The Development of a HIV vector system to explore drug susceptibility and fitness of clinical isolates

Arinder Kohli

Submitted to the University of London for the degree of Doctor of Philosophy

September 2007

Centre for Virology
Department of Infection and Immunity
Windeyer Institue of Medical Sciences
University College London

ProQuest LLC
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106-1346

Abstract

Following the identification of a highly resistant protease containing a novel insertion (35QN), together with a series of additional *gag* mutations. We wished to explore the *gag*-protease interrelationships, and associated replication capacity deficit/compensation. A replication competent NL4-3 based HIV-1 vector system was developed allowing ligation of full length *gag* with or without protease derived from patient plasma virus. Drug susceptibility assays were undertaken using the MTT assay, and replication capacity assessed by a GHOST indicator cell line. Recombinant viruses containing the protease alone or protease with full length *gag* demonstrated reduced susceptibility to PIs. The recombinants containing the *gag* from the patient isolated virus were shown to have a much reduced replication capacity compared to wild-type, even when cloned with the cognate protease. However, the virus containing the patient virus derived *gag* and protease, demonstrated a 2-fold increase in resistance to some of the PIs tested. The production of high titred stocks was shown to require multiple reversions in *gag*. This occurred in both viruses containing this region derived from the original patient virus. The reversion pattern was shown to be dependant on whether or not the recombinant contained the highly resistant protease. The lack of reversions from the input genotype in the protease only virus isolate, suggest that this virus is not growth impaired, which was also confirmed by a growth comparison study.

The interaction of protease with the *gag* demonstrates a co-dependence regarding the evolution under drug selection pressure. The removal of the pressure causes reversion of some, but not all of the *gag* mutations. The lack of reversion in the protease gene suggests that the virus mutates *gag* to accommodate the protease. The reversion pattern in *gag* was particularly evident in the p2 protein and or the p2/p7 cleavage site.

Acknowledgements

I would like to express my thanks to Professor Deenan Pillay for the provision of research facilities in the Department of Infection and Immunity at University College London and the Health Protection Agency for financial support during the course of this work.

I am indebted to my second supervisor Dr. Chris Parry for his input throughout, particularly during the writing of this thesis.

I would like to take this opportunity to thank the many people who have made the completion of this thesis possible, not only through donation of reagents and advice with practical aspects of this thesis, but for providing a fun and vibrant atmosphere...mostly. This includes all of the members of the Wohl Virion Centre, the Viral Genomics and Bioinformatics group and 'Team Towers'. I would also like to thank Dr. Richard Myers for his help in constructing Table 5.3.

Finally, but definitely not least, I would like to thank my family and friends, for their support, patience and guidance, especially during the difficult times when they have provided a much needed outlet for my rants.

Abbreviations

AIDS	Acquired Immune Deficiency Syndrome
APOBEC3G	Apolipoprotein B mRNA-editing enzyme, catalytic polypeptide-like 3G
ATP	Adenosine triphosphate
AZT	3'-Azido-3'-deoxythymidine
BAF	barrier-to-autointegration
bp	Base pair
CA	Capsid
CCR5	chemokine (C-C motif) receptor 5
CD4	Cluster of Differentiation
CDC	Centers for Disease Control and Prevention
CPE	Cytopathic effect
cpz	Chimpanzee
CRFs	Clinical recombinant forms
Crm1	Chromosome Region Maintenance
CSMs	cleavage site mutations
CTL	Cytotoxic lymphocyte
CTS	Central Termination Sequence
CypA	Cyclophillin A
d4T	2'-3'-didehydro-2'-3'-dideoxythymidine (Stavudine)
Da	Dalton
DC	Dendritic cells
DC-SIGN	DC-Specific Intercellular adhesion molecule 3 (ICAM-3)-Grabbing Non-integrin
DIS	Dimer Inititiation Site
dNTP	Deoxyribonucleotide triphosphate
EM	Electron Microscopy
etc	etcetera
FACS	Fluorescence-Activated Cell Sorting
FDA	Food and Drug Administration
Gag/*gag*	Group specific Antigen (Protein/*gene*)
GALT	Gut Associated Lymphoid Tissue
GFP	Green Fluorescent protein
gp	Glycoprotein
HA	Hemagglutinin
HAART	Highly Active Antiretroviral Therapy
HEK	Human Embryonic Kidney
HGM I(Y)	High Mobility Group protein
hisD	histidinol dehydrogenase
HIV	Human Immunodeficiency Virus
HLA	Human Leukocyte Antigen
HLA	Human Leukocyte Antigen
HOS	Human Osteosarcoma
HR	Heptad Repeat
HSA	Heat Stable Antigen

HTA	heteroduplex tracking assay
HTLV-III	Human T-cell Lymphotropic virus type 3
Ig	Immunoglobulin g
IL	Interleukin
INI 1	integrase interactor 1
kDa	Kilodalton
KS	Kaposi's Sarcoma
LAV	Lymphadenopathy-associated virus
LC	Langherhans cells
LEDGF/p75	Lens Epithelium-Derived Growth Factor/p75
LTNP	Long Term Non-Progressor
LTR	Long Terminal Repeat
MA	Matrix
MDR	multi-drug resistance
MHC	Major Histocompatibility Complex
MHR	Major Homology Region
MIP-1	Macrophage Inflammatory Protein-1a
MLV	Murine Leukemia Virus
MSM	Men who have Sex with Men
MTT	3-(4,5-Dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide
Nef	Negative Factor
NES	Nuclear Export Signal
NC	Nuclear Capsid
NF- B	Nuclear Factor kappa B
NSI	Non-Syncytium Inducing
ORF	Open Reading Frame
PBMCs	Peripheral Blood Mononuclear Cells
PCP	Pneumocystis carinii pneumonia
PIC	Pre-Integration Complex
PIs	Protease Inhibitors
PLAP	placental alkaline phosphatase
PMPA	1-(6-Aminopurin-9-yl)propan-2-yloxymethylphosphonic acid
PPT	Polypurine Tract
P-TEF-b	Positive Transcription Elongation Factor b
RANTES	Regulated on Activation, Normal T cell Expressed and Secreted
Rev	Regulatory factor for HIV expression
RMVA	recombinant marker virus assay
RNase H	Ribonuclease H
RRE	Rev responsive element
RT	Reverse Transcriptase
RTC	Reverse Transcription complex
RTIs	Reverse Transcriptase Inhibitors
RVA	recombinant virus assay
SDF-1	Stromal-Derived Factor 1
SI	Syncytium Inducing
SIV	Simian Immunodeficiency Virus
SL	Stem Loop
sm	sooty mangabeys
SU	Surface Subunit
TAMs	Thymidine Analogue Mutations

TAR	Transactivation Response element
Tat	Transactivator of HIV gene expression
$TCID_{50}$/ml	50% Tissue Culture Infectivity Dose per millilitre
TRIM	Tripartite Motif
UN	United Nations
UTR	Untranslated Region
Vif	Viral infectivity factor
VL	Viral Load
Vpr	Viral protein R
Vpu	Viral protein u
WHO	World Health Organisation
ψ	Packaging Sequence

Contents

Figures

Tables

Chapter 1 Introduction

1.1 AIDS Epidemic and Viral Etiology

AIDS (Acquired Immune Deficiency Syndrome) was first defined by the Centers for Disease Control and Prevention (CDC) in September 1982 [35] as a disease, at least moderately predictive of a defect in cell-mediated immunity, occurring in a person with no known cause for diminished resistance to that disease. Such diseases include Kaposi's sarcoma (KS), *Pneumocystis carinii* pneumonia (PCP), and other serious "opportunistic" infections predictive of cellular immunodeficiency. The causative agent of AIDS was first isolated at the Institute Pasteur [14] and later at the National Cancer Institute [177]. The virus had been known under two names, lymphadenopathy-associated virus (LAV) and Human T-cell Lymphotropic virus type 3 (HTLV-III). In May 1986, members of a subcommittee empowered by the International Committee on the Taxonomy of Viruses published an article suggesting a new name for the etiological agent of AIDS; Human Immunodeficiency Virus (HIV) [44]. HIV belongs to the *Lentiviridae* family of retroviruses. These viruses are able to infect a number of different animal species, including cats, sheep, horses, cattle and monkeys. Studies to determine the origins of HIV focused on cross-species (zoonotic) infections with viruses closely related to Simian Immunodeficiency Virus (SIV). The cross species jump required for viruses to establish in either a simian or human host is not an uncommon phenomenon it has been seen with monkey smallpox, simian/human T cell leukaemia virus and simian monkey virus [163].

The viruses causing the HIV epidemic have been subdivided into HIV-1 and HIV-2. A strain of SIV isolated from a frozen sample taken from a captive member of the sub-group of chimpanzees known as *Pan troglodytes troglodytes* (*P. t. troglodytes*), indicated that *P. t. troglodytes* was the primary reservoir for HIV-1 and had been the source of at least three independent introductions of SIVcpz into the human population [102]. The primate reservoir of HIV-2 has been clearly identified as the sooty mangabey (*Cercocebus atys*) [102,126],classified as an old world monkey. A strain of SIV, (SIVsm) that is genetically indistinguishable and closely related phylogenetically to HIV-2 was found in substantial numbers of wild-living sooty mangabeys. The natural habitat of these primates coincides with the epicentre of the HIV-2 epidemic located in Western Africa. These monkeys and chimpanzees are commonly hunted for food, this close contact between the two species represents a ready source for

potential zoonotic transmissions of SIVsm and SIVcpz to humans. Using calculations based on a fixed mutation rate i.e. a constant rate of evolution, researchers estimated that the zoonotic transmission from chimpanzee to human responsible for the HIV-1 epidemic, most likely occurred around the 1930's, with a confidence interval of ± 20 years [126,163].

The number of people living with HIV continues to grow, in 2006 a total of 39.5 million people were living with HIV, which represents an increase of 2.6 million compared to figures for 2004 [320]. Approximately two thirds (63%) of all adults and children infected with HIV worldwide live in sub-Saharan Africa. Figure 1.1 shows the global distribution of the HIV-1 epidemic.

Figure 1.1: Adults and children estimated to be living with HIV in 2006. Total: 39.5 (34.1 – 47.1) million. Global summary of the AIDS epidemic December 2006. Figures and diagram reproduced from the 2006 AIDS epidemic update. A joint United Nations Programme on HIV/AIDS (UNAIDS) and the World Health Organization (WHO) report [320].

1.1.2 HIV Diversity

The majority of HIV infections are derived from HIV-1. HIV-2 infections have been mainly located in countries such as the former Portuguese colony Guinea Bissau, where it was first isolated in a 29 year old African AIDS patient from an area bordering Senegal [41]. HIV-2 was originally referred to as lymphadenopathy associated virus type 2. Some cases of HIV-2 have also been reported in India, and Brazil and some other countries which had past socio-economic links with Portugal. Portugal itself has the highest rate of HIV-2 infection in Europe, HIV-2 accounts for around 10-13% of total HIV infections [41,274].There are 3 distinct groups of HIV-1, group M for 'major', so called as the vast majority of infections worldwide stem from this group, group O for 'outlier' and N for 'new' [330]. The diversity afforded to HIV by its evolutionary flexibility has allowed the virus to establish different strains throughout the world. These 'founder' infections are established when a single strain of HIV populates a particular area. Over time the virus continues to infect several different hosts and establishes a population, subsequent sequence analysis of these viruses shows a higher sequence similarity within that population, than compared to other HIV populations present within different geographical locations. Within group M there is a further level of classification referred to as subtypes or clades, which refer to the major clades found within HIV-1 group M viruses. These clades are designated A to K. The distribution of the clades is shown in Figure 1.2. Another clade or assembly of clades are the CRF's Circular recombinant forms, which are essentially viruses composed of genes derived from two or more different clades within group M that arise by recombination. Examination of clade distribution reveals that subtype C is predominant in Southern and eastern Africa, India, Nepal and China. A and A/G recombinants are mainly located in and around Central and West Africa [330]. The predominant subtype in North America and Western Europe and Australia is subtype B [148]. The basis for subtype classification is usually alignment of sequences in the *env* gene, where differences of 20-50% can be seen and more recently classification based on protease and reverse transcriptase (RT) sequences which can show variation of 10-12% at the nucleotide level or 5-6% at the amino acid level [148]. The impact of subtype on rates of transmission, disease progression and other clinical consequences are not clearly understood, but one area of research aimed at determining the difference in response to antivirals seems to imply that subtype does not adversely affect response

to treatment and that mutational patterns that arise in subtype B also appear in non-B subtypes [148]. Differences have been reported that suggest the divergent evolutionary paths taken by different HIV-1 clades may play a role in virus pathogenesis. Centlivre *et al.* reported that the subtype C virus is better adapted to replication in gut associated lymphoid tissue (GALT) [36]. It is thought that the primary cellular targets for HIV are the activated cluster of differentiation 4 (CD4+) T lymphocytes expressing the chemokine receptor for HIV entry chemokine (C-C motif) receptor 5 (CCR5). This population of T cells is thought to be in higher concentration in the GALT than in peripheral blood, spleen or lymph nodes. This could explain why the gut is the preferred target for viral replication and early CD4 T cell depletion. Increased activation of the subtype C virus promoter by the GALT cytokine microenvironment would favour it's replication during primary infection and could have resulted in enhanced transmission rates. This in turn would afford a growth advantage to this subtype and may explain the observation that the subtype C virus dominates HIV-1 epidemiology, accounting for 50% of infected individuals in the world [36].

Figure 1.2: Geographical distribution of HIV-1 clades. Letters indicate the predominant HIV-1 clade circulating in selected countries. A/B, B/E, and B/F indicate co-circulation of two dominant clades; AE and AG indicate circulating recombinant forms.
Figure taken from UNAIDS/WHO [320].

Sequencing of virus isolates from the clade C subtype has revealed that they posses a uniquely truncated regulatory protein, Regulator of expression of viral proteins (Rev) and an enlarged viral protein u (Vpu) gene. Analysis of subtype D samples has shown that they contain a C-terminally truncated version of another regulatory protein, Tat protein [101,316,330]. The functions of these proteins will be discussed more extensively in the Life cycle section 1.3. However, it could be said that variation in the two regulatory genes Tat and Rev and variation in Vpu, which affects viral infectivity, could have resulted due to virus-host adaptation and resulted in increased virus production and therefore transmissibility. Undoubtedly, the aim of a virus when it infects its host is to replicate and spread its progeny virus to nearby cells and ultimately to a new host and therefore continue the spread of the virus among the host species. The adaptation to host genetics is the initial evolutionary pressure placed on the virus when it first infects the host. The ability of HIV to adapt quickly to its host immune system and other environmental pressures that it may be exposed to is a consequence of several factors which are discussed in the next section.

1.1.2a Mechanisms for HIV sequence variation

An estimated 10^9 to 10^{10} virions are generated in each infected patient daily [140,333]. As early as one week after the onset of clinical symptoms approximately, 10^5 to 10^7 HIV RNA copies/ml are present in the initial burst of viral replication associated with primary infection. However, this only represents 10^2 to 10^4 infectious units of HIV [134,263]. This reduction in actual replication competent virus compared to viral RNA measurements, represents a substantially impaired viral population which is primarily due to two unique steps within the Retroviridae Life cycle which contribute to the genetic diversity of HIV. The first is the process carried out by the enzyme RT (discussed further in Life cycle section 1.3). The RT of HIV is devoid of any error repair mechanisms and hence it has a very high mutation rate (3×10^{-5} substitutions/ site/ generation) [316], which is thought to generate every possible point mutation, every day in an infected person [45,204].The second process also occurs during genome duplication and involves two template jumps which are required to replicate the terminal repeat regions during proviral DNA synthesis. Retroviral RNA is a dimer, held together by sequences near the 5′ends of both strands. This diploid RNA allows

recombination to occur in virions that contain two RNAs encoding different genetic information. Genetic recombination is one of the mechanisms that facilitates the rapid diversification of HIV populations, a single round of replication in T lymphocytes *in vitro* generating an average of nine recombination events per virus, and infection of macrophages led to approximately 30 crossover events [178]. Recombination occurs when an individual becomes infected with another strain of HIV, this can be by co-infection, which is when two viruses infected within a timescale of each other, such that the immune response is directed simultaneously at both. Another possibility is re-infection, also termed super-infection, this occurs when an individual who has previously been exposed to an infection and has established an immune response to that infection then becomes infected by another strain [300]. Recombination can result in the generation of intra-subtypes, which are composed of viruses from within the same subtype or inter-subtypes which are generated when the RNAs of two subtypes of viruses are co-packaged into the same virion. Intra-subtype recombination occurs at a much higher frequency than inter-subtype recombination. A proposed mechanism for the lower recombination rates detected across subtypes was suggested by Chin *et al.*, where they investigated the barriers to HIV-1 inter-subtype recombination, between subtypes B and C. The data demonstrated that mismatched sequences in the 5' untranslated region (5'UTR) which are involved in dimerisation initiation and packaging of viral RNA, altered the ability of HIV to form heterozygous virions, this in turn caused a reduction in the observed recombination rate [37].

It is thought that HIV recombinants contribute to 10–40% of the HIV infections in Africa and 10–30% of those in Asia. Many of the inter-subtype recombinants circulate with high prevalence in certain geographical regions, such as A/E recombinants in Thailand and B/C recombinants in parts of Southeast Asia and China [37].

1.1.3 Course of Infection

HIV can be transmitted by exposure of the oral, rectal, or vaginal mucosa during sexual intercourse, which accounts for more than 90% of HIV infections worldwide [89,298]. It can also be transmitted by inoculation with contaminated blood or blood products, such as through the sharing of equipment during intravenous drug injection, through blood transfusion products and through perinatal transmission from mother to

newborn infant. Approximately half the cases of primary infection remain asymptomatic whereas, the remainder of patients experience flu-like symptoms within a month of infection.

Clinically HIV-1 infection of typical progressors, who represent the majority of HIV-infected individuals, can be divided into 3 phases, see Figure 1.3. During primary infection, the initial phase, the virus load is extremely high (up to 10^8 HIV-1 RNA copies/ml of plasma) [298], this also coincides with a decline in the number of CD4+ T lymphocytes. After the initial burst of viremia, virus levels in the blood start to decline. This reduction is thought to be due to the development of an immune response to HIV-1 infection through the activation of cytotoxic CD8+ T lymphocyte (CTL) response. The CTL response exerts its antiviral effect through several different means, including the production of specific antiviral chemicals such as, interferon gamma, which inhibits viral replication.

The second phase of infection is also known as the long asymptomatic period; this phase is associated with a gradual loss of CD4+ cells which correlates with levels of viremia. During this long period of sustained replication that precedes the final stage, CD4+ T cells are constantly being infected, eliminated and regenerated. The elimination of cells can occur by infection related cell death and through immune destruction of infected or by-stander cells i.e. cells that are not infected but may have viral proteins attached, due to cell to cell contact of an infected cell. This phase is also known as clinical latency and generally can last eight to ten years in typical progressors [242].

The final stage is known as clinical immunodeficiency or AIDS; this is preceded by a rapid decline in CD4+ cells and increase in viral load. Individuals who are at this stage of HIV infection usually succumb to one of the many opportunistic infections associated with the late stage disease status.

Figure 1.3: Phases of HIV-1 infection. Typical course of HIV infection. Patterns of $CD4^+$ T-cell decline and virus load increase vary greatly from one patient to another, as do the actual values of viral RNA load. Figure taken from Coffin *et al.* [46].

1.2 HIV Replication and Structure

1.2.1 HIV-1 Virion structure

The lentiviruses like other retroviruses are enveloped viruses, the mature virions have a spherical morphology of 110nm in diameter [360]. The outer envelope is composed of a lipid bilayer with viral envelope glycoproteins. Since the envelope is mainly derived from the infected cell, and so some cellular components are also incorporated. The envelope surrounds an internal layer of matrix protein (MA), which is anchored to the internal surface of the lipid bilayer by N-terminal myristoyl associated groups [89,149]. This matrix shell contains a dense cone-shaped nucleocapsid core in which the viral genomic RNA molecules, viral proteins and some cellular factors are packaged.

The genome of all retroviruses consists of three genes; Group specific Antigens (*gag*), Polymerase (*Pol*) and Envelope (*env*), these encode the structural proteins. The more complex retroviruses such as SIV and HIV encode additional proteins, Tat, Rev, Negative factor (Nef), Viral infectivity factor (Vif), Vpu and Viral protein r (Vpr) that have regulatory or accessory roles.

The genome size of HIV is approximately 9.2kb; it is present as two identical single stranded RNA molecules within the virion, but in persistently infected cells it exists as proviral double-stranded DNA [298]. Figure 1.4 shows a linear representation of the HIV-1 genome. The polyprotein products of the 3 main genes are processed into mature particle-associated proteins by either viral protease, as is the case with the 55-kDa Gag precursor $Pr55^{Gag}$ and the 160-kDa Gag-Pol precursor $Pr160^{Gag\text{-}Pol}$. Or by host cell proteases, as occurs with the 160-kDa *env* precursor. The additional proteins, Tat, Rev, Vpr, Vif, Vpu and Nef are primary translation products of messenger RNA (mRNA) splicing, although HIV protease also cleaves a single site in Nef [66,89].

Figure 1.4: Organization of the HIV-1 genome. The location of the long terminal repeats (LTRs) and the genes encoded by HIV-1 are indicated. *gag*, Pol and Env proteins are initially synthesized as polyprotein precursors. The Gag precursor is cleaved by the viral protease into the mature Gag proteins: matrix (MA), capsid (CA), nucleocapsid (NC) and p6. The GagPol precursor undergoes protease mediated processing to generate the Gag proteins and the Pol enzymes: protease, reverse transcriptase (RT) and integrase (IN). The Env glycoprotein precursor gp160 is cleaved by a cellular protease during transport to the cell surface to generate the mature surface glycoprotein gp120 and the transmembrane glycoprotein gp41. The sizes of the genes and encoded proteins are not to scale. Figure taken from Freed (2004) [96].

1.3 Life cycle

The HIV Life cycle can be divided into two phases, early and late, as depicted in Figure 1.5. The early stages of replication involve binding, fusion, entry, reverse transcription and integration. The late phase of the Life cycle begins with the regulated expression of the integrated proviral genome, leading to virus budding and finally maturation.

1.3.1 Early Phases

1.3.1a HIV Entry

Fusion of HIV to CD4+ T cells requires interaction of viral gp120, the surface subunit (SU) envelope glycoprotein and CD4 on the cellular plasma membrane; this triggers a conformational change in gp120 which reveals the previously occluded coreceptor binding site. The binding of the cellular coreceptor induces another conformational change in the transmembrane region of viral glycoprotein, exposing the hydrophobic N-terminal portion of gp41 known as the fusion peptide, which is inserted into the plasma membrane of the target cell. Specific heptad repeat regions within gp41 fold to form a stable six helix bundle structure which interacts with the plasma membrane leading to fusion and pore formation [39,47,199,345].

Figure 1.5: HIV-1 Life Cycle. General features of the HIV-1 replication cycle. The early phase (upper portion of the diagram) begins with CD4 recognition and involves events up to and including integration of the proviral DNA, and the late phase includes all events from transcription of the integrated DNA to virus budding and maturation. Figure fom Turner *et al.* [318].

1.3.1b Target Cell Receptors

Entry into the target cell requires that the viral membrane must be fused with the plasma membrane of the cell. This process is facilitated by an interaction of the envelope protins of HIV, gp160 with cellular receptors. Two host receptors are usually required for entry of HIV into the cell. The primary receptor CD4, is a type 1 integral membrane glycoprotein [195], which is a 55-kDa member of the immunoglobulin (Ig) superfamily [57]. CD4 is a ligand for Major Histocompatibility Complex (MHC) class II molecules and is expressed predominantly on the surface of T-helper cells. It is a receptor for IL-16, which is a cytokine with chemoattractant activity for CD4+ lymphocytes, monocytes and eosinophils [39].

The second cellular receptor HIV interacts with was first identified in 1996 and is a member of the G protein-coupled receptor superfamily of seven-transmembrane domain proteins. By facilitating entry into cells, these receptors play a major role in determining viral tropism. HIV cellular tropism is described in terms of its ability to replicate in macrophages or T-cells lines. Viruses that are able to infect and establish infection in macrophages and peripheral blood mononuclear cells (PBMCs) are classified as M-tropic. These strains also fail to induce syncytia (giant multinucleated cells) and so are also termed non-syncytium inducing (NSI) strains. Viruses that are able to infect T-cells and PBMCs and also induce synctia (SI) are referred to as T-tropic. The α-chemokine receptor CXCR4 previously known as Fusin [87] was identified as the primary coreceptor required for infection by T-tropic isolates. The infection of macrophages and PBMCs required the presence of CD4 and the β-chemokine receptor CCR5 [43]. Dual tropic viruses are able to gain entry using either CCR5 or CXCR4. Early in infection, a homogenous population which utilises the CCR5 coreceptor dominates, studies have demonstrated that the transition from CCR5 utilizing (R5) phenotype to a CXCR4 (X4) phenotype is associated with increased viral replication kinetics and progression to AIDS [39,208,316,323]. The association of R5 viruses with primary infection is also demonstrated by the resistance to infection of individuals who possess two copies of a non-functional CCR5 allele, Δ CCR5. The 32 bp deletion results in a premature stop codon and a truncated receptor protein; as a result no CCR5 is expressed on the surface. Homozygotes are strongly, but not

absolutely protected against infection by HIV-1; heterozygosity does not protect against HIV-1 infection but is possibly associated with an improved prognosis in the form of prolonged onset of symptoms and overall survival periods [39].

Chemokine receptors form rod like structures on the cell membrane with a central pore surrounded by the 7 transmembrane regions, which contain four domains exposed on the surface and three extracellular loops. The sites involved in HIV entry for R5 viruses require interaction of CCR5 with gp120, this is thought first to involve an initial interaction between sulfated tyrosines in the CCR5 N-terminus and gp120, followed by a second interaction with the second extracellular loop [317].
Chemokines are the natural ligands of the chemokine receptors and are 8 to 10-kDa proteins made up of 70-120 amino acid residues. They adopt a common folding pattern and show 20 to 70 percent homology in amino acid sequences with each other [186]. The family of chemokines contains four subsets (CXC, CXXXC, CC and C), which are differentiated on the relative position of their cysteine residues. In the α-chemokines, the first two cysteine residues are separated by a single amino acid (CXC), whereas in the β-chemokines, the first two cysteine residues are adjacent to each other (CC) [220] .

The natural ligands for CCR5 are Regulated on Activation, Normal T cell Expressed and Secreted (RANTES), Macrophage Inflammatory Protein-1α (MIP-1α), and MIP-1β. The homozygous ΔCCR5 allele is tolerated in individuals who express it due promiscuity of its ligands which are able to bind with other receptors. MIP-1α can bind to CCR1 and RANTES can bind CCR1 or CCR3, although MIP-1β binds only to CCR5 [317]. CXCR4 is the only receptor for stromal-derived factor 1 (SDF-1). SDF-1 was originally isolated from a bone marrow stromal cell line and was found to support the proliferation of bone marrow B-cell progenitors in the presence of interleukin (IL)-7 [220]. The SDF-1/CXCR4 system plays an important role in the migration of progenitors during embryologic development; knock-out of either CXCR4 or SDF-1 is lethal in mouse embryos causing marked cardiac and cerebellar defects [317].

1.3.1c Virus Receptors

The HIV envelope glycoproteins are synthesised as a single polypeptide precursor, gp160, which undergoes cleavage in the Golgi apparatus. The processing of gp160 is thought to involve one or more proteases from the subtilisin family of endoproteases which have shown the ability to correctly cleave HIV-1 gp160 into biologically active gp120 and gp41 [127,317]. HIV undergoes extensive glycosylation i.e. the addition of N-linked polysaccharide chains called glycans, also known as N-linked glycosylation. This occurs most abundantly on gp120, where nearly half the molecular mass is attributable to N-linked glycans [127,267]. The role of glycosylation in HIV infection is thought to have two functions; firstly it is thought essential for correct folding and the overall conformation of gp120. Secondly, as a mechanism to conceal antibody epitopes on gp120; glycans illicit no immunogenic response as they are produced by the host cell. Immune evasion may also be achieved by repositioning the glycans so that they cause a structural change in the viral receptor and so affect the binding sites for neutralizing antibodies. The rearrangement effects of the glycans could conceivably compensate for localised amino acid changes caused by neutralizing antibody evasion [243]. The amount of glycosylation has also been implicated in the compartmentalisation of viruses, those viruses originating from the semen were shown to contain less surface glycans [265].

During the HIV budding process, the gp120-gp41 complexes, which are initially expressed at the surface of infected cells, are incorporated into the virus envelope and displayed on its surface as viral spikes. Both subunits of the viral envelope proteins are connected through non-covalent interactions [39,243]. Wild-type HIV virions showed 14 ± 7 Env spikes per virion, randomly distributed, while SIV has a much higher density of surface receptors with around 73 ± 25 spikes per virion which are also unevenly distributed. Each envelope spike is composed of three molecules of gp120 and gp41, formed as trimer. The transmembrane regions, or 'stalks', project from the 'head', the gp120 trimers, in a manner similar to the legs of a tripod .[360]

The sequence of gp120 can be divided into five conserved regions (C1-C5) and five variable regions (V1-V5), its structure has been determined through the use of truncated proteins that lacked V1/V2 and V3 regions [243]. The interactions of gp120

and coreceptors that are involved in tropism are located in V1/V2 and V3 loops in an area known as the bridging sheet [39,243].

It has been suggested that the amount of receptors expressed on the surface of the virus influences the infectivity. Using SIV, Yuste *et al.* demonstrated that deletions in the cytoplasmic tail of gp41 increased levels of Env incorporated into mature virions, however, only a modest increase in infectivity was demonstrated. Surface expression of *env* was increased from an initial range of 7 to 16 spikes for wild type SIV and 200 to 417 for the gp41 truncated mutant. However, a comparable increase in infectivity was not observed, only an increase of 2-3 fold in infection was detected [354]. Such a mechanism of increasing infectivity, by expressing a truncated gp41 transmembrane protein has been seen *in vivo* in a virus isolated from a patient with primary symptomatic infection and experiencing rapid CD4 cell decline. This patients virus also contained an amino acid substitution that mutated a highly conserved tyrosine residue in *env*, which is also thought to increase surface expression of envelope glycoproteins. [100]

1.3.1d Primary Infection

The primary target for HIV is the CD4+ T lymphocytes, although on establishing primary infection the virus is first thought to interact with dendritic cells (DC). DCs are antigen presenting cells that capture, transport and present antigens to CD4+ and CD8+ T lymphocytes. Although HIV can infect certain dendritic cells, such Langerhans cells (LCs) [294], DCs role in establishing primary infection is thought to be involved in the capture and transport of HIV-1 to the lymph nodes for viral propagation and dissemination [287]. A subset of DC's present in dermal and mucosal tissues express the C-type (i.e calcium dependent) lectin, DC-specific intercellular adhesion molecule 3 (ICAM-3)-grabbing non-integrin, (DC-SIGN). This is able to bind to ICAM-3 on resting T cells and it has been reported that DC-SIGN binds gp120 with a higher affinity than CD4 [121]; however, the virus is unable to utilise DC-SIGN for entry. Instead it is thought to function as an attachment factor for HIV-1 by binding to the mannose residues on the heavily glycosylated gp120 protein and facilitating DC-mediated viral transmission. [124,332].

HIV infection through sexual contact is the major route of transmission [354] with herterosexual intercourse accounting for more than 90% of all infections globally [55]. Infections in women account for approximately half the infections worldwide and transmission from male to female has also been shown to be more effective than female to male [55,294]. In females the initial barrier to the incoming virus is in the vaginal epithelium, the events that lead to entry and establishment of infection are not fully understood. Recent research has suggested that both activated and resting CD4+ cells as well as LCs are the primary source of infection [139]. LCs cells are situated in epidermis of the skin and in all mucous membranes including oral, vaginal and cervical epithelium. They do not express DC-SIGN, but are capable of supporting infection through CD4 and CCR5 mediated entry, CCR5 is expressed at low levels, present only on 24-49% of cells, with CXCR4 being undetectable [294]. Novel *in situ* studies were carried out using an 'ex vivo organ culture system'; where vaginal epithelial sheets were removed from donors, grown and used to monitor infection with fluorescently labelled HIV-1 virions [139]. Using inhibitors to CCR5, monoclonal antibodies to CD4 and C-type lectin inhibitors, it was shown that entry into LCs was partially mediated by CD4 and CCR5. The role of DC-SIGN was only marginal and the majority of infection was through endocytic means in the LCs. The endocytosed virions in the LCs remain viable for up 5 days after infection [294] and were able to infect CD4+ T-cells through LC-T-cell conjugates. This mechanism of capture, internalisation and transport of virus to CD4+ T-cell rich regions of the lymph nodes, an area primed for viral replication, has become known as the 'Trojan Horse' scenario, taken from Greek mythology [294]. It was also shown that upon the first encounter with HIV-1, simultaneous entry occurred in both LCs and CD4+ T cells in the outer vaginal epithelium [139].

Receptor expression on the primary target cells for HIV has shown an equal distribution of both coreceptors, CCR5 and CXCR4. However, as mentioned earlier the coreceptor usage of HIV isolated during most asymptomatic infections has been CCR5. The preferential replication of the R5 virus early in infection could be explained by differential expression of the coreceptors on the most prolific virus producing cell i.e. the activated T-lymphocyte. Analysis of expression levels of chemokine receptors in HIV-1 infected and uninfected individuals revealed CCR5 expression on CD4+ T cells is up-regulated and CXCR4 is down-regulated in HIV-

infected individuals, compared with uninfected controls. The primary target for HIV, the activated T cells, predominantly express CCR5, lower expression of CXCR4 and higher expression of CCR5 on CD4+ T cells correlated with advancing disease [240]. CXCR4 utilising viruses emerge during the late stages of disease and can be isolated from up to 50% of AIDS cases although infrequently from subtype C infected patients [39]. Other reasons for this disproportionate level of R5 could be chemokines and post entry events. CCR5 is thought to promote post entry steps, one possible mechanism is through the increased ability to mobilise calcium in monocyte-derived macrophages, which promotes the transport of the capsid through the highly viscous cytoplasm [157,301]. This data is also supported by the observations of Schmidtmayerova *et al.* They demonstrated that CXCR4 binding and entry was possible in macrophages, but that replication was restricted at a post entry level, predominantly at the step of nuclear importation of the pre-integration complex (PIC) [288]. The prevention of CXCR4 virus infection could also be affected by the cells along the mucosal membrane, which express high levels of SDF-1. CCR5 may also out compete CXCR4 as it binds soluble gp120 with greater affinity [317]. Although many research articles have demonstrated X4 infection of cells associated with initial infection can occur. These may be partly explained by the '*in vitro*' nature of the work which was carried out away from the chemokine rich environment of the mucosa. Data from one study which looked at the receptor usage of patients with early stage disease showed that all the samples were either R5 (70-80%) or dual tropic, no X4 utilising viruses were found [317]. It has also been reported that virus present in male semen is predominantly of R5 phenotype, even if the majority of circulating virus from the individual was X4 [265].

