M, King NM, Nalivaika EA, Heroux A, Kakizawa J, Sugiura W and Schiffer CA. 2010. The effect of clade-specific sequence polymorphisms on HIV-1 protease activity and inhibitor resistance pathways. *J. Virol*. 84: 9995-10003.

15. Bandarnayake RM, Prabu-Jebayalan M, Kakizawa J, Sugiura W and Schiffer CA. 2008. Structural analysis of human immunodeficiency virus type 1 CRF01-AE protease in complex with the substrate P1-P6. *J. Virol*. 82: 6762-66.

16. Bangsberg DR, Moss AR and Deeks SG. 2004. Paradoxes of adherence and drug resistance to HIV antiretroviral therapy. *J.Antiomicrob Chemother*. 53: 696-9.

17. Banke S, Lillemark MR, Gerstoft J, Obel N and Jørgensen LB. 2009. Positive selection pressure introduces secondary mutations at Gag cleavage sites in human immunodeficiency virus type 1 harbouring major protease resistance mutations. *J.Virol*. 83: 8916-24.

18. Barlow KL, Tatt ID, Cane PA, Pillay D and Clewley JP. 2001. Recombinant strains of HIV type 1 in the United Kingdom. *AIDS. Re. Hum. Retroviruses*. 17: 464-74.

19. Barnes WM. 1994. PCR amplification of up to 35 KB DNA with high fidelity and high yield from lambda bateriophage templates. *Proc. Natl. Acd. Sci. U. S. A.* 91: 2216-20.

20. Barre-Sinoussi F, Chermann JC, Rey F, Nugeyre MT, Chamaret S, Gruest J, Dauguet C, Axler-Blin C, Vezinet-Brun F, Rouzioux C, Rozenbaum and Montagnier L. 1983. Isolation of a lumphotropic retrovirus from a patient at risk for acquired immune deficiency sindrome (AIDS). *Science*. 220: 868-71.

21. Baxter JD, Mayers DL, Wentworth DN, Neaton JD, Hoover ML, Winters MA, Mannheimer SB, Thompson MA, Abrams DI, Brizz BJ, Loannidis JP and Merigan TC. 2000. A randomized study of antiretroviral management based on plasma genotypic antiretroviral resistance testing in patients failing therapy. *AIDS*. 14: F83-F93.

22. Bebenek, K., J. Abbotts, J. D. Roberts, S. H. Wilson, and T. A. Kunkel. 1989. Specificity and mechanism of error-prone replication by human immunodeficiency virus-1 reverse transcriptase. *J. Biol. Chem*. 264: 16948-56.

23. Berger EA, Doms RW, Fenyo EM, Korber BT, Littman DR, Moore JP, Sattentau QJ, Schuitemaker H, Sodroski J and Weiss RA. 1998. A new classification for HIV-1. *Nature*. 391: 240.

24. Berkowitz R, Fisher J and Goff SP. 1996. RNA packaging. *Curr.Top. Microbiol. Immunol*. 214: 177-218.

25. Bieniasz PD. 2009. The cell biology of HIV virion genesis. *Cell Host & Microbe*. 5: 550-8, 18 June 2009.

26. Bilello JA, Bilello PA, Stellrecht K, et al. 1996. Human serum α1-acid glycoprotein reduces uptake, intracellular concentration, and anti-viral activity of A-80987, an inhibitor of the human immunodeficiency virus type 1 protease. *Antimicrob Agents Chemother*. 40: 1491-1497.

27. Billich S, Knoop MT, Hansen J, Strop P, Sedlacek J, Mertz R and Moelling K. 1988. Synthetic peptides as substrates and inhibitors of human immune deficiency virus-1 protease. *J. Biol. Chem*. 263:17905-8.

28. Booth C L, Garcia-Diaz A M, Youle M S, Johnson M A, Phillips A and Geretti AM. 2007. Prevalence and predictors of antiretroviral drug resistance in newly diagnosed HIV infection. *J.Antimicrob. Chemother*. 59: 517-24.

29. Boom R, Sol C and Wertheim-Van Dillen P. 1990. Rapid purification of ribosomal RNAs from neutral agarose gels. *Nucleic. Acids. Res*. 18: 2195.

30. Borsetti A, Ohagen A and Göttlinger HG. 1998. The C-terminal half of the human immunodeficiency virus type 1 Gag precursor is sufficient for efficient particle assembly. *J. Virol*. 72: 9313-7.

31. Boucher C.A, O'Sullivan E, Mulder J.W, Ramautarsing C, Kellam P, Darby G, Lange J.M, Goudsmit J and Larder B.A. 1992. Ordered appearance of zidovudine resistance mutations during treatment of 18 human immunodeficiency virus-positive subjects. J. *Infect. Dis*. 165: 105-110.

32. Boyer P. L, Sarafianos SG, Arnold E and Hughes SH. 2001. Selective excision of AZTMP by drug-resistant human immunodeficiency virus reverse transcriptase. *J. Virol*. 75: 4832-4842.

33. Bracho MA, Moya A and Barro E. 1998. Contribution of Taq polymerase-induced errors to the estimation of RNA virus diversity. *J. Gen. Virol*. 79: 3921-2928.

34. Brann TW, Dewar RL, Jiang MK, Shah A, Nagashima K, Metcalf JA, Falloon J, Lane HC and Imamichi T. 2006. Functional correlation between a novel amino acid insertion at codon 19 in the protease of human immunodeficiency virus type 1 and polymorphism in the p1/p6 Gag cleavage site in drug resistance and replication fitness. *J. Virol*. 80: 6136-45.

35. Brass AL, Dykxhoorn DM, Benita Y, Yan N, Engelman A, Xavier RJ, Lieberman J and Elledge SJ. 2008. Identification of host proteins required for HIV infection through a functional genomic screen. *Science*. 319: 921-26.

36. Brehm, J. H., D. Koontz, J. D. Meteer, V. Pathak, N. Sluis-Cremer, and J. W. Mellors. 2007. Selection of mutations in the connection and Rnase H domains of human immunodeficiency virus type 1 reverse transcriptase that increase resistance to 3'-azido-3'-dideoxythymidine. *J. Virol*. 81: 7852-7859.

37. Brehm JH, Mellors JW, and N. Sluis-Cremer. 2008. Mechanism by which a glutamine to leucine substitution at residue 509 in the ribonuclease H domain of HIV-1 reverse transcriptase confers zidovudine resistance. *Biochemistry* 47: 14020-14027.

38. Briones MS, Dobard CW and Chow SA. 2010. Role of human immuodeficiency virus type 1 integrase in uncoating of the viral core. *J. Virol*. 84: 5181-5190.

39. Brown AL and Monaghan P. 1988. Evolution of the structural proteins of human immunodeficiency virus: selective constriants on nucleotide substitutions. *AIDS. Res. Hum. Retrovirus*. 4: 399-406.

40. Brumme ZL, Chan KJ, Dong WW, Wynhoven B, Mo T, Hogg RS, Montaner JS, O'Shaughnessy MV and Harrigan PR. 2003. Prevalence and clinical implications of insertions in the HIV-1 p6Gag N-terminal region in drug-naïve individuals initiating antiretroviral therapy. *Antivir. Ther*. 8: 91-6.

41. Bryant M and Ratner L. 1990. Myristoylation-dependent replication and assembly of human immunodeficiency virus 1. *Proc. Natl. Acad. Sci. U. S. A*. 87: 523-527.

42. Bukrinsky M, Sharova N and Stevenson M. 1993a. Human immunodeficiency virus type 1 2-LTR circles reside in a nucleoprotein complex which is different from the preintegration complex. *J. Virol*. 67: 6863-5.

43. Bukrinsky MI, Haggerty S, Dempsey MP, Sharova N, Adzhubel A, Spitz L, Lewis P, Goldfarb D, Emerman M and Stevenson M. 1993b. A nuclear localization signal within HIV-1 matrix protein that governs the infection of non-dividing cells. *Nature*. 365: 666-9.

44. Cairns TM and CraveN RC. 2001. Viral DNA synthesis defects in assembly-competent rous sarcoma virus CA mutants. *J. Virol*. 75: 242-250.

45. Casado JL, Hertogs K, Ruiz L, Dronda F, Van Cauwenberge A, Arno A, Garcia-Arata I, Bloor S, Bonjoch A, Blanquez J, Clotet B and Larder B. 2000. Non-nucleoside reverse transcriptase inhibitor resistance among patients failing a nevirapine plus protease inhibitor-containing regimen. *AIDS.* 14: F1.

46. CDC Weekly. 1982a. A cluster of Kaposi's sarcoma and Pneumocystis carinii pneumonia among homosexual male resisdents of Los Angeles Orange Counties, California. *MMWR Morb. Mortal. Wkly. Rep*. 31: 305-307.

47. CDC Weekly. 1982b. Pneumocystis carinii pneumonia among persons with hemophilia A. *MMWR Morb. Mortal. Wkly. Rep*. 31: 365-367.

48. CDC Weekly. 1982c. Update on acquired immune deficiency syndrome (AIDS)--United States. *MMWR Morb. Mortal. Wkly*. *Rep*. 31: 507-8.

49. Charneu P, Mirambeu G, Roux P, Paulous S, Buc H and Clavel F. 1994. HIV-1 reverse transcription. A termination step at the centre of the genome. *J.Mol. Biol*. 241: 651-62.

50. Charpentier C, Dwyer DE, Mammano F, Lecossier D, Clavel F and Hance AJ. 2004. Role of minority populations of HIV type 1 in evolution of viral resistance to protease inhibitors. *J.Virol*. 78: 4234-4247.

51. Chazal N, Carrière C, Gay B and Boulanger P. 1994. Phenotypic characterization of insertion mutants of the human immunodeficiency virus type 1 Gag precursor expressed in recombinant baculovirus-infected cells. *J.Virol*. 68: 111-22.

52. Chrusciel RA and Strohbach JW. 2004. Non-peptidic HIV protease inhibitors. *Curr.Top. Med.Chem*. 4: 1097-114.

53. Clark SA, Shulman NS, Bosch RJ, and Mellors JW. 2006. Reverse transcriptase mutations 118I, 208Y, and 215Y cause HIV-1 hypersusceptibility to non-nucleoside reverse transcriptase inhibitors. *AIDS*. 20: 981-84.

54. Clavel F, Guetard D, Brun-Vezinet F, Chmaret S, Ray MA, Santos-Ferreira MO, Laurent AG, Dauguet C, Katlama C, Rouzioux C. 1986. Isolation of a new human retrovirus from West African patients with AIDS. *Science*. 233: 343-6.

55. Clavel F, Race E and Mammmano F. 2000. HIV Drug resistance and viral fitness. *Adv. Pharmacol*. 49: 41-66.

56. Clewley J P, Arnold C, Barlow K L, Grant P R and Parry J V. 1996. Diverse HIV-1 genetic subtypes in UK. *Lancet*. 347: 1487.

57. Cline J, Braman JC and Hogrefe H. 1996. PCR fidelity of Pfu DNA polymerase and other thermostable DNA polymerases. *Nucleic. Acids. Res*.18: 3546-51.

58. Cohen CJ, Hunt S, Sension M, Farthing C, Conant M, Jacobson S, Nadler J, Verbiest W, Hertogs K, Ames M, Rinehart AR, Graham NM; VIRA3001 Study Team. 2002. A randomized trial assessing the impact of phenotypic resistance testing on antiretroviral therapy. *AIDS*. 16: 579-588.

59. Condra JH, Holder DJ, Schleif WA, Blahy OM, Danovich RM, Gabryelski LJ, Graham DJ, Laird D, Quintero JC, Rhodes A, Robbins HL, Roth E, Shivaprakash M, Yang T, Chodakewitz JA, Deutsch PJ, Leavitt RY, Massari FE, Mellors JW, Squires KE, Steigbigel RT, Teppler H and Emini EA.1996. Genetic correlates of in vivo viral resistance to indinavir, a human immunodeficiency virus type 1 protease inhibitor. *J. Virol*. 70: 8270-6.