The binding of virus envelope with the cell has been shown to induce many chemotactic events. These may be solely related to virus entry such as, the ability of HIV/SIV to recruit susceptible target cells to sites of active viral replication; this represents an important mechanism for viral dissemination. Chemotaxis may result in the aggregation of receptors to facilitate fusion [157]. As mentioned earlier the receptor density is relatively low and there appears to be no uniform arrangement on the viral membrane [360]. Therefore, the receptors may not be bound to specific positions and may be free to coalesce to a specific site to facilitate fusion [111]. This would seem to be an efficient means of compensating for either low density of receptors or inefficient fusion due to immune evasion. Env engagement has also been shown to signal the

recruitment of receptors and Gag on the membrane of effector cells during cell to cell direct transfer [147]. This chemotaxis 'priming' of the host cell may also be used for entry and aggregate cellular proteins required for events further downstream in the replication cycle [95].

1.3.2 Virus Uncoating

Fusion allows the transference of the virus core and capsid containing two copies of the unspliced viral genome along with the enzymes protease, RT and integrase through the fusion pore [199]. Also packaged into the capsid is the nucleocapsid protein (NC) p7, which helps to stabilise the viral RNA and the accessory proteins Nef, Vif and Vpr [89].

The initial step in uncoating is the dissociation of the matrix coat from the lipid membrane; this is attached to the matrix through the N-terminal myristyl moeity. Dissociation is thought to be triggered by basic residues within the matrix proteins, which connect to the lipid membrane through electrostatic interactions. Myristylation and phosphorylation are crucial regulatory events for matrix function. The phosphorylation of specific tyrosine residues in matrix which are thought to interact directly with the lipid membrane, are thought to be required for successful dissociation [95,149]. Uncoating is the process that occurs directly after entry into the cytoplasm, but before the completion of reverse transcription and nuclear entry [81]. The uncoating process is still not very clearly understood. Experimental determination of exactly what processes are required for successful uncoating have been hampered by the need to purify cores in the presence of detergent to remove the outer envelope region, this in turn causes instability of the purified cores. However, some studies have been able to determine the requirement for cellular factors to assist in uncoating. This was demonstrated using sucrose density gradients and ultracentrifugation to separate the intact virus cores, which were then exposed to the lysate from activated or unactivated CD4+ lymphocytes. Only cores that were incubated with the lysate from activated cells were able to complete the uncoating process and initiate reverse transcription [10]. Many other cellular factors have been implicated in assisting, or are required, in order for uncoating and reverse transcription to occur. The ATPase heat shock protein 70 (Hsp70) is a constituent of the HIV-1 virion and is thought to play a key role in stabilisation of the core structure [227]. It has also been implicated in nuclear

localisation of the HIV-1 viral PIC to the nucleus in the absence of HIV-1 Vpr protein. The Hsp70 protein family members perform a variety of functions that could implicate their role in early entry events of HIV-1. Proteins in this group facilitate assembly and disassembly of oligomeric protein complexes as well as their folding and intracellular transport [123].

The dissociation of the matrix complex exposes the viral capsid, bound to it are approximately 200 molecules of the human cellular proline isomerase, cyclophilin A (CypA), [318] although HIV-2 and SIV do not package CyPA [339]. The incorporation of CypA occurs during assembly and is thought to be required to abrogate host restriction factors [96] and may also play a role in uncoating by helping to destabilise the mature capsid; acting as an uncoating factor [339]. Structural studies have revealed that CypA binds in Gag to a loop structure within capsid, residues 85-99, it specifically interacts with Gly89 and Pro90 [339]. Mutation of either of these two residues or treatment with the immunosuppressive agent Cyclosporin A, abolishes the capsid-CypA interaction and confers a substantial loss of infectivity on the virus [132]. The block to infection is post-entry and is thought to occur before completion of uncoating [81]. The host restriction factor thought to be responsible for diminished replication in CypA deficient viruses is the human form of the α spliced variant of the tripartite motif (TRIM) 5 gene. The mechanism by which TRIM5 exerts its host restriction effect is still unknown, it may disrupt either or both of the highly ordered events of uncoating or reverse transcription [81]. The binding of CyPA may also obscure capsid residues or cause structural changes that prevent detection or binding by the restriction complex [302].

Another cellular defence mechanism that has been shown to be active against retroviruses is a member of a family of RNA-editing enzymes that deaminates cytosine residues in viral DNA. Apolipoprotein B mRNA-editing enzyme, catalytic polypeptide-like 3G (APOBEC3G), also known as CEM15, catalyses the replacement of cytosine by uracil in the minus strand DNA during reverse transcription. This leads to the replacement of the uracil with a thymidine and the possible degradation by host uracil-DNA glycosylases in the uracil-based excision pathway. The other possibility is G →A hypermutation, this results from the uracils being replaced for thymidine in the minus-strand DNA, which is an adenosine in the complementary strand [96,308].

However, HIV also packages a host a derived enzyme that removes uracil bases from DNA, called uracil DNA glycosylase hUNG-2; this may counteract the activity of APOBEC3G [196]. Although the potential role for the uracil removal is unclear, one mechanism HIV has to counteract APOBEC3G is Vif. The mechanism Vif employs to inhibit G→A hypermutation is still unclear, it has been shown to bind to APOBEC3G and induce its degradation and may also prevent its incorporation into virions [96]. The inhibition of APOBEC3G packaging into the virion by Vif does not affect mRNA levels of APOBEC3G, which suggests that Vif effects protein expression, post-transcriptionally [96,308].

Uncoating can be seen as a two stage process, firstly as described above the dissociation of matrix occurs, which is followed by a slower secondary stage of capsid protein disassembly [75]. The cone shaped capsid of HIV-1 is composed of a polymer of approximately 1500 capsid proteins [75]; the association of capsid with the reverse transcription complex (RTC) is still not clearly understood. RTC's which are complexes of viral proteins and cellular components, are reverse transcription competent and this activity has been shown regardless of whether or not capsid was still bound. The capsid of HIV-1 has also been shown to contain pores that may allow for the diffusion of nucleotides and so increase the possibility that active reverse transcription takes place in the cytoplasm [75,203].

HIV-1 like all other lentiviruses is able to infect non-dividing cells; this is achieved through its ability to transport the PIC to the nucleus. The PIC is a high molecular weight nucleoprotein complex, containing the viral proteins integrase, matrix, RT, NC and Vpr, it is thought to contain very little or possibly no capsid [75]. The PIC differs from the RTC in that PICs are defined as integration-competent complexes, whereas RTCs are thought to contain incomplete components of reverse transcription [233]. Nuclear targeting of the PIC could be achieved by several methods, research has suggested that many if not all of the viral proteins present in the PIC contain nuclear localisation signals. The Vpr may facilitate nuclear localisation through invoking the cellular nuclear import machinery [318]. The nuclear import signal may also come from an associated cellular protein, as mentioned earlier many cellular proteins are packaged with the virion during assembly. The cellular protein lens epithelium-

derived growth factor/p75 (LEDGF/p75) is thought to target the PIC to the nucleus through its interaction with integrase [75]

The core is thought to utilise the host cell cytoskeleton by travelling on the system of actin microfilaments network within the cell to aid transport to the nucleus through the highly viscous cytoplasm. [111,196]. This event is thought to be triggered by fusion which, it is proposed, may activate chemotatic signals also causing remodelling of the actin cytoskeleton [147].

1.3.3 Reverse Transcription

HIV genome contains the highly conserved region that encodes for the viral enzymes, Protease, RT and IN [89]. The production of the aforementioned proteins is the result of HIV protease cleavage of the $Pr160^{Gag\text{-}Pol}$, the major protein product is the $Pr55^{Gag}$, which is expressed approximately 90%-95% of the time. The mechanism for production of the virally encoded enzymes is through a slippage or ribosomal frameshift during the translation of $Pr55^{Gag}$ [89]. RT of HIV and all other retroviruses carry out both DNA polymerase and Ribonuclease H (RNase H) activities; the RNase H is nuclease specific to the RNA strand of an RNA-DNA hybrid [46].

1.3.3a Reverse Transcriptase Structure

Mature virions contain RT in the form of a heterodimer made up of two subunits, p66 and p51. The p51 subunit is generated by protease cleavage of p66; therefore both subunits share identical sequences which are located at the N-terminal region. The p55 subunit lacks the RNase H domain and does not contribute to either of the catalytic activities of RT. Although isolated RNase H is not active and requires the presence of p51 to restore activity, it is thought that residues located in the C-terminal region of p51 and p66 are important for RNase H cleavages [113].
Crystal structures of p66 domain of RT initially bound to inhibitors and also unbound have revealed a structure that resembles a right hand, see Figure 1.6.

Figure 1.6: Structure of HIV-1 RT. Cartoon overview of the subdomains of HIV-1 RT. The subdomains of the polymerase domain, which are named based on analogy to a human right hand, are colour coded. Both subunits are shown. Fingers (blue), palm (red), thumb (green), connection (yellow). The RNase H domain, which is part of the p66 subunit, is shown in light brown. Figure adapted from Coffin *et al.* [46].

This has lead to the use of nomenclature based on that similarity; the subdomains have been named fingers, palm, thumb and the connection subdomain, which links the polymerase domain to the RNase H domain [46]. The palm domain of p66 contains the polymerase active site, which forms a cleft, exposing the three catalytic aspartic acid residues 110, 185 and 186. The latter two aspartic acid residues are located in a conserved region which contains the motif YXDD. Mutagenesis of any of the catalytic residues results abolishment of polymerase activity. This motif is conserved among numerous viral and cellular polymerases and believed to be essential for polymerase function [318,331]. The p51 subunit has a more rigid structure and despite having identical amino acid sequence does not form a polymerising cleft, the aspartic acid active site residues in the palm subdomain are buried underneath a more closed conformation adopted by the fingers subdomain [318].

The structural analyses of RT with various templates bound have suggested that the overall structure is flexible. The proposed mechanism of action suggests the fingers close in around the primer-template and deoxyribonucleotide triphosphate (dNTP) prior to formation of the 3'→5' phosphodiester bond, this causes the lengthening of the growing chain, resulting in a relaxation in the fingers subdomain. The release of the pyrophosphate allows the enzyme to advance to incorporate the next dNTP [89,318]. The majority of residues in RT that are in contact with the template DNA are located in the p66 subdomain, with the palm and thumb subdomains clamping the DNA and 3'- hydroxyl group (OH) of the primer in the correct conformation relative to residues in the active site [318]. The primer-template chain extends for 18 bases from the active site of RT to that of RNase H [113].

1.3.3b Initiation of Reverse Transcription

Reverse transcription requires a primer with a free 3'-OH group to initiate extension; this is provided by the binding of host encoded cellular $tRNA^{lys}$ primer, the tRNA primer interacts with the thumb subdomain of RT. The 18 nucleotides located at the 3'-terminal end of $tRNA^{lys}$ are complementary to a region located downstream from the 5'-end of the unspliced viral RNA, known as the primer binding site [113]. There are three major types of $tRNA^{lys}$ isoacceptors, $tRNA^{lys1,2}$ that differ from each other by only a single base pair in the anticodon stem and $tRNA^{lys3}$, which has a much more diverse sequence with 14-16 changes. All three isoacceptors are packaged although in lentiviruses only $tRNA^{lys3}$ serves as the tRNA primer [113,158]. The enzyme that

aminoacylates $tRNA^{lys}$ to the charged active form is lysyl-tRNA synthetase, which is also packaged into the virion during assembly. The $tRNA^{lys}$ is associated with Gag-Pol during assembly, although the exact nature of the complex is not fully known. The lysyl-tRNA synthetase is thought to bind to Gag and several other interactions are required between Gag, Gag-Pol, viral RNA, $tRNA^{lys}$ and lysyl-tRNA synthetase [158]. The annealing of the $tRNA^{lys}$ primer to the primer binding site does not appear to require Gag maturation [279].

1.3.3c Overview of Reverse Transcription

Following initiation of reverse transcription, synthesis of minus-strand DNA proceeds eventually leading to the formation of a double-stranded provial DNA copy, prior to integration into the host genome. To accomplish this successfully, HIV-1 undergoes two template switches that occur by virtue of sequence homology. These occur within the primer binding site and the repeat regions located at both the 5' and 3' ends of the viral genome [46,89,113]. The majority of first strand transfers occur after the entire 5'-repeat region has been synthesised [237], these may be either inter-strand or intra-strand, both are thought to occur at similar frequency, while second strand transfer is thought to be solely intra-molecular i.e. a transfer from the 5' repeat region of one viral RNA to the 3' repeat region of the same RNA molecule [113].

A brief description of the steps involved in reverse transcription is outlined below and in a diagram of the process is presented in Figure 1.7.

1. Reverse transcription is initiated with $tRNA^{lys}$ primer annealing to primer binding site in the viral RNA template. Minus-strand DNA synthesis is initiated from the 3'-OH of the $tRNA^{lys}$ primer and proceeds through terminal repeat region to the 5' end of the genome.
2. Concomitant with DNA synthesis of the minus-strand, the RNA strand of the RNA/DNA hybrid is degraded by RNase H, this produces a short single stranded fragment called minus-strand strong stop DNA. The length of this fragment is approximately 100-150 bases due to the location of the $tRNA^{lys}$ binding site. The sequence of this fragment is complementary to the repeat region located on 3'-end of the genome and this leads to the first translocation, termed first strand transfer.

3. The formation of the RNA/DNA hybrid allows continuation of minus-strand DNA to the primer binding site at the 5' end; this is accompanied by RNase H degradation of the RNA strand of the RNA/DNA hybrid.
4. This degradation by RNase H proceeds up to a purine rich region termed the polypurine tract (PPT), this stretch of sequence is relatively resistant to RNase H degradation.
5. The synthesis of plus-strand DNA is initiated from the 3' end of the PPT using the single-stranded minus DNA as a template. This generates plus-stranded strong stop DNA which includes sequence of the primer binding site of the $tRNA^{lys}$ that initially served as the template for minus-strand DNA synthesis.
6. RNase H removes the $tRNA^{lys}$ primer, exposing the sequences in the plus-stranded DNA primer binding site that are complementary to those in the 3'-end on minus-strand DNA, promoting the second strand transfer.
7. Plus and minus-strand synthesis continues until completion; with each strand able to use the other as a template.

Plus-strand DNA terminates at the end of the minus-strand or when RT encounters the central termination sequence (CTS). The CTS is essentially a centrally located PPT and results in the linear DNA molecule bearing in its centre a stable 99 nucleotide long, plus strand overlap referred to as the central DNA flap. This flap along with other viral proteins is incorporated into the PIC and is thought to facilitate its nuclear import [233].

Figure 1.7: Process of reverse transcription of the retroviral genome. (*Black line*) RNA; (*light colour*) minus-strand DNAs; (*dark colour*) plus-strand DNA. See text for a description of this process. Figure adapted from Coffin *et al.* [46].

1.3.4 Integration

As previously mentioned the PIC is an integration competent complex, the migration of the PIC to the nucleus is thought to be achieved through viral components and also through the packaging of cellular components. Other cellular proteins are required to help with integration, as *in vitro* work on integration revealed that purified integrase was only able to partially complete the process [95,209]. The high mobility group protein, HMG I(Y) has also been shown to be important for successful completion of integration and is thought to associate with the PIC during its journey to the nucleus [233]. Other cellular proteins include integrase interactor 1 (INI 1), which targets integration to transcriptionally active regions of the host chromosome. Another host protein, barrier-to-autointegration (BAF) protein, prevents autointegration events that may result in the production of aberrant forms of proviral DNA [89,233]. The PIC is estimated to be 50nm in diameter [203], it contains a linear copy of proviral DNA from which two to three nucleotides are cleaved from the 3' end by integrase, to form a pre-integration substrate [46].

The integrase enzyme of retroviruses allows the insertion the newly synthesised proviral DNA into the host genome, which then serves as a template for production of new virions for the lifetime of the infected cell. Full-length integrase is 288 amino acids long, it is a 32 kDa protein that has three domains: the central/catalytic core, the C-terminal, and the N-terminal domains. The catalytic core domain contains the highly conserved active site residues (Asp64, Asp116, and Glu152). Mutation of these residues generally leads to a loss of all the catalytic activity of the enzyme and thus, the catalytic triad is thought to be an essential component of the integrase active site. The N-terminal domain (amino acids 1–50) is thought to be involved in protein multimerisation, it contains a zinc binding motif that stabilises protein folding. The C-terminal domain (amino acids 213–288) has a very high affinity for DNA and thus, is thought to play a role in binding to host DNA [172]. Crystal structures indicate a dimeric model in which two monomers interact with each other, although it is not clear whether integrase works *in vivo* as a monomer, a dimer, or a tetramer [189,266].

This primary step in the integration process is called terminal cleavage, this event probably occurs in the cytosol and results the terminal addition of recessed 3'-OH groups to the viral DNA, providing the sites of attachment into the host chromosomal DNA. The second step is the staggered cleavage of the target cellular DNA, this

occurs in the nucleus and is not based on target sequences, but primarily on the overall structure, with highly bent DNA sites being the preferred target for integration [172]. The 3' recessed ends of the viral DNA are joined to the 5' cellular DNA by the integrase-viral DNA complex. Cellular repair complexes in conjunction with integrase and possibly RT repair the gaps completing the integration process. [46,89,209,318].

1.3.5 Late Stages

1.3.5a Transcription, Translation and Export

Following integration of proviral DNA into the host chromosome, the provirus serves as a template for the synthesis of viral mRNAs. Transcription and splicing of HIV mRNA is a complex process, and is initiated in the U3 of the 5' long terminal repeat (LTR). The viral regulatory proteins Tat, Rev and Nef are the first proteins to be produced early on in infection. Accumulation of Tat and Rev has a profound effect on transcription and both these proteins play an essential role in its regulation [89].

1.3.5b Long Terminal Repeats

Regulation of basal transcription levels is controlled by the interaction of cellular proteins with specific sequences within the HIV LTRs. Initial rounds of transcription are carried out by cellular RNA polymerase II, the induction of the polymerase complex requires the binding of several cellular transcription factors, including transcription factor II D (TFII D) which binds to the TATA sequence upstream of the transcription start site [54]. Other protein/DNA interactions with the HIV LTR such as, the binding of nuclear factor kappa B (NF-κB) and binding of Sp1 are also thought to contribute to basal transcription levels [95].

1.3.5c Tat

The initial transcriptional activity of the LTRs is very low; this increases dramatically when the transcriptional activator protein Tat binds to the RNA element known as the transactivation response element (TAR). The 5' TAR element is located in the 5'repeat region of all HIV RNA transcripts; it is 59 nucleotides long and forms a based paired stem loop structure, with a tri-nucleotide unpaired bulge and a six nucleotide G-rich loop. The primary sequence information is contained in the terminal

loop and bulge of TAR [54]. The Tat gene is divided into two coding exons and produces a protein of 101 amino acids containing several functional domains. The N-terminal activation domain spans residues 1-47, the hydrophobic core element, a basic segment which is essential for recognition and binding of TAR and a glutamine rich C-terminal domain, thought to be involved in inducing post-translational capping of mRNA [308,318]. Tat binds at the top of the hairpin to the bulge structure recruiting several host proteins to form a complex known as the positive transcription elongation factor b (P-TEF-b) [96]. The association between Tat, TAR and P-TEF-b causes the hyperphosphorylation of the C- terminal domain of RNA polymerase II, this induces conformational changes which results in an increase in processivity and initiation of full-length transcript production. [237]

1.3.5d Rev

Activation of the infected cell by HIV or other stimulatory chemokines results in transcription of the viral genome. Three classes of mRNA are produced in the nucleus, (i) completely spliced which encodes the proteins Tat, Rev and Nef (ii) partially spliced which is around 5kb in size and contains the *env* glycoproteins, Vif, Vpu and Vpr and (iii) unspliced resulting in Gag and Gag-Pol polyproteins [308]. Initially, only the completely spliced, 2-kb class of viral mRNA is produced. Then due to the binding of Tat, viral gene expression increases (by greater than two logs [95]) resulting in a switch to synthesis of viral structural gene mRNAs [54]. The transition from early regulatory gene expression phase to that of the later structural phase is thought to be promoted by critical levels of the Rev protein [54,95,209,308,338] i.e when Rev protein levels reach a certain level, the switch to late phase expression occurs, thus Rev down regulates its own expression.

Once inside the nucleus Rev interacts with a *cis*-acting element that is located in the *env* gene, present only in the partially spliced and unspliced viral RNA transcripts [298]. The binding of Rev to the Rev response element (RRE) facilitates the transport of the aforementioned RNAs through the nuclear membrane, thus preventing their translation in the nucleus [298].

Rev is a 19kDa protein encoded by two exons, resulting in a protein of 116 amino acids [95,318]. It contains two functional domains; a basic domain which is an arginine-rich RNA binding motif that binds to the RRE [308] and is also involved in nuclear

localisation. The second domain is a hydrophobic leucine-rich domain involved in nuclear export and is known as the effector or activation domain [95,318].

The RRE is 250 nucleotides in length and folds into a series of base paired stem loop structures emanating from a central unpaired region. Initially Rev binds to the second stem loop and this promotes the multimerisation of Rev on the RRE, eventually binding up 12 Rev monomers [95].

The main function of Rev is to shuttle between the cytoplasm and the nucleus initiating nuclear export of unspliced and incompletely spliced mRNA's [338]. As previously mentioned, Rev encodes both nuclear import and export signals which facilitate this chaperone activity. The RNA binding motif of Rev is thought to be involved in the recruitment of host nuclear import factors, importins to aid its nuclear localisation functions [308]. The nuclear export of Rev-RRE also requires the formation of complexes with host proteins, such as chromosome region maintenance 1 (Crm1). This is an evolutionary conserved protein that acts as a receptor for proteins containing nuclear export signals, which localises to the nuclear pore [96,308].

1.3.6 Assembly Maturation and Budding

Viral assembly and budding, takes place on plasma membranes in T cells and on endosomal membranes in macrophages [112,255]. These final processes of the HIV viral Life cycle include encapsidation of the viral RNA, congregation of the viral proteins, budding and release from the infected cell membrane and maturation of the virus particle. These steps of the HIV Life cycle are tightly regulated by the Gag precursor polyproteins, $Pr55^{Gag}$ and $Pr160^{Gag\text{-}Pol}$, which are synthesised by unspliced transcripts in the cytoplasm [89]. The envelope proteins are synthesised by the same cellular machinery as is used by the host to process its own surface proteins. The envelope glycoproteins are translated on ribosomes associated with the rough endoplasmic reticulum, into the precursor protein gp160. The Env polyprotein contains a signal peptide region located in the N-terminal portion which facilitates its insertion into the bilipid layer. Trafficking of the gp160 through the Golgi apparatus results in the attachment of N-linked glycans and also the cleavage of the precursor polypeptide by host proteases into the mature envelope glycoproteins gp120 and gp41 [30,94,95,149,155]. Lipid rafts play an essential role in HIV particle production, a definitive role for rafts and cholesterol association of HIV is suggested through comparisons of the

composition of the virion lipid bilayer and that of the host cell membrane from which it is mainly derived. The virion bilayer is significantly enriched in cholesterol and sphingolipids compared to that of the host [238]. The final steps in the production of an infectious virion are also controlled by the viral proteins encoded in the unprocessed polypeptide, $Pr55^{Gag}$ [46,95].

1.3.6a HIV Protease

The HIV Gag proteins matrix (MA), capsid (CA), nucleocapsid (NC), p6 and the spacers p1 (also known as SP2), p2 (also known as SP1) are formed from the proteolytic cleavage of the $Pr55^{Gag}$ precursor by the virally encoded protease [238]. Protease also cleaves a site in Nef, although the requirement for cleavage of Nef does not seem to be functionally important [89]. Protease is 99 amino acids long and is active as a symmetrical homodimer, the dimer is stabilised by interactions between N and C terminal ends of each monomer. Aspartic acid residues of each monomer are involved in catalysis in the active site; therefore it is termed an aspartyl proteinase [89]. The active site is formed at the interface of the two subunits by three residues from each monomer; aspartic acid 25, threonine 26 and glycine 27. The residues 49 to 52 of each monomer form a flexible β hairpin structure or 'flap' region, that extends over the substrate-binding cleft, the flexibility is thought to be essential for enzyme activity [318]; see Figure 1.8. Structural and proteolytic studies indicate that the binding cleft of protease can accommodate peptides of approximately seven amino acids in length. Three to four amino acids either side of the peptide bond that is to be cleaved; referred to as the scissile bond, with the N-terminal flanking amino acids denoted as P1, P2 P3 etc. The C-terminal flanking bonds of the substrate peptide are known P1', P2', P3' etc and the subsites of the enzyme that interact with the peptide are termed, N-terminal S1, S2, S3 etc and C-terminal S1', S2', S3' etc [20,46,89]. During cleavage of the substrate the aspartic acid residues at position 25 of both monomers are involved in the activation of a water molecule located between the active site aspartates, which is required for the hydrolysis of the scissile bond [25]. The specificity towards the substrate appears to be dictated by protein secondary structure which unlike the primary amino acid sequence is well conserved [349].

The initial steps in the processing of the Gag-Pol fusion protein involve the autocatalysis of protease in the $Pr160^{Gag\text{-}Pol}$. The Gag and Pol proteins are encoded by overlapping reading frames; *gag* contains its own initiation (AUG) and termination (TAA) codons whereas, processing of *Pol* requires a ribosomal -1 frameshift during translation. The slip site is located at C-terminus of the NC coding sequence and forms a stem structure that stalls the ribosome during translation [295]. Protease cleavage of the Gag-Pol fusion protein results in the production of the four N-terminal Gag proteins (MA, CA, p2 and NC) as well as the transframe protein, protease, RT and integrase [337].

The mechanism by which protease becomes activated is thought to involve primary intra-molecular cleavages that are carried out by the Gag-Pol embedded immature protease in *cis*. As with the enzyme, which is active as a dimer, the immature protease requires the dimerisation of the Gag-Pol precursors, this creates a partial active site and the initial cleavage in Gag-Pol occurs at the N-terminus of protease and leads to the release of a catalytically competent, C-terminally extended protease species [30,258-260,313]. This autocatalysis of protease occurs either at the plasma membrane or just after the virus particle has been released from the infected cell membrane [298].

Figure 1.8: Structure of HIV-1 Protease. Colour indicates distinct regions. Flaps: residues 43–58, red; flap tips: residues 49–52, yellow; flap elbow: residues 37–42, magenta; residues 59–75, green; residues 10–23, orange and β-strand motif forming the dimer interface: residues 1–4 and 96–99, blue/cyan. Figure adapted from Hornak, Viktor *et al.* [141].

1.3.6b Gag Cleavage

The cleavage of the $Pr55^{Gag}$ precursor and the formation of the internal structural proteins, is referred to as maturation, and is thought to take place after the budding event has occurred. A linear representation of the *gag* gene is shown in Figure 1.9, along with the location of domains present within each product and below a summary of the functions of each domain.

The exact order of cleavage of the Gag substrate *in vivo* is still not fully understood, but the *in vitro* order of cleavage has been elucidated [24,262]. The initial cleavage occurs at the N-terminus of NC (p2/NC); the liberation of the NC is thought to play a vital role in core formation and stability of the RNA dimers. The subsequent cleavage step releases the MA protein allowing it to form membrane associations [3,30]. The second cleavage event by the mature HIV protease releases the matrix protein from the capsid/p2 polyprotein, this is thought to induce major structural and morphological changes in the virion, during condensation of the core protein [30,94]. The next cleavage event takes place at the p1/p6 site followed by the penultimate cleavage at p1/p7 [24]. The final Gag cleavage event is the release of the mature capsid protein, this cleavage occurs between CA/p2, even though this is one of the fastest HIV-1 processing sites, the negative regulation by the p2 protein slows the cleavage, so that this becomes one of the slowest Gag sites to be cleaved [262]. The order of cleavage is fundamentally important to the production of infectious viruses. Incorrectly cleaved Gag leads to the formation of immature virions with viral cores that appear doughnut-shaped by electron microscopy, as opposed to the conical-shaped cores adopted by mature retroviruses [89].

Figure 1.9: Linear representation of the HIV-1 gag region. (a) entire Gag polyprotein (b) matrix and the domains located within (c) capsid domain (major homology region (MHR)) (c) p6 protein. Figure adapted from Freed [94].

1.3.6c Gag Domians

1.3.6d Matrix (MA)

MA is largely responsible for membrane targeting and binding, as mentioned previously the N-terminus of MA is co-translationally modified by N-terminal myriystolyation, this modification is essential for the membrane transport and binding function of MA [30,95]. The positively charged basic residues on MA are thought to interact with the negatively charged residues on the inner surface of the plasma membrane, thus stabilising membrane association [30,95]. The overall structure of matrix is thought to be trimeric and this may be functionally relevant for assembly. Matrix is also involved in the incorporation of full-length envelope into virions, a direct interaction takes place between the cytoplasmic tail of gp41and matrix [94].

1.3.6e Capsid (CA)

Both CA and MA have functional roles during both early and late events of the HIV-1 Life cycle. The CA proteins form a cone-shaped structure located centrally in the mature virus, surrounding the complex required for proviral DNA synthesis and integration. As mentioned earlier a region located in CA forms an exposed loop structure which is responsible for the binding of cyclophilin A. The C-terminal domain contains a region which displays significant homology amongst retroviruses, known as the major homology region (MHR), mutations in this area produce assembly, maturation and infectivity defects in progeny viruses [94]. The dimerisation domain of CA also interacts with the MHR region, forming hydrogen bonds which contribute by stabilising the structure of CA [89].

1.3.6f Nucleocapsid (NC)

The NC protein like other Gag proteins serves a multifunctional role with HIV Life cycle. Its major role is in the binding of RNA, particularly to an area in the 5'-UTR, which contains a specific stretch of RNA that forms several stem loop structures thought to be essential for the packaging of viral RNA into the virion. The packaging signal or Ψ, is 120 nucleotides long and is located between the 5'-LTR and the *gag* start codon. The analogous region to the packaging signal in the 3' repeat region

contains the polyadenylation signal [237]. The RNA/NC interaction occurs primarily through secondary structure interactions, as there is little sequence conservation of this region [131]. The two zinc finger motifs of NC are joined by a flexible basic linker region. Mutations that abolish zinc binding also result in loss of infectivity and diminished genome encapsidation [30]. The NC/RNA interactions also include the coating of viral RNA in the mature virion and binding of the $tRNA^{lys}$ primer. All these interactions can be attributed to the 'nucleic acid chaperone' function of NC, which may also carry out functional roles in reverse transcription and integration [89]. The N-terminal domain of mature NC is thought to play a role in Gag multimerisation by binding to $Pr55^{Gag}$, and through interactions with host proteins is also thought to be involved in assembly and transport [89].

1.3.6g p6 Protein

The C-terminal protein to be released from the $Pr55^{Gag}$ precursor is the proline rich p6, which is only found in primate lentiviruses [89]. It plays an essential role in the final budding step and also by incorporating Vpr. A motif within Gag known as the late domain is thought to interact with host endosomal sorting machinery. The amino acid sequence P(T/S)APP near the N-terminus of the protein has been shown to promote budding of the virion from the membrane by binding to components of the ESCRT (endosomal sorting complex required for transport) I, II and III pathway [96]. The transframe protein $p6^{pol}$ has been implicated in the regulation of protease activity, truncation of the C-terminal portion of p6 are thought to increase *gag* processing [246,337].

1.4 Inhibitors of HIV Replication

HIV-1 primarily infects CD4+ T-lymphocytes, monocytes/macrophages, astrocytes and cells of the central nervous system such as brain microglial cells [89,255,294]. Infection spreads to the lymphatic tissue, it is thought that follicular dendritic cells act as a site for virus latency. Continued virus replication leads to the slow and progressive destruction of the immune system and in some cases, the central nervous system, leading to dementia [184,242]. The development of small molecules to block the

replication of HIV-1 in infected individuals and delay the course of disease progression has been the subject of dedicated research efforts over the past decades.

1.4.1 RT Inhibitors

1.4.1a Nucleoside RT Inhibitors

Inhibition of HIV using antivirals has centred around using specific targets within the replication cycle. The first inhibitors to be shown to have an effect on HIV were those directed at the reverse transcription step of the Life cycle, termed the RT inhibitors (RTIs). The first reported demonstration of the efficacy of inhibiting reverse transcription was shown by Mitsuya *et al.* in 1984; the drug Suramin, originally used for the treatment of African sleeping sickness, demonstrated protection of T-cells from the cytopathic effect of HTLV-III *in vitro*. The mechanism of action of this compound was through the inhibition of retroviral reverse transcription and this lead to the initial research and development of agents which could inhibit this replication process [210].

1.4.1b Nucleoside Analogues

The first compound to be licensed as an antiviral against HIV was 3'-azido-3'-deoxythymidine (AZT, Zidovudine), which was originally developed in the 1960s to treat cancer [80,211,212]. AZT is an analogue of the pyrimidine, thymidine, one of the natural cellular substrates for RT and host DNA polymerases. In order for nucleoside analogues to be effective as antivirals they require three sequential phosphorylation steps to convert the dideoxynucleoside into the active triphosphate form. The triphosphate form then competes with natural nucleoside triphosphates and when incorporated, results in chain termination of the viral DNA, due to the lack of a 3′-OH on their ribose or ribose mimic moiety [212]. The initial phosphorylation event is thought to be the most discriminating step, as a result only some nucleoside analogues are phosphorylated and are able to demonstrate reduced HIV replication [236]. The discovery that nucleoside analogues could provide suitable compounds for HIV-1 inhibition lead to the subsequent development of other similar antivirals, there are now at least 7 approved for use in HIV therapy, see Figure 1.10a.