60. Corbet S, Muller-Trutwin MC, Versmisse P, Delaure S, Ayouba A, Lewis J, Brunak S, Martin P, Brun-Vezinet F, Simon F, Barre-Sinousi F and Mauclere P. 2000. Env sequences of simian immunodeficiency viruses from chimpanzees in Cameroon are strongly related to those of human immunodeficiency virus group N from the same geographic area. *J. Virol*. 74: 529-34.

61. Côté HC, Brumme ZL and Harrigan PR. 2001. Human immunodeficiency virus type 1protease cleavage site mutations associated with protease inhibitor cross-resistance selected by indinavir, ritonavir, and/or saquinavir. *J. Virol*. 75: 589-94

62. Craven RC, Bennett RP and Wills JW. 1991. Role of the avian retroviral protease in the activation of reverse trancriptase during virion assembly. J Virol. 65: 6205-6217.

63. Croteau G, Doyon L, Thibeault D, McKercher G, Pilote L and Lamarre D. 1997. Impaired fitness of human immunodeficiency virus type 1 variants with high level resistance to protease inhibitors. *J. Virol*. 71: 1089-1096.

64. Dam E, Quercia R, Glass B, Descamps D, Launay O, Duval X, Krausslich HG, Hance AJ and Clavel F. 2009. Gag mutations strongly contribute to HIV-1 resistance to protease inhibitors in highly drug-experienced patients beside compensating for fitness loss. *PLoS. Path*. 5: e1000345.

65. Darke JW. 1993. Rates of spontaneous mutation among RNA viruses. *Proc. Natl. Acad. Sci. U.S.A.* 90: 4171-5.

66. D'Arminio MA, Lepri AC, Rezza G et al. 2000. Insights into thereasons for discontinuation of the first highly active antiretroviral therapy (HAART) regimen in a cohort of antiretroviral naïve patients. I.C.O.N.A. Study Group. Italian cohort of antiretroviral naïve patients. *AIDS*, 14; 499-507.

67. Dauber DS, Ziermann R, Parkin N, Maly DJ, Mahrus S, Harris JL, Ellman JA, Petropoulous C and Craik CS. 2002. Altered substrate specificity of drug resistance human immunodeficiency virus type 1 protease. *J. Virol*. 76: 1359-68.

68. Delaugerre C, Flandre P, Chaix ML, Ghosn J, Raffi F, Dellamonica P, Jaeger H, Shürmann D, Cohen-Codar I, Van PN, Norton M, Taburet AM, Delfraissy JF, Rouzioux C; MONARK Study Group. 2009. Protease inhibitor resistance analysis in the MONARK trial comparing first-line lopinavir-ritonavir monotherapy to lopinavir-ritonavir plus zidovudine and lamivudine triple therapy. *Antimicrob. Agents. Chemother*. 53: 2934-9.

69. Delaugerre C, Rohban R, Simon A, Moroux M, Tricot C, Agher R, Huraux JM, Katlama C and Calvez V. 2001. Resistance profile and cross-resistance to HIV-1 among patients failing a non-nucleoside reverse transcriptase inhibitor-containing regimen. *J. Med. Virol*. 65: 445.

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

71. Delelis O, Thierry S, Subra F, Simon F, Malet I, Alloui C, Sayon S, Calvez V, Deprez E, Marcelin AG, Tchertanov L and Mouscadet JF. 2010. Impact of Y143 HIV-1 integrase

mutation in resistance to raltegravir in vitro and in vivo. *Antimicrob. Agents. Chemother*. 54: 491-50.

72. Demaison C, Parsley K, Brouns G, Scherr M, Battmer K, Kinnon C, Grez M and Thrasher AJ. 2002. High level transduction and gene expression in hematopoietic repopulating cells using human immunodeficiency virus type-1 based lentiviral vector containing an internal spleen focus forming virus promoter. *Hum. Gen. Ther*. 13: 803-13.

73. Derdeyn CA, Decker JM, Sfakianos JN, Wu X, O'Brien WA, Ratner L, Kappes JC, Shaw GN and Hunter E. 2000. Sensitivity of human immunodeficiency virus type 1 to the fusion inhibitor T-20 is modulated by coreceptor specificity defined by the V3 loop of gp120. *J. Virol*. 74: 8358-67.

74. Derdeyn CA, Decker JM, Sfakianos JN, Zhang Z, O'Brien WA, Ratner L, Shaw GM and Hunter E. 2001. Sensitivity of human immunodeficiency virus type 1to fusion inhibitors targeted to the gp41 first heptad repeat involves distinct regions of gp41 and is consistently modulated by gp120 interactions with the coreceptor. *J. Virol*. 75: 8605-14.

75. Derooo S, Robert I, Fontaine E, Lambert C, Plesseria J M, Arendt V, Staub T, Hemmer R, Schneider F and Schmit J C. 2002. HIV-1 subtypes in Luxembourg, 1983-2000. *AIDS*. 16: 2461-7.

76. Deval J, White KL, Miller MD, Parkin NT, Courcambeck J, Halfon P, Selmi B, Boretto J and Canard B. 2004. Mechanistic basis for reduced viral and enzymatic fitness of HIV-1 revrse transcriptase containing boith K65R and M184V mutations. *J. Biol. Chem*. 279: 509-16.

77. Devereux H, loveday C, Burke A, Dann L, Johnson M and Phillips A. 1999. The prevalence of non-B subtype HIV-1 in a London HIV/AIDS outpatient clinic. *AIDS*. 13: 142.

78. Devereux, HL, Emery VC, Johnson MA, and Loveday C. 2001. Replicative fitness in vivo of HIV-1 variants with multiple drug resistance associated mutations. *J. Med. Virol*. 65: 218-24.

79. Dierynck I, De Wit M, Gustin E, Keuleers I, Vandersmissen J, Hallenberger S and Hertogs K. 2007. Binding kinetics of darunavir to human immunodeficiency virus type 1 protease explain the potent antiviral activity and high genetic barrier. *J. Virol*. 81: 13845-51.

80. Ding J, Das K, Moereels H, Koymans L, Andries K, Janssen PA, Hughes SH and Arnold E. 1995. Structure of HIV-1 RT/TIBO R 86183 complex reveals similarity in the binding of diverse nonnucleoside reverse transcriptase inhibitors. *Nat. Struct. Biol*. 2: 407-15.

81. Dingwall C, Ernberg I, Gait MJ, Green SM, Heaphy S, Karn J, Lowe AD, Singh M and Skinner MA. 1990. HIV-1 tat protein stimulates transcription by binding to a U-rich bulge in the stem of the TAR RNA structure. *EMBO J*: 4145-54.

82. Domingo E and Holland J. 1997a. RNA virus mutations and fitness for survival. *Ann. Rev. Microbiol*. 51: 151-78.

83. Domingo E, Menendez-Arias L and Holland JJ. 1997b. RNA virus fitness. *Rev. Med. Virol*. 7: 87-96.

84. Dorfman T, Luban J, Goff SP, Haseltine WA and Göttlinger HG. 1993. Mapping of functionally important residues of a cysteine-histidine box in the human immunodeficiency virus type 1 nucleocapsid protein. *J. Virol*. 67: 6159-69.

85. Dorfman T, Mammano F, Haseltine WA and Gottlinger HG. 1994. Role of the matrix protein in the virion association of the human immunodeficiency virus type 1 envelope glycoprotein. *J. Virol*. 68: 1689-1696.

86. Dorr P, Westby M, Doobs S, Griffin P, Irvine B, Macartney M, Mori J, Rickett G, Smith-Bruchnell C, Napier C, Webster R, Armour D, Price D, Stammen V, Wood A and Perros M. 2005. Maraviroc (UK-427,857), a potent, orally bioavailable, and selective small-molecule inhibitor of chemokine receptor CCR5 with broad-spectrum anti-human immunodeficiency virus type 1 activity. *Antimicrob. Agents. Chemother*. 49: 4721-32.

87. Doyon L, Croteau G, Thibetault D, Poulin F, Pilote L and Lamarre D. 1996. Second locus involved in human immunodeficiency virus type 1 resistance to protease inhibitors. *J. Virol*. 70: 3763-9.

88. Draenert R, Allen TM, Liu Y, Wrin T, Chappey C, Verril CL, Sirera G, Eldridge RL, Lahaie MP, Reiz L, Clotet B, Petropoulos CJ, Walker BD and Martinez-Picado J. 2006. Constraints of HIV-1 evolution and immunodominance revelaed in monozygotic adult twins infected with the same virus. *J. Exp Med*. 203 (3): 259-39.

89. Dragic T, Trkola A, Thompson DA, Cormier EG, Kajumo FA, Maxwell E, Lin SW, Ying W, Smith SO, Sakmar TP and Moor JP. 2000. A binding pocket for a small moleculae inhibitor of HIV-1 entry within the transmembrane helices of CCR5. *Proc. Natl. Acad. Sci. U.S.A.* 97: 5639-44.

90. Dronda F, Casado JL, Morenao S, Hertogs K, Garcia-Arata I, Antela A, Perez-Elias MJ, Ruiz L, Larder B; NELSANE Study. 2001. Phenotypic cross-resistance to Nelfinavir: the role of prior antiretroviral therapy and the number of mutations in the protease gene. *AIDS Res Hum Retroviruses.* 17: 93-8

91. Durant J, Clevenbergh P Halfon P, Delgiudice P, Porsin S, Simonet P, Montagne N, Boucher CA, Schapiro JM and Dellamonica P. 1999. Drug-resistance genotyping in HIV-1 therapy: the VIRADAPT randomised controlled trial. *Lancet*. 353: 2195-99.

92. Dykes C. and Demeter LM. 2007. Clinical significance of human immunodeficiency virus type 1 replication fitness. *Clin. Microbiol. Rev*. 20: 550-78.

93. Eastman PS, Mittler J, Kelso R, Gee C, Boyer E, Kolberg J, Urdea M, Leonard JM, Norbeck DW, Mo H and Markowitz M. 1998. Genotypic changes in human immunodeficiency virus type q associated with loss of suppression of plasma viral RNA levels in subjects treated with ritonavir (Norvir) monotherapy. *J. Virol*. 72: 5154-64.

94. Ebbets-Reed D, Scarlata S and Carter CA.1996. The major homology region of the HIV-1 gag precursor influences membrane affinity. *Biochemistry*. 12. 35: 14268-75.

95. Eckert KA and Kunkel TA.1991. DNA polymerase fidelity and the polymerase chain reaction. *Genome. Res*. 1: 17-24.

96. El SY, Vivet-Boudou V and Marquet R. 2007. HIV-1 reverse transcriptase inhibitors. Appl. Microbiol. *Biotechnol*. 75: 723-37.

97. Elzi L, Marzolini C, Furrer H, et al. 2010. Treatment modification in human immunodeficiency virus-infected individuals starting combination antiretroviral therapy between 2005 and 2008. *Arch Intern Med*; 170: 57-65.

98. Emerman M, Vazeux R, Peden K. 1989. The rev gene product of human immunodeficiency virus affects envelope-specific RNA location. *Cell*. 57: 1155-65.

99. Facke M, Janetzko A, Shoeman, R L and Krausslich H G. 1993. A large deletion in the matrix domain of the human immunodeficiency virus gag gene redirects virus particle assembly from the plasma membrane to the endoplasmic reticulum. *J. Virol*. 67:4972-80.

100. Fang G, Zhu G, Burger H, Kethly JS and Weiser B. 1998. Minimizing DNA recombination during long RT-PCR. *J. Virol. Met*. 76: 139-49.

101. Farnet CM and Haseltine WA. 1991. Determination of viral proteins present in human immunodeficiency virus type 1 pre-integration complex. *J.Virol*. 65: 1910-15.

102. Farnet CM and Bushman FF. 1996. HIV cDNA integration and inhibitor development. *AIDS*. 10: S3-S11.

103. Fassati A. 2006. HIV infection of non-dividing cells: a divise problem. *Retrovirology*. 3: 74.

104. Feher A, Weber IT, Bagossi P, Boross P, Mahalingan B, Louis JM, Copeland TD, Torshin IY, Harrison RW and Tozser J.2002. Effect of seqeunce polymorphism and drug resistance on two HIV-1 gag processing sites. *Eur. J. Biochem*. 269: 4114-20.

105. Flexner C. 1998. HIV-1 protease inhibitors. *N. Engl. J. Med*. 338: 1281-1292.

106. Fox J, Castro H, Kaye S, McClure M, Weber J N, Fidler S; UK Collaborative group on HIV Drug resistance. 2010. Epidemiology of non-B clade forms of HIV-1 in men who have sex with men in the UK. *AIDS*. 15: 2397-401.

107. Frahm, N., B. Baker, and C. Brander. 2008. Identification and optimal definition of HIV-derived cytotoxic T lymphocyte (CTL) epitopes for the study of CTL escape, functional avidity and viral evolution, p. 3-24. In B. T. M. Korber, ed. *HIV molecular immunology. Los Alamos National Laboratory, Theoretical Biology and Biophysics, Los Alamos, NM.*

108. Franke E K, Yuan HE, Bossolt KL, Goff SP and Luban J. 1994. Specificity and sequence requirements for interactions between various retroviral Gag proteins. *J. Virol*. 68:5300-5.

109. Fransen S, Gupta S, Danovich R, Hazuda D, Miller M, Witmer M, Petropoulos CJ and Huang W. 2009. Loss of raltegarvir susceptibility of HIV-1 is conferred by multiple non-overlapping genetic pathways. *J. Virol*. 83: 11440-6.

110. Freed EO, Orenstein JM, Buckler-White AJ and Martin MA. 1994. Single amino acid changes in the human immunodeficiency virus type 1 matrix protein block virus particle production. *J. Virol*. 68: 5311-20.

111. Frentz D, Boucher CA, Assel M, De Luca A, Fabbiani M, Incardona F, Libin P, Manca N, Muller V, Nuallain B, Paredes R, Prosperi M, Quiros-Roldan E, Ruiz L, Sloot PM, Torti C, Vandamme AM, Van Laethem K, Zazzi M and Van de Vijver DA. 2010. Comparison of HIV-1 genotypic resistance test interpretation systems in predicting virological outcomes over time. *PLoS. One*. 5: e11505.

112. 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.

113. Gallego O, De Mendoza C, Corral A and Soriano V. 2003. Changes in human immunodeficiency virus P7-P1-P6 gag gene in drug-naïve and pretreated patients. *J. Clin. Microbiol*. 41: 1245-47.

114. Gamble TR, Yoo S, Vajdos FF, von Schwedler UK, Worthylake DK, Wang H, McCutcheon JP, Sundquist WI and Hil CP. 1997. Structure of the carboxyl-terminal dimerization domain of the HIV-1 capsid protein. *Science*. 278:849-853.

115. Gao F, Bailes E, Robertson DL, Chen Y, Rodenburg CM, Michael SF, Cummins LB, Arthur LO, Peeters M, Shaw GM, Sharp PM and Hahn BH. 1999. Origin of HIV-1 in the chimpanzee Pan troglodytes troglodytes. *Nature*. 397: 436-41.

116. Gao WY, Shirasaka T, Johns DG, Broder S Mitsuya H. 1993. Differential phosphorylation of azidothymidine, dideoxycytidine, and dideoxyinosine in resting and activated peripheral blood mononuclear cells. *J. Clin. Invest*. 91: 2326-2333.

117. Garcia JV and Miller AD. 1991. Serine phosporilation-independent downregulation of cell surface CD4 by nef. *Nature*. 350: 508-11.

118. Gatanaga H, Suzuki Y, Tsang H, Yoshimura K, Kaulick NF, Nagashima K, Gorelick RJ, Mardy S, Tang C, Summers MF and Mitsuya H. 2002. Amino acid substitutions in gag protein at non-clevage sites are indispensable for the development of a high multitude of HIV-1 resistance against protease inhibitors. *J. Biol. Chem*. 277: 5952-5961.

119. Gatlin J, Arrigo SJ and Schmidt MG. 1998. HIV-1 protease regulation: the role of the major homology region and adjacent C-terminal capsid sequences. *J. Biomed. Sci*. 5:305-8.

120. Gazzard B, on behalf of the BHIVA treatment guidelines writing group. 2008. British HIV association guidelines for the treatment of HIV-infected adults with antiretroviral therapy. *HIV. Med*. 9: 563-608.

121. Geretti AM, Fox ZV, Booth CL, Smith CJ, Phillips AJ, Johnson M, Li JF, Heneine W and Johnson JA. 2009. Low-frequency K103N strengthens the impact of transmitted drug resistance on virologic responses to first-line efavirenz or nevirapine-based highly active antiretroviral therapy. *J. Acquir. Immune. Defic. Syndr*. 52: 569-73.

122. Geretti AM, Smith C, Haberl A, Garcia-Diaz A, Nebbia G, Johnson M, Phillips A and Staszewski S. 2008. Determinants of virological failure after successful viral load suppression in first-line higly active antiretroviral therapy. *Antivir. Ther*. 13: 927-36.

123. Geretti AM, 2006. HIV-1 subtypes: epidemiology and significance for HIV management. *Curr. Opin. Infect. Dis.* 19:1-7.

124. Ghosn J, Delaugerre C, Flandre P, Galimand J, Cohen-Codar I, Raffi F, Delfraissy JF, Rouzioux C and Chaix ML. 2011. Polymorphims in Gag gene cleavage sites of HIV-1 non B subtype and virological outcome of a first line lopinavir/ritonavir single drug regimen. *PLos One*. 6: e24798.

125. Gifford R, De Oliveira T, Rambaut A, Myers R E, Gale C V, Dunn D, Shafer R, Vandamme AM, Kellam P, Pillay D; UK collaborative Group on HIV Drug resistance. 2006. Assessment of automated genotyping protocols as tools for surveillance of HIV-1genetic diversity. *AIDS*. 13: 152-9.

126. Gorelick RJ, Nigida SM, Bess JW, Arthur LO, Henderson LE and Rein A. 1990. Noninfectious human immunodeficiency virus type 1 mutants defficient in genomic RNA. *J. Virol*. 64: 3207-11.

127. Görlich D, Henklein P, Laskey RA and Hartmann E. 1996. A 41 amino acid motif in importin-alpha confers binding to importin beta and hence transit into the nucleus. *EMBO J*. 15: 1810-7.

128. Göttlinger HG, Dorfman T, Sodroski JG and Haseltine WA. 1991. Effect of mutation affecting the P6 gag protein on human immunodeficiency virus particle release. *Proc.Natl .Acad. Sci.U.S.A.* 88: 3195-9.

129. Gottlinger HG, Sodroski J and Haseltine WA. 1989. Role of capsid precursor processing and myristoylation in morphogenesis and infectivity of human imunodeficiency virus type 1. *Proc. Natl. Acad. Sci.U.S.A.* 86: 5781-5.

130. Graham FL, Smiley J, Russel WC, Nairn R. 1977. Characteristics of a human cell line transformed by DNA from human adenovirus type 5. *J. Gen. Virol*. 36: 59-74.

131. Griffiths JT, Phylip LH, Konvalinka J, Strop P, Gustchina A, Wlodawer A, Davenport RJ, Briggs R, Dunn BM and Kay J. 1992. Different requirements for productive interaction between the active site of HIV-1 proteinase and substrates containing –hydrophobic hydrophobic- or –aromatic pro- cleavage sites. *Biochemistry*. 31:5193-200.

132. Groom HCT, Anderson EC and Lever AM. Rev beyond nuclear export.2009. *J. Gen. Virol.* 90: 1303–18.

133. Gross I, Hohenberg H, Wilk T, Wiegers K, Grattinger M, Muller B, Fuller S and Krausslich HG. 2000. A conformational switch controlling HIV-1 morphogenesis. *Embo J*. 19: 103-13.

134. Gulick RM, Mellors JW, Havlir D, Eron JJ, Gonzalez C, McMahon D, Richman DD, Valentine FT, Jonas L, Meibohm A, Emini EA and Chodakewitz JA. 1997. Treatment indinavir, zidovudine and lamivudine in adults with human immunodeficiency virus infection and prior antiretroviral therapy. *N. Engl. J Med*. 337: 734-9.

135. Gupta R, Hill A, Sawyer AW and Pillay D. 2008. Emergence of drug resistance in HIV type 1-infected patients after receipt of first-line highly active antiretroviral therapy. A sytematic review of clinical trials. *Clin. Infec.t Dis*. 47: 712-22.

136. Gupta R, Kohli A, McCormick A, Towers G, Pillay D and Parry C. 2010. Full length HIV-1 Gag determines protease inhibitor susceptibility within in vitro assays. *AIDS*. 24: 1651–55.

137. Haddad M, Napolitano LA, Paquet AC. 2011. Impact of HIV-1 reverse transcriptase E138 mutations on rilpivirine drug susceptibility. [Abstract 10.] *Antivir Ther*. 16: A18.

138. Haddad M, Stawiski E, Benhamida J, Coakley E. 2010. Improved genotypic algorithm for predicting etravirinbe susceptibility: Comprehensive predicting list of mutations identified through correlation with matched phenotype. [Abstract 574] *17th Conference on Retrovirus and oportunistic infections (CROI). San Francisco, CA.*

139. Hammer SM, Eron JJ Jr, Reiss P, Schooley RT, Thompson MA, Walmsley S, Cahn P, Fischl MA, Gatell JM, Hirsch MS, Jacobsen DM, Montaner JS, Richman DD, Yeni PG and Volberding PA; International AIDS Society-USA. 2008. Antiretroviral treatment of adult HIV infection. 2008 recommendations of the IAS-USA panel. *JAMA*. 300: 555-70.

140. Hammer SM, Squires KE, Hughes MD, Grimes JM, Demeter LM, Currier JS, Eron JJ Jr, Feinberg JE, Balfour HH Jr, Deyton LR, Chodakewitz JA and Fischl MA. 1997. A control trial of two nucleoside analogues plus indinavir in persons with human immunodeficiency virus infection and CD4 cell counts of 200 cubic millimeter or less. AIDS clinical Trials Group 320 Study Team. *N. Engl. J. Med*. 11: 725-33.

141. Hare S, Gupta SS, Valkov E, Engelman A and Cheperanov P. 2010. Retroviral intasome assembly and inhibition of DNA strand transfer. *Nature*. 464:232-6.

142. Harrigan PR, Kinghorn I, Bloor S, Kemp SD, Najera I, Kohli, A and Larder BA. 1996. Significance of amino acid variation at human immunodeficiency virus type 1 reverse transcriptase residue 210 for zidovudine susceptibility. *J. Virol*. 70: 5930-4.

143. Havlir DV, Tierney C, Friedland GH, Pollard RB, Smeaton L, Sommadossi JP, Fox L, Kessler H, Fife KH and Richman DD. 2000. In vivo antagonism with zidovudine and stavudine combination therapy. *J. Infect. Dis*. 182: 321-25.

144. Hazuda DJ, Felock P, Witmer M, Wolfe A, Stillmock K, Grobler JA, Espeseth A, Gabryelski L, Schleif W, Blau C and Miller MD. 2000. Inhibitors of strand transfer that prevent integration and inhibit HIV-1 replication in cells. *Science*. 287: 646-50.

145. Hemelaar J. 2012. The origin and diversity of the HIV-1 pandemic. *Trends. Mol. Med*.18: 182-92.

146. Henderson LE, Benveniste RE, Sowder R, Copeland TD, Schultz AM and Oroszlan S. 1988. Molecular characterization of gag proteins from simian immunodeficiency virus (SIVMne*). J. Virol*. 62: 2587-95.

147. Henderson LE, Bowers MA, Sowder RC, Serabyn SA, Johnson DG, Bess JW, Arthur LO Jr, Bryant DK and Fenselau C. 1992. Gag proteins of the higly replicative MN strains oh human immunodeficiency virus type 1: posttranslational modifications, proteolytic processing, and complete amino acid sequences. *J. Virol*. 66: 1856-1865.