1.4.1c Nucleotide Analogues

AZT is a prodrug in that it requires phosphorylation by cellular kinases to generate the active form, AZT-5'-triphosphate [212]. The use of the active triphosphates as antivirals has not proved successful, primarily due to cleavage of the phosphate groups by extracellular enzymes and reduced intracellular levels due to poor uptake by the cell [236]. The development of nucleotide analogues i.e. a nucleoside with a phosphate group attached to the carbon 5 instead of the hydroxyl group; avoids the need for the intracellular triphosphorylation. The only nucleotide analogue to be approved by the Food and Drug Administration (FDA) is an acyclic nucleoside monophosphate diester analogue of adenosine, 1-(6-Aminopurin-9-yl) propan-2-yloxymethylphosphonic acid (PMPA) or Tenofovir disoproxil fumarate [255], see Figure 1.10a. This compound is more stable than the nucleoside RT inhibitors and is hydrolysed by intracellular esterases, which results in the monophosphate form, PMPA. This undergoes further phosphorylation to the active triphosphate form [80].

1.4.1d Pyrophosphate Analogues

The only pyrophosphate analogue to be marketed is Foscarnet, although it has not been approved for treatment of HIV-1 infection due to very low bioavailability and toxicity, it has been used for treating cytomegalovirus retinitis in AIDS patients [80]. It acts by interfering with the exchange of pyrophosphate from deoxynucleoside triphosphate during viral replication by binding to a site on the HIV RT [166].

1.4.1e Non-Nucleoside RT Inhibitors

This final group of compounds that are classified as RT inhibitors are unlike the other compounds within this class; they are non-competitive inhibitors that bind to a hydrophobic pocket near the polymerase active site and exert thier antiviral effects allosterically i.e. they prevent the enzyme from functioning optimally via a non-substrate binding mechanism. NNRTI binding induces structural changes in RT, which result in termination of the polymerisation reaction, caused by the displacement of active site residues, relative to the polymerase binding site [62,304]. This causes the active site in the p66 subdomain to be locked in an inactive conformation, similar in structure to the inactive p51 subunit [61]. This mechanism of inhibition only affects the

RT of HIV-1, both RTs of HIV-2 and SIV are insensitive to NNRTI's since they share only 60% homology with HIV-1 RT [143]. The elucidation of the structure of HIV-1 RT facilitated the design of NNRTIs; although the discovery of NNRTIs was initially through the screening of compound libraries [61]. Currently, there are 3 NNRTIs licensed for the treatment of HIV-1 infection, Nevaripine, Efavirenz and Delavirdine, see Figure 1.10b. Nevaripine was the first NNRTI to be licensed in 1997, although the discovery of this compound class of inhibitor occurred nearly a decade previous to that with the discovery of the HEPT(1-(2-hydroxyethoxymethyl)-6-(phenylthio) thymine) and TIBO (tetrahydroimidazo [4,5,1 - jkj] [1,4]benzodiazepin - (1*H*)-one and –thione) compound series [61].

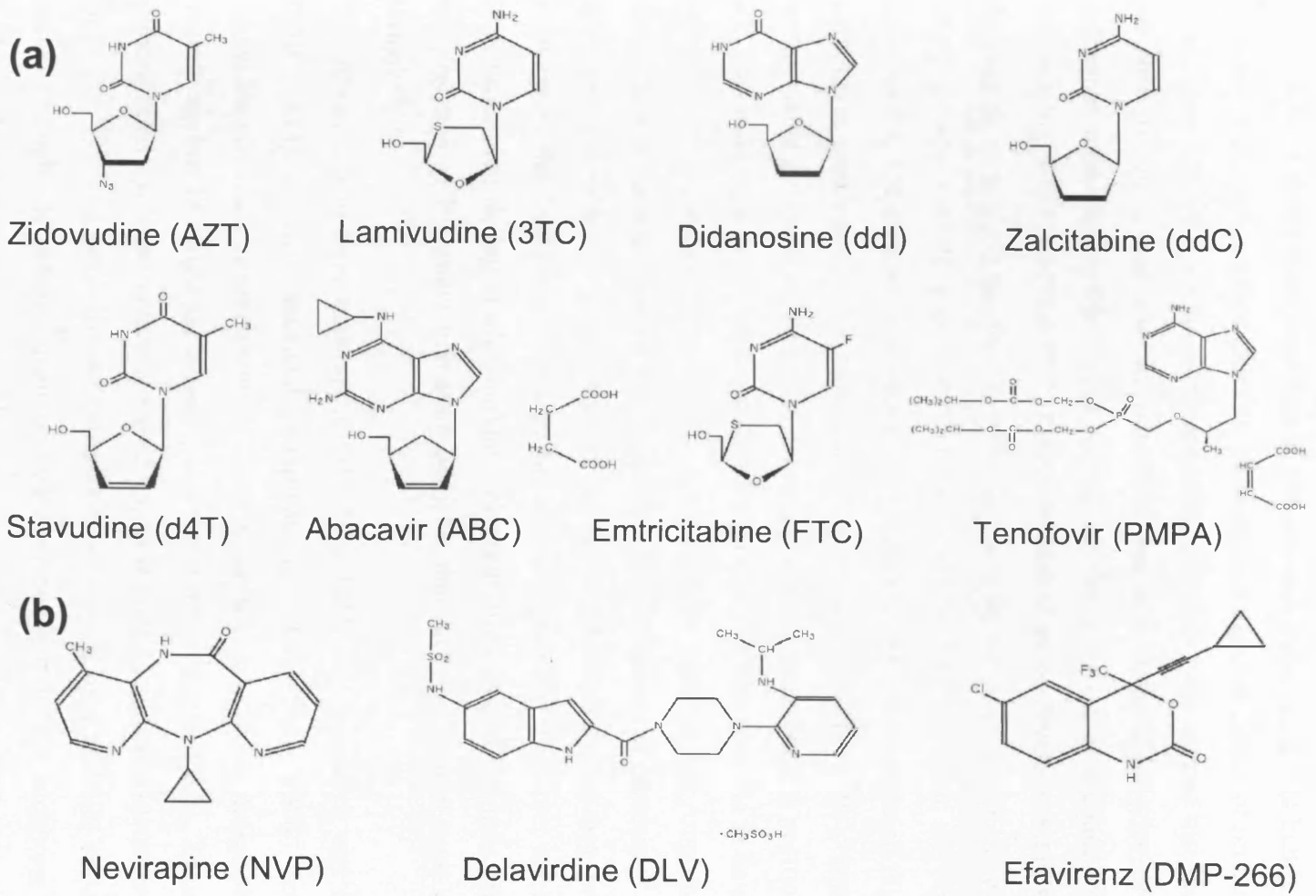


Figure 1.10: Clinically approved RT inhibitors. (a) Nucleoside and nucleotide RT inhibitors **(b)** Non- nucleoside RT inhibitors

1.4.2 Protease Inhibitors

The use of rational design in the discovery of potent inhibitors of HIV-1 was also applied to the development of protease inhibitors (PIs). Roberts *et al.* used the knowledge of the unusual cleavage of Phe-Pro and Tyr-Pro peptide bonds by HIV-1 protease in the design of the first PIs. The cleavage of amide bonds of proline residues is not seen with mammalian endopeptidases, therefore, this seemed like an excellent potential target. It was also hypothesized that due to the small size of protease, resistance mutations to PIs would be difficult for the virus to accommodate whilst maintaining protein function [20,280]. Research resulted in the development of the first PI Saquinavir, which was licensed in 1995, followed by Ritonavir and Indinavir in 1996. There are now nine PIs approved for the treatment of HIV infection, see Figure 1.11. Like NRTIs, PIs are also competitive inhibitors. The first PIs were small oligopeptide derivatives which mimic the natural substrate of protease. These PIs contained non-hydrolysable groups at the positions either side of the scissile bond at positions P1 and P1', this modification prevents cleavage by protease [89]. The next PIs to be developed and approved were Nelfinavir which is a non-peptidic inhibitor and Amprenavir, both these PIs were designed based on the protein crystal structure of protease [20]. PIs exert their antiviral effect at a late stage in the HIV replication cycle; they prevent the cleavage of the Gag polyproteins by the viral protease and therefore, interfere with viral maturation during or after budding. Electron microscopy of viruses produced in the presence of PIs show immature particles that lack the conical core of mature virions [52].

The newest PIs to have been approved by the FDA are, Tipranavir and Darunavir (TMC 114), which received approval in 2005 and 2006 respectively (www.fda.gov/oashi/aids/virals.html). Both these inhibitors were designed to show activity against PI resistance viruses (see section on resistance 1.6.3), Tipranavir has demonstrated excellent antiviral activity against HIV-1 clinical isolates resistant to other PIs [165,222]. Tipranavir and Darunavirs activity against multi PI resistant HIV occurs through the establishment of hydrogen bonds with the backbone atoms of protease, these interactions more closely resemble those that occur between substrate and enzyme [165]. Older generation PIs form hydrogen bonds with the side chains of the amino acids which are readily changeable by without affecting function. Tipranavir is the only PI that establishes a direct hydrogen bond interaction to the flap region,

traditional peptidomimetics utilise a ubiquitous water molecule to mediate this bond [222,349]. Direct hydrogen bond interactions are considered to be more energetically favourable than those mediated by water [222].

Many of the antiretrovirals are substrates for the P-glycoprotein cellular transport system, which facilitates their transport out of the cell, resulting in reduced intracellular levels of the compounds and therefore decreased virus inhibition [305]. To abrogate this effect, regimens routinely contain Ritonavir, used as a pharmalogical enhancer it has two effects. Firstly, it inhibits the P-glycoprotein cellular transport of PIs and it also inhibits cytochrome-P450 mediated metabolism of antivirals; regimens that utilise Ritonavir this way are known as Ritonavir boosted therapies [67].

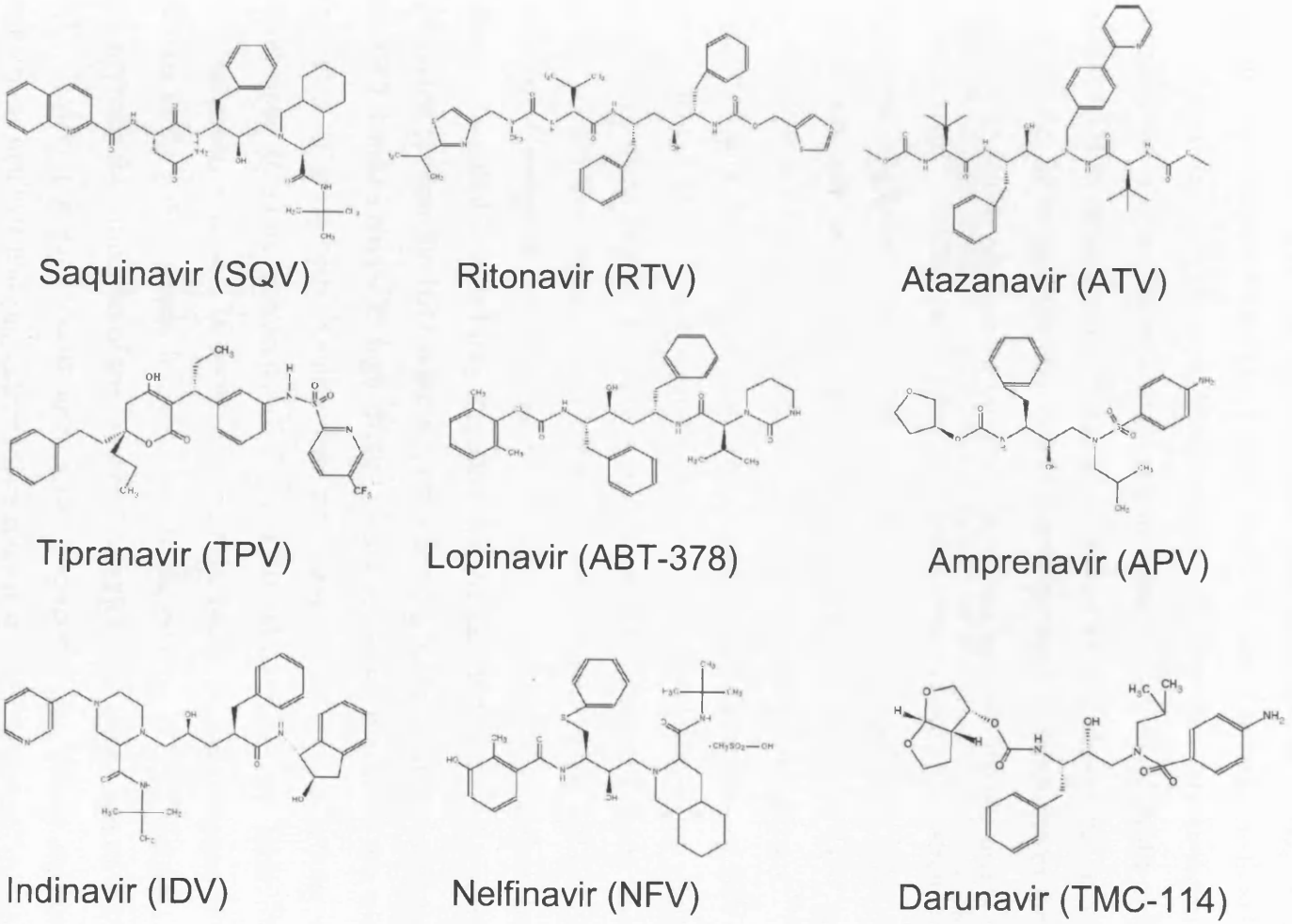


Figure 1.11: Clinically approved PI inhibitors.

1.4.3 Fusion and Entry Inhibitors

One of the first approaches to inhibit the fusion process centred around the primary receptor for entry, CD4. During this time it was understood that CD4 was not the only requirement for virus receptor mediated entry, but the discovery of chemokine receptor involvement was some 4 years away [212]. Initial attempts with soluble CD4 proved very successful *in vitro*, however these excellent viral inhibition profiles were not mirrored *in vivo*. The sensitivity seen *in vitro* was attributed to the ability of the soluble CD4 to remove gp120 from the surface of virions. However, this was not the case *in vivo*, as here failure was attributable to the increased stability of the envelope complex in primary isolate virions and also to the reduced affinity levels of soluble CD4 *in vivo* [39]. Another one of the early approaches to fusion inhibition was the use of various sulphated sugars, such as low mass dextran sulphate. This and other compounds such as heparin were shown to block the binding of HIV to CD4 by binding to sites on gp120, however, clinical trials suggested little efficacy [39,212].

The first fusion inhibitors to be developed and used for treatment of HIV infection were synthetic peptides that mimicked a region within gp41. The extracellular domain of gp41 contains four major regions, a hydrophobic N-terminal fusion domain, two leucine zipper-like repeat regions known as heptad repeats (HR). HR1 is N-terminal and HR2 C-terminal; and a hinge region containing two cysteine residues which forms a disulphide-bonded loop structure, linking the HR1 and HR2 regions. During the fusion process the HR1 regions form a trimeric coiled-coiled structure, whereas the HR2 domians bind with high affinity to grooves on the outside of this trimer. This formation of a six-helix bundle, 'hair pin' structure brings the cellular and viral membranes in close proximity [39,118,275]. T-20, also known as Enfuvirtide is a subcutaneously injected, 36-amino-acid peptide, based on the sequence of the HR2 domain in the gp41 subunit. It interferes with the binding of HR2 and HR1 domains and prevents the formation of the 'hair pin' complex. The interaction of T-20 and the HR1 domain can only occur after CD4 interaction and before engagement of coreceptors in a 'pre-hair pin' intermediate stage that exists for several minutes [118,276].

As mentioned earlier the requirements for entry necessitates engagement of both CD4 and coreceptor, the discovery of coreceptor mediated entry provided a new therapeutic target [87]. The viability of CCR5 as a target was further enhanced by the

redundancy associated with CCR5, especially in light of the ΔCCR5 allele. Several CCR5 chemokine receptor antagonists have been developed which bind to the coreceptor in a cavity formed between CCR5 transmembrane helices [317]. These compounds are small organic molecules with molecular weights of 800–1000 Da, are orally bioavailable and inexpensive to manufacture. The redundancy of CCR5 affords its use as an inhibitor. CXCR4 however, is a primordial chemokine receptor and its requirement is indispensable. Nonetheless, inhibitors of CXCR4 have been developed and have shown clinical efficacy, although some individuals enrolled in trials for one of the earliest compounds AMD3100, suffered from cardiac toxicities which prohibited further development of this compound [39,317]. Although CCR5 inhibitors have proved to be successful *in vitro* and in early *in vivo* trials [317], the consequences of a CCR5 blockade and the possibility of selection of more virulent CXCR4 variants remains a possibility. Also, the use of CCR5 inhibitors in individuals homozygous for the ΔCCR5 allele may be less toxic than for those who possess the functional CCR5 allele or are heterozygous [26].

Other compounds that inhibit CD4–gp120 binding include, PRO-542(CD4–IgG2) a soluble recombinant antibody-like fusion protein, which mimics the CD4 receptor. A disadvantage of PRO-542, is that like T-20 it cannot be orally administered. Another CD4–gp120 interaction inhibitor, BMS-806 has been shown to bind with high affinity to HIV gp120, blocking the conformational changes induced in gp120 after CD4 binding which allow the subsequent interaction with the chemokine co-receptor. An advantage of this compound is that it demonstrated good oral bioavailability [187,276].

1.4.4 Integrase Inhibitors

The integrase protein has no known functional mammalian homologue and the coding region is highly conserved among HIV-1 clinical isolates [32] and so, potentially provides an excellent therapeutic target for inhibitors. Integrase inhibitors can be divided into two categories based on their mode of action; (i) inhibitors of 3' processing, which act by docking at the HIV-1 DNA binding site of integrase and (ii) selective strand transfer inhibitors which compete with the target DNA in the strand transfer step [189,266]. The diketo acids were amongst one of the first series of compounds to be explored as integrase inhibitors, they selectively inhibit the strand transfer reaction and are the most clinically advanced compounds in this class [172,266].

These compounds bind to the catalytic triad at the core of integrase, they do not inhibit the 3' processing step, and instead it is thought that these inhibitors prevent the viral DNA-integrase complex from binding to cellular DNA. The unintegrated viral DNA is then degraded by cellular enzymes and creates dead-end circular by-products called 2-LTRs. The increasing formation of 2-LTRs has been shown to correspond with escalating concentrations of strand transfer integrase inhibitors [172]. Recently, two compounds from the strand transfer class of inhibitor have entered clinical trials and both demonstrated potent antiretroviral activity in short-term monotherapy [69,201].

1.4.5 Maturation/Capsid Inhibitiors

A new approach to HIV-1 inhibition has been to look at obstructing the assembly process of the HIV Life cycle. In one approach, small molecules were identified through a phage display screening process; peptides were selected that targeted the CA and NC proteins in Gag. These inhibitors are termed CA assembly inhibitors or maturation inhibitors and showed *in vitro* inhibition of both immature and mature virions. The mechanism of inhibition in immature virions is thought to be by means of preventing cleavage of p24/p2 site in the infected cells. Inhibition in mature HIV virions is achieved by binding to CA, preventing any further involvement in the virus Life cycle [307]. PA-457 [3-*O*-(3′,3′-dimethylsuccinyl)betulinic acid] is another potent HIV drug candidate, it acts through inducing a defect in Gag processing by disrupting p25 to p24 conversion and preventing formation of mature CA protein [179].

1.5 Treatment of HIV Infection

The advancement in the understanding of the pathogenesis of HIV disease and the development of compounds that could, for a limited time control the rate of disease acceleration, led to a great accumulation in the knowledge concerning the impact of inhibitors. Initially single drug dose regimens were evaluated and showed delayed progression of disease in symptomatic individuals. Subsequent trials involved the use of dual nucleoside analogues and indivduals on this treatment regimen were shown to suppress viral replication better than the monotherapy nucleoside group [212,278]. The predictors of therapy success were initially based on CD4+ T cell counts, although now the most predictive marker is the patients' actual plasma viral load (VL) [169,206].

The introduction of new classes of antiretrovirals, namely the NNRTIs and PIs increased the number of treatment options available to HIV infected individuals. Combinations of different classes of inhibitors soon became the standard treatment option for patients in the developed world; the use of this potent combination of antirerovirals became known as highly active antiretroviral therapy (HAART) [80,184]. By 1996 mortality rates amongst HIV infected individuals had dramatically declined in western countries, the advent of HAART had changed the prognosis of HIV infection from a lethal to a potentially chronic disease [347]. Successful HAART reduces viral load to undetectable levels, as measured by the most sensitive, routinely available viral load assays (typically below 50 copies/ml) and causes a reduction in mortality and opportunistic events, even in patients with very advanced HIV disease [221]. Also associated with successful therapy is a reconstitution in the number of CD4+ cells [27]. This increased life expectancy of HIV infected patients led to the identification of adverse events related to HIV antiretroviral therapy. Initial toxic effects had been described from the onset of antiretroviral use, bone marrow suppression which required transfusion was reported after 6 months therapy in one third of patients enrolled on a phase II clinical AZT trial in 1986 [212]. One of the main toxicities associated with AZT is damage to the mitochondria and loss of mitochondrial function, known as mitochondrial myopathy. Investigations into the observed mitochondrial toxicity implicated the mitochondrial DNA polymerase, DNA polymerase γ, although DNA polymerases β and γ also appear to be sensitive to dideoxynucleoside-5'-phosphates [212,347]. More recently patients receiving either therapies including NNRTIs and PIs have been associated with elevated triglyceride, cholesterol levels and accumulation of central fat, also known as lipodystropy [89]. The toxicity associated with HAART and other related complexities of drug regimens such as, pill burden, dosing frequency and food constraints at time of dosing, all combine to make compliance to a particular regimen a problem. Infrequent dosing of antivirals would lead to reduced levels of inhibitor in the patient, these concentrations may not be enough to sustain antiviral efficacy and may lead to the generation of resistant variants (discussed in the next section). The level of adherence required to maintain regimen efficacy in a PI based regimen has been estimated to be greater than 95%, although, still resulting in 22% of patients experiencing virological failure. Substantially more regimen failure occurred in 80% of patients with compliance

levels of less than 80% [247]; unfortunately, population based studies of adherence have estimated compliance levels to be only around the 70% mark [67]. The *in vivo* efficacy also depends on access of the antivirals to the intracellular sites where HIV-1 replication occurs. As mentioned previously, P-glycoprotein (multi-drug resistance protein) interacts with antivirals causing its activation and pumps the antivirals out across the plasma membrane, thereby reducing intracellular accumulation of drugs [305,347]. This mechanism for exclusion of antiretrovirals occurs in lymphocytes and macrophages as both these cell types express the multi-drug resistance protein. All these mechanisms can severely impair the efficacy of regimens due to sub-optimal drug concentrations, this in turn leads to resistance and the generation of variants that require even greater drug concentrations in order to successfully suppress any further virus replication.

1.6 Resistance to Antiretrovirals

The use of antivirals has resulted in reduced viral burden in infected individuals and also prolonged life expectancy. As mentioned previously, the replication strategy of HIV includes certain events that facilitate the development of heterogeneity, such as, inter-strand transfer and error-prone polymerase activity during viral genome replication. These events coupled with a high viral turn-over allow the virus to generate an enormous swarm of mutants known as viral quasispecies [72,88], which may or may not generate fully-functional virions. In the context of this virus species being present in an individual receiving therapy, then it is likely that these same mechanisms coupled with the potential for non-compliance to drug regimens, could lead to the development of viruses that replicate better in this environment, due to some mutation(s) incorporated during error-prone replication. The development of resistance to all the classes of drugs that have been approved is the main reason for the need for new antiretrovirals and is also the reason most therapeutic regimens require constant evaluation regarding their effectiveness. The genetic barrier is a quantity that refers to the difficulty for the virus to escape from the selective pressure of the drug by developing escape mutations [15]. Some antivirals show a low genetic barrier, such as 3TC, which requires only a single base substitution; **A**TG-**G**TG (M184V) in order to negate antiviral efficacy. Due to the high viral turnover and the

error prone replication of the HIV genome (see section 1.3), such single mutations are more likely to pre-exist in the viral quasispecies and thus, are more likely to be selected. Other compounds such as PIs have high genetic barriers with several mutations required in order to reduce susceptibility to the compound and so are less likely to pre-exist in the viral swarm within an infected individual. A high genetic barrier compound(s) for the initial regimen may be preferable, in order to minimise the risk of resistance and treatment failure.

The resistance mutations of each class of inhibitor are shown in Figures 1.12 (a), (b), (c) and Figure 1.13.

Figure 1.12 Mutations in RT associated with NRTI resistance. (a) and (b) Nucleoside and nucleotide RT inhibitors (c) Non- nucleoside RT inhibitors. Figure adapted from Update of the Drug Resistance Mutations in HIV-1: 2007 [146].

Figure 1.13 Mutations in Protease associated with PI resistance. Figure adapted from Update of the Drug Resistance Mutations in HIV-1: 2007 [146].

1.6.1 Resistance to NRTIs

As AZT was the first compound to be licensed for use as an anti-HIV compound, mutations were first reported in patients receiving Zidovudine monotherapy [212]. Virus isolated from patients who were on prolonged AZT therapy showed the presence of mutations in the RT gene, the order of appearance of mutations that confer increasing resistance to AZT was also determined [21]. This indicated that an initial resistance mutation at codon 70 in RT appeared, transiently and conferred only low level resistance. This mutation was then replaced by the primary resistance mutation at codon T215Y/F which conferred increased resistance levels to AZT, followed by additional mutations at 67 and 219 in some patients. However, it was demonstrated that the development of highly resistant isolates was not required for treatment failure and progression to AIDS in treated individuals [21]. Trials involving the use of dual NRTI therapy showed promising results when virus isolated from patients enrolled in a dual NRTI trail demonstrated reduced *in vitro* sensitivity to AZT or ddI. However, when mutations conferring resistance to both NRTIs were present on the same genome the virus displayed increased sensitivity to AZT [306]. The potential for suppressing viral replication completely was further enhanced by *in vitro* work that demonstrated the effectiveness of combinations of NRTIs. In this study, researchers failed to generate drug resistance variants, through step-wise increases in the concentrations of NRTIs over a 10-week period [103]. However, these combination therapies eventually failed in patients and the development of further mutations that resulted in dual NRTI resistance were eventually identified [9,152]. Research efforts to determine the role of mutations identified in patients receiving antiretrovirals was greatly enhanced by early *in vitro* work using a novel vector system, developed by Kellam *et al.* This was composed of a live HIV vector with a deleted RT gene. Introduction of the RT deleted genome and separate RT gene, allowed for the recombination with RT genes containing specific mutations conferring resistance to NRTIs introduced through site directed mutagenesis. This enabled the study into the effects of resistance mutations on the growth ability and sensitivity of these recombinant viruses to NRTIs. The system became known as the recombinant virus assay system, RVA [153] and has been widely used and adapted to further advance the study of HIV resistance mutations and their modulation of sensitivity to antiretrovirals.

1.6.1a Thymidine Analogue Mutations

The mutational pattern associated with AZT resistance was also shown to occur with another thymidine nucleoside analogue, Stavudine or 2'-3'-didehydro-2'-3'-dideoxythymidine (d4T). Accordingly the mutations, M41L, D67N, K70R, L210W, T215Y/F and K219Q, are grouped as the thymidine analogue mutations (TAMs); TAMs may also be selected for by ddI [80,329,349]. The mechanism by which these mutations reduce the sensitivity of the mutated RT to NRTIs is by removal of the blocking nucleoside analogue, by a hydrolytic mechanism, termed primer unblocking. The excision of the nucleoside analogue is an Adenosine triphosphate (ATP)-dependent process and TAMs cluster around the ATP-binding pocket and enhance the binding of ATP to boost the excision process [349]. The chain termination mechanism of NRTIs firstly requires the competition for binding with natural dNTPs, if successful the NRTI-monophosphate enters the binding site in the catalytic core of RT. The incorporation of the NRTI-monophosphate causes a structural shift in the enzyme resulting in the translocation of the NRTI-monophosphate to the priming site in the active site, which frees the binding site to accept the next incoming dNTP. However, the lack of the 3'-OH group prevents the binding of the incoming nucleotide, which now occupies the binding site, resulting in a dead end complex, terminating chain extension. TAMs are thought to facilitate the excision of the bound NRTI, by either preventing the binding of the next dNTP, which in turn allows the removal of the blocking NRTI. Or TAMs may increase the affinity of the NRTI for the binding site, preventing its translocation to the priming site [80,329,349]. Either of these mechanisms contributes to the development of resistance by TAMs to NRTIs.

1.6.1b Other NRTI Resistance Pathways

Another set of mutations associated with NRTI therapy, exert their resistance effect through the increased discrimination of the incoming dNTP by the mutant RT. Resistance mutations appear to alter the contact between the enzyme, the template and the incoming dinucleotide in such away as to allow the natural substrate to be preferentially incorporated. However, the bulky side chains of the nucleoside analogues, prevents their incorporation [349]. The mutations that are involved in this

type of resistance pattern are located around the dNTP-binding site. Residues that are in close proximity to the incoming dNTP's and are also associated with this mechanism of resistance include, K65R L74V, D113, Y115, F116 and Q151 [64,80,329]. The K65R mutation which is located in the fingers domain of RT, causes reduced sensitivity to ddC, ddI and Abacavir, it is also the only mutation to be selected for by Tenofovir *in vitro*. The development of multi-drug resistance (MDR) has been reported and involves a Q151M change that appears initially, followed by the step-wise accumulation of other mutations located in or close to the dNTP binding pocket; V62A, V75I, F77I and F116Y. The initial appearance of the Q151M mutation causes a moderate increase in resistance to NRTIs, however, accumulation of additional mutations results in the development of high level resistance [64,329,349]. The MDR complex causes resistance to all NRTIs except Tenofovir [80]; Lamivudine provides partial suppression of viral replication although high level resistance soon emerges [289]. The M184V/I change in the highly conserved YMDD region, which contains two catalytically important aspartic acid residues, results in the development of high level resistance to 3TC and FTC. In the case of 3TC, the appearance of the M184V change results in the development of up to 1000 fold resistance and it can develop rapidly *in vivo*, in as little as two weeks [11]. These compounds are analogues of cytosine and contain a sulphur atom located on the ribose sugar moiety, this bulky ring and the presence of the M184V/I mutation leads to increased contact of the enzyme with the dNTP or NRTI and therefore increases discrimination in terms of nucleotide incorporation [64,80,329].

1.6.1c Insertions and Deletions leading to NRTI resistance

The long-term use of NRTIs has lead to the appearance of divergent strains of HIV with distinctive genotypes causing variation of inhibitor activity. Insertions in RT in response to NRTI therapy have been reported; in all cases the HIV-RT contains a dipeptide insertion between codons 69 and 70 usually SG, SS or SA, as well as a T69S mutation, leading to the nomenclature T69S+XX [22,64,342]. This insertion occurs in the fingers region of RT, in a loop which makes contact with the incoming dNTP during polymerisation [64]. The use of site directed mutagenesis showed that the insertions alone cause only a moderate increase in resistance to NTRIs which explains their 95% association with other TAMs. The insert containing RT with additional

TAMs may have higher excision rates for thymidine analogues [342]. Nucleotide sequence analysis of the insertions demonstrated that they are duplications of neighbouring sequences, probably generated by RT stalling or slippage during reverse transcription. Such insertions in RT have been linked to NRTI exposure as there are no reports of insertions in therapy naïve individuals [342].

Deletions that occur in the same fingers region of RT as the insertions have also been reported, these single peptide deletions are located at codon 67 and are also associated with TAMs. These deletion variants may have a replication advantage in quiescent cells, as it has been shown they were able to effectively carry out polymerisation in low dNTP levels and were also efficient at excision of both AZT-monophosphate and PMPA at lower ATP concentrations [22,342] .

Other potential mechanisms for NRTI resistance have been attributed to the RNase H domain of RT in conjunction with other TAMs and also to the p6 protein in Gag [64,256,349].

1.6.1d Hypersensitivity to NRTIs

As discussed earlier, the mutations that cause resistance to ddI, caused hypersensitivity or hypersusceptibility, i.e. increased sensitivity to AZT [306]. This effect has also been seen with TAM containing RTs and the addition of M184V, L74V and K65R mutations. RTs that contain a resistance mutation Y181C, which emerges in NNRTI therapy, are also associated with increased sensitivity to AZT, with additional TAMs [329]. Hypersensitivity to NRTIs is through accumulation of these mutations and is associated with AZT, Tenofovir and Stavudine [80]. This re-sensitisation to AZT with the M184V mutation occurs by reducing the ATP-mediated excision of the AZT-monophosphate, through alteration in the polymerase active site [23]. The re-sensitisation to NNRTIs is also thought to occur in HIV-RTs which contain the T69S+XX mutations [342].

1.6.2 Resistance to NNRTIs

NNRTIs exert their effect through steric hindrance rather than participation in a biochemical event as is the case in NRTIs, due to this their effect can be acutely negated by the addition of a single mutation that prevents binding in the hydrophobic pocket. The hydrophobic pocket does not exist in RT that is not bound to a ligand, in the bound state the pocket is formed through side chain interactions of the surrounding amino acids [80]. Due to the structural diversity of the NNRTIs (see Figure 1.10b), the binding and subsequent structural constraints inferred on RT by each of the compounds differs; therefore each exhibits slightly different interactions with the enzyme. As binding of all currently approved NNRTI requires interactions with the hydrophobic pocket, there is extensive clustering of residues that are involved in NNRTI resistance [64]. Common NNRTI resistance mutations include L100I, K103N, V106A, Y181C, Y188L and G190A. The K103N is the most clinically significant NNRTI mutation, acquisition of this mutation results in a substantial loss in efficacy of the regimen [56]. The asparagine side chain of this mutant forms a hydrogen bond with the amino acid at codon Y188, causing a structural rearrangement of the pocket, similar to that of the unbound state [349].

Recently, alternative methods for the mechanism of action of NNRTIs have been proposed. One study suggests that the susceptibility of HIV-1 to NNRTIs is influenced by the level of RT activity and that NNRTI hypersensitivity is related to a reduction in RT activity. This mechanism did not affect NRTIs such as Lamivudine which were not affected by the level of RT per virus particle [7]. Another more recent study suggests the mechanism of action of NNRTIs involves the preferential inhibition of the initiation of plus-strand viral DNA synthesis. This *in vitro* effect was seen only in wild-type RT enzymes and not in those that contained NNRTI mutations. The mechanism for inhibition of viral DNA synthesis appears to involve competing with incoming dNTP for the binding site in RT. This data is contrary to previous reports; this opposing conclusion could be due to the different template used in the different experiments. Grobler *et al.* demonstrated that NNRTIs do compete with dNTPs for binding, but only with RNA-DNA template substrates and not with DNA-DNA as most previous studies had used [120].

1.6.3 Resistance to PIs

Protease cleavage of $Pr55^{Gag}$ is an essential step in the replication cycle of HIV, both the correct order of cleavage and the cleavage products themselves are a requirement for the formation of a mature infectious virion [20,77]. The X-ray crystal structures of several HIV protease-PI complexes have been solved and used to assist the process of inhibitor development [84]. The elucidation of these structures has facilitated the understanding of the interplay in the many associated drug binding and resistance mechanisms at an atomic level. Such as, the pattern of hydrogen bonding between inhibitor and enzyme backbone atoms, compared to that with the natural substrates. The identification of these types of interactions can aid the discovery of mutations at specificity determining residues that could impact on inhibitor binding but not that of the substrate, and would therefore be a potential mechanism for the development of resistance to PIs. These types of mutations that envelop the active site and usually interfere with PI binding are referred to as primary resistance mutations and cause a reduction in sensitivity to PIs [20,84]. PIs were designed to fit into the protease binding site, the appearance of primary mutations changes the shape of this substrate cavity, causing the drug to significantly lose its inhibiting potency. Although most primary mutations are located near the protease catalytic site, each PI has its' own primary resistance mutation. The first PI, Saquinavir initially selects for the G48V primary mutation, although the appearance of the L90M is also seen [20,84,292]. Nelfinavir selects for the D30N, Amprenavir and the new PI Darunavir, an analogue of Amprenavir, both select for the I50V mutation, which resulted in a rapid increase in viral replication during *in vitro* selection experiments with Duranavir [351]. Two primary mutations that are located in the active site and are associated with failing PI therapy are the I84V and V82A/I/F/T. The V82A change occurs primarily with the older PIs Ritonavir and Indinavir, it results in significant changes in the flap and disruption of side chain interactions with PI, but not the backbone-backbone interactions with the substrate [349]. The I84V mutation is selected for by all failing PI therapies except Nelfinavir, its role in mediating resistance is unclear, but it has been suggested that it may protect the protease from autocatalysis [64,349].

Resistance mutations associated with PI resistance have been mapped to terminal domains, core domain and flap domain, PI resistance mutations have been reported at

45 of the 99 amino acids positions of protease [344]. Many of the PI associated mutations are not located within the substrate binding site and do not affect resistance levels to PIs [84]. These mutations are referred to as secondary resistance mutations because they appear with primary mutations and seem to restore the deleterious effects of the primary mutations on protease; these changes are also known as compensatory mutations. The effects of these secondary resistance mutations are largely unknown, theories as to how they mediate their effects could be through a mutated enzyme with enhanced catalysis [84]. Studies into the effects of non-active site mutations by measuring the binding affinities of resistant proteases to substrate and inhibitors, suggests that they can directly affect the binding of PIs without the presence of primary PI mutations. The development of non-active site mutations also revealed increased stability of the mutants, with or with out the primary resistance mutation over the wild-type enzyme [223]. However, another similar study which employed a less heavily mutated protease, demonstrated a solely compensatory role for the flap residues M46I and I47V [253]. The appearance of compensatory mutations is not limited to the targeted protein; PIs have been shown to induce changes in the protease substrate, Gag. The development of high level resistance to PIs requires the accumulation of multiple resistance mutations; multi-PI resistant protease, containing several resistance mutations has been shown to contain an expanded active site cavity [349]. Such effects of PI mutations not only impact on the binding of PIs, the processivity of protease, but also impact on the binding of the substrate. Therefore, another proposed mechanism for PI resistance was through the enhanced processing by mutant enzymes for a mutant substrate. The appearance of compensatory cleavage site mutations (CSMs) was first shown to occur *in vitro* using a peptidomimetic PI and a highly resistant protease with multiple resistance mutations [77]. The initial CSM was a Leu to Phe change at position P1' of p1/p6 and was present in nearly all the clones isolated from the PI passaged virus. This was followed by a second double mutation at NC (p7)/p1, which also acquired additional protease mutations; removal of the *gag* mutations severely impaired the growth of the PI resistant protease. The initial CSM was shown to appear with the I84V mutation and resulted in 100-1000 fold increase in resistance to the PI [77]. The first *in vivo* appearance of CSMs was reported during Indinavir therapy. The emergence of mutations at position P2 of the p7/p1 cleavage site was paralleled by the serial acquisition of protease mutations, unlike the *in vitro* work of Doyon *et al.*, which demonstrated a requirement for the

appearance of multiple PI resistance mutations prior to the CSMs [358]. The emergence of mutations in cleavage sites is thought to modify the substrate for the mutant proteases and generates a better peptide for the protease *in vitro* [77]. Due to the multi-functional role of $Pr55^{Gag}$ and its cleaved proteins in HIV replication, it is also possible that mutations could arise elsewhere in the HIV genome leading to improved viral growth in the presence of PIs.

Newer PIs such as Darunavir (TMC114) require several novel protease mutations in order to cause an increase in resistance to the compound. Viruses generated by *in vitro* passage which contained mutations R41T and K70E showed reduced susceptibility to TMC114 of approximately 10-fold compared to wild-type, but replicated poorly. However, when these mutations were introduced into a molecular clone through SDM, no reduction in susceptibility to TMC-114 was detected. This implies that other mutations had arisen in regions other than protease, most likely the Gag polyprotein and may play a key role in determining susceptibility to TMC114 [65].

A recent publication suggests that Darunavir use in patients with prior PI therapy and specific baseline PI resistance mutations (V11I, V32I, L33F, I47V, I50V, I54L/M, G73S, L76V, I84V, and L89V) may be associated with a poor *in vivo* response to Darunavir [133,170].

Recently, *in vitro* selection experiments using a novel PI (RO033-4649) demonstrated the selection of PI-resistant viruses. However, unlike older generation PIs, no substitutions were shown to occur in the viral protease. Sequencing of the Gag polyprotein revealed the presence of mutations at the NC/p1 cleavage site (K436E and/or I437T/V) in all resistant viruses. The mechanism leading to PI resistance is thought to be through increased mutant Gag polyprotein processing by wild-type protease [231].

1.6.3a Insertions and Deletions leading to PI resistance

Like RT inhibitors, PIs also select for insertions in the protease gene, several insertions have been reported ranging from 1 to 6 amino acids in length. The location of the inserts also varies and has been mapped at or between codons 35 and 38, 17 and 19, 21 and 25, 70, or 95 and 96. Like the insertions in RT, sequence analysis demonstrated the insertions were duplications of proximal sequences, possibly due to

RT slippage during reverse transcription. Construction of recombinant clones in which the inserts were removed, but primary and secondary PI resistance mutations remained, revealed they were no more resistant than the viruses carrying PI mutations and inserts, but provided a borderline advantage in replication. [24,156,309,314,341,342].

Emergence of CSMs in the *gag* substrate can occur with the acquisition of PI resistance mutations *in vivo*, resulting in reduced sensitivity to PIs, insertions and deletions have also been seen in *gag*. Two independent studies have revealed the duplication of the PTAPP motif after prolonged exposure to PIs [171,311]. As mentioned earlier, the PTAPP motif has been implicated in facilitating the late budding event from the cell membrane. The first study demonstrated the appearance of either partial or complete duplication of the PTAPP motif in virus isolated from patients after Amprenavir therapy. The duplications were significantly associated with the V82A/F/T/S mutation and a reduced response to therapy [171]. Insertions in *gag* occur in several other regions other than the PTAPP motif. This was looked at in the second study where insertions at the N- terminus of *gag*, around the p17/p24 cleavage site and also PTAPP insertions at C-terminus were found in multi-antiretroviral experienced patients failing therapy. Construction of full-length clones revealed the contribution of these insertions was required in order to rescue efficient Gag processing of the mutant viruses [311].

1.6.4 Resistance to other Antivirals

1.6.4a Fusion Inhibitor Resistance

The development of resistance to T-20 (Enfuvirtide) was determined by early *in vitro* passage experiments, these selected for resistant viruses carrying mutations located in a highly conserved region of the N terminus of HR1. Initially changes at codons G36S and V38M were detected, but it appears that the development of resistance is associated over a greater region of gp41, amino acids 36–45 with most changes in an amino acid triad (GIV) at positions 36–38 in the HR1 region of gp41 [118,273]. The location of these mutations supports the proposed involvement of HR1/HR2 binding in the fusion event, and that homologues of HR2, like Enfuvirtide, block this interaction. Changes in the HR2 region have also been implicated in reduced susceptibility to T-20, residues N126K, E137K and S138A. The N126K change is thought to eliminate a glycosylation site and changes at positions 137 and 138 are

thought to coexist with changes in HR1, predominantly at codons 42 and 43. Structure predictions of the HR1-HR2 six helix bundle structure indicate these residues lie in close proximity to each other [118,119]

Single substitutions such as the G36D substitution exhibit a 5 to 10 fold reduction in susceptibility to Enfuvirtide *in vitro* and double amino acid substitutions displayed >100-fold reduction [119]. The impact that this may have in a clinical setting is less clearly understood as variation in T-20 susceptibilities of primary isolates containing single mutations varied in resistance levels from 4 to 450 fold. The *in vivo* incidence of genotypic changes at amino acids associated with Enfuvirtide resistance is low in naïve populations, implying that this is a conserved region, critical for viral fusion [118,119].

Research has also determined Enfuvirtide resistance does not preclude the use of other classes of entry inhibitors, as *in vivo* derived resistance to T-20 did not appear to affect the activity of other coreceptor inhibitors [273]. In a separate study it was found that some HR1 mutations conferred increased sensitivity to a subset of neutralising monoclonal antibodies. The mechanism for this enhanced neutralisation could be due to reduced fusion kinetics [276].

Effectiveness of this class of compound is likely to abrogated by the natural gp120 variability among different HIV-1 subtypes. CD4–gp120 binding inhibitors although not clinically approved have showed effectiveness in *in vitro* work. Resistance to these compounds is primarily caused by changes within gp120 surrounding the 'Phe-43 cavity' [187]; a cavity created by the binding of the three gp120 core domains and CD4, in which Phe 43 of CD4 is the only CD4 residue that binds to this cavity.

1.6.4b Chemokine Inhibitor Resistance

Resistance to CCR5 and CXCR4 antagonists can possibly result from a switch in coreceptor usage. Incidences of this occurring are rare and viruses that are resistant to CCR5 inhibitors, tend to continue use of this receptor. Selection experiments performed in increasing concentration of a CCR5 inhibitor resulted in the development of a phenotype conferring >20,000 fold resistance by 19 passages, this virus continued to display an R5 entry preference [317]. Resistance to CCR5 antagonists can be mapped to both the variable and conserved regions in HIV envelope, whereas, CXCR4 resistance can mainly attributed to changes in the variable regions. Changes

in V3 account for resistance to both CXCR4 and CCR5 inhibitors, which can result in the development of cross-resistance to inhibitors of this class [26].

Maraviroc, a recently developed CCR5 antagonist failed to select for resistant variants after serial passage experiments using laboratory adapted and clinical isolates of HIV-1. Although, high-level maraviroc resistant viruses were eventually selected after passaging of primary isolates in PBMCs [336]. As mentioned in the previous section, natural gp120 variability among different HIV-1 subtypes may account for differences in baseline susceptibility to these types of inhibitors and may thus affect the mutational pathway to resistance [76]. It has also been proposed with this class of inhibitor that most resistance mutations are specific for each of the different compounds, which may hopefully limit cross-resistance [26]. However, further research with other CCR5 antagonists has demonstrated that cross resistance between structurally related molecules such as, AD101, SCH-C, and vicriviroc, can exist [336]. The most common mutations associated with maraviroc are changes at amino acid positions A316V and I323V, a three-amino-acid deletion in the V3 crown at codons 315–317 (QAI), T163K in V2, N355Y in C3, S405A in V4, D389N in V4, N442K in C4, E509K in the gp120/gp41 junction and S132N in V1 [336].

1.6.4c Integrase Inhibitor Resistance

Resistance to this class of inhibitor results from changes in the enzyme structure that cause changes in the integrase catalytic site and can also affect metal ion coordination, which can be detrimental to enzyme activity, often producing non-infectious progeny [172]. Development of resistant strains *in vitro* can take up to 6 months and requires the accumulation of multiple changes in order to confer high level resistance. Changes at codon T66I, L74M, F121Y, T125K, G140S, S230R, V249I and C280Y have been reported in conjunction with resistance in this class and these have not been reported to occur as natural polymorphisms [122]. Raltegravir (MK-0518) and elvitegravir (GS-9137) are two integrase inhibitors in late-phase development; both compounds are classed as integrase strand transfer inhibitors. In a recent study, raltegravir was used as salvage therapy for nine heavily pre-treated patients. Resistance was detected and four different mutational profiles were identified containing combinations of the mutations E92Q, G140S + Q148H, N155H and E157Q. The study also reported the observation that the HIV-1 viral load was able to rebound to the pre-therapeutic levels

for viruses harbouring only a single mutation [190]. This would suggest that the genetic barrier to resistance of this compound is low.

Although the aforementioned compounds, raltegravir and elvitegravir, are the only two compounds to have been developed for clinical use, several others have also been tested and have shown potent antiviral activity. The diketo acid derivatives L-731,988, L-870,810 and S-1360, have been described and in *in vitro* resistance selection experiments were shown to display a similar resistance profile to the two most advanced compounds. Active site mutations, T66I, S153Y, and M154I were shown to appear with L-731,988 and S-1360. Whereas, L-870,810 produced a unique resistance profile in *in vitro* resistance passage experiments, with mutations V72I, F121Y, T125K, and V151I, also located in and around the active site [297].

1.6.4d Capsid / Maturation Inhibitor Resistance

The potent maturation/capsid inhibitor PA-457 prevents the late stage cleavage event that releases capsid from the spacer protein p2. Resistance to this compound was first described after serial passage experiments selected for resistance mutations that were mapped to the p24/p2 site. A single A to V change at codon 231 of capsid or P1' of p24/p2, resulted in resistance to PA-457, as was determined by correction of cleavage patterns and also restoration of morphological impairments visualised by electron microscopy (EM) [179]. In a separate study an L230F mutation at the P1 site was also found to confer resistance to PA-457 [359]. More recent work has identified an additional two more sites located slightly more distal from the scissile bond that also confer resistance to this compound [2,192].

1.7 Transmission and Fitness

There are many biological determinants that influence the transmissibility of different viral variants from within the quasispecies present in the original host or donor. As mentioned in the Life cycle section 1.3, co-ordination is needed by virus and host factors in order to establish a productive infection. A virus which contains a high density of CCR5 receptors stands a better chance of gaining entry [354], although if the host contains a ΔCCR5 allele, then entry will be blocked or diminished. Even before

entry into the cell, other host and viral factors are again required to work in tandem, such as CyPA binding to capsid, glycoslylation of Env and Vif inhibition of APOBEC3G packaging; these and many other interactions will determine the success of an infection.

1.7.1 Transmisson

The immediate bottle-neck imposed on the infecting virus, is entry, this is demonstrated by the homogeneity of *env* V3 sequences and the observed heterogeneity of other regions of the genome in viruses isolated from recently infected individuals [71]. In one study, the V3 regions from viruses isolated during primary infection from five separate patients demonstrated complete homogeneity in all intra-patient samples and only showed variability in 6 out of the 35 residues of V3 [357]. These results suggest that viruses share a genetic signature across individuals due to site-specific selective pressures that are common across hosts, specifically located in the V3 region. The mode of transmission may also influence the diversity of the genotypes that infect the recipient. Infections termed as vertical transmissions are defined as those that are transmitted from donor to the donors progeny, as occurs from mother-to-child. Horizontal transmissions are those that occur from donor to another host within the species, such as through sexual contact and blood transfusion products [16]. Vertical transmission in thirteen mother and infants pairs demonstrated that more than half of the samples had a heterogeneous virus population; this is more diversity than detected in the previous mentioned horizontally transmitted study. The increased inoculum levels and elevated exposure times associated with vertical transmissions were thought to contribute to this observed increase in heterogeneity [328]. However, Zhang *et al.* demonstrated that the inoculum size may not necessarily contribute to viral diversity during primary infection, as several patient samples used in this study were from transmissions through contaminated factor VIII [357]. Another factor that has been associated with a greater heterogeneity in pre-seroconverters is gender; several studies have reported the that virus samples isolated from women recently infected have contained a more diverse quasispecies when compared to those in men [71,183]. This may possibly account for the higher transmission rates seen in women, who are up to 5 times more likely to become infected than men [294].

The worldwide predominant HIV subtype is subtype C (see Figure 1.2), it accounts for approximately 50% of worldwide infections possibly through the enhanced replication in the GALT [36]. A consequence of this high viral turnover is the increased detection levels in semen. Viral loads of subtype C are 3-4 fold higher than subtype B, leading to estimated transmission frequency increases of 2% to 6% for subtype B compared with 7% to 24% for subtype C [264].

The impact of host genetics on selection and divergence of infecting strains over time was illustrated in an interesting article in which two individuals acquired HIV through blood transfusion from a donor who was classified as a long term non-progressor (LTNP). Both the recipients developed AIDS and samples taken over the course of infection from all three individuals revealed that the main variation in HIV sequence appeared to be an insertion in the Nef gene, present only in the LTNP. However, both Nef alleles, with and without the insertion, were present in the LTNP; this suggests that there may be a host effect elicited by the LTNP donor, which suppressed the more pathogenic strain in this host. The major contributing factor driving the transition to a more virulent form was specifically thought to be Human Leukocyte Antigen (HLA) type [86,207]. Another demonstration of the modulation of virulence or pathogenicity was established using the primate lentivirus, SIV. This infects African green monkeys and sooty mangabeys, producing high levels of viremia, with no immune pathogenesis to the primate host. However, SIV infection in macaques results in high titre virus and the development of AIDS like symptoms [112].

1.7.2 Drug Resistance and Transmission

Effective HAART reduces VL to undetectable levels and increases the number of circulating CD4+ T cells [228], significantly delaying clinical and immunological deterioration related to HIV infection The effectiveness of regimens is halved with prior exposure to antiretrovirals, presumably through the lack of decay of previous resistant viruses and their subsequent selection on recommencement of therapy [184]. This finding has implications for the transmission of such resistant variants and the potential failing of therapies initiated.

Transmission of resistant viruses was first reported by Erice *et al.*; primary infection with an AZT resistance strain carrying the T215Y/F mutation was detected in a

recently infected individual, without any prior exposure to AZT. Initiation of Zidovudine treatment 7 days after reporting symptoms resulted in no clinic benefit and the patient was eventually switched to a ddI containing regimen. The identification of this case and the subsequent failure of treatment, prompted the researchers to postulate that, AZT treatment of primary infection with a resistant virus may have favoured a faster progression of the disease, possibly to due inhibition of immune functions by antiviral drugs and the lack of suppression selecting for a resistant strain with a higher virulence [83].

Further studies of transmission of multi drug resistant virus has demonstrated that these variants are less transmissible than drug sensitive (wild-type) viruses [175,348]. These studies were the first to demonstrate that HAART can significantly reduce the possible spread of the HIV infection by reducing the potential of transmissions.

The reduced rate of transmission of drug resistant variants may also be as a result of the apparent compartmentalisation that occurs between plasma or blood derived virus and that from the genital tract [265]. The lack of penetration of some antiretrovirals into seminal fluids [312] leading to reduced levels of antivirals could promote the appearance of resistance mutations in this compartment. However, most resistance testing is carried out on blood derived virus, which may be more suppressed and thus, carry a divergent resistant pattern to that of the seminal fluids. This could produce some discordant results when comparing viruses isolated from sexually transmitted, epidemiologically linked couples. Estimates of prevalence of transmitted drug resistance range from 10-20%, with some studies reporting rates of up to 27% [56].

1.8 Viral Fitness

Survival of the fittest is a Darwinian concept that was originally applied to explain the impact of evolutionary pressures on organisms [59]. These evolutionary forces either result in the adaptation of the organism through acquisition of beneficial mutations or result in its extinction. Viral fitness can be defined as the ability of a virus to replicate and produce infectious progeny in a host environment [56,67,323]. This definition includes the concept of survival of the fittest as viruses are required to respond to evolutionary

pressures exerted by host selection/adaptive forces and they may also have to respond to drug pressure, through use of antiretrovirals. The term replication capacity defines the ability of a virus to replicate in a fixed environment. It has also been used in the place of fitness; however, this is essentially one of the many components that make-up viral fitness and lacks the myriad selection forces within a host environment.

The bottle-neck effect of transmission is caused by the transference of a small fraction of the quasispecies present within the donor [285]. It is a chance event and due to the high error rate and viral turnover of RNA viruses, it is predicted that the successful virus will likely harbour a deleterious mutation compared to the original wild-type virus population. This results in the loss of the fit variant from the population and causes an overall replication deficit to the viral quasispecies. This principle which describes the tendency of small populations of asexual organisms to lose fitness, due to the accumulation of deleterious mutations is called Muller's ratchet [16].

A few *in vitro* bottle-neck experiments have shown Muller's Ratchet effect in dramatically reducing the fitness of HIV during repeated bottle-neck, plaque to plaque passage experiments. The high fitness heterogeneity among the initial passage was subsequently lost through serial passaging, resulting in the loss of 4 of the initial 10 clones before completion of passaging [355]. The observed mutation frequency did not correlate with the lack of growth and all clones demonstrated conservation of the V3 sequence; the most variability in the 'unfit' clones was found to occur in *gag* [353]. However, Muller's ratchet is not likely to have such a dramatic effect *in vivo*, especially once infection has been established; this is possibly due to the high rates of replication and availability of target cells [16,355].

The disease course in horizontal HIV transmission can also be explained in terms of maintaining fitness in the host. First HIV establishes an R5 infection phenotype, which is less pathogenic than X4 and this allows the intra-host infection to spread. For a long period of time the host remains asymptomatic and produces high amounts of virus that would facilitate any transmission through horizontal means which tend to be associated with a lower viral dose inoculum [16]. The switch to an X4 phenotype could be due to the limiting amount of target cells, as the initial infection with R5 virus has reduced the supply, therefore forcing the coreceptor switch to maintain infection. The X4 virus accelerates the disease status, leading to death. The most successful *in vivo* HIV is the subtype C virus and has been shown not to carry-out the coreceptor switch,

instead maintaining an R5 virus, possibly increasing survival of host and leading to increased transmission potential [39].

1.8.1 Drug Resistance and Fitness

The interplay among the various amino acid substitutions appears to be a highly intricate process in which the order of appearance of mutations plays a key role. The requirement for replication in the presence of inhibitors is paramount for the development of resistance; this would seem to be the first counter measure taken by the virus to adapt to its new environment. This first move results in the acquisition of resistance mutation(s), which may slightly impair the viruses' enzymatic processes, targeted by the inhibitor [228]. The acquirement of further mutations, leads to greater reductions in sensitivity to the antiretroviral(s) and additional impairment of enzyme function. The targeting of another protein, or in the case of NRTIs and NNRTIs, another functional region of the same protein, can result in the appearance of a multi-drug resistant virus, which incorporates non-synergistic mutations. The initial hope of this kind of multi-drug therapy was to force the virus into a series of fatal combinations leading to lethal mutagenesis. This theory called 'convergent combination therapy', was likened to a game of chess in which HIV moves through the acquisition of mutations to avoid the counter measures of the opponent i.e. antiretrovirals [278].

Another theory similar to the 'chess' paradigm, views all the possible mutations as a three-dimensional space; this 'landscape' contains 'peaks' and 'troughs' which represent levels of fitness. The space occupied by the wild-type virus may be at the top of a local peak, representing its' high level of fitness. Changes in environmental pressures, such as drug therapy will result in a change in the landscape occupied by the virus. Globally, the new landscape may now contain other peaks which are greater in height and therefore represent an improved level of fitness. This change in landscape may also result in the wild-type virus now occupying a relatively lower peak. In order to traverse this landscape and reach the higher peaks, the virus must contain the correct combination of mutations to move through the landscape. This may be achieved through the acquisition of new mutations, or the loss of mutations already present and possibly conferring a mutational pattern that results in the development of a maladapted virus, compared to the original strain. However, if this

reduction in fitness is not achievable i.e. unlikely to be selected for, then the virus will remain in this unfit state. Or the maladapted stage may only be a transitory state that is compensated for by other mutations, either lost or gained, leading to the generation of a more fit species [16,230,353,355].

Studies have revealed the transmission of viruses that carry 'relics' of a previous high level resistance phenotype. The viruses that were transmitted only conferred a low level resistance to NRTIs and therefore the growth characteristics displayed were less impaired. These viruses contained either single or double amino acid substitutions and were able to replicate better than wild-type virus in low and high levels of AZT. The transmission of these viruses would result in limited treatment options for the recently infected individual [104].

As mentioned earlier, viruses that display high resistance tend to have a reduced replication capacity and therefore have lower circulating VL, leading to reduced transmissibility [175]. The use of *in vitro* methods have demonstrated that several MDR HIV-1 viruses containing resistance mutations were more fit than wild-type virus in the presence and absence of drugs [164]. There have also been several *in vivo* reported cases where the transmitted virus has been multi-drug resistant and also showed high pathogenicity [200,269]. Recently, a case involving a newly infected homosexual male received a great deal of media attention, as it raised the possibility of a sudden widespread outbreak of a new 'supervirus'. This highly pathogenic, MDR HIV strain was feared to have spread among a very promiscuous, sexually active group of Men who have Sex with Men (MSM) and resulted in the rapid progression to AIDS of the infected individual. The virus coreceptor usage demonstrated the presence of equal proportions of both R5 and X4 i.e. dual tropic (X4/R5) variants, VL was 280,000 copies/ml and the genotypic and phenotypic data suggested Enfurvitide (T20) and the NNRTI, Efavirenz remained the only viable treatment options.

It is possible that due to the highly sexually active behaviour seen in some cases of infection of MSM, this virus had under gone serial bottle-neck events in several other hosts before infecting this individual. However, instead of Muller's ratchet selecting for the accumulation of deleterious mutations, a Darwinian selection mechanism may already have occurred selecting for a highly fit virus. The phenotype of this virus, MDR and highly pathogenic, could have been achieved if this virus had been exposed

to multiple MSM, who were also antiretroviral experienced and possibly, already infected, which would allow diversity to accumulate through recombination. This 'filtration' through several hosts resulted in the selection for an entry competent, highly resistant, pathogenic strain. The development of pathogenic strains is more likely to occur through horizontal means than vertical means, as the former does not require the host to reproduce [16]. Thus, the development of a highly virulent strain that infects and kills target cells, releasing equally or more pathogenic progeny, is not harmful to the overall success of horizontal infections [16]. Such a potential for the emergence of an extremely virulent virus has been demonstrated using an SIV model. The appearance of a highly pathogenic strain of SHIV (a chimeric SIV/HIV virus) was reported *in vivo* after rapid serial passage in rhesus macaques. The chimeric clone $SHIV_{SF162}$, was initially a non-pathogenic, M-tropic, CCR5 primary strain which was poorly transmissible by the intra-vaginal route. Upon serial passage through several macaques, the virus became highly pathogenic, maintained R5 phenotype and infected 2 out of 3 macaques via the vaginal mucosa [142].

Fitness differences can to some extent explain the complex patterns of resistance and compensation that are seen *in vivo* and *in vitro*. Viruses that confer resistance but demonstrate poor growth characteristics are not directly selected for, as is seen by the gradual accumulation of resistance to most antiretrovirals. Such an understanding of the complexities of resistance pathways have been derived from *in vitro* work that has tried to dissect the mechanisms associated with fitness and determine their impact on viral phenotype.

1.8.2 Resistance and Fitness Assays

The appearance of mutations associated with antiretroviral therapy impacts on the virus protein function, overall protein levels and virus replication capacity. All these effects have been utilised as a means of measuring the effects of therapy associated mutations, by comparing the abilities of viruses with mutations to perform against reference strains. The potential examination of fitness raises many questions such as the cell type to be used, PBMCs vs T-cell lines; it has been demonstrated that the debilitating effects of protease mutations were accentuated in T-cells compared to PBMCs [19]. The use of primary isolates over recombinant clones is also another factor

that should be considered, as is demonstrated by the effects of PIs on the appearance of *gag* mutations and the possible association of p6 mutations with NRTI therapy [256]. Also, the identification of pleiotropic effects is more likely to be observed if the whole genome is used [19]. Therefore, fitness studies require a precise understanding of the possible limitations of the methods used to answer the questions being addressed.

Figure 1.14 Different approaches used for measuring HIV-1 fitness and replication capacity. In vivo assays use viral kinetics such as RNA viral load and together with sequence analysis are able to determine the fitness of the viral population. All other assays are ex vivo and measure virus replication capacity or viral enzyme activity.
Adapted from [272].

1.8.2a Assessment of Enzyme function

Measurement of RT activity of mutations associated with AZT resistance, demonstrated that these mutations conferred only small changes in enzyme efficiencies which did not reflect the levels of resistance seen in the viruses [167,168]. The study into the effects of Lamivudine mutations M184V/I, revealed a reduced fitness of these mutations compared to the wild-type strain, primarily in primary cells or those where dNTP pools are limited. These were the first studies to report an impaired replicative fitness of mutations associated with therapy [11].

In the context of PIs, the impact of primary resistance mutations on catalytic activity and affinity for substrates was used as the first *in vitro* measure of fitness and was termed enzyme vitality [84]. The vitality value is the cleavage efficiency of mutant compared to that of the wild-type protease in the presence or absence of inhibitors. Vitality is essentially a measure of enzymatic or biochemical fitness. These studies highlighted the reduced catalytic effects of HIV protease containing primary mutations and the compensatory effects of other mutations [230,253]. However, these methods are unable to address the role of such modulations in enzyme activity on other processes within the life-cycle, such as increased proteolytic activity, which has been shown to be detrimental to infectious virion production [337].

1.8.2b Replication Assays

The development of the RVA by Kellam *et al.*, was initially used to asses the impact of changes in RT on RTI sensitivity [152,153]; since its' conception it has been utilised and adapted further and forms the basis for many of the *in vitro* fitness assays involving virus replication. These assays are performed by cloning the coding region(s) of interest into a standard viral backbone, usually from a sub-type B molecular clone such as, HXB2 or NL4.3. The source for the coding region to be cloned can be either a wild-type gene with specific mutations introduced through site-directed mutagenesis or it can be derived from a patient virus with prior exposure to antiretrovirals. These methods provide a constant background in the context of minimising the effects of other coding regions on the recombinant virus, other than those outside the coding region of interest. Assessment of viral replication kinetics

can be carried out by measuring the rate of growth of the different viruses over time, usually through differences of p24 levels between the cultures or differences in RT activity [323]. These methods are only able to detect large fitness differences in populations, as they are subject to a great deal of variation due to each virus being grown independently in a parallel infection [272].

1.8.2c Growth Competition

Quantification of initial input of viruses to be tested in fitness assays is of extreme importance, as any small differences in the initial phases will be several magnitudes higher after multiple rounds of replication. Competing two viruses together in the same culture also known as growth competition studies, allows for the detection of much smaller fitness differences [272]. By monitoring the starting ratio and the ratio of the two competing viruses in subsequent passages, divergence in the proportion of each virus after passaging is used to indicate the differences in fitness. The amount of each virus present in each passage can be determined by sequencing of viral DNA/RNA and comparing relative peak heights at nucleotides positions known to be different between the two competing viruses [185]. These assays provide an accurate means of distinguishing small fitness differences, but are more labour intensive due to the extensive amount of sequencing required. Comparing of relative peak heights is a relatively imprecise method of end-point determination and more accurate ways to determine the amount of viruses present in competition growth studies have been developed (see below). This growth competition type of assay system was employed to elucidate the possible mechanisms for the preferred appearance of the T215Y mutation pathway to NRTI resistance over the Q151M; even though the latter confers greater resistance. In competition growth cultures the Q151M virus was able to out compete the T215Y virus in the presence or absence of AZT. Both these mutations require a two base change Q151M (CAG→ATG) and T215Y (ACC→TAC). However, it was demonstrated that the Q151M intermediates competed poorly compared to the T215Y intermediates. The T215Y intermediates have been found *in vivo* and have also been shown to replicate better [115,116]. This study may also explain the late emergence of the Q151M viruses and the acquisition of compensatory mutations [164].

Several methods have been developed to circumvent the imprecise and labour intensive use of peak heights to determine virus ratios. The heteroduplex tracking assay (HTA) utilises the fact that DNA heteroduplexes formed between related sequences have a reduced mobility in polyacrylamide gels and the migration rate is proportional to their degree of divergence [272]. This assay involves the hybridisation of a radioactively labelled probe to a PCR product, derived from the total viral population. This generates a probe-PCR product heteroduplex which is specific for the two competing viruses. The heteroduplexes can be separated by gel electrophoresis; a drawback to this approach is that the design of probes limits the HTA to systems with known mutations [72,73]. A more recent approach has been to develop vectors that contain marker regions in the genome of the virus. This method called the recombinant marker virus assay (RMVA) generates recombinant virus clones of HIV-1 in which the *Nef* gene contains either the *Salmonella typhimurium* histidinol dehydrogenase (hisD) or human heat-stable placental alkaline phosphatase (PLAP) gene. The HisD and PLAP containing viruses then can be grown in competition and the proportion of each virus in the population can be assessed by quantifying the linked hisD or PLAP marker gene, using real-time PCR [185]. A similar approach was used by Ali *et al.*, in which two variants of the murine heat stable antigen (HSA) protein were used as reporter genes [4]. One variant of the protein contained nine amino acids from the extracellular domain of influenza hemagglutinin (HA) antibody epitope. This enabled the use of immunofluorescent antibodies against HSA and the modified HSA, to specifically detect cells infected with each virus. End-point determination was by real time PCR quantitation or by fluorescence-activated cell sorting (FACS) analysis.