148. Henriet S, Mercennet G, Bernacchi S, Paillart J-C and Marquet R. 2009. Tumultuous relationship between the human Immunodeficiency Virus type 1 viral infectivity factor (Vif) and the human APOBEC-3G and APOBEC-3F restriction factors. *Microbiol. Mol. Biol. Rev*. 73: 211-32.

149. Hermida-Matsumoto L and Resh MD. 1999. Human immunodeficiency virus type 1 protease triggers a myristoyl switch that modulates membrane binding of Pr55(gag) and P17MA. *J. Virol*. 73: 1902-8.

150. Hirsh MS, Gunthard HF, Schapiro JM, Brun-Vezinet F, Clotet B, Hammer SM, Johnson VA, Kuritzkes DR, Mellors JW, Pillay D, Yeni PG, Jacobsen DM and Richman DD. 2008. Antiretroviral drug testing in adult HIV-1 infection. Recommendations of an international AIDS Society Panel. *Clin. Infect. Dis*. 47: 266-85.

151. HIV in the United Kingdom: 2011 report. *Health Protection Agency*. Nov. 2011.

152. Ho D D, Toyoshima T, Mo H, Kempf D J, Norbeck D, Chen CM, Wideburg NE, Burts, SK, Erickson JW and Singh M K. 1994. Characterization of human immunodeficiency virus type 1 variants with increased resistance to a C2-symmetric protease inhibitor. *J. Virol*. 68: 2016-20.

153. Ho DD, Neumann AU, Perelson AS, Chen J, Leonard JM and Markpwitz M. 1995. Rapid turnover of plasma virions and CD4 lymphocytes in HIV-1 infections. *Nature*. 373: 123-6.

154. Hogg RS, Bangsberg DR, Lima VD, Alexander C, Bonner S, Yip B, Wood E, Dong WW, Montaner JS and Harrigan PR.2006. Emergence of drug resistance is associated with an increased risk of death among patients first starting HAART. *PloS Med*. 3: e356.

155. Holland JJ, De la Torre JC, Clarke DK, and Duarte E. 1991. Quantitations of relative fitness and great adaptability of clonal populations of RNA viruses. *J. Virol*. 65: 1960-67.

156. Hu ZF, Giguel H, Hatano P, Reid J, Lu, and Kuritzkes DR. 2006. Fitness comparison of thymidine analog resistance pathways in human immunodeficiency virus type 1. *J. Virol*. 80: 7020-7.

157. Huang M, Orenstein JM, Martin MA and Freed EO. 1995. P6Gag is required for particle production from full-length human immunodeficiency virus type 1molecular clones expressing protease. *J. Virol*. 69: 6810-8.

158. Japour AJ, Mayers DL, Johnson VA, Kuritzkes DR, Beckett LA, Arduino JM, Lane J, Blacj RJ, Reichelderfer PS and D'Aquila RT. 1993. Standardized peripheral blood mononuclear cell clture assay for determination of drug susceptibilities of clinical immunideficiency virus type 1 isolates. The RV-43 Study Group, the AIDS Clinical Trials Group Virology Committe Resistance Working Group. *Antimicrob. Agents. Chemother*. 37: 1095-101.

159. Javanbakht H, Diaz-Griffero F, Yuan W, Yeung DF, Li X, Song B and Sodroski J. 2007. The ability of multimerized cyclophilin A to restrict retrovirus infection. *Virology*. 367: 19-29

160. Jegede O, Babu J, Di Santo R, McColl J, Weber J and Quinones-Mateu M. 2008. HIV type 1 integrase inhibitors: from basic research to clinical implications. AIDS Review. 10: 172-189.

161. Jetzt AE, Yu H, Klarmann GJ, Ron Y, Preston BD and Dougherty JP. 2000. High rate of recombination throughout the human immunodeficicency virus type 1 genome. *J. Virol*. 74: 1234-40.

162. Ji JP and Loeb LA. 1992. Fidelity of HIV-1 reverse transcriptase copying RNA in vitro. *Biochemistry*. 31: 954-8.

163. Jinnopat P, Isarangkura-na-ayuthaya P, Utachee P, Kitagawa Y, de Silva UC, Siripanyaphinyo U, Kameoka Y, Tokunaga K, Sawanpanyalert P, Ikuta K, Auwanit W and Kameoka M. 2009. Impact of amino acid variations in Gag and protease of HIV type 1 CRF01_AE strains on drug susceptibility of virus to protease inhibitors. *J. Acquir. Immune. Defic. Syndr*. 52: 320-8.

164. Johnson JA, Li JF, Wei X, Lipscomb J, Irlbeck D, Craig C, Smith A, Bennett DE, Monsour M, Sandstrom P, Lanier ER and Heneine W. 2008a. Minority HIV-1 drug resistance mutations are present in antiretroviral treatment naive populations and assocaited with reduced treatment efficacy. *PLoS Med*. 5: e158.

165. Johnson VA, Brun-Vezinet F, Clotet B, Gunthard HF, Kuritzkes DR, Pillay D, Schapiro JM and Richman DD. 2008b. Update of the drug resistance mutations in HIV-1. *Top. HIV Med*. 16: 138-145.

166. Jones NA, Wei X, Flower DR, Wong M, Michor F, Saaq MS, Hahn BH, Nowak MA, Shaw GM and Borrow P. 2004. Determinants of human immunodeficiency virus type 1 escape from primary CD8 cytotoxiv T lymphocyte responses. *J.Exp. Med*. 200: 1243-56.

167. Kaplan AH, Zack JA, Knigge M, Pau DA, Kempf DJ, Norbeck DW and Swanstrom R. 1993. Partial inhibition of the human immunodeficiency virus type 1 protease results in aberrant virus assembly and the formation of noninfectious particles. *J. Virol*. 67:4050-5.

168. Katzman DJ,Odirizzi G and Erm SD. 2002. Receptor downregulation and multi-vesicular body sorting. *Nat. Rev. Mol. Cell. Biol*. 3: 893-905.

169. Kauffman GR, Suzuki K, Cunningham P, Mukaide M, Kondo M, Jnai M, Zaunders J and Cooper DA. 2001. Impact of HIV type 1 protease, reverse transcriptase, clevage site, and p6 mutations on the virological response to quadruple therapy with saquinavir, ritonavir and two nucleoside analogs. AIDS Res Hum. *Retroviruses*. 17: 1293-6.

170. Keele BF, Van Heuverswyn F, Li Y, Bailes E, Bibollet-Ruche F, Chen Y, Wain LV, Liegeois F, Loul S, Ngole EM, Bienvenue Y, Delaporte E, Brookfield JF, Sharp PM, Shaw GM, Peeters M and Hahn BH. 2006. Chimpanzee reservoirs of pandemic and nonpandemic HIV-1. *Science*. 313: 523-26.

171. Kellam P, Boucher CA and Larder BA. 1992. Fifth mutation in human immunodeficiency virus type 1 reverse transcriptase contributes to development of high-level resistance to zidovudine. *Proc. Natl. Acad. Sci. U.S.A.* 89: 1934-8.

172. Kempf DJ, King MS, Bernstein B, Cernohous P, Bauer E, Moseley J, Gu K, Hsu A, Brun S and Sun E. 2004.. Incidence of resistance in a double-blind study comparing lopinavir/ritonavir plus stavudine and lamivudine to nelfinavir plus stavudine and lamivudine. *J. Infect. Dis*. 189: 51-60.

173. King NM, Prabu-Jeyabalan M, Nalivaika EA, Wigerinck P, de Béthune MP and Schiffer CA. 2004. Structural and thermodynamic basis for the binding of TMC114, a next generation human immunodeficiency virus type 1 protease inhibitor. *J. Virol*. 78: 12012-21.

174. Kim EY, Bhattacharya T, Kuntsman K, Swantek P, Koning FA, Malim MH, Wolinsky SM. 2010. Human APOBEC3G-mediated editing can promote HIV-1 sequence diversification and accelerate adaptation to selective pressure. *J Virol* 84: 10402-10405.

175. Klimkait T, Strebel K, Hoggan MD, Martin MA and Orenstein JM. 1990. The human immunodeficiency virus type 1 specific protein vpu is required for efficient maturation and release. *J. Virol*. 64: 621-9.

176. Kobayashi M, Nakahara K, Seki T, Miki S , Kawauchi S, Suyama S, Wakasa-Morimoto C, Kodama M, Endoh T, Oosugi E, Matsushita Y, Murai H, Fujishita T, Yoshinaga T, Garvey E,. Foster D, Underwood M, Johns B, Sato A, and Fujiwara T. 2008. Selection of diverse and clinically relevant integrase inhibitor-resistant human immunodeficiency virus type 1mutants. *Antiviral Res*. 80: 213-222.

177. Koch N, Yahi N, Fantini J and Tamalet C. 2001. Mutations in HIV-1 gag cleavage sites and their association with protease mutations. *AIDS*. 15: 526-8.

178. Kohl NE, Emini EA, Schleif WA, Davis LJ, Heimbach JC, Dixon RA, Scolnick EM and Siga I S. 1988. Active human immunodeficiency virus protease is required for viral infectivity. *Proc. Natl. Acad. Sci. U. S. A.* 85:4686-4690.

179. Kohlstaedt LA, Wang J, Friedman JM, Rice PA, and Steitz TA. 1992. Crystal structure at 3.5 A resolution of HIV-1 reverse transcriptase complexed with an inhibitor. *Science* 256: 1783-90.

180. Kondo E and Göttlinger HG. 1996. A conserved LXXLF sequence is the major determinant in p6Gag required for the incorporation of human immunodeficiency virus tyep a Vpr. *J. Virol*. 170: 159-64.

181. Kondo E, Mammano F, Cohen EA and Göttlinger HG.1995. The p6Gag domain of human immunodeficiency virus type 1 is sufficient for the incorporation of Vpr into heterologous viral particles. *J. Virol*. 69: 2759-64.

182. Korber B, Nuldoon M, Theiler J, Gao F, Gupta R, Lapedes A, Hahn BH, Wolinsky S, Bhattacharya T. 2000. Timing the ancestor of the HIV-1 pandemic strains. *Science*. 288: 1789-96.

183. Kräusslich HG, Ingraham RH, Skoog MT, Wimmer E, Pallai PV and Carter CA. 1989. Activity of purified biosynthetic proteinase of human immunodeficiency virus on natural substrates and synthetic peptides. *Proc. Natl. Acad. Sci. U.S.A*. 86: 807-811.

184. Kuhmann S and Moore JP. 2005. The HIV-1 phenotypic variants: deadly and deadlier. *J Viral Entry*. 1: 4-16.

185. Mo H, King MS, King K, Molla A, Brun S and Kempf DJ. 2005. Selection of resistance in protease inhibitor-experienced, human immunodeficiency virus type 1-infected subjects failing lopinavir-and ritonavir-based therapy: mutation patterns and baseline correlates. *J. Virol*. 79: 3329-38.

186. Larder BA 1995. Viral resistance and the selection of antiretroviral combinations. *J. Acquir Immune Defic Syndr Hum Retrovirol*. 10: S28.

187. Larder, BA. 1992. 30-Azido-30-deoxythymidine resistance suppressed by a mutation conferring human immunodeficiency virus type 1 resistance to nonnucleoside reverse transcriptase inhibitors. *Antimicrob. Agents Chemother*. 36: 2664-9.

188. Larder BA, Hertogs K, Bloor S, van den Eynde CH., DeCian W, Wang Y, Freimuth WW and Tarpley G. 2000. Tipranavir inhibits broadly protease inhibitor-resistant HIV-1 clinical samples. *AIDS*. 14: 1943-8.

189. Laspia M.F, Rice A.P and Mathews M.B. 1990. Synergy between HIV-1 Tat and adenovirus E1A is principally due to stabilization of transcriptional elongation. *Genes Dev*. 4: 2397-408.

190. Lathouwers E, De Meyer S, Dierynck I, Van de Casteele T, Lavreys L, de Béthune MP and Picchio G. 2011. Virological characterization of patients failing darunavir/ritonavir or lopinavir/ritonavir treatment in the ARTEMIS study: 96 weeks analysis. *Antivir Ther*. 16: 99-108.