1.8.2d Virus Production Assays

These assays were originally utilised for drug susceptibility purposes [257], but have now been adapted to measure replication capacity of virus isolates derived from the protease and RT regions of patients on antiretroviral therapy [31]. This assay uses a replication defective HIV clone containing a luciferase reporter gene inserted within a deleted region of *env*. The PR-RT region is cloned into the DNA vector and viral stocks are prepared by cotransfecting cell cultures with the cloned protease-RT vector and an expression vector that produces the envelope proteins of murine leukemia

virus (MLV). Pseudotyped virus particles harvested from the transfected cell cultures are used to infect fresh cells. Because the protease-RT vectors contain a reporter gene within the *env* region, the production of luciferase in target cells is limited to one round of virus replication. The amount of luciferase activity detected in the infected cells is used as a direct measure of replication capacity. Differences between the amount of virus produced from the cloned protease-RT vectors and a recombinant reference virus derived from NL4-3 are expressed as the percent replication compared to the reference strain. [31]. This method is unable to be used to determine the comparative fitness of viruses, as these assays do not measure the competitive growth of different viral species. However, they are less labour intensive than competition experiments and also very accurate [323].

1.8.2e *In vivo* Fitness Determinations

The development of an SIV model for examination of *in vivo* resistance to NNRTIs has been described [6]. As SIV-RT in not susceptible to NNRTI inhibitors, RT substitutions were introduced; tyrosines at codons 181 and 188 were added to confer NNRTI susceptibility on SIV-RT and also another SIV clone was created that contained the whole of HIV-1-RT. Both these viruses produced decreased levels of RT activity compared to wild-type SIV and HIV and were also susceptible to the NNRTIs Efavirenz and Nevirapine. Passage of these viruses also resulted in the appearance of mutations associated with NNRTI therapy. Although this system has not yet been used to infect macaques and asses the rate of progression of these viruses compared to wild-type SIV, it may be a useful model for the preclinical evaluation of NNRTIs and for the development of drug resistance to these compounds *in vivo*.

Many *in vivo* studies of fitness have focused on the relationship between resistance mutations present during transmission and their evolution. Transmission cases possibly represent the best cohorts for the study of fitness, as the initial virus population size and diversity is small and it is forced to adapt to the milieu of the host. Such ideal environments can not be replicated *in vitro* and have enabled a greater understanding of the genetic barriers associated to the development of resistance. Delaugerre *et al.*, reported the transmission of MDR virus in two individuals who

maintained a homogeneous resistant population that persisted for 2 years without any treatment intervention. Due to the highly homogenous populations detected in the both donor and recipients, it is possible that the ability of the virus to traverse through the sequence landscape was limited to the development of mutations through error, as opposed to any potential recombination mechanism. This genotype represents an example where the virus is unable to back mutate due to genetic constraints and is forced to maintain an unfit genotype [70]. The fusion inhibitor Enfuvirtide, represents a compound with a low genetic barrier to resistance, i.e. a compound that has a relatively easy pathway to resistance, this was demonstrated in primary isolates containing single amino acid substitutions which displayed 450-fold resistance [118]. The fitness of viruses containing T-20 mutations is thought to be lower than wild-type in the absence of drug. Viruses with double amino acid substitutions were less fit than those with single substitutions, possibly due to reduced fusion kinetics. The effects of T-20 mutations *in vivo* have also suggested that these mutations are unfit. Following discontinuation of Enfuvirtide a rapid reversion to wild-type virus occurs, leading to outgrowth of the resistant virus and restoring replicative capacity [68,118].
One study also suggests that maintaining suppression of R5 virus may prevent the acquisition of MDR viruses [228]. In an *in vivo* study both R5 and X4 viruses demonstrated similar replication capacities, however, the R5 viruses were associated with a higher number of resistance associated mutations. This may suggest that R5 viruses respond better to environmental changes i.e. have better adaptability, through a mechanism other than increased replication capacity.

One of the main assets of studying fitness has been to demonstrate the possible advantages continuing with a potentially failing regimen in order to preserve an unfit but resistant virus phenotype. This reduced fitness approach as a therapy strategy is possibly the only way such research will benefit those individuals who are infected with a quasispecies, pre-exposed to antiretrovirals. However, in a antiretroviral experienced cohort, this strategy could possibly result in the development of highly resistant strains(s). It also is hoped that continued research into this area will lead to better management of patients who have limited treatment options and could lead to more effective treatment regimens with lower toxicities and costs. Although, the ultimate goal would be a form of 'convergent combination' therapy leading to lethal mutagenesis, this is highly unlikely with the arsenal of drugs available at the moment.

Ultimately, there is a need for more new compounds, either at targets already addressed or at new targets all together. These compounds need to combat resistance and ideally lead to non-synergistic mutational pathways; there is also need for reduced pill burden and less complicated regimens to address compliance.

The best hope for a'reduced fitness' antiretroviral therapy strategy in controlling viraemia may be early on in infection when the population is thought to be more homogeneous. Treating with the appropriate cocktail of drugs from the onset to force the virus to mutate to a more controllable form is a more realistic possibility.

1.9 Project Overview

The treatment of HIV infection with antiretrovirals in most cases successfully reduces viral burden and may lead to improved immune response. Prolonged therapy selects for resistant virus species with reduced sensitivity to antiretrovirals. Individuals infected with HIV in the developed world often undergo routine testing to determine the genotype of the virus. Most genotypic and phenotypic tests only include the protease and RT regions, as these are the proteins targeted by the antivirals. The acquisition of mutations by HIV in response to PIs, leads to the emergence of viruses with compensatory mutations in the *gag* gene. The effects of these mutations are not clearly understood, especially those that occur distal from the cleavage sites.

This study was undertaken to try and develop tools for characterising drug resistant viruses and to determine which gene products were required for the development of a replication competent virus in the presence of drug resistant and compensatory mutations. The initial approach was to utilise a single cycle replication assay system in which the *gag* gene products from patient samples were used to replace the wild-type *gag* and to determine the effects on viral production. The study also details the production of a live HIV vector system, derived from the sub-type B molecular clone NL4.3. This live virus cassette system enabled the transfer of single or multiple coding sequences; *gag*, protease and RT from HIV positive, antiviral treated patients. Vectors containing patient derived virus gene products were used to generate replication capable recombinant virus clones. During screening of samples to identify patients with prior PI experience and possible rebound virus, a highly resistant isolate with a novel two amino acid insertion was identified. This report describes the production of chimeras constructed from this virus isolate containing, *gag*, *gag* and protease and protease only genes derived from the patient virus. These recombinant clones were used to asses the impact of the resistant and compensatory mutations not only to antiviral sensitivity, but also to the growth characteristics of the virus itself. Drug susceptibility was determined by MTT assay. Growth kinetics were measured by comparing growth rates of the viruses in MT-4, SupT1 and GHOST cell lines. Sequencing of integrated proviral DNA was also carried out on passaged virus stocks to assess stability of the genotype. An attempt to compare replication capacity through

dual competition experiments was instigated to determine the contribution to replication capacity of each of the gene products.

The importance of understanding *gag*'s contribution to PI resistance through phenotypic analysis could assist in the development of algorithms to create better therapies based solely on genotypic data, therefore, resulting in cheaper and less time consuming analysis. As Gag performs a multifunctional role throughout the HIV life cycle, all regions of Gag would need to be included if a clearer understanding of its contribution to viral fitness is to ascertained. The aim of this study was to develop a system that would allow the whole *gag* region and the region upstream from *gag*, involved in packaging of viral RNA, to be analysed. Although the primary aims were to asses *gag*'s contribution to replication capacity and drug resistance with and without the cognate protease, divergence from the clonal input sequence was also assessed. Adaptation to a new environment is a key concept in the Darwinian definition of fitness, this effect has also been examined in this study as a measure of proviral DNA sequence divergence from the input clone sequence. In this way any changes in sequence after passage of the virus would represent positions which are detrimental to virus growth and so are lost. Comparing sequences from viruses containing different combinations of wild-type and patient derived *gag* and protease regions may facilitate a greater understanding of the adaptive changes required to produce a replication competent virus.

Chapter 2 Materials and Methods

2.1 General Molecular biology Techniques

Standard molecular biology techniques as described in Molecular Cloning: a Laboratory Manual, Maniatis *et al.*[197] were used throughout.

2.1.1 Preparation of competent bacteria

The following *Escherichia coli (E. coli)* strains were used: SCS110, DH5-α, HB101's, MAX Efficiency® Stbl2™ (Invitrogen), XL-GOLD and XL-1 Blue (Stratagene). Bacteria were streaked onto S.O.B. (2% (wt/vol) bacto typrone, 0.5% (wt/vol) bacto yeast, 0.05% (wt/vol) NaCl, 0.02% (wt/vol) KCl, 4ml 1M NaOH/L) after autoclave and prior to use 5ml 2M MgCl2/L) (Hanahan's Broth [129]) agar plates from a frozen stock and incubated at 37°C for 16 hours. A colony was picked and used to inoculate 5 ml Luria-Bertani (LB)-broth (1% (wt/vol) bacto typrone, 0.5% (wt/vol) bacto yeast, 0.5% (wt/vol) NaCl) with 50μg/ml (wt/vol) ampicillin and shaken at 37°C for 16 hours. This was added to 500 ml LB-broth and shaken at 37°C until the absorbance at 600 nm was 0.6 (around 3 hours). The culture was then incubated on ice to cool for 10 minutes. The bacteria were pelleted at 3000g for 10 minutes at 4°C and resuspended in 15 ml of 100mM calcium chloride (4°C). After incubation on ice for 30 minutes, the bacteria were centrifuged again and then resuspended in 2.5 ml of 100 mM calcium chloride containing 15% glycerol by volume (4°C). The bacterial suspension was frozen on dry ice in 400 μl aliquots.

2.1.2 Introduction of plasmid DNA into *E. coli*

1-200 ng of plasmid was mixed with 200 μl of competent *E. coli* and incubated on ice for 15-30 minutes. The cells were then heat shocked for 45 seconds at 42°C and then cooled on ice. 300 μl of S.O.C. media (S.O.B. media plus 20 mM glucose) was added to the reaction and incubated at 37°C for 1 hour. The cells were then plated on LB-agar plates containing 50 μg/ml ampicillin or kanamycin and incubated overnight at 37°C.

2.1.3 Double Stranded Plasmid DNA purification

2.1.3a Mini-preps

Mini preps of plasmid DNA were produced using the QIAprep Spin Miniprep Kit (Qiagen) according to the manufacturer's instructions. DNA was extracted from 3ml of LB-broth containing 50μg/ml ampicillin overnight cultures, inoculated using a single transformed *E. coli* colony.

2.1.3b Midi-preps

Midi preps of plasmid DNA were produced from 50ml of LB-broth containing 50μg/ml ampicillin overnight cultures of single transformed *E. coli* colony, using the Plasmid Midi Kit (Qiagen) according to the manufacturer's instructions.

2.1.3c Plasmid DNA maxi-preps

Maxi preps of plasmid DNA were derived by inoculating 500ml of LB-broth containing 50μg/ml ampicillin using a 10ml overnight culture, produced from a single transformed *E. coli* colony. The 500ml culture was grown a further 16-24 hours and the maxi prep was produced using the Plasmid Maxi Kit (Qiagen) according to the manufacturer's instructions.

2.1.4 DNA Quantification

DNA quantification was performed using 1μl DNA sample and a NanoDrop® ND-1000 UV-Vis Spectrophotometer (NanoDrop).

DNA was also quantified by visualization using Ethidium Bromide (Invitrogen) staining and agarose gel electrophoresis of a 5μl aliquot of purified DNA run along with a DNA molecular marker, Marker X (Roche).

2.1.5 Molecular Cloning

DNA insert and vector were digested with appropriate restriction enzymes, according to the manufacturer's protocols and each product was separated and visualised using Ethidium Bromide (Invitrogen) staining and agarose gel electrophoresis. They were then gel purified using the StrataPrep® Gel Extraction Kit (Stratagene) according to

the manufacturer's instructions. Also, prior to ligation, the vector was treated with shrimp alkaline phosphatase (Roche), to prevent self-ligation.

A typical ligation reaction contained vector:insert molar ratios of 1:3, 1:5 or 1:8, , in a final volume of 21µl and total DNA content of 200ng. Ligations were performed primarily using the Rapid Ligation Kit (Roche) and the ligation mixture incubated at room temperature for 5 to 15 minutes. Other ligation kits and T4 DNA ligases were also tested. These were Quick Ligation Kit and T4 DNA Ligase (New England biolabs), Fast-Link™ DNA Ligation (Epicentre), T4 DNA Ligase (Stratagene), all used in accordance with the manufacturer's specifications.

2.16 Polymerase Chain Reaction (PCR)

2.1.7 High Fidelity PCR's

Expand Long Template PCR System (Roche) and Phusion (New England Biolabs) polymerase kits were used according to manufacturers instructions.

2.1.7a Expand Long Template PCR System

The extension time, steps 4 and 8, varied depending on the length of amplicon generated; for small amplicons up to 1.5kb a 2minute extension was deemed sufficient, amplicons up to 3kb required 4 minute extension time and amplicons up to 10kb had 8 minute extensions. The annealing temperature also varied depending on which oligos were paired and was determined by the melting temperature (Tm) of the DNA oligo, which was derived from the PC application 'VecNTi' (Invitrogen).

Typical PCR cycling condition were:

1. 2 minutes at 94°C
2. 15 seconds at 94°C
3. 30 seconds at 55°C
4. 2 min at 72°C

Steps 2-4 an additional 10 times

5. 15 seconds at 94°C
6. 30 seconds at 55°C
7. 2 min at 72 °C (+5seconds per cycle)

Steps 5-7 an additional 20 times

8. Hold at 4°C

2.1.7b Phusion High Fidelity PCR

The annealing temperature also varied depending on which oligos were paired and was determined by the Tm of the DNA oligo, which was derived from the PC application 'VecNTi' (Invitrogen).

Typical PCR cycling conditions were:

1. 1 minutes at 98 °C
2. 10 seconds at 98 °C
3. 20 seconds at 55 °C
4. 1 min at 72 °C

Steps 2-4 an additional 35 times

5. 10 min at 72 °C
6. Hold at 4 °C

2.1.8 'A' Addition ('A' Tailing)

To facilitate the TA cloning of products generated with high fidelity polymerases, PCR products were heated to 95°C for 20 minutes, then to the 50µl PCR, 15µl of 2mM dATP and 5 units of *Taq* Polymerase was added and the mixture incubated a further 10 minutes at 70°C. 2µl of the 'A' tailed reaction was used in the subsequent TA cloning experiment

2.1.9 PCR TA Cloning using the pGEM-T-Easy vector system

The desired region was PCR-amplified using a high fidelity polymerase and 'A'tailed, see sections 2.1.13 and 2.1.10. The correct molecular weight product was separated and visualised using Ethidium Bromide (Invitrogen) staining and agarose gel electrophoresis. The desired product was subsequently excised and extracted using the StrataPrep® Gel Extraction Kit (Stratagene) in accordance with the manufacturer's instructions.

The pGEM-T-Easy vector system (Promega) ligation reaction was set-up according to the manufacturers protocol as below.

5 µ l 2X ligation buffer
1 µ l T4 DNA ligase
1 µ l pGEM-T-Easy linear DNA
1-3 µ l PCR product
0-2 µ l H_2O

Ligation reactions were transformed into *E. coli* as described in section 2.1.2 and positive colonies were identified by blue-white-screening, PCR-screening or restriction enzyme digest of purified plasmid DNA as required .

2.1.10 DNA sequencing

DNA sequencing was carried out by Lark Technologies (now Cogenics (http://www.nucleics.com/DNA_sequencing_support/sequencing-service/lark-technologies.html http://www.nucleics.com/DNA_sequencing_support/sequencing-service/lark-technologies.html)

Typically 100ng of plasmid DNA or gel purified PCR DNA was sent for DNA sequencing along with the required primers at 10 µM concentration. Sequencing was performed with BigDye terminator chemistry and a 3730xl Analyzer (Applied Biosystems).

2.1.11 Mutagenesis

Mutagenesis was performed with *Pfu* Polymerase (Stratagene), according to the Quickchange™ mutagenesis protocol (Stratagene). The site-directed PCR mutagenesis reaction contained 100-200 ng of plasmid DNA and 125ng of each forward and reverse primer containg the required mutation and incorporated the following PCR cycle:

1. 3 minutes at 94°C
2. 30 seconds at 94°C
3. 30 seconds at 60°C
4. 12 min at 68°C
5. Steps 2-4 an additional 18 times
6. 20 minutes at 70°C

The extension time, step 4, was determined by the size of the vector on which the mutagenesis was carried out. The protocol recommended an extension time of 2 minutes per kb, so for a 10 Kb plasmid the extension time was extended to 20 minutes. This was followed by addition of *DpnI* enzyme (1µl, 10units/µl) to the reaction which was incubated for 2 hours at 37°C. 2 µl of the reaction was then transformed into *E. coli* bacterial strains (section 2.1.2).

2.1.12 DNA Oligonucleotides

Table 2.1 (below) lists the oligonucleotides (oligos) used in this research project, the oligos listed were used for cDNA synthesis, PCR amplification and DNA sequencing, those oligos used for mutagenesis are contained with in the text of the section in which they are discussed.

Oligo Name	Oligo Sequence (5'→3')	3'/5'
3'Gag-*Pvu1*	ATACAGTTCCTTGT**CGATCG**GCTCC (2242)	3'
3'Integrase-*Xma1*	AGCTGCCATAT**CCCGGG**GCTACAGTCTACTTGTCC (4418)	3'
3'Pro-*SexA1*	GGGCCATCCAT**ACCTGGT**TTTA (2605)	3'
3'*Bgl*.seq	GGGAACCGGAGCTGAATGAAGCC (12651)	3'
3'p8.91*EcoR*-down	AGTCTAGGATCTACTGGCTCC (5852)	3'
3'pos: 13985	ATCTCCATTTTGATGTGAAGG (13985)	3'
5'Gag-*NarI*	TAGCAGT**GGCGCC**CGAACAGGGACTTGAAAGC (630)	5'
5'Gag-*Pvu1*	AGGAGC**CGATCG**ACAAGGAACTGTATCC (2216)	5'
5'Gag-*XbaI*	GCCCGAACAGGG**TCTAGA**AAGCG (640)	5'
5'Amp-R	AATGCTTAATCAGTGAGGCACC (12283)	5'
5'*Bgl*.seq	GGCTTCATTCAGCTCCGGTTCCC (12629)	5'
AQGAG5	ATACTGACGCTCTCGCACCCATC (810)	3'
AS3LTR	CGCCGGCCCGGGTGCTAGAGATTTTCCACACTGAC (9721)	3'
ASGAG 5A+	CTCTCTCCTGCTAGCCTCCGCTAGTC (789)	3'
F122271	TAATCCTCATCCTGTCTACTT (5095)	3'
Gag1	ATGGGTGCGAGAGCGTCAGTATT (790)	5'
Gag 3	CATCTAGTATGGGCAAGCAGGGA (886)	5'
GAG AS1	CATCTGCTCCTGTATCTAATAGAGC (2340)	3'
GAG SE2	GCACAGCAAGCAGCAGCTGACAC (1132)	5'
Gag SE3	GCAGGAACTACTAGTACCCTTCAGGA (1498)	5'
GAG SE4	GTTGGAAATGTGGAAAGGAAGGACACC (2027)	5'
HG/F/*NotI*	ATA**GCGGCCGC**TGGTGAGAGATGGGTGCGAGAGCGTCAGTATT A (781)	5'
HG/R/*BclI*	CTATGAGTATC**TGATCA**TACTGTCTTACTTTGATAA (2445)	3'
pos:1104(rev)	CTTACTTTTGTTTTGCTCTTC (1125)	3'
pos:1104	GGAAGAGCAAAACAAAAGTAAG (1104)	5'
pos:1620	TAGCCCTACCAGCATTCTGG (1620)	5'
pos: 2122	CCAGGGAATTTTCTTCAGAG (2122)	5'
pos: 2619	TGGCCATTGACAGAAGAAAA (2619)	5'
pos: 3102	GATGATTTGTATGTAGG (3102)	5'
pos: 3687	ATAGTAATATGGGGAAAGACTCCTAA (3687)	5'
pos: 4150	CATGGGTACCAGCACACAAAGG (4150)	5'
pos: 4644	TTTGGAATTCCCTACAATCCC (4644)	5'
pos: 5071	TGGCAAGTAGACAGGATGAGGAT (5071)	5'

Table 2.1: Oligonucleotides used in this Study. Restriction sites are shown in bold. Numbers in parenthesise indicate the position of the 5' nucleotide, according to those of HIV-1 vector pNL4-3, complete sequence (AF324493), except for oligo HG/F/NotI , where numbers in parenthesise indicate the position of the first HIV sequence nucleotide. 3' or 5' indicates if the oligo is on the forward or complementary strand.

2.1.13 cDNA synthesis

Viral RNA extraction had previously been carried out during routine resistance testing. cDNA was prepared from 7µl of patient viral RNA (that had been stored at -80 °C) using SuperScript™ II Reverse Transcriptase (Invitrogen) according to manufacturers instructions and using the oligonucleotide F122271.

2.1.14 Clinical Patient samples

Patient samples were identified from a database search of patient samples undergoing genotypic resistance testing at UCL, London, UK. The criteria for selection was extensive prior PI therapy, high viral load and the presence of multiple PI mutations. As control samples some patients were also selected that had no prior therapy or no primary resistance mutations. Samples were originally submitted for resistance testing and AMPLICOR® HIV-1 Test (Roche).

The CD4 counts and viral loads for patient samples used in this study are shown in figure 2.1(a and b). The data is also given in tables 2.1a, b, c, d and e, the sample time point in bold indicates the sample used in this study. Regimen highlighted in grey box indicates a switch to new regimen.

Patient sample 1 dated 07/01/2003 had a viral load of 91300 copies/ml and had previous exposure to antivirals, although the dates of these were not known. Antivirals this patient was previously exposed to were Indinavir, Efavirenz, Ritonavir, Zidovudine, Saquinavir, Stavudine and Lamivudine. Antiviral therapy over the 1 year period during the time of sampling was Kaletra (Lopinavir, RTV boosted), Lamivudine, Nevirapine and Didanosine. This sample was selected as the patient showed no response to therapy and had extensive previous exposure to PIs. This patient also showed a genotype that indicated resistance to PIs, NRTIs and NNRTIs (see Table 3.1).

Patient 2 had been off antiviral therapy for 10-12 months and previous antiviral history was unavailable. This patient appeared to controlling viral load in the absence of therapy and was selected as a 'no previous exposure to antiretrovirals control'.

For the remainder of the patient samples all other previous antiviral history is shown in figure 2.1(a and b) along with viral load and CD4 counts at the various sample time points available.
No adherence to regimen information was available for patients 1 and 2. Patient 3 was also selected as a non-responder. This patient had excellent adherence, however, the regimen of Nelfinavir and Combivir (AZT and 3TC) was switched due to increasing viral load. The switch to ddI, Abacavir and Nelfinavir was also unsuccessful at controlling viral load.

Patient 4 had poor adherence to the regimen and also showed no reduction in viral load during therapy and was also classed as a non-responder. Patient 5 had excellent adherence, and also showed possible rebound when therapy was switched from Trizivir (Abacavir, AZT and 3TC) and Efavirenz to Tenofovir, ddI and Nevirapine, which was also unsuccessful at controlling viraemia. (you keep on switching between generic names and abbreviations, use and stick to one or the other)

All samples were from patients attending clinic and obtaining routine viral load and genotype resistance data. The only patient material stored was RNA used in patient virus genotype identification. Routinely, patient serum or infected PBMC samples are not stored and so were not available and so all work was carried out with RNA that had been stored at -80C.

Table 2.1(a) Patient sample 1 Viral load, CD4 and previous antiviral history data.

Sample date	VL copies/ml	CD4 cells/ml	Therapy
07/02/2002	1900	159	Nevirapine, Lopinavir, Didanosine, Lamivudine
02/04/2002	50000	317	"
28/06/2002	24300	141	"
29/08/2002	61300	183	"
07/01/2003	**91300**		"

N.B. Prior to 2002: Indinavir,Efavirenz, Ritonavir, Zidovudine, Saquinavir, Stavudine, Lamivudine.

Table 2.1(b) Patient sample 2 Viral load data.

Sample date	VL copies/ml
07/02/2002	58200
30/04/2002	<50
20/08/2002	27900
08/05/2003	**34900**
12/08/2003	<50
18/11/2003	<50

Table 2.1(c) Patient sample 3 Viral load, CD4 and previous antiviral history data.

Sample date	VL copies/ml	CD4 cells/ml	Therapy
01/03/2001	155000	120	Nelfinavir and Combivir
01/03/2002	<50	200	Nelfinavir and Combivir
01/05/2002	400	226	Nelfinavir and Combivir
25/07/2002	1000	183	Nelfinavir and Combivir
01/10/2002	210	269	Nelfinavir and Combivir
08/11/2002	2700	212	Nelfinavir and Combivir
09/01/2003	2900	211	Nelfinavir and Combivir
03/04/2003	3700	155	Nelfinavir and Combivir
06/06/2003	14700	198	**ddI, Abacavir and nelfinavir**
05/08/2003	**25100**		**Not known if continued**

Table 2.1(d) Patient sample 4 Viral load, CD4 and previous antiviral history data.

Sample date	VL copies/ml	CD4 cells/ml	Therapy
13/01/2003	25600	195	
04/03/2003	77500	160	Nevirapine and Combivir
12/06/2003	69300	207	Nevirapine and Combivir
10/09/2003	**52600**	**186**	**Nevirapine and Combivir**

Table 2.1(e) Patient sample 5 Viral load, CD4 and previous antiviral history data.

Sample date	VL copies/ml	CD4 cells/ml	Therapy
14/01/2003	110700	7	Trizivir and Efavirenz
11/04/2003	<50	68	Trizivir and Efavirenz
23/05/2003	<50	79	Trizivir and Efavirenz
04/06/2003	2100		**Tenofovir, ddI and nevirapine**
08/07/2003	700		Tenofovir, ddI and nevirapine
13/10/2003	**73100**		**Tenofovir, ddI and nevirapine**

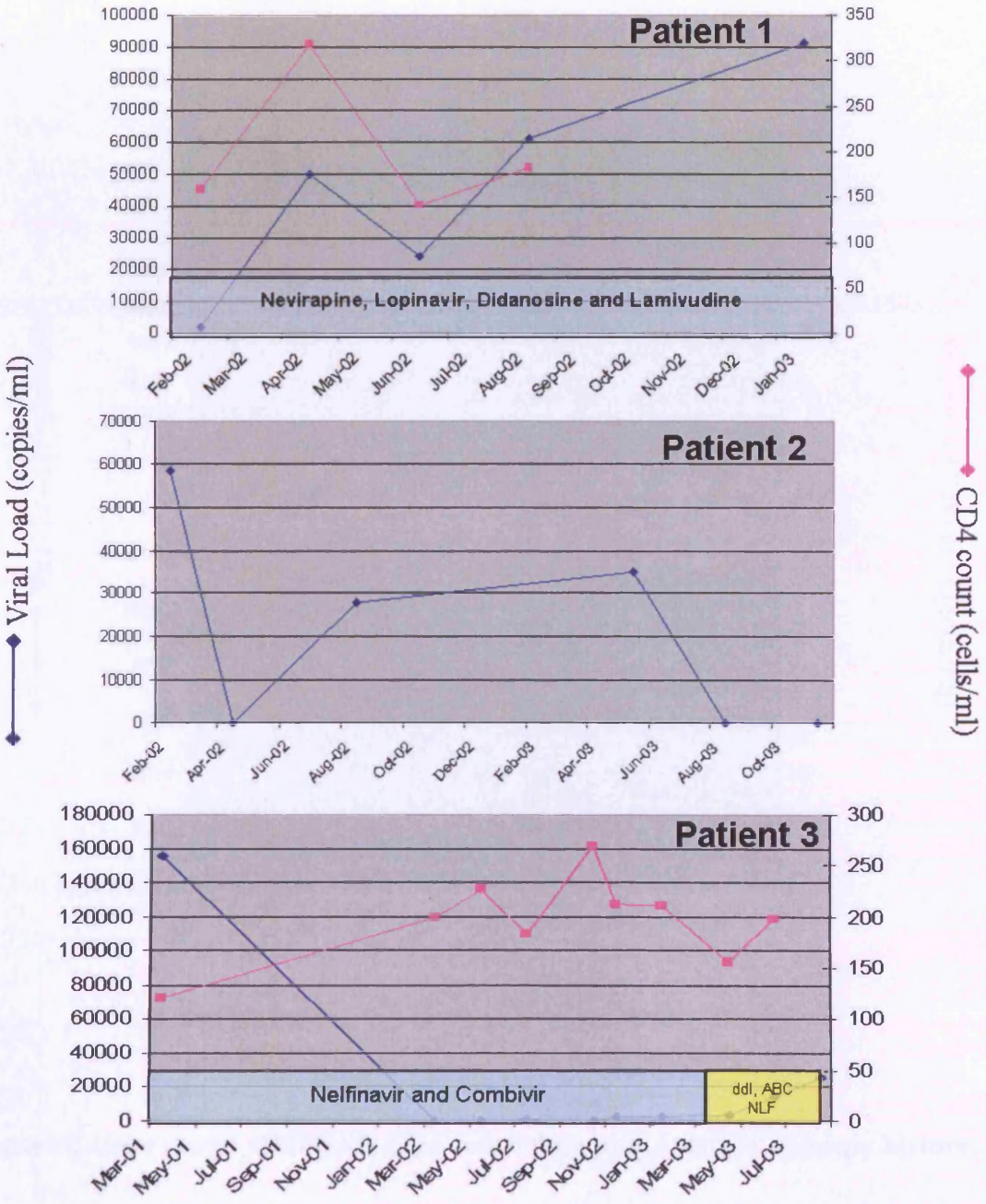


Figure 2.1(a) Patient viral load, CD4 count data and Antiviral therapy history.

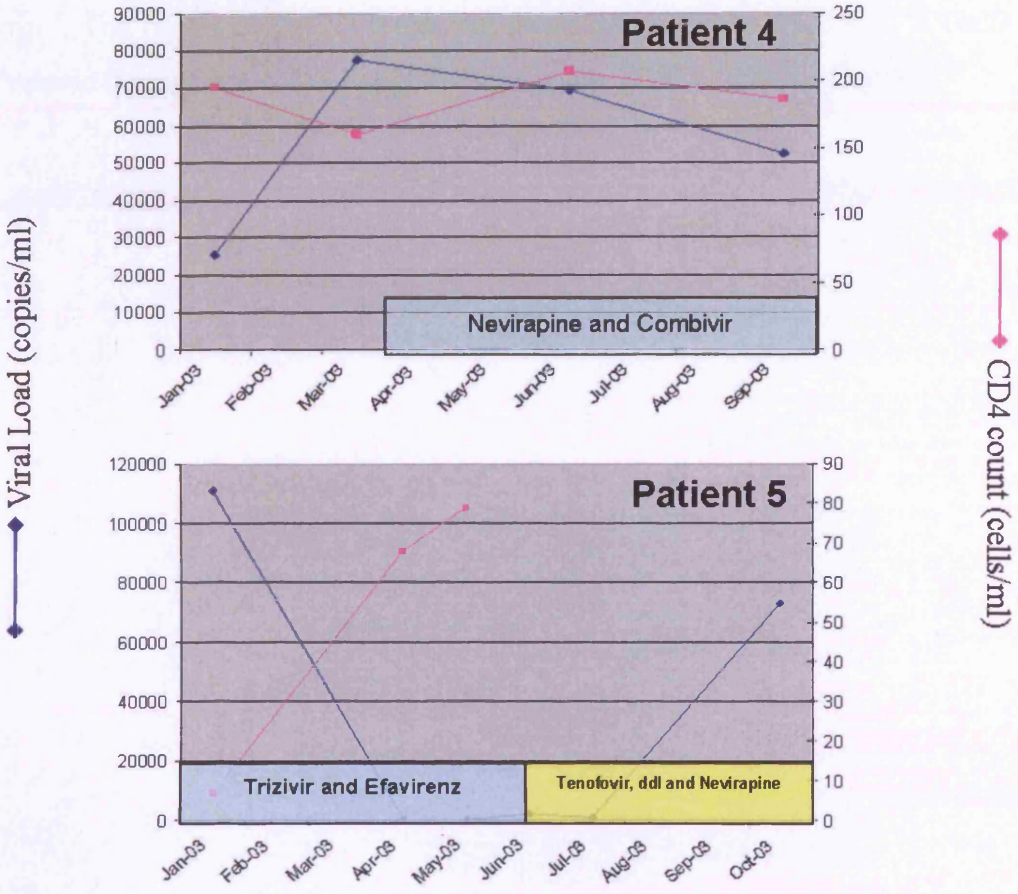


Figure 2.1(a) Patient viral load, CD4 count data and Antiviral therapy history.

2.1.15 Vectors

The molecular clone, pNL4.3 was obtained through the AIDS Research and Reference Reagent Program, Division of AIDS, NIAID, NIH: from Dr. Malcolm Martin [1]. The HIV-1 pHXB2D molecular clone was obtained from Drs. R Gallo and M Popovic through the AIDS Reagent Programme, NIBSC, Potters Bar, UK [91].

2.2 General Tissue Culture Techniques

All cell and virus cultures were grown in humidified 37°C incubators with 5% CO_2 (10% CO_2 for HEK 293T cells) in varying volumes and passaged as required.

2.2.1 Thawing cells

Cells were removed from liquid nitrogen and thawed rapidly at 37°C. MT-4 cells, a human T-cell line bearing human T-cell leukaemia virus type I (HTLV-I) and SupT1 cells were obtained through the AIDS Research and Reference Reagent Program, Division of AIDS, NIAID, NIH: from Dr. Douglas Richman and Dr. James Hoxie, respectively. Both cell lines were added to 10 ml of Roswell Park Memorial Institute (RPMI)-1640 medium (Invitrogen) with 10% foetal calf serum (FCS), (Biowest, France), 100 units (U)/ml penicillin and 100 μ g/ml streptomycin (Invitrogen) $RMPI_{10}$. The cells were then pelleted at 325g for 5 minutes, resuspended in 15 ml media and transferred to a $T25cm^3$ flask. The following day the media was exchanged for fresh media.

HeLa, TE 671, and HEK 293-T cells were added to 10 ml of Dulbecco's Modified Eagle Medium (DMEM, Invitrogen) with 10% FCS, (Biowest, France), 100 U/ml penicillin and 100 μ g/ml streptomycin (Invitrogen). The cells were then pelleted at 325g for 5 minutes, washed once in 10ml of media and then resuspended in 15 ml media in a 10cm dish. The following day the media was exchanged for fresh media.