191. Learn GH, Korber BT Jr, Foley B, Hahn BH, Wolinsky SM and Mullins JI.1996. Maintaining the integrity of human immunodeficicency virus sequence databases. *J. Virol*. 70: 5720-30.

192. Lecossier D, Shulman NS, Morand-Joubert L, Shafer RW, Joly V, Zolopa AR, Clavel F and Hance AJ. 2005. Detection of minority populations of HIV-1 expressing the K103N resistance mutation in patients failing nevirapine. *J. Acquir Immune Defic Syndr*. 38:37.

193. Lee K, Ambrose Z, Martin TD, Oztop I, Mulky A, Julias JG, Vandegraff N, Baumann JG, Wang R, Yuen W, Takemura T, Shelton K, Taniuchi I, Li Y, Sodroski J, Littman DR, Coffin JM, Hughes SH, Unutzan D, Engelman A and KewalRamani VN. 2010. Flexible use of nuclear import pathways by HIV-1. *Cell Host and Microbe*. 7: 221-3.

194. Levesque K, Zhao YS and Cohen EA. 2003. Vpu exerts a positive effect on HIV-1 infectivity by downmodulating CD4 receptor molecules at the surface of HIV-1 producing cells. *J. Biol. Chem*. 278: 28346-53.

195. Levy JA, Hoffman DA, Kramer SM, Landis JA, Shimabukuro JM Aand Oshiro LS. 1984. Isolation of cytopathic retroviruses from San Francisco patients with AIDS. *Science*. 4664 24: 840-42.

196. Loeb DD, Hutchison CA, Edgell MH, Farmerie WG and Swanstrom R. 1989. Mutational analysis of human immunodeficiency virus type 1 protease suggests functional homology with aspartic proteinases. *J. Virol*. 63:111-21.

197. Logsdon BC, Vickrey JF, Martin P, Proteasa G, Koepke JI, Terlecky SR, Wawrzak Z, Winters MA, Merigan TC and Kovari LC. 2004. Crystal structure of a multi-drug resistant human immunodeficiency virus type 1 protease reveals an expanded active-site cavity. *J. Virol*. 78: 3123-32.

198. Lospistao E, Alvarez A, Soriano V and Holguin A. 2005. HIV-1 subtypes in Spain: a retrospective analysis from 1995 to 2003. *HIV Med*. 6: 313-30.

199. Louis JM, McDonald RA, Nashed NT, Wondrak EM, Jerina DM, Oroszlan S and Mora, PT. 1991. Autoprocessing of the HIV-1 protease using purified wild-type and mutated fusion proteins expressed at high levels in Escherichia coli. *Eur. J. Biochem*. 199:361-9.

200. Louwagie J, McCutchan F.E, Feeters M, Brennan T.P, Saunders-Buell E, Eddy G.A, Van der Groen G, Fransen K, Gershy-Damet GM, Deleys R and Burke D.S. 1993. Phylogenetic analysis of gag genes from 70 international HIV-1 isolates provides evidence for multiple genotypes. *AIDS*. 7: 769-80.

201. Lu, J., Sista, P., Giguel, F., Greenberg, M., and Kuritzkes, D. R. 2004. Relative replicative fitness of human immunodeficiency virus type 1 mutants resistant to enfuvurtide (T-20). *J. Virol*. 78: 4628-4637.

202. Lu YL, Bennett RP, Wills JW, Gorelick R and Ratner L. 1995. A leucine triplet repeat sequence (LXX)4 in p6gag is important for Vpr incorporation into human immunodeficiency virus type 1 particles. *J. Virol*. 69: 6873-9.

203. Lundberg KS, Shoemaker D.D, Adams MW, Short JM, Sorge JA and Mathur EJ. 1991 High fidelity amplification using a thermostable DNA polymerase isolated from Pyrococcus furiosus. *Gen*. 108: 1-6.

204. MacArthur RD and Novak RM. 2008. Maraviroc: The first of a new Class of Antiretroviral Agents. *Clin. Infect. Dis*. 47:236–41.

205. Machuca R, Bogh M, Salminen M, Gerstoft J, Kuinesdal B, Pedersen C, Obel N, Nielsen H and Nielsen C. 2001. HIV-1 subtypes in Denmark. *Scand J Infect Dis*. 33: 697-701.

206. Mackie EN, Dustans S, McClure MJ, Weber JN and Clartre JR. 2004. Detection of HIV-1 antiretroviral drug resistance from patients with persistently low but detectable viremia. *J. Virol. Met*. 119: 73-81.

207. Mackie EN, Phillips AN, Kaye S, Booth C and Geretti AM on behalf of the UK HIV drug resistance database and the UK collaborative HIV cohort study. 2010. Antiretroviral drug resistance in HIV-1infected patients at low-level viremia. *J. Infect. Dis*. 201: 1303-7.

208. Maguire MF, Guinea R, Griffin O, Macmamus S, Elston RC, Wolfram J, Richards N, Hanlon MH, Portor DJ, Wrin T, Parkin N, Tisdale M, Furfine E, Petropoulos C, Snowdon BW and Kleim JP. 2002. Changes in human immunodeficiency virus type 1 gag at positions L449 and P453 are linked to I50V protease mutants in vivo and cause reduction of sensitivity to amprenavir and improved viral fitness in vitro. *J.Virol*. 76: 7398-406.

209. Malet I, Roquebert B, Dalban C, Wirden M, Annellel B, Agher R, Simon A, Katlama C, Costangliola D, Calvez V and Marcelin AG. 2007. Association of gag clevage site to protease mutations and to virological response in HIV-1 treated patients. J. Infect. 54: 367-74.

210. Malim MH, Emerman M. 2008. HIV-1 accessory proteins-ensuring viral survival in a hostile environment. *Cell Host Microb*. 3: 338-98.

211. Malim MH, Hauber J, Les Y, Maizel JV and Cullen BR. 1989. The HIV-1 rev trans-activator acts through a structured target sequence to activate nuclear export of unspliced viral mRNA. *Nature*. 338: 254-7.

212. Mammano F, Ohagen A, Höglund S and Göttlinger HG. 1994. Role of the major homology region of human immunodeficiency type 1 in virus morphogenesis. *J. Virol*. 68: 4927-36.

213. 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.

214. Maniatis T, Fritsch FF and Sambrook J. 1986. *Molecular cloning a laboratory manual. Cold Spring Harbor Laboratory.*

215. Marcelin AG, Flandre P, Pavie J, Schmidely N, Wirden M, Lada O, Chiche D, Molina JM, and Calvez V. 2005. Clinically relevant genotype interpretation of resistance to didadosine. *Antimicrob. Agents Chemother*. 49: 1739-44.

216. Marciniak RA, Calnan BJ, Frankel AD and Sharp PA.1990. HIV-1 tat protein transactivates transcription in vitro. *Cell*. 16: 63: 791-802.

217. Marinello J, Marchand C, Mott BT, Bain A, Thomas CJ, and Pommier Y. 2008. Comparison of realtegravir and elvitegravir on HIV-1 integrase catalytic reactions and on series of drug-resistant integrase mutants. *Biochemistry*. 47: 9345-54.

218. Markham RB, Yu X, Farzagedan H, Ray SC and Vlahov D. 1995. Human immunodeficiency virus type 1 env and P17 gag sequence variation in polymerase chain reaction-positive, seronegative injection drug users. *J. Infect. Dis*. 171: 797-804.

219. Marshall HM, Ronen K, Berry C, Llano M, Sutherland H, Saenz D, Bickmore W, Poeschla E and Bushman FD. 2007. Role of PSIP1/LEDG/p75 in lentiviral infectivity and integration targeting. *Plos One*. 19: e1340.

220. Martinez-Cajas JL, Wainberg M, Oliveira M, Asahchop EL, Doualla-Bell F, Lisovsky I, Moisi D, Mendelson E, Grossman Z and Brenner BG. 2012. The role of polymorphism at posistion 89 in the HIV-1 protesae gene in the development of drug resistance to HIV-1 protease inhibitors. *J. Antimicrob. Chemother*. 67: 988-94.

221. Martinez-Picado J, Savara AV, Sutton L and D'Aquila RT. 1999. Replicative fitness of protease inhibitor-resistant mutants of human immunodeficicency virus type 1. *J. Virol*. 73: 3744-52.

222. Mathers BM, Degenhardt L, Phillips B, Wiessing L, Hickman M, Strathdee SA, Wodak A, Panda S, Tyndall M, Toufik A and Mattick RP.2008. Global epidemiology of injecting drug use and HIV among people who inject drugs: a sytematic review. *Lancet*. 372: 1733-45.

223. McColl DJ and Chen X. 2010. Strand transfer inhibitors of HIV-1 integrase: brining a new era of antiretroviral therapy. *Antivir. Res*. 85:101-18.

224. Melikyan GB, Markosyan RM, Hemmati H, Delmedico MK, Lambert DM and Cohen FS. 2000. Evidence that the transition of HIV-1 gp41 into a six-helix bundle, not the bundle configuration, induces membrane fusion. *J. Cell. Biolog*. 151: 413-23.

225. Mervis RJ, Ahmad N, Lillehoj EP, Raum MG, Salazar FH, Chan HW, and Venkatessan S. 1988. The gag gene products of human immunodeficiency virus type 1: alignment within the gag open eading frame, identification of postranslational mdifications and evidence for an alternative gag precursors. *J. Virol*. 62: 3993-4002.

226. Metzner KJ, Giulieri SG, Knoepfel SA, Rauch P, Burgisser P, Yerly S, Gunthard HF and Cavassini M. 2009. Minority quasispecies of drug resistant HIV-1 that lead to early therapy failure in treatment-naive and-adherence patients. *Clin Infect Dis*. 48: 239-47.

227. Meyer PR, Matsuura SE, Mian AM, So AG., and Scott WA. 1999. A mechanism of AZT resistance: An increase in nucleotide-dependent primer unblocking by human HIV-1 reverse transcriptase. *Mol. Cell. Biol*. 4: 35-43.

228. Meyerhans A, Cheyner R, Albert J, Seth M, Kwok S, Sninsky J, Morfeldt-Manson L, Asjo B and Wain-Hobson S. 1989. Temporal fluctuations in HIV qusispecies in vivo are not reflected by sequential isolations. *Cell*. 58: 901-10.

229. Miller V. 2001. International perspectives on antiretroviral resistance. Resistance to protease inhibitors. *J. Acquir Immune Defic. Syndr*. 26: S34-50.

230. Miller, MD. 2004. K65R, TAMs and tenofovir. *AIDS Rev*. 6:22-33.

231. Miller MD, Farnet CM and Bushman FD. 1997. Human immunodeficiency virus type 1 pre-integration complexes: studoes of organization and composition. *J. Virol*. 71: 5382-90.

232. Misumi S, Inoue M, Dochi T, Kishimoto N, Hasegawa N, Takamune N and Shoji S. 2010. Uncoating of HIV-1 requires prolyl isomerase Pin 1. *J Biol Chem*. 285: 25185-95.

233. Mo H, King MS, King K, Molla A, Brun S and Kempf DJ. 2005. Selection of resistance in protease inhibitor-experienced, human immunodeficiency virus type 1-infected subjects failing lopinavir-and ritonavir-based therapy: mutation patterns and baseline correlates. *J Virol*. 79: 3329-38.

234. Molla A, Korneyeva M, Gao MQ, Vasavanonda S, Schipper PJ, Mo HM, Markowitz M, Chernyavskiy T, Niu P, Lyons N, Hsu A, Granneman GR, Ho GR, Boucher CA, Leonard JM; Morbeck DW and Kempf DJ. 1996. Ordered accumulation of mutations in HIV protease confers resistance to ritonavir. *Nat. Med*. 2: 760-6.