The GHOST (3) R3/X4/R5 GFP Indicator cells used for HIV-1 infection were obtained through the NIH AIDS Research and Reference Reagent Program, Division of AIDS, NIAID, NIH: from Dr. Vineet N. KewalRamani and Dr. Dan R. Littman. These cells were thawed as described above although the media contained additional 2μg/ml puromycin (Sigma) and G-418 500 μg/ml (Gibco) selection.

2.2.2 Passaging cells

MT-4 and SupT1 cell lines were sub-cultured twice a week at a density of approximately 3 x 10^5 cells/ml in $RMPI_{10}$ and grown in 5% CO_2 at 37°C.

HEK 293-T, TE671 and HeLA were maintained in complete Dulbeccos minimal essential medium (DMEM) (Gibco) supplemented with 10% FCS (Gibco), 100 U/ml of penicillin, and 100 μg of streptomycin per ml.(Gibco). The GHOST (3) R3/X4/R5 GFP Indicator were maintained in complete DMEM (Gibco) supplemented with 10%

FCS (Gibco), 100 U/ml of penicillin, and 100 μg/ml of streptomycin (Gibco) 2μg/ml puromycin (Sigma) and G-418 500 μg/ml (Gibco). Cells were washed with phosphate buffered saline (PBS) (137 mM NaCl, 2.7 mM Kcl, 4.3 mM Na_2HPO_4, 1.4 mM KH_2PO_4), incubated with trypsin (0.5 g/l) EDTA (0.2g/l) (Invitrogen) until detached from the tissue culture plastic and then resuspended in media. Cells were split 1:4 to 1:8, depending on cell density and rate of growth, two or three times per week and grown in 5% CO_2 (HEK293T cells were grown in 10% CO_2) at 37°C.

2.2.3 Freezing cells

Cells were centrifuged at 325g for 5 minutes and resuspended at $1x10^7$ cells/ml in 40% complete media (as described above), 50% FCS and 10% dimethyl sulphoxide (DMSO, Sigma, UK). Cells were aliquoted into cryovials (Nunc, USA) and gradually cooled to –80°C in an isopropanol-containing cryo-container (Nalgene, USA) before being transferred to liquid nitrogen.

2.3 Pseudotype Lentiviral Vector related Techniques

2.3.1 Transient transfection of HEK 293-T cells to make lentiviral vectors.

Lentiviral vectors were produced as described by Besnier *et al.*, [17] 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) was added to 200 μl of Opti-MEM (Invitrogen). 1 μg of pHR-SIN-CSGW (a generous gift from Adrian Thrasher, Institute for Child Health, UCL, UK), 1 μg of pMDG (a generous gift from Didier Trono, Ecole Polytechnique Federale de Lausanne (EPFL), Switzerland) and 1.5 μg of the Gag-Pol expression vector p8.91 (*Not I*) (generous gift from Greg Towers UCL, UK), were made up to 15 μl volume in H_2O and added to the FuGENE-6 and Opti-MEM mixture. The transfection mixture was incubated at room temperature for 30 min before being added dropwise to the confluent HEK 293-T cells in 8 ml DMEM medium (Invitrogen) with 10% FCS (Biowest, France) 100 U/ml penicillin and 100 μg/ml streptomycin (Invitrogen).

The HEK 293-T cells and the transfection reaction were incubated overnight at 37°C and 10% CO_2 and the following day the media was exchanged for fresh DMEM

containing 10% FCS. The lentiviral vector containing supernatants were harvested at 48, 72 hours post transfection and passed through a 0.45 μ m filter and stored at -80°C in 1 ml aliquots.

2.3.2 Lentiviral vector titration of infectious units by GFP expression

HEK 293-T or TE671 cells were seeded at $1x10^5$ per well of a 24-well plate and cultured overnight at 37°C and 10% CO_2 in 500 μl of HEK 293-T medium (DMEM containing 10% FCS, 100 U/ml penicillin and 100 μg/ml streptomycin). 100 μl, 80 μl, 60μl, 40 μl, 20 μl and 1 μl of lentiviral vector containing supernatant were mixed with HEK 293-T medium to a total volume of 100 μl. The diluted lentiviral vector containing supernatants were added to HEK 293-T or TE671 cells in the 24 well plates and incubated overnight at 37°C and 10% CO_2. The cells were washed once with HEK 293-T medium and incubated for a further 36 hours.

After incubation the HEK 293-T or TE671cells were trypsinised and the percentage of GFP expressing cells determined by flow cytometry relative to non-infected cells. The percentage of GFP expressing cells was converted to infectious units by assuming a total of $2x10^5$ cells at the time of infection. Infectious units were plotted against μl of input supernatant. The absolute titre of the lentiviral vector was determined by extrapolating the infectious units per ml (IU/ml) using the straight-line region of the above plot.

2.3.3 Flow cytometry

Ten thousand events were collected using a FACScan flow-cytometer with Cellquest software (Becton Dickinson, UK). Data was analysed using Windows Multiple Document Interface Flow Cytometry Application (WinMDI, J. Trotter, http://facs.scripps.edu).

2.4 Live Virus Propagation and Harvesting

The Saquinavir resistant molecular clone virus derived from HXB2, containing G48V and L90M mutations in protease, was obtained from the Programme EVA, and the Centralised Facility for AIDS Reagents (NIBSC). Sources, J Rose, A Kohli and Dr. C Craig. The cell-free supernatant was used to infect $5x10^6$ MT-4 cells, which were incubated at 37 °C and then resuspended in a total volume of 10ml $RMPI_{10}$. The cultures were maintained by addition of more $RMPI_{10}$ if the colour of the media

appeared too yellow and incubated further until initial signs of cytopathic effect (CPE) were observed. At this point the infection was supplemented with a further $4x10^7$ MT-4 cells and the culture maintained as described above, until full CPE was observed.

The infections were harvested by pelleting the culture (1500 rpm at 4°C); stocks of virus were aliquoted and stored as cell free supernatants at -80 °C. The infected cell pellet was washed twice in sterile PBS, then resuspended in two 400µl aliquots and stored at -20°C.

2.4.1 Extraction of infected cell genomic DNA

A 200µl aliquot of the infected cell pellet DNA resuspended in PBS, was extracted using the QIAamp® DNA Mini Kit (Qiagen), according to manufacturers instructions.

2.4.2 Transfections of Infectious Recombinant Clones

2.4.3 MT-4 and SupT1 Transfections

The day before transfection MT-4 cells (or SupT1) were split 1:2 to ensure an actively growing population. Transfection of HIV molecular clone DNA was carried out with $2x10^6$ MT-4 cells and 8µl DMRIE-C ((1,2-dimyristyloxypropyl-3-dimethyl-hydroxy ethyl ammonium bromide) and cholesterol) (Gibco) transfection reagent and 2µg vector DNA. Cells were incubated in 6-well plates and media changed every 1-2 days, and observed for any signs of CPE in the culture. When CPE was observed the infected cells and supernatant were transferred to a $75cm^3$ tissue culture flask and supplemented with $2x10^7$ cells in order to create large stocks of each clone. When full CPE was observed in these cultures the cell free supernatant was harvested and stored at –80°C. This stock was used to create and sub-stock by infecting $5x10^6$ MT-4 cells with 500µl of cell free supernatant in a $25cm^3$ flask. When most of the cells in this culture displayed signs of CPE, a further $4x10^7$cells were added and the culture transferred to a $75cm^3$ flask. The cell free supernatant and infected pellet were harvested and treated as described in section 2.4.

2.4.4 293T Transfections

HEK 293-T cells were seeded so that 10cm dishes were confluent on the day of transfection. 18 µl of FuGENE-6 (Roche) was added to 200 µl of Opti-MEM (Invitrogen) containing 2µg of HIV molecular clone DNA and the transfections were continued as outlined in section 2.3.1.

2.5 Titration of Live Recombinant Virus Stocks

2.5.1 MTT Titration Assay

Virus titrations were carried out in MT-4 cells using 3-(4,5-dimethylthiazol)-2,5-diphenyltetrazolium bromide (MTT) to determine cell viability. Frozen cell free virus supernatant stocks were thawed and serially diluted (0.5 $\log_{10}$) in clear $RPMI_{10}$ (without phenol indicator) (Gibco) and plated out onto a low evaporation 96 well plate (Corning Costar), in a total of 50µl. MT-4 cells, 2 x 10^6 per assay were resuspended at 7.5 x 10^5/ml and then plated out 50µl per well.

After 5 days incubation 20 µl of MTT (Sigma-Aldrich) (5 mg/ml in PBS) was added and the plates were incubated for 2 hours before the addition of 170 µl of acidified isopropanol (86% Isopropanol, 11% dH_2O, 2.7% NP-40 and 0.3% concentrated HCl). The absorbance was measured at 540nm and the $TCID_{50}$ (Tissue culture infectivity dose) determined by Spearman-Karber method [90]. Each assay contains four replicates and the $TCID_{50}$/ml is determined from the average of all four data sets.

2.5.2 GHOST Cell Titration Assay

A day prior to titration experiments, GHOST cells (5 × 10^4 cells/well) cells were seeded in 24-well plates to obtain a sub-confluent cell layer by the time of infection. Infections were set-up with cell free supernatants and virus was applied in the presence of 20 µg/ml polybrene to enhance infection efficiency. Serial two fold dilutions from 100µl of virus supernatant in 500µl total volume of non selective DMEM (minus Puromycin and G-418) were plated out and incubated overnight. The following day, the media in each well was replaced with fresh media. The plates were incubated for a further 24 hours, after which the cells were trypsinised and fixed with formaldehyde (1% final concentration) the percentage of GFP expressing cells determined by flow cytometry relative to non-infected cells and the IU/ml was determined as detailed in section 2.3.2.

2.5.3 HIV-1 Antiviral assay

The sensitivity of the recombinant viruses to a panel of marketed inhibitors, was determined by MTT cell viability assay with MT-4 cells [250]. Saquinavir, Indinavir, Ritonavir and Nelfinavir, were all obtained through the AIDS Research and Reference Reagent Program, Division of AIDS, NIAID, NIH. Amprenavir and AZT were provided by GlaxoSmithkline, Stevenage, UK.

MT-4 cells, 2 x 10^6 per assay, were incubated with virus (MOI 0.0001) for 1 hour at 37 °C and then resuspended at 7.5 x 10^5/ml in clear $RPMI_{10}$ (Gibco) and plated out onto a low evaporation 96 well plate (Corning Costar) containing serial two fold dilutions of inhibitor.

After 5 days incubation the plates were processed as mentioned above and the $TCID_{50}$ determined by Spearman-Karber method [90]. Each assay contains four replicates and the $TCID_{50}$/ml is determined from the average of all four data sets.

2.6 Growth Rate Comparison

Each infection was set-up to contain the amount of virus that would provide 10% infection of GHOST cells as determined by duplicate titrations on GHOST cells as described in section 2.5.2. A stock infection containing virus and 5.25 x10^6 SupT1 cells in a total volume of 8ml $RPMI_{10}$ was set-up for each of the viruses. To minimise any variation in each of the time points, the stock infection was aliquoted across 12 wells of a 24 well plate, each well containing in 500µl of media. The infections were incubated at 37°C, except for the zero time point which was harvested immediately, and stored at -80°C. In all 13 time points were taken at 0, 6, 12, 24, 30, 36, 49, 57, 73, 81, 103, 127 and 150 hours post infection and all time point were stored prior to parallel analysis.

Each time point was titrated on GHOST cells on three separate occasions and the number of infected cells per ml, as detected by FACS analysis of GFP expressing cells, was calculated for each time point.

Chapter 3 Development of a Single Cycle HIV-1 Assay to Study Drug resistance

3.1 Introduction

The prolonged use of antiretrovirals for treatment of HIV selects for viruses that contain mutations which confer reduced sensitivity to the inhibitors they have been exposed to [20,122,306]. The development of resistance to most antiretrovirals occurs in a step-wise manner [21,49,194,213] and the highly error prone mechanisms of HIV replication enhance this selection process. The detection of resistance is primarily by sequence evaluation of the isolated virus and the identification of known resistance mutations, which leads to a genotypic estimation of resistance [138]. These methods are quick and provide an insight into any potential appearance of resistance and may avert a failing therapy.

Phenotypic methods require the amplification of viral sequences from patient plasma and the introduction of sequences into the molecular clone, followed by culturing of the resulting recombinant virus. The minimum time to produce chimeric patient derived viruses and determine the antiviral sensitivity can be as little a 15 days in some assays [125], although longer culturing times are normally required and generally extend this period to 4 to 6 weeks [257].

The need of these assay systems for virus isolation adds labour, time and cost and so has limited their usefulness in screening large sample cohorts. More recently the development of new phenotypic assays which circumvent the requirement for virus isolation by generating replication defective vectors that contain patient virus sequences and a reporter gene, such as luciferase have been described. Amplified protease and RT gene segments are inserted into the reporter gene viral vector, in which the luciferase gene was inserted into a deleted region of *env*. Measurement of virus replication is carried out by luciferase detection of pseudotyped virus particles after co-transfection of cells with an MLV *env* vector and a single round of infection in fresh target cells. Virus particles are only able to complete a single round of replication, due to the lack of an envelope protein and the entire assay including the assessment of antiviral activities is completed in 8 to 10 days [257]. The assay described by Petropoulos *et al.*, utilises a PCR amplicon that spans the p7-p1-p6 protease cleavage sites in *gag* and also whole of protease protein and the RT coding region from amino acids 1 to 313.

Many publications have highlighted the importance of using the whole of *gag* in evaluating its contribution to resistance compensation and restoring the growth ability of viruses resistant to PIs [107,191,224,358]. An attempt has been made to study replication capacity, a component of fitness, by constructing recombinant viruses containing *gag* and protease sequences derived from virus infecting patients treated with antiretrovirals. A single cycle reporter assay was used to determine differences in replication rates of recombinants with mutant *gag* in a wild-type background. The reporter system utilised was the lentiviral vector system designed by Naldini *et al.* [225,226]. This three-plasmid expression system was used to generate HIV-derived vector particles that could mediate stable *in vivo* gene transfer into non-dividing cells. Segregation of the coding regions required to produce an infectious particle is carried out by a split-genome expression system, composed of a packaging vector, an *env* vector and a HIV Gag-Pol vector [78,361]. The system is similar to that described by Petropoulos *et al.*, in that only a single round of replication can occur after transfection of the viral genomic sequences.

Specific Aims and Questions

At present no system exists to explore the pleiotropic effects of PIs, this would require a vector system that encompasses the whole *gag* region and would also allow for insertion of the cognate protease. The research work detailed in this chapter aims to create such a system with primary focus on insertion of *gag*, as the effects of this region from virus isolated from PI treated patients has not been thoroughly explored.

The overall aim of this work was to create patient derived HIV viral vectors which could be used to determine the effects on replication capacity of the inserted patient virus derived *gag* regions.

I. Identify patient derived virus isolates which may contain CSM or *gag* compensatory changes from a sample cohort
II. Amplify *gag* sequences and clone into retroviral vector
III. Determine if the *gag* region affects virus growth ability
IV. Is the *gag* from PI resistant viruses more or less efficient at replication than those from naïve sources?
V. And to see if the system can be used for drug sensitivity assays

3.2 Results

3.2.1 Patient Derived Virus Sample Sequencing:

Anonymised patient plasma were identified from a database search of patients receiving anti-retroviral therapy which were submitted for routine resistance testing at UCLH, London, UK. The criteria for selection was extensive prior antiretroviral therapy, high viral load and the presence of multiple PI mutations, including primary, those which directly effect inhibitor sensitivity, and secondary which are those mutations that alone do not affect drug sensitivity, but, may increase the mutant enzymes ability to function. As a control sample a patient derived virus was also selected that had no prior therapy or no primary resistance mutations. Sequencing of these protease and RT regions revealed the presence of multiple drug resistance mutations. Only sample 2, the control sample, had no resistance mutations, but contained a variety of other changes compared to the HXB2 consensus, however these are not thought to be associated with resistance to antiretrovirals [40]. (See Table 3.1)

Stored patient viral RNA was reverse transcribed as outlined in the Materials and Methods section 2.1.15 using the primer F122271. A PCR product encompassing the entire *gag* region and containing the first 60 amino acids of protease was amplified with primers HG/F/*NotI* and HG/R/*BclI* containing *NotI* and *BclI* restriction sites using the Expand Long Template PCR System (Roche).

Only four out of the five samples generated a 1676bp product containing the restriction sites necessary for cloning into the Gag-Pol expression vector.
Sequencing was carried out on the patient virus derived 1.6kb fragments that were cloned into the vector p8.91(*NotI*). The following primers were used to verify the sequence of the clones, the numbers in parenthesis denotes the position of the primers in vector pNL4.3. The sequences to the oligonucleotides used are given in Materials and Methods Table 2.1. AQ GAG5 (789), Gag1 (790), GAG SE2 (1132), GAG SE3 (1498), GAG SE4 (2027) and GAG AS1 (2316).
The amino acid translations of the *gag* sequences derived from the 1.6kb fragments are shown in Figure 3.1 aligned with reference sequences, HXB2, pNL4.3 and p8.91(*NotI*). Sequencing revealed that 5 clones derived from patient isolated virus samples 2, 4 and 5

had in frame stop codons in one or more of the chromatograms. These were not studied further. One clone isolated from virus obtained from patient 5 had a premature stop in the protease gene and 2 others from the same patient virus isolate also had stops in *gag*, as did 2 clones from patient 4. All of the sequenced *gag*s showed sequence differences from lab reference strain HXB2. Two of the sequences from virus derived from patient 2 had the most changes. The only CSM detected was from virus obtained from patient 5 at amino acid position 449. However, this was a Leucine to Proline change rather than the usual Phenylalanine change which is associated with drug treatment [107]. The *gag* isolated from this patients virus is also slightly extended, compared to the wild type, with an insertion after the PTAPP motif.

Viral clones isolated from patient 2 have a complete duplication of the PTAPP, preceded by a Leucine and Glutamic acid insertion. These two virus clones from patient 2 are overall four amino acids shorter than wild-type *gag*, with a four amino acid deletion prior to the p17/p24 cleavage site, a two amino acid deletion prior to the p2/p7 cleavage site and two separate amino acid deletions in the p6 protein (Figure 1).

Figure 2 shows an alignment of the protease genes from the patient virus derived clones. The sequencing of the protease gene revealed a more homogenous population than that found in *gag*. The only significant mutation observed generated a Leucine to Isoleucine change at amino acid position 10 in Pt5-c17, which is a secondary resistance mutation. The *BclI* site used for cloning is located at amino acid position 60 in the protease gene; therefore, the recombinants contained the C-terminal 39 amino acids of the wild-type protease.

Patiert	VL (copies /ml)	Sensitivity to Drug Class			Mutations from WT	
		PI	RTI	NNRTI	Protease	RT
1	9'30C	R	R	R	**L10I**, I113V, L24I, R41K, **M46L**, **I54V**, K55R, G57K/R, 62V, I63Q, I64V, **A71V**, **V82T/A**	**M41L**, K43N, E44D, V60 , **A98G**, C102<, V118 , D123E, I135_, K173Q, **M184V**, **R211K**, **T215Y**, **K219R**, K223E, **F227L**, **L228R**, A272P, R277K, V293I, E297R.
2	34900	S	S	S	K20M, L63P	V35T, E36A, T39E, S48T, K82R, Q102K, K104R, K'22E, D123G, C'62S, K173A, D177E, T200A E249Q, Q207D, V245Q, E248D, A272P, Q278H, T286A, V292I, E297K, V317A
3	2510C	R	S	S	V3I, **L10I/V**, L19Q K20M, S37N, **M46V**, I62V, L63P, I64V/L, **L90M** I93_	A98S, K103R, E122K, Q174K, **M184V**, Q207E, R211K, L214F, T215S, V245E, P272A, R277K, A288S, I293V, K311R, D324E, I329V
4	5260C	S	R	R	V3I, I13V, S37N R41K, D60E, I62V, L63P, H69Y	V35T/I, V60I, **S68G**, R83<, **F116Y**, D123E, **Q151M**, I178_, V179D, **Y188L**, T200A, Q207D, R211K, L214F, V245Q, D250E, R277K, T286A, I293V, E297K, D320E, S322A, I329V, **G333D**
5	7310C	S	R	R	V31, **L10V**, I12V, I13V, S37N, I63P, I64V	**D67N**, **K70R**, V90I, A98S, K102N, E122K, S162Y, Q174H, **Y181I**, **M184I**, T200A, E203K, Q207E, **R211K**, L214F, L283I, 293V, E297K, I329V

Table 3.1: Viral Load and mutations in Protease and RT coding regions of viruses isolated from patient samples. See Materials and Methods section 2.1.16. Mutations in bold are those known to be associated with resistance to antiretrovirals 40 . Sensitivity to drug class (R=Resistant) (S=Sensitive) was determined using the Stanford Genotypic Resistance Interpretation Algorithm, version 4.2.6.

Figure 3.1: Aligned gag amino acid sequences of Pseudotyped Clones (arrow indicates start codon ►◄ and bold letters indicate cleavage site and shaded amino acids represent changes from p8.91_*Not1*).

Gag Start
▼

```
                 ....|....|  ....|....|  ....|....|  ....|....|  ....|....|  ....|....|
                      5          15         25         35         45         55
Pt2-1            MGARASVLSG GELDRWEKIR LRPGGKKKYK LKHIVWASRE LERFAVNPGL LETSEGCRQI
Pt2-2            MGARASVLSG GELDRWEKIR LRPGGKKKYK LKHIVWASRE LERFAVNPGL LETSEGCRQI
Pt2-10           MGARASVLRG GKLDAWEKIR LRPGGKKTYM LKHIVWASRE LEKFALNPGL LETSEGCKQI
Pt2-12           MGARASVLKG GKLDTWEKIR LRPGGKKTYM LKHIVWASGE LERFALNPGL LETAEGCRQI
Pt3-9            MGARASVLSG GELDRWEKIR LRPGGKKKYK LKHIVWASRE LERFAVNPGL LETSEGCRQI
Pt3-10           MGARASVLSG GELDRWEKIR LRPGGKKKYK LKHIVWASRE LERFAVNPGL LETSEGCRQI
Pt4-13           MGARASVLSG GELDRWEKIR LRPGGKKKYK LKHIVWASRE LERFAVNPGL LETSEGCRQI
Pt5-c17          MGARASVLSG GELDKWEKIR LRPGGKKKYK LKHIVWASRE LERFAVNPGL LETTEGCRQI
WT-C13           MGARASVLSG GELDKWEKIR LRPGGKKQYK LKHIVWASRE LERFAVNPGL LETSEGCRQI
NL4.3            MGARASVLSG GELDKWEKIR LRPGGKKQYK LKHIVWASRE LERFAVNPGL LETSEGCRQI
HXB2             MGARASVLSG GELDRWEKIR LRPGGKKKYK LKHIVWASRE LERFAVNPGL LETSEGCRQI
p8.91_Not1       MGARASVLSG GELDRWEKIR LRPGGKKKYK LKHIVWASRE LERFAVNPGL LETSEGCRQI
Clustal          ******** * *:** ***** ******* *  ******** * **:**:**** ***:***:**

                 ....|....|  ....|....|  ....|....|  ....|....|  ....|....|  ....|....|
                      65         75         85         95        105        115
Pt2-1            LGQLQPSLQT GSEELRSLYN TVATLYCVHQ RIEIKDTKEA LDKIEEEQNK SKKKAQQAAA
Pt2-2            LGQLQPSLQT GSEELRSLYN TVATLYCVHQ RIEIKDTKEA LDKIEEEQNK SKKKAQQAAA
Pt2-10           IKQLHPALQT GTEELRSLFN TVATLYCVHE GIEVRDTKEA LDKIEEEQNK CQQKTQQAAD
Pt2-12           IKQLQPALQT GTEELRSLFN TVATLYCVHA GIEVRDTKEA LDKIEEEQNK CQQKIQQAAD
Pt3-9            LGQLQPSLQT GSEELRSLYN TVATLYCVHQ RIEIKDTKEA LDKIEEEQNK SKKKAQQAAA
Pt3-10           LGQLQPSLQT GSEELRSLYN TVATLYCVHQ RIEIKDTKEA LDKIEEEQNK SKKKAQQAAA
Pt4-13           LGQLQPSLQT GSEELRSLYN TVATLYCVHQ RIEIKDTKEA LDKIEEEQNK SKKKAQQAAA
Pt5-c17          LAQLQPSLQT GSEELRSLYN TVATLYCVHQ KIEVKDTKEA LDKVEEEQNK SKKQAQQAAA
WT-C13           LGQLQPSLQT GSEELRSLYN TIAVLYCVHQ RIDVKDTKEA LDKIEEEQNK SKKKAQQAAA
NL4.3            LGQLQPSLQT GSEELRSLYN TIAVLYCVHQ RIDVKDTKEA LDKIEEEQNK SKKKAQQAAA
HXB2             LGQLQPSLQT GSEELRSLYN TVATLYCVHQ RIEIKDTKEA LDKIEEEQNK SKKKAQQAAA
p8.91_Not1       LGQLQPSLQT GSEELRSLYN TVATLYCVHQ RIEIKDTKEA LDKIEEEQNK SKKKAQQAAA
Clustal          : **:*:*** *:******:* *:*.*****   *:::***** ***:****** .::: ****

                      p17-MATRIX ►◄ p24 CAPSID
                 ....|....|  ....|....|  ....|....|  ....|....|  ....|....|  ....|....|
                     125        135        145        155        165        175
Pt2-1            DTGHSNQ-VS QNYPIVQNIQ GQMVHQAISP RTLNAWVKVV EEKAFSPEVI PMFSALSEGA
Pt2-2            DTGHSNQ-VS QNYPIVQNIQ GQMVHQAISP RTLNAWVKVV EEKAFSPEVI PMFSALSEGA
Pt2-10           KGK-----VS QNYPIVQNLQ GQMVHQAISP RTLNAWVKVI EEKAFSPEVI PMFTALSEGA
Pt2-12           KGK-----VS QNYPIVQNLQ GQMVHQAISP RTLNAWVKVI EEKAFSPEVI PMFTALSEGA
Pt3-9            DTGHSNQ-VS QNYPIVQNIQ GQMVHQAISP RTLNAWVKVV EEKAFSPEVI PMFSALSEGA
Pt3-10           DTGHSNQ-VS QNYPIVQNIQ GQMVHQAISP RTLNAWVKVV EEKAFSPEVI PMFSALSEGA
Pt4-13           DTGHSNQ-VS QNYPIVQNIQ GQMVHQAISP RTLNAWVKVV EEKAFSPEVI PMFSALSEGA
Pt5-c17          NTGNNSPPVS QNYPIVQNLQ GQMVHQALSP RTLNAWVKVV EEKAFSPEVI PMFSALSEGA
WT-C13           DTGNNSQ-VS QNYPIVQNLQ GQMVHQAISP RTLNAWVKVV EEKAFSPEVI PMFSALSEGA
NL4.3            DTGNNSQ-VS QNYPIVQNLQ GQMVHQAISP RTLNAWVKVV EEKAFSPEVI PMFSALSEGA
HXB2             DTGHSNQ-VS QNYPIVQNIQ GQMVHQAISP RTLNAWVKVV EEKAFSPEVI PMFSALSEGA
p8.91_Not1       DTGHSNQ-VS QNYPIVQNIQ GQMVHQAISP RTLNAWVKVV EEKAFSPEVI PMFSALSEGA
Clustal          .       ** ********:* *******:** *********: ********** ***:******
```

Figure 3.1(continued): Aligned gag amino acid sequences of Pseudotyped Clones (arrow indicates start codon ►◄ and bold letters indicate cleavage site and shaded amino acids represent changes from p8.91_*Not1*).

```
                                                                    ►CYP A BINDING◄
                    ....|....|  ....|....|  ....|....|  ....|....|  ....|....|  ....|....|
                      185         195         205         215         225         235
Pt2-1               TPQDLNTMLN  TVGGHQAAMQ  MLKETINEEA  AEWDRVHPVH  AGPIAPGQMR  EPRGSDIAGT
Pt2-2               TPQDLNTMLN  TVGGHQAAMQ  MLKETINEEA  AEWDRVHPVH  AGPIAPGQMR  EPRGSDIAGT
Pt2-10              TPQDLNTMLN  TVGGHQAAMQ  MLKETINDEA  AEWDRLHPVH  AGPIAPSQLR  EPRGSDLAGT
Pt2-12              TPQDLNTMLN  TVGGHQAAMQ  MLKETINDEA  AEWDRLHPVH  AGPIAPGQLR  KPRGSDIAGT
Pt3-9               TPQDLNTMLN  TVGGHQAAMQ  MLKETINEEA  AEWDRVHPVH  AGPIAPGQMR  EPRGSDIAGT
Pt3-10              TPQDLNTMLN  TVGGHQAAMQ  MLKETINEEA  AEWDRVHPVH  AGPIAPGQMR  EPRGSDIAGT
Pt4-13              TPQDLNTMLN  TVGGHQAAMQ  MLKETINEEA  AEWDRVHPVH  AGPIAPGQMR  EPRGSDIAGT
Pt5-c17             TPQDLNTMLN  TVGGHQAAMQ  MLKETINEEA  AEWDRVHPVH  AGPIAPGQMR  EPRGSDIAGT
WT-C13              TPQDLNTMLN  TVGGHQAAMQ  MLKETINEEA  AEWDRLHPVH  AGPIAPGQMR  EPRGSDIAGT
NL4.3               TPQDLNTMLN  TVGGHQAAMQ  MLKETINEEA  AEWDRLHPVH  AGPIAPGQMR  EPRGSDIAGT
HXB2                TPQDLNTMLN  TVGGHQAAMQ  MLKETINEEA  AEWDRVHPVH  AGPIAPGQMR  EPRGSDIAGT
p8.91_Not1          TPQDLNTMLN  TVGGHQAAMQ  MLKETINEEA  AEWDRVHPVH  AGPIAPGQMR  EPRGSDIAGT
Clustal             **********  **********  *******:**  *****:****  ******.*:*  :******:***

                    ....|....|  ....|....|  ....|....|  ....|....|  ....|....|  ....|....|
                      245         255         265         275         285         295
Pt2-1               TSTLQEQIGW  MTHNPPIPVG  EIYKRWIILG  LNKIVRMYSP  TSILDIRQGP  KEPFRDYVDR
Pt2-2               TSTLQEQIGW  MTHNPPIPVG  EIYKRWIILG  LNKIVRMYSP  TSILDIRQGP  KEPFRDYVDR
Pt2-10              TSTLQEQIGW  MTIPHLSPVG  DIYKRWIILG  LNKIVRMYSP  VSILDIRQGP  KEPFRDYVDR
Pt2-12              TSTLQEQIAW  MTNNPPVPVG  DIYKRWIILG  LNKIVRMYSP  VSILDIRQGP  KNPFRDYVDR
Pt3-9               TSTLQEQIGW  MTHNPPIPVG  EIYKRWIILG  LNKIVRMYSP  TSILDIRQGP  KEPFRDYVDR
Pt3-10              TSTLQEQIGW  MTHNPPIPVG  EIYKRWIILG  LNKIVRMYSP  TSILDIRQGP  KGPFRDYVDR
Pt4-13              TSTLQEQIGW  MTHNPPIPVG  EIYKRWIILG  LNKIVRMYSP  TSILDIRQGP  KEPFRDYVDR
Pt5-c17             TSTLQEQIAW  MTNNPPIPVG  EIYKRWIILG  LNKIVRMYSP  TSILDIRQGP  KEPFRDYVDR
WT-C13              TSTLQEQIGW  MTHNPPIPVG  EIYKRWIILG  LNKIVRMYSP  TSILDIRQGP  KEPFRDYVDR
NL4.3               TSTLQEQIGW  MTHNPPIPVG  EIYKRWIILG  LNKIVRMYSP  TSILDIRQGP  KEPFRDYVDR
HXB2                TSTLQEQIGW  MTNNPPIPVG  EIYKRWIILG  LNKIVRMYSP  TSILDIRQGP  KEPFRDYVDR
p8.91_Not1          TSTLQEQIGW  MTHNPPIPVG  EIYKRWIILG  LNKIVRMYSP  TSILDIRQGP  KEPFRDYVDR
Clustal             ********.*  **     ***  :*********  **********  .*********  * ********

                    ....|....|  ....|....|  ....|....|  ....|....|  ....|....|  ....|....|
                      305         315         325         335         345         355
Pt2-1               FYKTLRAEQA  SQEVKNWMTE  TLLVQNANPD  CKTILKALGP  GATLEEMMTA  CQGVGGPGHK
Pt2-2               FYKTLRAEQA  SQEVKNWMTE  TLLVQNANPD  CKTILKALGP  GATLEEMMTA  CQGVGGPGHK
Pt2-10              FFKTLRAEQA  TQEVKNWMAD  TLLVQNANPD  CKTILKALGP  GATLEEMMTA  CQGVGGPGHK
Pt2-12              FFKTLRAEQA  TQEVKNWMTD  TLLVQNANPD  CKTILKALGP  GATLEEMMTA  CQGVGGPGHK
Pt3-9               FYKTLRAEQA  SQEVKNWMTE  TLLVQNANPD  CKTILKALGP  GATLEEMMTA  CQGVGGPGHK
Pt3-10              FYKTLRAEQA  SQEVKNWMTE  TLLVQNANPD  CKTILKALGP  GATLEEMMTA  CQGVGGPGHK
Pt4-13              FYKTLRAEQA  SQEVKNWMTE  TLLVQNANPD  CKTILKALGP  GATLEEMMTA  CQGVGGPGHK
Pt5-c17             FYKTLRAEQA  SQEVKNWMTE  TLLVQNANPD  CKTILKALGP  GATLEDMMTA  CQGVGGPGHK
WT-C13              FYKTLRAEQA  SQEVKNWMTE  TLLVQNANPD  CKTILKALGP  GATLEEMMTA  CQGVGGPGHK
NL4.3               FYKTLRAEQA  SQEVKNWMTE  TLLVQNANPD  CKTILKALGP  GATLEEMMTA  CQGVGGPGHK
HXB2                FYKTLRAEQA  SQEVKNWMTE  TLLVQNANPD  CKTILKALGP  AATLEEMMTA  CQGVGGPGHK
p8.91_Not1          FYKTLRAEQA  SQEVKNWMTE  TLLVQNANPD  CKTILKALGP  GATLEEMMTA  CQGVGGPGHK
Clustal             *:********  :*******::  **********  **********  .****:****  **********
```

Figure 3.1(continued): Aligned gag amino acid sequences of Pseudotyped Clones (arrow indicates start codon ►◄ and bold letters indicate cleavage site and shaded amino acids represent changes from p8.91_*Not1*).

```
p24 CAPSID ►◄        p2        ►◄ p7 NUCLEOCAPSID
           365       375       385       395       405       415
Pt2-1      ARVLAEAMSQ VTN---PATI MIQKGNFRNQ RKTVKCFNCG KEGHIAKNCR APRKKGCWKC
Pt2-2      ARVLAEAMSQ VTN---PATI MIQKGNFRNQ RKTVKCFNCG KEGHIAKNCR APRKKGCWKC
Pt2-10     ARVLAEAMSQ ANS-----NI MMQRSNFKGP KRTVKCFNCG KEGHIAKNCR ASRKKGCWKC
Pt2-12     ARVLAEAMSQ ANS-----NI MMQRSNFKGP KRIVKCFNCG KEGHIAKNCR APRKKGCWKC
Pt3-9      ARVLAEAMSQ VTN---PATI MIQKGNFRNQ RKTVKCFNCG KEGHIAKNCR APRKKGCWKC
Pt3-10     ARVLAEAMSQ VTN---PATI MIQKGNFRNQ RKTVKCFNCG KEGHIAKNCR APRKKGCWKC
Pt4-13     ARVLAEAMSQ VTN---PATI MIQKGNFRNQ RKTVKCFNCG KEGHIAKNCR APRKKGCWKC
Pt5-c17    ARVLAEAMSQ VTNSAASAAI MAQRGNFRGQ RKIIKCFNCG KEGHIAKNCR APRRKGCWKC
WT-C13     ARVLAEAMSQ VTN---PATI MIQKGNFRNQ RKTVKCFNCG KEGHIAKNCR APRKKGCWKC
NL4.3      ARVLAEAMSQ VTN---PATI MIQKGNFRNQ RKTVKCFNCG KEGHIAKNCR APRKKGCWKC
HXB2       ARVLAEAMSQ VTN---SATI MMQRGNFRNQ RKIVKCFNCG KEGHTARNCR APRKKGCWKC
p8.91_Not1 ARVLAEAMSQ VTN---PATI MIQKGNFRNQ RKTVKCFNCG KEGHIAKNCR APRKKGCWKC
Clustal    ********** ...      * * *:.**:.  :: :****** **** *:*** *.*:******

p7 NUCLEOCAPSID ►◄       p1        ►◄       p6
           425       435       445       455       465       475
Pt2-1      GKEGHQMKDC TERQANFLGK IWPSHKGRPG NFLQSRPEPT AP------PE ESFRIGEETT
Pt2-2      GKEGHQMKDC TERQANFLGK IWPSHKGWPG NFLQSRPQPT AP------PE ESFRFGEETT
Pt2-10     GKEGHQMKDC TERQANFLGK LWPSNKGRPG NFLQNRPEPT APLQPTAPPA ESFRF--EET
Pt2-12     GKEGHQMKDC TERQANFLGK LWPSNKGRPG NFLQNRPEPT APLQPTAPPA ESFRF--EET
Pt3-9      GKEGHQMKDC TERQANFLGK IWPSHKGRPG NFLQSRPEPT AP------PE ESFRFGEETT
Pt3-10     GKEGHQMKDC TERQANFLGK IWPSHKGRPG NFLQSRPEPT AP------PE ESFRFGEETT
Pt4-13     GKEGHQMKDC TERQANFLGK IWPSHKGRPG NFLQSRPEPT AP------PE ESFRFGEETT
Pt5-c17    GREGHRMKDC TERQANFLGK IWPSLKGRPG NFPQSRPEPT APPA---PPE ESFRFGEGTT
WT-C13     GKEGHQMKDC TERQANFLGK IWPSHKGRPG NFLQSRPEPT AP------PE ESFRFGEETT
NL4.3      GKEGHQMKDC TERQANFLGK IWPSHKGRPG NFLQSRPEPT AP------PE ESFRFGEETT
HXB2       GKEGHQMKDC TERQANFLGK IWPSYKGRPG NFLQSRPEPT AP------PE ESFRSGVETT
p8.91_Not1 GKEGHQMKDC TERQANFLGK IWPSHKGRPG NFLQSRPEPT AP------PE ESFRFGEETT
Clustal    *:***:**** ********** :*** ** ** ** *.**:** **      *  ****     *

                                          Gag STOP
                                             ▼
           485       495       505
Pt2-1      TPSQKQEPID KELYPLASLR SLFGSDPSSQ*
Pt2-2      TPSQKQEPID KELYPLASLR SLFGSDPSSQ*
Pt2-10     TPAPKQESKD RE--PLISLK SLFGNDPLSQ*
Pt2-12     TPAPKQEPKD RE--PLISLK SLFGNDPLSQ*
Pt3-9      TPSQKQEPID KELYPLASLR SLFGSDPSSQ*
Pt3-10     TPSQKQEPID KELYPLASLR SLFGSDPSSQ*
Pt4-13     TPSQKQEPID KELYPLASLR SLFGSDPSSQ*
Pt5-c17    TPSQKQEPID KELYPLASLR SLFGNDQSR-*
WT-C13     TPSQKQEPID KELYPLASLR SLFGSDPSSQ*
NL4.3      TPSQKQEPID KELYPLASLR SLFGSDPSSQ*
HXB2       TPPQKQEPID KELYPLTSLR SLFGNDPSSQ*
p8.91_Not1 TPSQKQEPID KELYPLASLR SLFGSDPSSQ*
Clustal    **. ***. * :*  ** **: ****.*
```

Figure 3.2: Aligned Protease amino acid sequences of Pseudotyped Clones (▼indicates start codon and shaded amino acids represent changes from p8.91_*Not1*).

```
                Protease Start                                                                  Bcl1 cloning site
                ▼...|....| ....|....| ....|....| ....|....| ....|....| ....|....▼
                     5          15         25         35         45         55
Pt2-1           PQITLWQRPL VSIKVGGQIK EALLDTGADD TVLEDINLPG KWKPKMIGGI GGFIKVRQYD
Pt2-2           PQITLWQRPL VTIKIGGQLK EALLDTGADD TVLEEMNLPG RWKPKMIGGI GGFIKVRQYD
Pt2-10          PQITLWQRPL VTIKVGGQIK EALLDTGADD TVLEDMNLPG RWKPKMIGGI GGFIKVRQYD
Pt2-12          PQITLWQRPL VSIKVGGQIK EALLDTGADD TVLEDMNLPG RWKPKMIGGI GGFIKVRQYD
Pt3-9           PQITLWQRPL VTIKIGGQLK EALLDTGADD TVLEEMNLPG RWKPKMIGGI GGFIKVRQYD
Pt3-10          PQITLWQRPL VTIKIGGQLK EALLDTGADD TVLEEMNLPG RWKPKMIGGI GGFIKVRQYD
Pt4-13          PQITLWQRPL VTIKIGGQLK EALLDTGADD TVLEEMNLPG RWKPKMIGGI GGFIKVRQYD
Pt5-C17         PQITLWQRPI VVVKIGGQLK EALLDTGADD TVLEEMNLPG RWKPKMIGGI GGFIKVRQYD
WT-C13          PQITLWQRPL VTIKIGGQLK EALLDTGADD TVLEEMNLPG RWKPKMIGGI GGFIKVRQYD
pNL4.3          PQITLWQRPL VTIKIGGQLK EALLDTGADD TVLEEMNLPG RWKPKMIGGI GGFIKVRQYD
HXB2            PQVTLWQRPL VTIKIGGQLK EALLDTGADD TVLEEMSLPG RWKPKMIGGI GGFIKVRQYD
p8.91_Not1      PQITLWQRPL VTIKIGGQLK EALLDTGADD TVLEEMNLPG RWKPKMIGGI GGFIKVRQYD
Clustal         **:******: * :*:***:* ********** ****::.**  ********** **********

                ....|....| ....|....| ....|....| ....|....
                     65         75         85         95
Pt2-1           QILIEICGHK AIGTVLVGPT PVNIIGRNLL TQIGCTLNF
Pt2-2           QILIEICGHK AIGTVLVGPT PVNIIGRNLL TQIGCTLNF
Pt2-10          QILIEICGHK AIGTVLVGPT PVNIIGRNLL TQIGCTLNF
Pt2-12          QILIEICGHK AIGTVLVGPT PVNIIGRNLL TQIGCTLNF
Pt3-9           QILIEICGHK AIGTVLVGPT PVNIIGRNLL TQIGCTLNF
Pt3-10          QILIEICGHK AIGTVLVGPT PVNIIGRNLL TQIGCTLNF
Pt4-13          QILIEICGHK AIGTVLVGPT PVNIIGRNLL TQIGCTLNF
Pt5-C17         QILIEICGHK AIGTVLVGPT PVNIIGRNLL TQIGCTLNF
WT-C13          QILIEICGHK AIGTVLVGPT PVNIIGRNLL TQIGCTLNF
pNL4.3          QILIEICGHK AIGTVLVGPT PVNIIGRNLL TQIGCTLNF
HXB2            QILIEICGHK AIGTVLVGPT PVNIIGRNLL TQIGCTLNF
p8.91_Not1      QILIEICGHK AIGTVLVGPT PVNIIGRNLL TQIGCTLNF
Clustal         ********** ********** ********** *********
```

Figure 3.2: Aligned Protease amino acid sequences of Pseudotyped Clones (▼indicates start codon and shaded amino acids represent changes from p8.91_*Not1*).

```
                Protease Start                                                                  BclI cloning site
                ▼...|....| ....|....| ....|....| ....|....| ....|....| ....|....▼
                     5          15         25         35         45         55
Pt2-1           PQITLWQRPL VSIKVGGQIK EALLDTGADD TVLEDINLPG KWKPKMIGGI GGFIKVRQYD
Pt2-2           PQITLWQRPL VTIKIGGQLK EALLDTGADD TVLEEMNLPG RWKPKMIGGI GGFIKVRQYD
Pt2-10          PQITLWQRPL VTIKVGGQIK EALLDTGADD TVLEDMNLPG RWKPKMIGGI GGFIKVRQYD
Pt2-12          PQITLWQRPL VSIKVGGQIK EALLDTGADD TVLEDMNLPG RWKPKMIGGI GGFIKVRQYD
Pt3-9           PQITLWQRPL VTIKIGGQLK EALLDTGADD TVLEEMNLPG RWKPKMIGGI GGFIKVRQYD
Pt3-10          PQITLWQRPL VTIKIGGQLK EALLDTGADD TVLEEMNLPG RWKPKMIGGI GGFIKVRQYD
Pt4-13          PQITLWQRPL VTIKIGGQLK EALLDTGADD TVLEEMNLPG RWKPKMIGGI GGFIKVRQYD
Pt5-C17         PQITLWQRPI VVVKIGGQLK EALLDTGADD TVLEEMNLPG RWKPKMIGGI GGFIKVRQYD
WT-C13          PQITLWQRPL VTIKIGGQLK EALLDTGADD TVLEEMNLPG RWKPKMIGGI GGFIKVRQYD
pNL4.3          PQITLWQRPL VTIKIGGQLK EALLDTGADD TVLEEMNLPG RWKPKMIGGI GGFIKVRQYD
HXB2            PQVTLWQRPL VTIKIGGQLK EALLDTGADD TVLEEMSLPG RWKPKMIGGI GGFIKVRQYD
p8.91_Not1      PQITLWQRPL VTIKIGGQLK EALLDTGADD TVLEEMNLPG RWKPKMIGGI GGFIKVRQYD
Clustal         **:******: * :*:***:* ********** ****::.**  ********** **********

                ....|....| ....|....| ....|....| ....|....
                     65         75         85         95
Pt2-1           QILIEICGHK AIGTVLVGPT PVNIIGRNLL TQIGCTLNF
Pt2-2           QILIEICGHK AIGTVLVGPT PVNIIGRNLL TQIGCTLNF
Pt2-10          QILIEICGHK AIGTVLVGPT PVNIIGRNLL TQIGCTLNF
Pt2-12          QILIEICGHK AIGTVLVGPT PVNIIGRNLL TQIGCTLNF
Pt3-9           QILIEICGHK AIGTVLVGPT PVNIIGRNLL TQIGCTLNF
Pt3-10          QILIEICGHK AIGTVLVGPT PVNIIGRNLL TQIGCTLNF
Pt4-13          QILIEICGHK AIGTVLVGPT PVNIIGRNLL TQIGCTLNF
Pt5-C17         QILIEICGHK AIGTVLVGPT PVNIIGRNLL TQIGCTLNF
WT-C13          QILIEICGHK AIGTVLVGPT PVNIIGRNLL TQIGCTLNF
pNL4.3          QILIEICGHK AIGTVLVGPT PVNIIGRNLL TQIGCTLNF
HXB2            QILIEICGHK AIGTVLVGPT PVNIIGRNLL TQIGCTLNF
p8.91_Not1      QILIEICGHK AIGTVLVGPT PVNIIGRNLL TQIGCTLNF
Clustal         ********** ********** ********** *********
```

3.2.2 Cloning of Patient Derived Virus Samples

The initial and preferred cloning strategy was to clone the patient virus derived PCR products into the Gag-pol expression vector p8.91 (*NotI*) (see Figure 3.3) after first digesting with the relevant restriction enzymes. This approach was attempted several times without success and so a second cloning strategy was adopted. This was to first clone the patient derived virus PCR products into a TA cloning vector then to transform plasmid DNA into methylase negative cells, which would allow the patient virus derived sequences to be digested out from the un-methylated plasmid DNA for ligation into the Gag–Pol expression vector. (Direct cloning of the TA ligation reaction into *E.coli*, M-cells was not possible due to the low transformation efficiency of these cells (Transformation efficiency of SCS110 & JM110 competent cells to generate unmethylated DNA is 5 x 10^6 transformants/μg pUC18 DNA, compared with 5 x 10^9 transformants/μg pUC18 DNA of XL10-Gold® Ultracompetent cells; from Stratagene product information sheet)). Patient viral DNA sequences were successfully cloned into the TA cloning vector. However, sub-cloning into the Gag-Pol expression vector failed. Several possible reasons were identified for the failure of the final cloning step: (a) general ligation conditions such as the stoichiometry of the vector to insert ratio; (b) incubation temperature; (c) the ligase used and alkaline phosphatase treatment of the vector; (d) over digestion of the PCR product could also account for the ligation failure. Many optimisation experiments were carried out to determine the reasons for the failure. Throughout the optimisation, the problem encountered was the failure to obtain colonies containing the correct sized insert and not the actual failure of the ligation to occur. The success of the ligations was confirmed by using PCR to amplify across the junction of the region ligated using an oligo specific only for the Gag-Pol expression vector, downstream of the *BclI* cloning site and another oligo located upstream of the *BclI* cloning site. Therefore, the amplicon generated could only be derived from a correctly ligated product. Bands were detected at the correct size in the ligation reactions, but the vector only control ligations remained negative. Colonies screened contained plasmids smaller than expected, suggesting some deletions had occurred during replication of the vector by *E.coli*. To overcome this, ligations were transformed into StBL2 cells (Invitrogen), which are more tolerant of DNA sequences containing repeat elements likely to cause recombination, such as those of HIV [92]. The products for ligation were tested for

complete digestion, by ligating the compatible ends to each other, thus forming heteromers if both ends were cut and dimers if only one end was digested. It was also proposed that UV exposure of vector DNA in the ligation caused nicks in the DNA which would reduce transformation efficiency. This was explored by making new vector DNA without exposing it to UV, thus, avoiding this possibility.

Eventually, correctly ligated clones were isolated using *E. coli* HB101 and the Roche Rapid ligation kit. HB101, like StBL2 is more tolerant of HIV sequences and a much higher transformation efficiency was achieved using this strain compared to StBL2. The ligation kit used a "5-minute quick" protocol, that had been tried previously (NEB and Fastlink) without success. It is not clearly understood why the Roche ligase would perform any better. Correct clones were originally derived by pre-cloning into the pGEM vector, but as previously mentioned this method diminished heterogeneity and direct cloning from the PCR product was eventually successful, albeit at a much lower frequency i.e. ratio of clones screened to clones with insert. A total of 14 clones were produced from patient virus derived samples, four virus clones from patient 2, two from patient 3 derived virus, three viruses derived from patient 4 and five virus clones from patient 5. However, five of these clones contained stops codons, resulting in truncated *gag*s or protease genes. A wild-type clone derived from pNL4.3 was also constructed.

3.2.3 Modification of Gag-Pol expression vector to include RT

The initial aim of this work was to examine the role of mutations in *gag* associated with drug resistance on viral replication capacity, using the Gag-Pol expression vector p8.91 (*NotI*). This was used to generate pseudotyped clones with the *gag* and part of the protease from patient derived virus samples, enabling the analysis of *gag* mutations. However, it was not understood how the substitution of 60 amino acids of protease would affect the growth characteristics of the pseudotyped viruses. As Gag is the substrate for protease, substituting a mutant *gag* and combining with a part mutant protease is less ideal than being able to insert the whole mutant protease gene. With this in mind, it was decided to modify the existing expression vector to allow for the insertion of a larger region from the patient's virus. Restriction analysis of the vector and HIV sequences identified a region that would allow the silent insertion of an *XmaI* restriction site downstream of the junction of RT and integrase, 44 amino acids into the integrase open reading frame (ORF). The aim was to make this vector using the cloning strategy outlined

in Materials and Methods, however, due to the difficulties of ligating after *Tth111* digestion, this approach was abandoned in favour of directly mutating the Gag-pol expression vector, p8.91 (*NotI*) . This approach yielded the p8.91-*Xma* vector containing the *XmaI* restriction site. To verify the presence of the restriction site, the plasmid was digested with *XmaI* and *EcoR1*, which created two fragments 11449bp and 702bp. However due to time limitations this vector construct was not used as it would require either complete sequence analysis to verify no spurious mutations were incorporated during PCR muatgensis. Or further sub-cloning to reduce the fragment size that would require sequence verification.

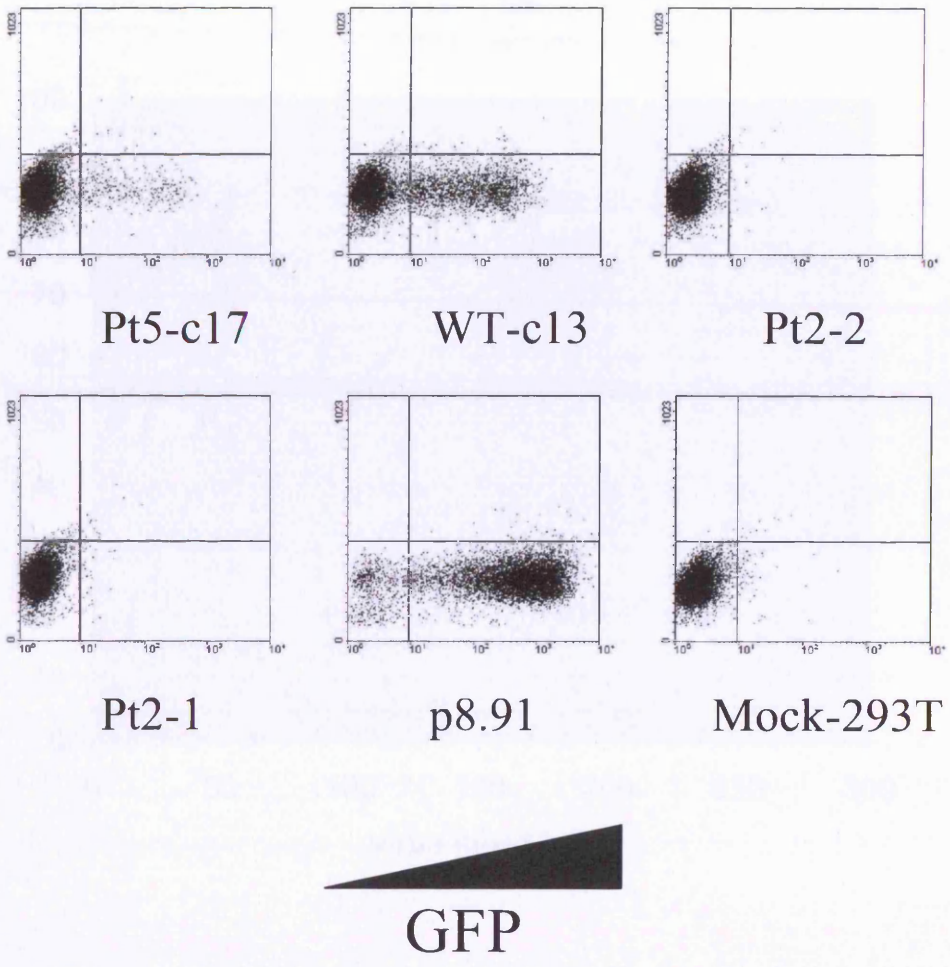
3.2.4 Generation of Pseudotyped viruses.

Co-transfection of the Gag-Pol expression vector p8.91 (*NotI*), along with the vector pMDG, which expresses the VSV-G envelope and the packaging vector pHR-SIN-CSGW, into the 293T cell line results in the production of a VSV-G pseudotyped virus particle (see Figure 3.3); capable of only a single round of replication. Infection of T cells with the harvested viral supernatant results in GFP expression, from which viral titres can be calculated. Initially, experiments were carried out to determine the levels of replication of clones carrying patient virus derived *gag* sequences. However, due to very low levels of GFP expression in the patient virus clones and lower than expected levels in the positive controls, some DNA stocks were remade and other reagents tested. Different cell lines, 293T, TE671 and HeLa were tested to see if they would give better results when infected with the pseudotyped retroviral vectors. With new 293T cells and new pMDG DNA, very high GFP levels were observed with the positive controls. However, patient virus clone 5c-17 was the only patient-derived viral clone in which any detectable levels of GFP were recorded. The only other clone constructed to give measurable values was the wild-type clone 13, which was amplified from pNL4.3 and cloned into the Gag-Pol expression vector. Data obtained from FACS analysis of pseudotypes is shown in Figure 3.4; Figure 3.5 shows the titration curves used to calculate the number of infectious units per ml. Production of recombinant pseudotypes from patient derived virus was eventually successful. However, due to the low titres obtained and the low level of the wild type pNL4.3 recombinant pseudotype, it was thought that the differences between fit and unfit viruses would be difficult to measure accurately in this single cycle system. It was also hoped to use this system to determine IC_{50}'s of the recombinant pseudotypes to

antiretrovirals, but due to the low titres, the dynamic range of the assay would be too small to measure inhibition by antivirals accurately.

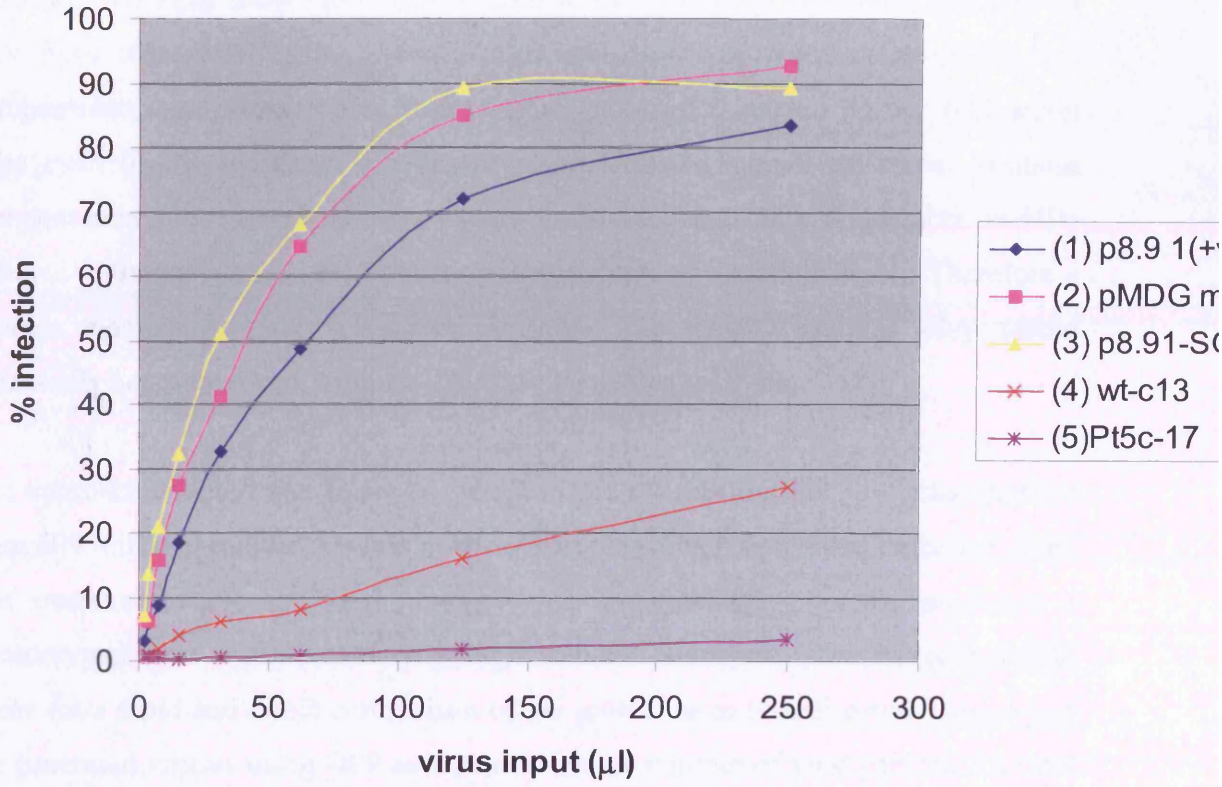
Figure 3.3: Schematic of the generation of lentiviral vector. Depicted are the relevant portions of the three plasmids cotransfected into 293T cells. The packaging construct, pCMVDR8.2 shown in the diagram was replaced by p8.91(Not1), which has additional HIV accessory genes removed, indicated by (X). The major 5' splice donor site (SD) has been conserved in both p8.91(Not1) and in pHR', which represents the transducing vector pHR-SIN-CSGW, this provides the vector genome. The viral LTRs are indicated, the *gag* gene is truncated after 350 bp and is out of frame, and it follows the RRE and a splice acceptor site (SA).
pMD.G encodes the heterologous hCMV promoter driven, VSV envelope that pseudotypes the vector. Figure adapted from Naldini *et al.* [225].

Figure 3.4: Level of GFP expression of Pseudotypes detected by using FACS. Table below lists the percentage cells expressing GFP in the right bottom quadrant.



	% Cells Expressing GFP
Pt5-c17	3.4
WT-C3	26.9
Pt2-2	0.3
p8.91	92.6
Pt2-1	0.4
Mock-293T	0.6

Figure 3.5: Titration curves of Pseudotypes viruses. Below are listed the calculated Infectious units per ml. (pMDGmnp and p8.91-SG were new reagent controls made and tested alongside positive control p8.91).



	Infectious Units/ml
(1) p8.9 1(+ve)	2355200
(2) pMDG mnp	3517440
(3) p8.91-SG	4124160
(4) wt-c13	215680
(5)Pt5c-17	27360

3.3 Discussion

The appearance of resistance mutations is one of the main causes of treatment failure in HIV infected individuals on antiretroviral therapy. The acquisition of these mutations has been shown to be detrimental to the growth of HIV. The genotype of HIV from individuals failing therapy often reveals the presence of resistance and compensatory mutations, which restore viral fitness and permit higher replication rates even in the presence of antiretrovirals. The mechanism of action of these compensatory mutations is not yet clearly understood and their appearance in HIV infected individuals can make predicting therapy outcome more difficult. Therefore a clearer understanding of the effects and the role these mutations play would potentially benefit patients by allowing better treatment regimens.

The approach adopted was to study the effects of substituting the *gag* gene, derived from HIV infected individuals into a wild-type background replication defective clone that upon co-transfection with an envelope and packaging vector produced a pseudotyped virus capable of only a single round of infection. This approach would allow for a rapid and direct comparison of the growth rates or replication capacity of the generated viruses using GFP as a marker for the number of viral particles. A total of eight recombinant virus clones were generated from the patient derived material and only one of these lead to expression of detectable levels of GFP. A wild-type recombinant virus was also constructed from the molecular clone pNL4.3, this also had very low GFP expression levels. Cloning into the Gag-pol expression vector p8.91 (*NotI*), proved to be problematic and several cloning strategies were employed. The initial strategy adopted was to clone the patient derived viral PCR products into the expression vector p8.91 (*NotI*) after first digesting with the appropriate restriction enzymes. This strategy had the advantages of generating a heterogeneous population i.e. direct from the population PCR, and also the methylase sensitive *BclI* enzyme was able to digest the PCR product. The problems encountered with cloning into the Gag-pol expression vector, p8.91 (*NotI*), could have been due to the instability of the HIV sequence, although fewer problems were encountered during the cloning and construction of the cassette vector pNL4.3-PSX; subject of the next chapter. The problems associated with the expression vector p8.91 (*NotI*), could be attributed to

the fact that it has the HIV sequence under the control of a CMV promoter, this may have caused expression of HIV sequences that caused the observed difficulties. This problem can arise when the gene product itself can interfere with metabolism in *E. coli*. Expression of HCV proteins by *E.coli* has been shown to occur by Forns *et al.* They demonstrated that spurious translation was initiated within the HCV sequence and that the sequence upstream of the initiating AUG codon of HCV might conceivably function as an initiation complex and allow translation of the HCV sequences in *E. coli*. They also showed that the introduction of a frameshift mutation upstream of the HCV translation–initiation codon dramatically increased the proportion of translationally competent clones isolated [92].

The reasons for the very low levels of infection detected by the recombinant virus pseudotypes, could be an effect of substituting in 'unfit' genes; however, this would not explain the disappointing reduction in titres seen by substituting the *gag* from another wild-type molecular clone. This reduction could be due to combining the *gag*s from two HIV molecular clones, HXB2 and NL4.3. The *gag* gene of the expression vector p8.91 is a combination of the first 239 amino acids from HXB2 and the remainder of the *gag* comes from pNL4.3, as does the whole Pol gene [225]. The wild-type clone constructed in this study has sequences solely derived from pNL4.3, so the main differences are the MA and nearly half of CA sequences. This amounts to 11 changes in the 239 N-terminal amino acids of *gag* in the constructed wild-type clone compared to the p8.91 (*NotI*) vector. Four of these changes are grouped close to the p17/p24 cleavage site, although their proximity to the scissile bond may undermine their relevance. It is thought these differences alone could not possibly account for the substantial loss seen in infectivity, as the newly formed *gag* of the wild-type clone is less chimeric. It should also be noted that during previous construction of the vector within this triple transfection system many of the sequences encoding critical virulence factors, such as Vif , Vpr, Vpu and Nef were removed from this expression system, in order to optimise the safe *in vivo* delivery of genes [225]. Therefore, substituting the whole Gag-Pol region with that derived from another source is likely to adversely affect virus production. What is also unclear is the effect of having only half of the protease from the donor inserted into the newly formed clone. It may be that the less efficient processing of Gag by the chimeric protease contributed to the titre reductions observed in the patient virus clones compared to the wild-type clones.

Cloning difficulties with the vector p8.91(*NotI*) limited the manipulation of this vector as did the lack of suitable, unique cloning sites. Therefore, this required either extensive modification of this clone, or a new HIV Gag-Pol vector. Modifying the molecular clone, pNL4.3 to insert gag, protease and RT separately would create a live virus system that could be made into a complementary single cycle system by deletion or truncation of *env*. The use of the live virus was in part due to the availability of Class III facilities and as adaptation of a virus to a new environment is a key component of fitness, it requires multiple replication cycles and so infectious progeny must be produced.

The sequencing of the cloned patient-derived virus *gag*s revealed the presence of cleavage and non-cleavage site mutations. Viral clones isolated from patient 5 and two viral clones from patient 2 demonstrated the most variation and both had duplications of the PTAPP motif, which is thought to be necessary for efficient budding from the cell surface [241]. Patient 2 was PI naïve and patient 5 PI experienced. Insertion mutations in *gag*, such as those described in this report have previously been detected at a moderately low frequency in therapy naïve subjects. However, all patients who had mixed PTAPP sequences, comprising of wild type and insertion mutants reverted to complete PTAPP insertion during therapy, suggesting a selective advantage for this genotype in the presence of antiretrovirals [241]. Lastere *et al.* [171] also reported that the presence of a duplicated PTAPP motif correlated with a poorer virological response to a PI-containing salvage regimen. The recombinant virus clone from patient 5 was the only patient derived isolate to produce any detectable levels of virus. It also was the only sample with known PI resistance associated mutations (Table 3.1). Therefore, it may have had the most time to evolve in a drug resistant selecting environment. This clone also has a single proline insertion near the p17/p24 cleavage site, which is also preceded by a Glutamine to Proline change. It is thought that the introduction of prolines can introduce flexibility or a turn in the polyprotein perhaps aiding cleavage by the less efficient drug resistant protease [343]. This clone also has three amino acid duplications near the p2/p7 cleavage site and a single deletion truncating the C-terminal amino acid of p6. With such a highly evolved or divergent genotype it could be expected that this *gag* and its associated protease would have a level of complementarity as this has been shown with the V82A and I50V mutations in protease [171]. The mutations in *gag* alone did not confer any level

of resistance. However, when present with substitutions in the protease, they not only had an impact on the *in vitro* replication capacity of the viruses but also increased their resistance to Amprenavir [188].

Patient 2 derived virus clones, 2-10 and 2-12 are very different from the other two clones isolated from the same patient derived viral RNA. From Table 3.1 it can be seen that this patient's viral genotype was PI sensitive. It is possible that these mutations could have evolved without any drug selection or that, the virus at seroconversion was PI resistant and the 'unfit' resistance mutations in protease were lost. This patients virus does have an M36I mutation in the protease, which has been observed in PI treated individuals, but this mutation is also present in wild-type viruses in most non-B clades [40,176]. It is also possible that the errors seen are PCR induced errors or that this is sample contamination, although no other samples amplified resemble this genotype.

The adaptation of this infection system to produce a safe means of delivering genes into non-dividing target cells required extensive modification. Even though the system has been shown to operate successfully without the accessory genes, the data presented by Zufferey *et al.*, shows an additive effect i.e. an increase in the number of transduced cells, when Vpr was included in the expression construct [361]. So it is possible that Vpr has a function in this system albeit at a reduced level. HIV infection and virus production is a complex mechanism that involves an intricate relationship with virus and host. Therefore it is possible that other interactions may play a role in the development of high titre pseudotyped particles, such as, NC interacting with integrase and stimulating DNA integration *in vitro* [34].