235. Mulder-Kapinga GA, Simonon A, Kuiken CL, Dekker J, Scherpbier HJ, Van Deferre p, Boer K and Goudsmit J. 1995. Similarity in env and gag genes between genomic RNAs of human immunodeficiency virus type 1 (HIV-1) from mother and infant is unreleated to time of HIV-1 RNA positivity in the child. *J. Virol*. 69: 2285-96.

236. Mullan B, Kenny-Walsh E, Collins JK, Shanahan F and Fanning LJ. 2001. Inferred hepatitis C Virus Quasispecies Diversity is influenced by choice of DNA polymerase in reverse-transcriptase-polymerase chain reaction. *Analytical Biochemistry*. 289: 137-46.

237. Myint L, Matsuda M, Matsuda Z, Yokomaku Y, Chiba T, Okano A, Yamada K and Sugiura W. 2004. Gag non-cleavage site mutations contribute to full recovery of viral fitness in protease inhibtor-resistant human immunodeficiency virus type 1. *Antimicrob Agents Chemother*. 48: 444-52.

238. De Oliveira T, Engebrecht S, Janse Van Rensburg E, Gardon M, Bishop K, Zur Megede J, Barnet SW and Cassol S. 2003. Variability at human immunodeficiency virus type 1 subtype C protease clevage sites: an indication of viral fitness?. *J. Virol*. 77: 9422-30.

239. Naldini L, Blomer U, Galley P, Ory D, Mulligan R, Gage FH, Verma IM and Trono D. 1996. In vivo gene delivery and stable transduction of nondividing cells by a lentiviral vector. *Science*. 272: 263-7.

240. Neil SJ, Zang T, Bieniasz PD. 2008. Tetherin inhibits retrovirus release and is antagonized by HIV-1 vpu. *Nature* 451: 425-430.

241. Nijhuis M, Deeks S and Bouder C. 2001. Implications of antiretroviral resistance on viral fitness. *Curr. Opin Infect Dis*. 14: 23-8.

242. Nijhuis M, Schuurman R, De Jong D, Erickson J, Gustchina E, Albert J, Schipper p, Gulnik S and Boucher CA. 1999. Increased fitness of drug resistant HIV-1 protease as a result of acquisition of compensatory mutations DURING suboptimal therapy. *AIDS.* 13: 2349-59.

243. Nijhuis M, Van Maarseveen NM, Lastere S, Schipper P, Coakley E, Glass B, Rovenska M, De Jong D, Chappey C, Goedegebuure IW, Heilek-Sneyder G, Dulude D, Cammack N, Brakier-Gingras L, Kovalinka J, Parkin N, Krausslich HG, Brun-Vezinet F and Boucher CA. 2007. A novel substrate-based HIV-1 protease inhibitor drug resistance mecahnism. *PLoS Med*. 4: 636.

244. Nijhuis M, Wensing AM, Biermes WF, De Jong D, Kagan W, Fun A, Jasper CA, Schurink KA, Van Agtmael MA and Boucher CA. 2009. Failure of treatment with first line lopinavir boosted with ritonavir can be explained by novel resistance pathways with protease mutation L76V. *J. Infect. Dis*. 200: 698-709.

245. Nikolenko GN, Delviks-Frankenberry KA, Palmer S, Maldarelli F, Fivash MJ Jr, Coffin JM and Pathak VK. 2007. Mutations in the connection domain of HIV-1 reverse transcriptase increase 30-azido-30-deoxythymidine resistance. *Proc. Natl. Acad. Sci. U.S.A.* 104: 317-22.

246. Nikolenko GN, Palmer S, Maldarelli F, Mellors JA, Coffin JM, and Pathak VK. 2005. Mechanism for nucleoside analog-mediated abrogation of HIV-1 replication: balance between Rnase H activity and nucleotide excision. *Proc. Natl. Acad. Sci. U.S.A.* 102: 2093-98.

247. Nitahara-Kasahara Y, Kamata M, Yamamoto T, Zhang X, Miyamoto Y, Muneta K, Iijima S, Yoneda Y, Tsunetsugu-Yokota Y and Aida Y. 2007. Novel nuclear import of Vpr promoted by importin alpha is crucial for human immunodeficiency virus type 1 replication in macrophages. *J. Virol*. 81: 5284-93.

248. Nomaguchi M, Fujita M and Adachi A. 2008. Role of HIV-1 vpu protein for virus spread and pathogenesis. *Microb and Infect*. 10: 960-67.

249. Nunberg JH, Chleif WA, Boots EJ, O'Brien JA, Quintero JC, Hoffman JM, Emini EA and Goldman ME. 1991. Viral resistance to human immunodeficiency virus type 1 specific pyrimidone reverse transcriptase inhibitors. *J. Virol*. 65: 4887.

250. O'Connor DH, McDermott AB, Krebs KC, et al. 2004. A dominant role for CD8þ-Tlymphocyte selection in simian immunodeficiency virus sequence variation. *J. Virol*. 78: 14012-22.

251. Op de Coul E L, Coutinho R A, Van der Schoot A, Van Doornum G J, Lukashov V V, Goudsmit J and Cornelissen M. 2001. The impact of immigration on env HIV-1 subtype distribution among heterosexual in the Netherlands: influx subtype B and non-B strains. *AIDS.* 15: 2277-86.

252. Orozslan, S. and Luftig, R. B. 1990. Retroviral proteinases. *Curr. Top. Microbiol. Immunol.* 157:153-85.

253. Palmer S, Kearney M, Malderelli F, Halvas EK, Bixby CJ, Bazmi H, Rock D, Falloon J, Davey RT Jr, Dewar RL, Metcalf JA, Mellors JW and Coffin JM. 2005. Multiple, linked human

immunodeficiency virus type 1 drug resistance mutations in treatment-experienced patients are missed by standard genotypes analysis. *J. Clin Microbiol*. 43: 406-13.

254. Panel on antiretroviral guidelines for adults and adolescents. Guidelines for the use of antiretroviral agents in HIV-1 infected adults and adolescents. 2008. *US Department of Health and human services. Http://www. Aidsinfo.nih.gov/contentfiles/adultandadolescentGL.pdf.*

255. Paredes R, Lalama CM, Ribaudo HJ et al. 2010. Pre-existing minority drug-resistant variants, adherence and risk of antiretroviral treatment failure. *J Infect Dis*; 201: 662-71.

256. Parker WB, White EL, Shaddix SC, McCaleb G, Secrist JA, Vince R, Shannon WM. 1991. Mechanism of inhibition of human immunodeficiency virus type 1 reverse transcriptase and human DNA polymerases alpha, beta, and gamma by the 5'-triphosphates of carbovir, 3'-azido-3'-deoxithymidine, 2',3'-dideoxiguanosine and 3'-deoxythimidine. A novel RNA template for the evaluation of antiretroviral drugs. *J Biol Chem*. 266: 1754-62.

257. Parry CM, Kohli A, Boinett CJ, Towers GJ, McCormick AL and Pillay D. 2009. Gag determinants of fitness and drug susceptibility in protease inhibitor-resistant human immunodeficiency virus type 1. *J. Virol*. 83: 9094-101.

258. Patick AK and Potts KE. 1998. Protease inhibitors as antiviral agents. *Clin Microbiol Rev*. 11: 614-627.

259. Paulus C, Hellebrand S, Tessmer U, Wolf H, Kräusslich HG and Wagner R. 1999. Competitive inhibition of human immunodeficiency virus type 1 protease by the Gag-pol transframe protein. *J. Biol. Chem*. 274: 21539-43.

260. Paxton WA, Martin SR, Tse D, O'Brien TR, Skurnick J, VanDevanter NL, Padian N, Braun JF, Kotler DP, Wolinsky SM and Koup RA. 1996. Relative resistance to HIV-1 infection of CD4 lymphocytes from persons who remain uninfected despite multiple high-risk sexual exposures. *Nat. Med*. 2: 412-7.

261. Pelemans H, Esnouf RM, Parniak MA, Vandamme AM, De Clercq E and Balzarini J. 1998. A proline-to-histidine substitution as position 225 of human immunodeficiency virus type 1 (HIV-1) reverse transcriptase (RT) sensitizes HIV-1 RT to BHAP U-90152. *J. Gen. Virol. 79:1347.*

262. Peng C, Ho BK, Chang TW, and Chang NT. 1989. Role of human immunodeficiency virus type 1-specific protease in core protein maturation and viral infectivity. *J. Virol*. 63:2550-6.

263. Perandin F, Pollara C.P, Gargulo F, Bonfauti C and Manca N. 2009. Performance evaluation of the automated Nuclisens easyMag nucleic acid extraction platform in comparison with QIAamp Mini Kit from clinical specimens. *Diagnostic Microbiol. Infect. Dis*. 64: 158-65.

264. Perelson AS, Neumann AU, Markowitz M, Leonard JM, and Ho DD. 1996. HIV-1 dynamics in vivo: virrion clearance rate, infected cell life-span, and viral generation time. *Science*. 271: 1582-86.

265. Petropoulos CJ, Parkin NT, Limoli KL, Lie YS, Wrin T, Huang W, Tian H, Smith D, Winslow GA, Capon DJ and Whitcomb JM. 2000. A novel phenotypic drug susceptibility assay for human immunodeficiency virus type 1. *Antimicrob Agents Chemother*. 44: 920-8.

266. Pettit SC, Henderson GJ, Schiffer CA and Swanstrom R. 2002. Replacement of the P1 amino acid of human immunodeficiency virus type 1 gag processing sites can inhibit or enhance the rate of cleavage by the viral protease. *J. Virol*. 76: 10226-33.

267. Pettit SC, Moody MD, Wehbie RS, Kaplan AH, Nantermet PV, Klein CA and Swanstrom R. 1994. The p2 domain of human immunodeficiency virus type 1 Gag regulates sequential proteolytic processing and is required to produce fully infectious virions. *J. Virol*. 68: 8017-27.

268. Pettit SC, Simsic J, Loeb DD, Everitt L, Hutchison CA and Swanstrom, R. 1991. Analysis of retroviral protease cleavage sites reveals two types of cleavage sites and the structural requirements of the P1 amino acid. *J. Biol. Chem*. 266:14539-47.

269. Pettit SC, Everitt L E, Choudhury S, Dunn BM and Kaplan A H. 2004. Initial Cleavage of the Human Immunodeficiency Virus Type 1 GagPol Precursor by Its Activated Protease Occurs by an Intramolecular Mechanism. *J. Virol*. 78:8477-85.

270. Phillips RE, Rowland-Jones S, Nixon DF, Gotch FM, Edwards JP, Ogunlesi AO, Elvin JG, Rothbard JA, Bangham CR, Rizza CR. 1991. Human immunodeficiency virus genetic variation that can escape cytotoxic T cell recognition. *Nature*. 354: 453-9.

271. Phylip LH, Griffiths JT, Mills JS, Graves MC, Dunn BM and Kay J. 1995. Activities of precursor and tethered dimer forms of HIV proteinase. *Adv. Exp. Med. Biol*. 362:467-72.

272. Plantier JC, Leoz M, Dickerson JE, De Oliveira F, Cordonnier F, Lemee C, Damond F, Robertson DL and Simon F. 2009. A new human immunodeficiency virus derived from gorillas. *Nat Med*. 15: 871-2.

273. Polyak SJ, Sullivan DG, Austin MA, Dai JY, Shuhart MC, Lindsay KL, Bonkousky HL, Di Bisceglie AM, Lee WM, Morrishima C, Gretch DR and the HALT-C Trial group. 2005. Comparison of amplification enzymes for hepatitis C virus quasispecies analysis. *Virology J*. 2: 41.

274. Pommier Y, Johnson AA, and Marchand C. 2005. Integrase inhibitors to treat HIV/AIDS. *Nat. Rev. Drug Disc*. 4: 236-48.

275. Poorman RA, Tomasselli AG, Heinrikson RL and Kezdy FJ. 1991. A cumulative specificity model for proteases from human immunodeficiency virus types 1 and 2, inferred from statistical analysis of an extended substrate data base. J. *Biol. Chem*. 266:14554-61.

276. Popovic M, Sarngadharan MG, Read E and Gallo RC. 1984. Detection, isolation, and continuous production of cytopathic retrovirus (HTLV-III) from patients with AIDS and pre-AIDS. *Science*. 224: 497-500.

277. Prabu-Jeyabalan M, King NM, Nalivaika EA, Heilek-Snyder G, Cammack N and Schiffer CA. 2006a. Substrate envelope and drug resistance: crystal structure of RO1 in complex with wild-type human immunodeficiency virus type 1 protease. *Antimicrob Agents Chemother*. 50: 1518-21

278. Prabu-Jeyabalan M, Nalivaika E and Schiffer C. 2000. How does a symmetric dimer recognize an asymmetric substrate? A substrate complex of HIV-1 protease. *J. Mol. Biol*. 301: 1207-20.

279. Prabu-Jeyabalan M, Nalivaika E and Schiffer C. 2002. Substrate shape determines specificity of recognition for HIV-1 protease: analysis of crystal structure of six substrate complexes. *Structure*. 10: 369-81.

280. Prabu-Jeyabalan M, Nalivaika EA, Romano K and Schiffer CA. 2006b. Mechanism of substrate recognition by drug-resistant human immunodeficiency virus type 1 protease variants revealed by a novel structural intermediate. *J. Virol*. 80: 3607-16.

281. Prado JG, Wrin T, Beauchaire J, Ruiz L, Petropoulos CJ, Krost SDW, Clotet B, C'Aquila RT and Martinez-Picado J. 2002. Amprenavir resistant HIV-1 exhibits lopinavir cross-resistance and reduced replication capacity. *AIDS*. 16: 1009-17.

282. Quinones-Mateu M.E and Arts EJ. 2001. HIV-1 fitness: implications for drug resistance, disease progression and global epidemic evolution. In: Kuiken, C; Foley, B; Hahn, B; Marx, P.A; McCutchan, F; Mellors, J (Eds), HIV sequence compendium. *Theoretical Biology and Biophysics Group, Los alamos National Laboratory.*

283. Randolph JT and De Goey DA. 2004. Peptidomimetic inhibitors of HIV protease. *Curr Top Med Chem.* 4: 1079-95.

284. Ratner L, Haseltine W, Patarca R, Liva KJ, Starchi B, Josephs SF, Doran ER, Rafalski JA, Whitehorn EA, Baumeister K. 1985. HTLV-III, LAV and ARV are variants of the same AIDS virus. *Nature*. 313: 277-84.

285. Reuman EC, Bachmann MH, Varghese V, Fessel WJ and Shafer RW. 2010. Panel of prototypical raltegravir-resistant infectious molecular clones in a novel integrase-deleted cloning vector. *Antimicrob Agents Chemother*. 54: 934-6.

286. Rhee S-Y, Taylor J, Wadhera G, Ben-Hur A, Brutlaq DL and Shafer R. 2006. Genotypic predictors of human immunodeficiency virus type 1 drug resistance. *Proc. Natl. Acad. Sci. U.S.A.* 103: 17355.

287. Richards A D, Phylip LH, Farmerres WC, Scarborough PE, Alvarez A, Dunn BM, Hirel Ph-H, Konvelinka J, Strop P, Paulickova L, Kostka V and Kay J.1990. Sensitive, soluble chromogenic substrates for HIV proteinase. *J. Biol. Chem*. 15. 265: 7733-6.

288. Rimsky LT, Shugars DC and Matthew TJ. 1998. Determinants of human immunodeficiency virus type 1 resistance to gp41-derived inhibitory peptides. *J. Virol*. 72: 986-93.

289. Rimsky LT, Vingerhoets J, Van Eygen V, Eron J, Clotet B, Hoogstoel A, Boven K and Picchio G. 2012. Genotypic and phenotypic characterization of HIV-1 isolates obtained from patients on rilpivirine therapy experiencing virologic failure in the phase 3 ECHO and THRIVE studies: 48-week analysis. *J. Acquir. Immune. Defic. Syndr*. 59: 39-46.

290. Roberts JD, Bebenek K and Kunkel TA. 1988. The accuracy of reverse transcriptase from HIV-1. *Science*. 242: 1171-73.

291. Robertson DL, Anderson JP, Bradac JA, Carr JK, Foley B, Funkhouser RK, Gao F, Hahn BH, Kalish ML, Kuiken C, Learn GH, Leitner T, McCutchan F, Osmanov S, Peeters M, Pieniazek D, Salminen M, Sharp PM, Wolinsky S and Korber B. 2000. HIV-1 nomenclature proposal. *Science*. 288: 55-6.

292. Robinson LH, Myers RE, Snowden BW, Tisdale M and Blair ED. 2000. HIV type 1 protease cleavage site mutations and viral fitness: implications for drug susceptibility phenotyping assays. *AIDS. Res. Human. Retroviruses*. 16: 1149-56.

293. Rodríguez-Barrios F, Balzarini J and Gago F. 2005. The molecular basis of resilience to the effect of the Lys103Asn mutation in non-nucleoside HIV-1 reverse transcriptase inhibitors studied by targeted molecular dynamics stimulation. *J.Am. Chem. Soc*. 127: 1570-8.

294. Romani B and Engelbrecth S. 2009. Human immunodeficiency virus type 1 Vpr: functions and molecular interactions. *J. Gen. Virol*. 90: 1795-805.

295. Roxrud I, Stenmark H and Malerod L. 2010. ESCRT & Co. Biol. *Cell.* 102: 293–318.

296. Sadler HA, Stenglein MD, Harris RS, Mansky LM. 2010. APOBEC3G contributes to HIV-1 variation through sublethal mutagenesis. *J Virol* 84: 7396-7404.

297. Samson M, Libert F, Doranz BJ, Rucker J, Liesnard C, Farber CM, Saragosti S, Lapoumeroulie C, Cognaux J, Forcellie C, Muyldermans G, Verhofstede C, Burtonboy G, George M, Imai T, Rana S, Yi Y, Smyth RJ, Collman RG, Doms RW, Vassart G and Parmentier M. 1996. Resistance to HIV-1 infection in Caucasian individuals bearing mutant alleles of the CCR5 chemokine receptor gene. *Nature*. 382: 722-25.

298. Sarafianos SG, Marchand B, Das K, Himmel DM, Parniak MA, Hughes SH and Arnold E. 1999. Structure and function of HIV-1 reverse transcripatse: molecular mechanism of polymerization and inhibition. *J. Mol. Biol*. 385: 693-713.

299. Sattentau QJ and Moore JP. 1991. Conformational changes induced in the human immunodeficiency virus envelope glycoprotein by soluble CD4 binding. *J. Exp. Med*. 174: 407-15.

300. Schmidt B, Korn K and Walter H. 2002. Technologies for measuring HIV-1 drug resistance. *HIV. Clin. Trials*. 3: 227-36.

301. Seminari E, Castagna A and Lazzarin A. 2008. Etravirine for the treatment of HIV infection. *Exp. Rev. Anti-Infect. Ther*. 6: 427.

302. Shafer RW, Kantor R and Gonzales MJ. 2000. The genetic bases of HIV-1 drug resistance to Reverse Transcriptase and protease inhibitors. *AIDS. Rev*. 2: 211-28.

303. Skalka, A. M. 1989. Retroviral proteases: first glimpses at the anatomy of a processing machine. *Cell.* 56:911-13.

304. Sluis-Cremer N, Sheen CW, Zelina S, Torres PS, Parikh UM, and Mellors JW. 2007. Molecular mechanism by which the K70E mutation in human immunodeficiency virus type 1 reverse transcriptase confers resistance to nucleoside reverse transcriptase inhibitors. *Antomicrob. Agents. Chemother*. 51: 48-53.

305. Sluis-Cremer N, Arion D, and Parniak MA. 2000. Molecular mechanism of HIV-1 resistance to nucleoside reverse transcriptase inhibitors (NRTIs). *Cell Mol. Life. Sci*. 57: 1408-22.

306. Snoeck J, Van Laethem K, Hermans P, Van Wijngaerden E, Derdelinckx J, Schrooten Y, Vijver D A van de, De Wit S, Clumeck N and Vandamme A M. 2004. Rising prevalence of HIV-1 non-B subtypes in Belgium: 1983-2001. *J. Acquir. Immune. Defic. Syndr*. 35: 279-85.

307. Spearman P, Wang JJ, Vander Heyden N and Ratner L. 1994. Identification of human immunodeficiency type 1 Gag proteins domain essential to membrane binding and particle assembly. *J. Virol*. 68: 3232-42.

308. Strack B, Calistri A, Craig S, Popova E and Gootlinger HG. 2003. A1P1/ALIX is a binding partner for HIV-1 P6 and EIAV P9 functioning in virus budding. *Cell*. 114: 689-99.

309. Suñé C, Brennan L, Stover DR and Klimkait T. 2004. Effect of polymorphisms on the replicative capacity of protease inhibitor-resistant HIV-1 variants under drug pressure. *Clin. Microbiol. Infect.* 10: 119-26.

310. Swanstrom R and Wills J. 1997. Retroviral gene expression: Synthesis, processing and assembly of viral proteins. In retrovirus edited by: Varmus HE, Coffin JM and Hughes SH. *Cold Spring Harbor, NY, Cold Spring Harbor Laboratory Press*: 263-334.

311. Takagi M, Nishioka M, Kakihara H, Kitabayashi M, Inoue H, Kawkani B, Oka M and Imanaka T. 1997. Characterization of DNA polymerase fromPyrococcus sp strain KOD1 and its application to PCR. *Appl. Environ. Microbiol*. 63: 4504-10.

312. Takebe Y, Kusagawa S and Motomura K. 2004. Molecular epidemiology of HIV: tracking AIDS pandemic. *Pediatr. Int*. 46: 236-44.

313. Tantillo C, Ding J, Jacobo-Molina A, Nanni RG, Boyer PL, Hughes SH, Pauwels R, Andries, K, Janssen PA and Arnold E. 1994. Locations of anti-AIDS drug binding sites and resistance mutations in the three-dimensional structure of HIV-1 reverse transcriptase. Implications for mechanism of drug inhibition and resistance. *J. Mol. Biol*. 243: 369-87.

314. Tessmer, U. and Krausslich, H. G. 1998. Cleavage of human immunodeficiency virus type 1 proteinase from the N- terminally adjacent p6* protein is essential for efficient Gag polyprotein processing and viral infectivity. *J. Virol*. 72:3459-63.

315. The British HIV Association Guidelines for the treatment of HIV-1 positive adults with antiretroviral therapy 2012.. *http://www.bhiva.org.*

316. Thomson MM and Najera R. 2005. Molecular epidemiology of HIV-1 variants in the global AIDS oandemic: an update. *AIDS. Rev*. 7: 210-24.

317. Thomson MM and Najera R. 2001. Travel and the introduction of human immunodeficiency virus type 1 non-B subtypes genetic forms into Western countries. *Clin. Infect. Dis*. 32: 1732-7.

318. Tomasseli AG and Heinrikson RL. 1994. Specificity of retroviral proteases: an analysis of viral and nonviral protein substrates. *Methods Enzymol*. 241:279-301.

319. Tozser J, Bagossi P, Weber IT, Louis JM, Copeland TD and Oroszlan S. 1997. Studies on the symmetry and sequence context dependence of the HIV-1 proteinase specificity. *J. Biol. Chem*. 272:16807-14.

320. Tozser J, Blaha I, Copeland TD, Wondrak EM and Oroszlan S. 1991. Comparison of the HIV-1 and HIV-2 proteinases using oligopeptide substrates representing cleavage sites in Gag and Gag-Pol polyproteins. *FEBS Lett*. 281:77-80.

321. Tozser J, Shulenin S, Kadas J, Boross P, Bagossi P, Copeland TD, Nair BC, Sarngadharan MG and Oroszlan, S. 2003. Human immunodeficiency virus type 1 capsid protein is a substrate of the retroviral proteinase while integrase is resistant toward proteolysis. *Virology*. 310:16-23.

322. Tozser J, Weber IT, Gustchina A, Blaha I, Copeland TD, Louis JM and Oroszlan S. 1992. Kinetic and modeling studies of S3-S3' subsites of HIV proteinases. *Biochemistry*. 31:4793-800.

323. Tritch RJ, Cheng YE, Yin FH and Erickson-Viitanen S. 1991. Mutagenesis of protease cleavage sites in the human immunodeficiency virus type 1 gag polyprotein. *J. Virol*. 65:922-30.

324. Turner SR, Strohbach JW, Tommasi RA, Aristoff PA, Johnson PD, Skulnick HI, Dolak LA, Seest EP, Tomich PK, Bohanon MJ, Horng MM, Lynn JV. 1998. Tipranavir (PNU-140690): A potent, orally bioavailable nonpeptidic HIV protease inhibitor of the 5,6-dyhydro-2-pyrone sulfonamide class. *J. Med. Chem*. 41: 3467-76.

325. U.S Department of Health and Human Services HIV Guidelines. *http://WWW.AIDSINFO.NIH.GOV/guidelines*.

326. Ueda T, Myint L, Shino T, Nishizawa M, Matsuda M and Sugiura W. 2005. Analysis of interference and co-evolution between protease inhibitor resistance mutations and gag mutations. *Antivir. Ther*. 10:S116.

327. UNAIDS 2010. *UNAIDS AIDS epidemic update*.

328. Usami Y, Popov S, Popova E, Inoue M, Weissenhorn W and Gottlinger H. 2009. The ESCRT pathway and HIV budding. *Biochem. Soc. Trans*. 37:181–84.

329. Van Heuverswyn HF, Li Y, Neel C, Bailes E, Keele BF, Liu W, Loul S, Butel C, Liegeois F, Yangda B, Sharp PM, Mpoudi-Ngole E, Delaporte E, Hahn BH and Peeters M. 2006. Human immunodeficiency viruses: SIV infection in wild gorillas. *Nature*. 444: 164.

330. Verheyen J, Litau E, Sing T, Daumer M, Balduin M, Oette M, Fatkenheuer G, Rockstroh J K, Schuldenzucker U, Hoffmann D, Pfister H and kaiser R. 2006. Compensatory mutations at the HIV clevage sites P7/P1 and P1/P6-gag in therapy-naïve and therapy-experienced patients. *Antivir. Ther*. 11: 879-87.

331. Verplank L, Bouamr F, LaGrassa TL, Agresta B, Kikonyogo A and Leis J. 2001. Tsg101, a homologue of ubiquitin-conjugating (E2) enzymes, binds the L domain in HIV type 1. Pr55Gag. *Proc. Natl. Acad. Sci. U.S.A*. 98: 7724-9.

332. viks-Frankenberry KA, Nikolenko GN, Barr R, and Pathak VK. 2007. Mutations in human immunodeficiency virus type 1 Rnase H primer grip enhance 3'-azido-3'-deoxythymidine resistance. *J. Virol*. 81: 6837-45.

333. viks-Frankenberry KA, Nikolenko GN, Boyer PL, Hughes SH, Coffin JM, Jere A and Pathak VK. 2008. HIV-1 reverse transcriptase connection subdomain reduces template RNA degradation and enhances AZT excision. *Proc. Natl. Acad. Sci. U.S.A*. 195: 10943-48.

334. Vodicka MA, Koepp DM, Silver PA and Emerman M. 1998. HIV-1 Vpr interacts with the nucelar transport pathway to promote macrophage infection. *Genes. Dev*. 12: 175-85.

335. Vogt VM. 1996. Proteolytic processing and particle maturation. *Curr.Top. Microbiol. Immunol*. 214: 95-131.

336. Wainberg MA. 2003. HIV resistance to nevirapine and other non-nucleoside reverse transcriptase inhibitors. *J. Acquir. Immune. Defic. Syndr*. 34: S2.

337. Wainberg MA, Zaharatos GJ and Brenner BG. 2011. Development of antiretroviral drug resistance. *N. Engl. J. Med*. 18: 637-46.

338. Wain-Hobson S, Alizon M and Montagnier L. 1985. Relationship of AIDS to other retroviruses. *Nature*. 313: 743.

339. Wang C, Mitsuya Y, Gharizadeh B, Ronaghi M and Shafer RW. 2007. Characterization of mutation spectra with ultra-deep pyrosequencing: application to HIV-1 drug resistance. *Genome. Res*. 17: 1195-201.

340. Wang CT, Zhang Y, McDermott J and Barklis E. 1993. Conditional infectivity of a human immunodeficiency virus matrix domain deletion mutant. *J. Virol*. 67: 7067-76.

341. Watson CS, Jenkinson C, Kazmierski W and Kenakin T. 2005. The CCR5 receptor-based mechanism of action of 873140, a potent allosteric noncompetitive HIV entry inhibitor. *Mol Pharmacol*. 67: 1268-82.

342. Wei X, Decker JM, Liu H, Zhang Z, Arani RB, Kilby JM, Saag MS, Wu X, Shaw GM and Kappes JC. 2002. Emergence of resistance human immunodeficiency virus type 1 in patients receiving fusion inhibitor (T-20) monotherapy. *Antimicrob. Agents. Chemother*. 46: 1896-905.

343. Wertheim JO and Worobey M. 2009. Dating the age of the SIV lineages that gave rise to HIV-1 and HIV-2. *PLoS. Comput. Biol*. 5: e1000377.

344. Westby M. 2007a. Resistance to CCR5 antagonists. *Curr Opin in HIV AIDS*. 2:137–44.

345. Westby M, Lewis M, Whitcomb JM, et al. 2006. Emergence of CXCR4-using human immunodeficiency virus type 1 (HIV-1) variants in a minority of HIV-1 infected patients following treatment with the CCR5 antagonist maraviroc in form pre-treatment CXCR4-using virus reservoir. *J. Virol*. 80: 4909-4920.

346. Westby M, Smith-Burchnell C, Mori J, Lewis M, Mosley M, Stockdale M, Dorr P, Ciaramella G and Perros M. 2007b. Reduced maximal inhibition in phenotypic susceptibility assays indicates that viral strains resistant to the CCR5 antagonist maraviroc utilize inhibitor-bound receptor for entry. *J. Virol*. 81: 2359-71.

347. White KL, Chen JM, Feng JY, Margot NA, Ly JK, Ray AS, Macarthur HL, McDermott MJ, Swaminathan S and Miller MD. 2006. The K65R reverse transcriptase mutation in HIV-1 reverses the excision phenotype of zidovudine resistance mutations. *Antivir. Ther*. 11: 155-63.

348. Wiegers K, Rutter G, Kottler H, Tessmer U, Hohenberg H and Krausslich HG. 1998. Sequential steps in human immunodeficiency virus particle maturation revealed by alterations of individual Gag polyprotein cleavage sites. *J. Virol*. 72: 2846-54.

349. Willey RL, Bonifacino JC, Potts BJ, Martin MA and Klausner RD. 1988. Biosynthesis, cleavage, and degradation of the human immunodeficiency virus 1 envelope glycoprotein gp160. *Proc. Natl. Acad. Sci. U.S.A*. 85: 9580-84.

350. Winters MA, Shafer R, Jellinger RA, Mamtora G, Gingeras T, and Merigan TC. 1997. Human immunodeficiency virus type 1 reverse transcriptase genotype and drug susceptibility chamges in infected individuals receiving dideoxyinosine monotherapy for 1 to 2 years. *Antimicrob. Agents. Chemother*. 41: 757-62.

351. Wlodawer A and Erickson J. 1993. Structure-based inhibitors of HIV-1 protease. *Annu. Rev. Biochem*. 62: 543-85.

352. Wood N, Bhattacharya T, Keele BF, Giorgi E, Liu M, Gaschen B, Daniels M, Ferrari G, Haynes BF, McMichael A, et al. 2009. HIV evolution in early infection: Selection pressures, patterns of insertion and deletion, and the impact of APOBEC. *PLoS pathog* 5: e1000414. Doi: 10.371/journal.ppat.10000414.

353. Wondrak EM, Nashed NT, Haber MT, Jerina DM and Louis JM. 1996. A transient precursor of the HIV-1 protease. Isolation, characterization, and kinetics of maturation. *J. Biol. Chem*. 271:4477-81.

354. Wu L, Gerard NP, Wyatt R, Choe H, Parolin C, Ruffing N, Borsetti A, Cardoso AA, Desjardin E, Newman W, Gerard C and Sodroski J. 1996. CD4-induced interaction of primary HIV-1 gp120 glycoproeins with the chemokine receptor CCR-5. *Nature*. 384: 179-83.

355. Wu TD, Schiffer CA, Gonzales MJ, Taylor J, Kantor R, Chou S, Israelski D, Zolopa AR, Fessel WJ and Shafer RW. 2003. Mutation patterns and structural correlates in human immunodeficiency virus type 1 protease following different protease inhibitor treatments. *J. Virol*. 77: 4836-47.

356. Xiang Y, Ridky TW, Krishna NK and Leis J. 1997. Altered Rous sarcoma virus Gag polyprotein processing and its effects on particle formation. *J. Virol*. 71: 2083-91.

357. Yap SH, Sheen CW, Fahey J, Zanin M, Tyssen D, Lima VD, Wynhiven B, Kuiper K, Sluis-Cremer N, Harrigan RP, and Tachedjian G. 2007. N348I in the connection domain of HIV-1 reverse transcriptase confers zidovudine and neviripine resistance. *PLos. Med*. 4: e335.

358. Yokomaku Y, Miura H, Tomiyama H, Kawana-Tachikawa A, Takiguchi M, Kojima A, Nagai Y, Iwamoto A, Matsuda Z and Ariyoshi K. 2004. Impaired processing and presentation of cytotoxic-T-lymphocyte (CTL) epitopes are major escape mechanisms from CTL immune pressure in human immunodeficiency virus type 1 infection. *J. Virol*. 78: 1324-32.

359. Yoshimura FK, Diem K, Leran GH, Riddell S and Corey L. 1996. Intrapatient sequence variation of the gag gene of human immunodefiency virus type 1 plasma virions. *J. Virol*. 70: 8879-87.

360. Yu X, Yu QC, Lee TH and Essex M. 1992. The C terminus of human immunodeficiency virus type 1 matrix protein is involved in early steps of the virus life cycle. *J. Virol*. 66: 5667-70.

361. Yuan X, Yu X, Lee TH and Essex M. 1993. Mutations in the N-terminal region of human immunodeficiency virus type 1 matrix protein block intracellular transport of the gag precursor. *J Virol*. 67: 6387-94.

362. Zaccarelli M, Tozzi V, Lorenzini P, Trotta MP, Forbici F, Visco-Comandini U, Gori C, Narcisso P, Perno CF, Antinori A; Collaborative group for clinical use of HIV Genotype Resistance Test (GRT) at national institute for infectious diseases Lazzaro Spallanzini. 2005. *AIDS*. 19: 1081-9.

363. Zdanowicz MM. 2006. The pharmacology of HIV drug resistance. *Am. J. Pharm. Educ*. 70: 100.

364. Zennou V, Mammano F, Paulous S, Mathez D and Clavel F. 1998. Loss of viral fitness associated with multiple Gag and Gag-pol processing defects in human immunodeficiency virus type 1 variants selected for resistance to protease inhibitors in vivo. *J. Virol*. 72: 3300-6.

365. Zennou V, Pettit C, Guetard D, Nerhbass U, Montagnier L, and Charneu P. 2000. HIV-1 genome nuclear import is mediated by a central DNA flap. *Cell*. 101: 173-85.

366. Zhang YM, Imamichi H, Imamichi t, Lane H C, Falloon J, Vasudevarachi M B and Salzman N P. 1997. Drug resistance during indinavir therapy is caused by mutations in the protease gene and its gag substrate clevage sites. *J. Virol*. 7: 6662-70.

367. Zhou W, Parent LJ, Wills J and Rehs MD. 1994. Identification of membrane binding domain within the amino-terminal region of human immunodeficiency virus type 1 Gag protein which interacts with acidic phospholipids. *J. Virol*. 68: 2556-69.