The production of these pseudotypes was essentially to explore the effects of the variation in patients virus *gag*s on viral replication. It was also hoped this system could be used to determine the phenotypic sensitivity to antiretrovirals of the pseudotyped viruses. These viruses could also have been used to look at the effects of drug resistant patient derived virus *gag*s on cyclophilin binding, as two of the virus clones both from Patient 2 had changes in the cyclophilin binding loop. However, one of these mutations has been reported to appear through *in vitro* passage of HIV-NL4.3 in the absence of any PIs [107]. Patient 5 was the only patient derived virus to produce

any detectable infectious units, although it contained only a secondary mutation as detected in the resistance test sequence data, see Table 3.1 and in the clone used to create the psuedotyped virus, figure 3.2. Patient 5 derived virus was also used in a later chapter of this thesis to construct a live virus, this also had the same PI mutation, see figure 5.1a. Further evaluation of the gag derived from this virus may be worthy as the patient derived virus has had prior exposure to NRTIs which can induce duplications, as is seen with the insertion near the PTAP motif see figure 3.1. It would have also been interesting to see if the PTAPP duplicated clones have a better replication capacity in the presence of NRTIs [256].

The approach chosen made use of a transfection system that segregated the coding regions required to produce an infectious particle over three separate plasmids. This also negated the need for class III virus handling facilitates, which were unavailable at the time of this work. Had class III facilities been available for completion of this work then a slightly different approach would have been adopted.

(i) Completion of the additional cloning sites would have allowed separate addition of *gag*, protease and RT coding regions. This chapter includes the cloning strategy for the addition of an *XmaI* restriction site at the end of RT. Had this been completed it would have only required the addition of another site in protease to facilitate the desired single and multiple gene cloning system which would be required to dissect the effects of mutations in the patient derived viruses.

(ii) The lack of class III virus handling laboratory required the use of a replication incompetent system. Had these facilities been available then this work would have been superseded by the live virus work, which is the subject of the next chapter. This is because the effects of continued virus replication may help to determine the importance or function of the mutations seen in the patient derived viruses e.g. loss of mutations in a drug-free environment would implicate their involvement in resistance or compensation to antiretrovirals.

Chapter 4 Construction of a Live Virus Vector System to Study Drug Resistance

4.1 Introduction

The single-cycle pseudotyped virus assays have proved to be a valuable tool in predicting the resistance of patient derived viruses to antiretrovirals and have also been adapted to study replication capacity of such recombinants [13,31,138,270]. However, in order to study replication capacity and evolution in the presence of antiretrovirals, a system which allows infection and re-infection to occur is likely to be more relevant. The initial RVA system developed by Kellam *et al.*. requires homologous recombination to occur to form a replication competent virus [153]. This event is a normal cellular event and is similar to that which occurs during meiosis and mitosis and was also shown to occur in tissue culture [286]. Homologous recombination requires a double cross-over event to occur in order to generate a viable viral genome. This was first described by Upcroft *et al.*, where an SV40 genome was recombined to produce viable progeny [321].

The limitation of this system is the need to have identical sequences in the recombination or overlap region, although this is not a requirement for the system to function. However, if the sequences are not completely homogeneous across the overlap regions, then this will result in the generation of a heterogeneous virus population. The population derived from such a transfection experiment will result in a virus pool that differs in the base sequence in the recombination region.

RVA systems that generate replication competent patient derived viruses utilise the heterogeneity present within the patient virus pool to allow growth of the majority species present. [272]. This does not preclude the presence of quasispecies which may exist at a lower frequency and play little part in the generation of resistance or differential growth capacity. However, it is possible for these viruses to contribute a genetic diversity which could alter the genetic make up of the majority virus and so lead to the generation an artificial predominant virus population, especially under the *in vitro* conditions of transfection systems [88].

The development of replication competent virus vectors which allow for the insertion of clonal genomes have been previously reported [19,296], but these systems do not incorporate the *gag* gene and therefore, cannot be used to determine its effects on replication and resistance of the recombinant viruses. RVA systems have been developed that include the regions coding for the p7/p1 and p1/p6 cleavage sites of the *gag* gene [281]. As do the majority of single cycle phenotypic assays employed for the routine demonstration of resistance in patient derived virus isolates [257,296]. However, the inclusion of only the C-terminal portion of the gene precludes monitoring the effects of the whole of *gag*, which has been shown to effect drug susceptibilities and replication capacity of viruses [107].

Specific Aims

This chapter details the production of a replication competent HIV vector system in which restriction sites were introduced at the termini of the *gag*, protease and RT coding regions. The overall aim was to develop a system that encompasses the whole *gag* region and would also allow for insertion of the cognate protease and RT from virus isolated from patients who had prior exposure to antiretrovirals. The system would allow the relationship of mutations seen in regions more distal from the protein targeted by the antivirals used in the patients regimen to be studied. The use of a clonal genome input may also help to clarify the phenotype of the derived viruses sequence i.e. with no contribution from undetectable populations, as routinely measured by the relatively insensitive method of sequence analysis of a patient derived virus. The vector, based on the subtype B molecular clone pNL4.3 was shown to generate virus which was able to grow to comparable levels of the parental virus. Several different methods were employed to determine the amount of virus present within the harvested virus stocks.

I. Introduce unique restriction sites into an existing vector system, so as to facilitate the removal of the *gag*, Protease and RT genes.
II. Confirm that no other changes were introduced into the HIV coding regions.
III. Determine if the introduced mutations impact on the phenotypic characteristics of the virus.
IV. Check if the inserted restriction site mutations are stable in the virion.

4.2 Results

The initial steps in the production of the molecular live virus clone involved the production of a plasmid containing a fragment from the HIV-1 live virus vector pNL4.3 encompassing the protease and RT genes, pGEM-*Spe-EcoR*. A 5.8kb fragment from pNL4.3 was amplified using the primer pair 5'Gag*NarI* and 3'p8.91*EcoR*-down and the Expand Long Template PCR System (Roche). After conformation of correct PCR amplification of the product by agarose gel electrophoresis and EtBr visualisation, the whole PCR product was purified through PCR product purification columns (Stratagene). In order to clone the cleaned PCR product into the TA vector cloning vector, pGEMTEasy (Promega), the addition of Adenosine overhangs was first required, see Materials and Methods section 2.1.8. The pGEM ligation reaction was incubated overnight at 4 °C, and then an aliquot was transformed into chemically competent *E. coli* cells. The confirmation of colonies containing the correct *NarI* to *EcoRI* insert was carried out by plasmid DNA digestion of the pGEM clones using the *NarI* and *SpeI* restriction enzymes.

This vector containing the pNL4.3 PCR derived fragment in the TA vector provided the backbone sequence for the introduction of restriction sites to facilitate the removal of the protease and RT genes.

4.2.1 Insertion of restriction sites in *gag*, Protease and RT

Analysis of the HIV sequence was carried out using Vector NTi (Invitrogen) to identify potential restriction sites that could be added at the junction of the protease and RT sequences, as native restriction sites that flanked these genes were not present. A *NarI* restriction site located 151bp from the ATG start of *gag* was to be utilised for the restriction endonuclease removal at the N-terminus of *gag*. The criteria for the selection of suitable restriction sites were primarily their absence in the HIV coding sequence of the vector pNL4.3. Also required was the ability to incorporate the site without changing the subsequent amino acid sequence i.e. silent or synomonous changes

Once suitable sites were identified their introduction in the HIV sequence contained within pGEM-*Spe-EcoR* was carried out using the Quick-Change mutagenesis kit

(Stratagene). The addition of each restriction site required was carried out sequentially and involved: (see Table 4.1)

(i) Quick Change mutagenesis

(ii) transformation in *E.coli* and growth of selected clones

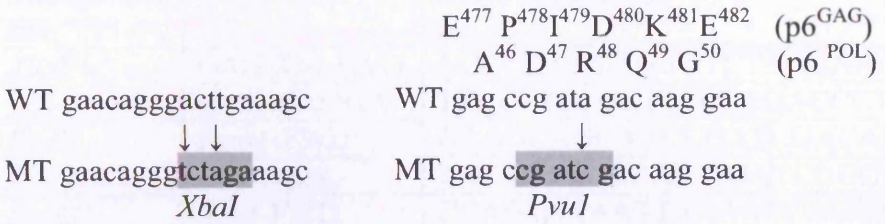
(iii) preparation of plasmid DNA from the picked clones

(iv) restriction enzyme digestion with the mutationally inserted restriction endonuclease and *EcoRI*; which is only present once in the pNL4.3 vector and thus, confirms the presence of mutation by production a double band upon EtBr visualisation.

The pGEM clone containing HIV *NarI* to *EcoRI* fragment with the *PvuI*, *SexAI* and *XmaI* restriction site was called pGEM(P+S+X).

The mutagenesis protocol utilises a high fidelity polymerase. Although these enzymes have demonstrated very low error rates compared to other polymerases, it was still thought necessary to adopt a cloning strategy that minimised the regions exposed to the mutagenesis system. This was overcome by transferring the smallest sized fragments possible through use of native restriction sites, into vectors that had not undergone mutagenesis. Thus, only the fragments that were cloned into the vector had been created by mutagenesis and were the only regions that required DNA sequencing.

Figure 4.1(a): Construction of pNL4.3+PSX (a) mutagenesis required for insertion of restriction sites..



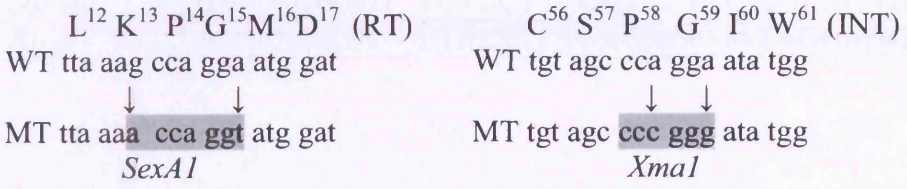


Figure 4.1(b): Construction of pNL4.3+PSX, exploded view of restriction sites around the ampicillin gene in pNL4.3. Numbers in parenthesis indicate base pair position in the molecular clone pNL4.3.

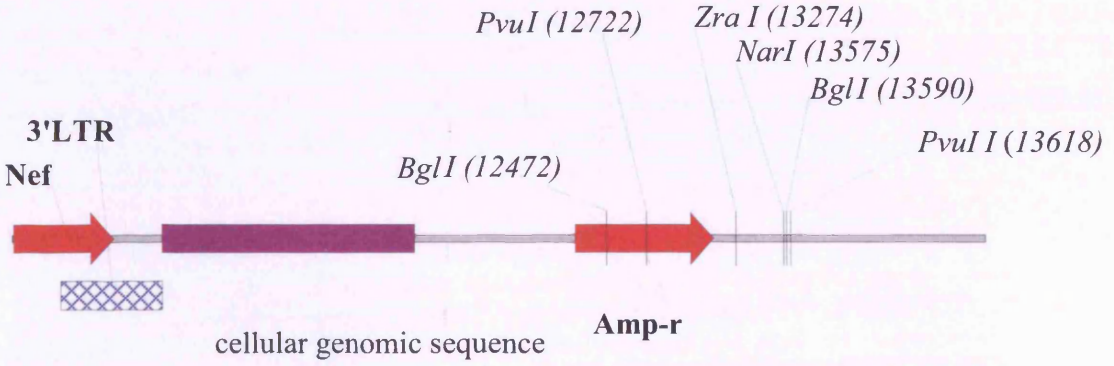


Table 4.1: Oligos used for mutagenesis

Mutagenesis site	Oligo Name	Oligo Sequence (5'→3')
XbaI	GAG-*Xba*-FWD	GCCCGAACAGGGTCTAGAAAGCGAAAGTAA
	GAG-*Xba*-REV	TTACTTTCGCTTTCTAGACCCTGTTCGGGC
PvuI	*Pvu*-1-FWD	GAAGCAGGAGCCGATCGACAAGGAACTGTA
	Pvu-1-Rev	TACAGTTCCTTGTCGATCGGCTCCTGCTTC
SexAI	*SexA*-FWD	CCAGTAAAATTAAAACCAGGTATGGATGGC
	SexA-REV	GCCATCCATACCTGGTTTTAATTTTACTGG
XmaI	8.91*Xma*-FWD	GACTGTAGCCCCGGGATATGGCAGCTAGATTG
	8.91*Xma*-REV	CAATCTAGCTGCCATATCCCGGGGCTACAGTC
(-)*PvuI* Amp^R	Blac-*Pvu*-FWD	CTCCTTCGGTCCTCCTATCGTTGTCAGAAGTA
	Blac-*Pvu*-REV	TACTTCTGACAACGATAGGAGGACCGAAGGAG

4.2.2 Removal of Inherent Restriction Sites *NarI* and *PvuI* from pNL4.3

As previously mentioned the existence of sites in the wild-type pNL4.3 HIV coding sequence was of primary concern, however, their presence in the coding or non-coding regions outside the HIV sequence did not exclude their selection.

The restriction site for *PvuI* was to be utilised for the excision of the C-terminal region *gag*. This site was also found to be present in the ampicillin resistance gene located in the pNL4.3 vector (see Figure 4.1b). To remove this site required the insertion of a silent change that would abolish the *PvuI* site, but retain the amino acid sequence. This was carried out with the Blac primer pair, see Table 4.1. Analysis of transformed *E.coli* colonies by *PvuI* and *EcoRI* digestion in correct clones produced a double band upon EtBr visualisation in the *PvuI* abolished clones. In those clones where the *PvuI* site remained, a triplet was produced, as another *PvuI* site is present in a non-coding region, located at base position 13618. A second single-cutter, *BglI* was also used in conjunction with *EcoRI* to confirm the removal of the *PvuI* site in the ampicillin gene. To ensure no spurious mutations were introduced after PCR mutagenesis, a *BglI* fragment from a clone containing the removed *PvuI* site was cloned back into pNL4.3 to create the plasmid pNL4.3-*Pvu*. This plasmid was then subjected to restriction digest with the enzymes *ZraI* and *PvuI* to remove a 346 bp fragment, which also contained a *NarI* restriction site. After gel purification the correctly digested plasmid molecules now contain protruding 3' overhangs from *PvuI* digestion and blunt ends from *ZraI* endonuclease action. The 3' overhangs were removed by treatment with *Pyrococcus furiosus* (*Pfu*) DNA polymerase, (Stratagene), which possesses 3′→5′ exonuclease activity. Thus, leaving a molecule with two blunt ends which were ligated together, creating the plasmid pNL4.3 (*-Pvu+Nar*).

Sub-cloning of the 4.2kb *SpeI* and *EcoRI* fragment which contained the inserted restriction sites *PvuI*, *SexAI* and *XmaI* from the pGEM-*Spe-EcoR* vector into pNL4.3(-*Pvu+Nar*) produced the molecular clone pNL4.3-PSX.

4.2.3 DNA sequencing of pNL4.3-PSX

DNA sequencing was carried out on the 4.2kb *SpeI* and *EcoRI* fragment which had undergone mutagenesis, to confirm the presence of the *PvuI*, *SexAI* and *XmaI* restriction sites and to verify no other errors were present. DNA sequencing was also carried out on the *BglI* fragment inserted from pNL4.3-Pvu to eliminate the *PvuI* site in the ampicillin resistance gene.

The following primers were used to verify the sequence of the clone pNL4.3+PSX.

3' pos: 13985; and 5'Amp-r were used for sequencing the ampicillin gene. 5'*Bgl*.seq and 3'*Bgl*.seq were used for sequencing the *BglI* fragment cloned back into pNL4.3. The following primers were used for sequencing the *SpeI* and *EcoRI* fragment that contained the three inserted restriction sites. The positions on the primers names are reference to their respective positions in the parental vector pNL4.3; pos: 1104, pos: 1620, pos: 2122, pos: 2619, pos: 3102, pos: 3687, pos: 4150, pos: 4644, pos: 5071 and pos: 1104(rev), see Materials and Methods Table 2.1.

4.2.4 Addition of Xba1 site to produce pNL4.3X-PSX

In order to generate viruses with the *gag* region derived from alternative sources required the utilisation of the N-terminal *NarI* site. Several approaches were tired to overcome the problem of cloning with this restriction enzyme. Although fragments could be correctly digested through the use of a modified enzyme, Turbo *NarI* (Promega), the major obstacle proved to be the ligation into the recombinant vector pNL4.3-PSX.

After several attempts with the modified enzyme and the use of an alternative enzyme *KasI*, which recognised the same restriction enzyme sequence, it was decided to introduce a new restriction site at the N-terminus of *gag*. Sequence analysis identified a possible *XbaI* site located upstream of the *NarI* site, 137 bp from the ATG start of *gag*. Mutagenesis on the pNL4.3-PSX vector and subsequent selection of *XbaI* containing clones was carried as described in Materials and Methods section 2.1.8. The oligos used for the mutagenesis are shown in Table 4.1.

Once a correct clone was identified, both this and the pNL4.3-PSX vector were digested with *Bst*Z171 releasing a 3.8kb blunt-ended fragment. This fragment from the *XbaI*

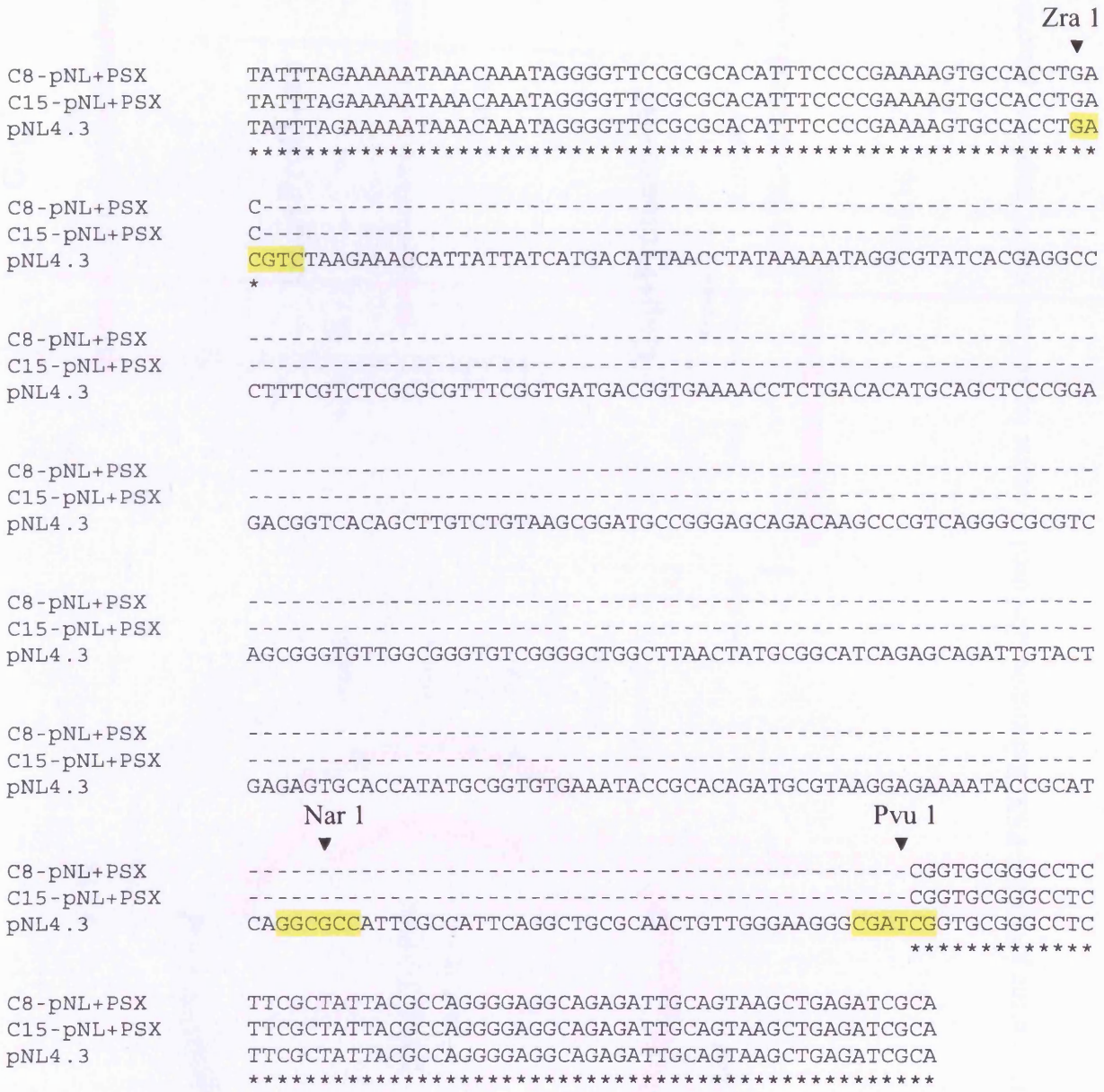
containing clone was blunt-ended ligated into the *Bst*Z171 cut pNL4.3-PSX vector and transformed into *E.coli*. The identification of clones containing the required *XbaI* site was achieved through *XbaI* and *EcoRI* restriction digest analysis and the clone displaying the correct restriction digest pattern was called pNL4.3X-PSX.

4.2.5 DNA sequencing of pNL4.3X-PSX

To confirm the presence of the site and also to ensure that no other mutations were inadvertently incorporated, the whole *Bst*Z171 fragment in pNL4.3X-PSX vector was sequenced. In addition to the primers previously mentioned in section 4.2.3, the following primers were used for sequencing the *Bst*Z171 fragment that contained the *XbaI* restriction site; AS3LTR, AS GAG 5A, Gag 3, Gag SE3, 5'Amp-r, 5'*Bgl*.seq and 3'*Bgl*.seq, see Materials and Methods Table 2.1.
Alignments for sequences after the removal of *NarI* and *PvuI* sites are shown in Figure 4.2.

The construction of the pNL4.3X-PSX vector produced a molecular clone in which the individual gene products *gag*, protease and RT could be removed by restriction enzyme digestion. Removal of the *gag* gene was accomplished by *XbaI* and *PvuI* digestion; the protease was excised by *PvuI* and *SexAI* digestion and removal of RT utilised the *SexAI* and *XmaI* sites. Combinations of genes could also be removed through correct restriction enzyme digest, allowing the insertion of linked *gag*, protease and RT genes from an alternative source. (See Figure 4.3)

Figure 4.2: Sequence alignment of pNL4.3+PSX clones 8(C8) and 15(C15) against pNL4.3; removal of 346bp fragment containing *Nar1* and *Pvu1* sites.



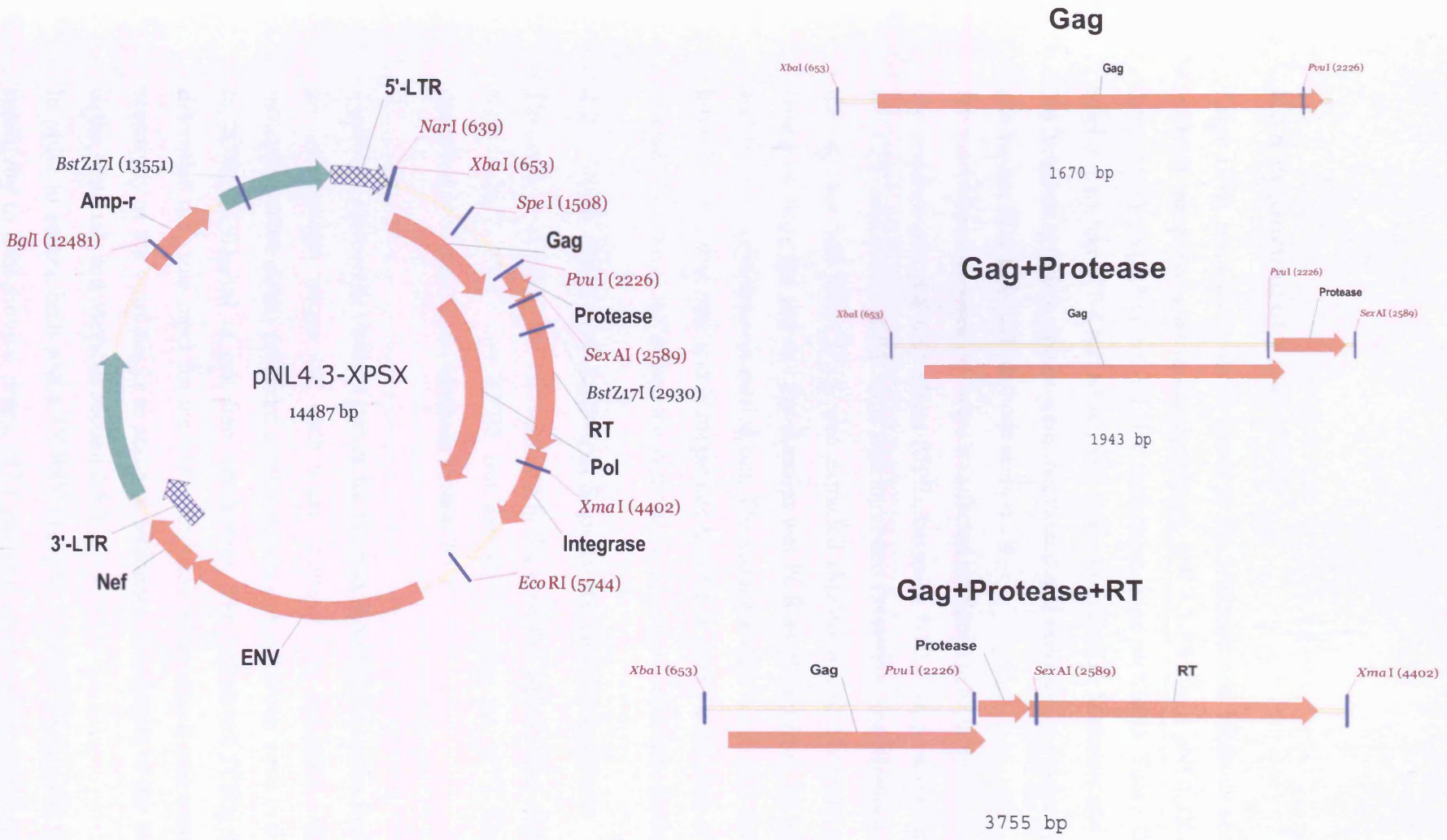


Figure 4.3: Schematic of the vector pNL4.3X-PSX. Restriction sites used to excise regions and size of fragments generated are also illustrated .

4.2.6 Production of Virus stocks

Large DNA stock preparations (maxi preps, Materials and Methods section 2.1.6) were made of the constructed molecular clones, pNL4.3-PSX and pNL4.3X-PSX along with the parental pNL4.3 clone and HIV molecular clone pHXB2D. These DNA stocks were used to transfect MT-4 and HEK 293T cells as outlined in Materials and Methods section 2.4.2 and transfected viruses were maintained and cultured to produce transfected virus stocks; see Materials and Methods section 2.4.

All the molecular clones that were transfected produced detectable virus as determined by the presence of cytopathic effect (CPE), between 7-10 days post-transfection. After the full CPE was observed the virus was harvested (Materials and Methods section 2.4) and the infected cell pellet DNA was extracted (Materials and Methods section 2.4.1). A region covering the *gag* and *Pol* domains was PCR amplified from the infected cell pellet and DNA sequencing was carried out. The sequence data confirmed the presence of the inserted restriction sites and homogenous nucleotide peaks were seen at these locations. The details of these techniques are given in the Materials and Methods section 2.1.13.

4.2.7 Virus Titre and Antiviral Sensitivity Determinations

The amount of infectious virus present in the prepared MT-4 virus stocks, NL4.3-PSX, NL4.3X-PSX, NL4.3 and HXB2 was determined using the MTT Titration Assay as described in Materials and Methods section 2.5.1.

Figure 4.4 shows the titration curves for the transfected virus stocks from a minimum of six independent assays and each assay contains four replicates. All the harvested molecular clone stocks produced similar amounts of infectious virus in the range of 9000 to 20000 $TCID_{50}$/ml of cell free supernatant. The calculated $TCID_{50}$/ml was used to determine the virus input for the MTT Antiviral Assay which was used to measure the sensitivity of the virus stocks to antiviral inhibitors. The details of the assay are outlined in the Materials and Methods section 2.5.3.

In order to ensure both pNL4.3X-PSX and pNL4.3-PSX vectors did not show altered sensitivity to antiretroviral drugs, MTT antiviral assays were carried out, see Materials

and Methods section 2.5.3. Figure 4.5 shows the IC_{50} values obtained for both viruses derived from pNL4.3X-PSX and the parental vector pNL4.3, in four PIs that are currently licensed. Several other stocks of different antiretrovirals were ordered from NIH, but failed to show any antiviral activity. Two different stocks of 3TC were ordered from the NIH, neither of these lots demonstrated any antiviral activity in several MTT assay determinations. The antiviral assays for the PIs SQV, APV and NLF were repeated on five separate occasions and demonstrate almost identical IC_{50} values in three out of the five assays. The two other assays in this data set show a less than 2-fold difference in the IC_{50}'s over the three PIs tested and also appear to have elevated IC_{50}'s compared to the other three assays. A less than 2-fold variation was seen in three out of the five assays with SQV, IDV and APV.

The fourth chart in Figure 4.5 represents the IC_{50} values obtained for three independent IDV antiviral assays. The NL4.3X-PSX virus stock demonstrates a 2-fold increase in IC_{50} compared to the parental virus NL4.3 and variation across the 3 assays also shows a 2-fold deviation.

The sensitivity of the virus stocks to the RT inhibitor AZT was also investigated and the percentage protection afforded by increasing AZT concentration is shown in Figure 4.6. Two separate assays were performed and X-PSX virus demonstrates slightly lower IC_{50} values in both assays compared to the NL4.3 virus stock. A difference of 19nM and 8nM was seen between the two virus stocks over the two assays and the overall variation between assays was less than 2-fold.

The double amino acid SQV resistance mutant, G48V+L90M virus stock was used as a control for the antiviral assay, to demonstrate an increase in IC_{50} to SQV. The growth of this virus stock was carried out as outlined in the Materials and Methods section 2.4. This mutant contains a double mutation which has previously been reported to confer approximately 40-fold resistance to SQV [20,144]. The assay demonstrates a >30-fold increase in resistance to SQV of the double mutant compared to wild-type, with IC_{50}'s of 200nM and 6nM respectively. The IC_{50}'s to the other two compounds AZT and APV were much more similar for the mutant and wild-type viruses; 33nM and 36nM for AZT and 80nM and 57nM for APV. A slight decrease in sensitivity to APV can be seen with

the SQV mutant, but this is less than 1-fold and so is within the error rate for the assay (Figure 4.7).

4.2.8 GHOST assays

An alternative assay system was also employed to determine the amount of virus present in the harvested viral supernatant stocks. This method involved the use of a reporter cell line, containing the green fluorescent protein (GFP). The GFP gene is under the control of a HIV-2 LTR; infection with HIV results in the production of the Tat protein which in turn induces GFP expression [217]. The correct plating density was represented by the plating density that resulted in a confluent layer of cells in each well of the 24 well plate on the day of FACS analysis, see Table 4.2 below. After 24 hours incubation only the highest cell density plate showed a full confluency level, further 24 hours incubation resulted in the next highest cell density also displaying 100% confluency. The two lowest cell densities remained sub-confluent even after 48 hours incubation.

Table 4.2: Plating density and confluency as visualised by light microscopy.

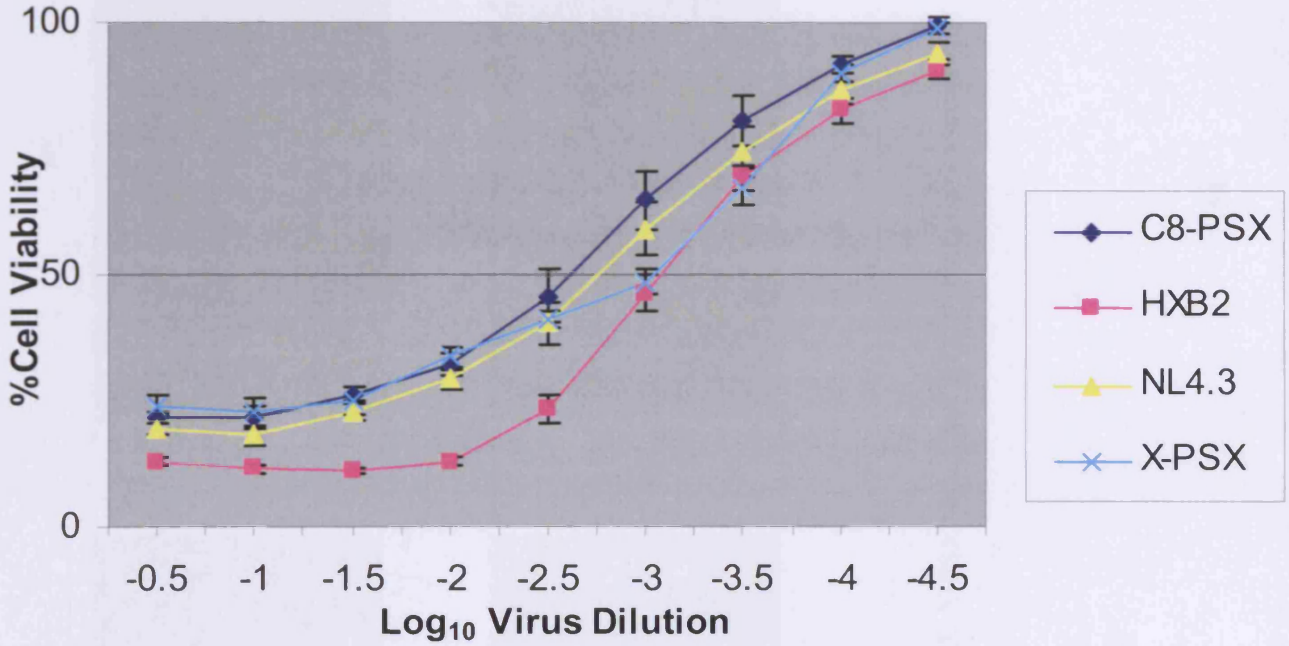
Cell Density per Well	Percentage Confluency	
	24Hr	48Hr
$1.25x10^4$	30%	50%
$2.5x10^4$	50%	75%
$5x10^4$	80%	100%
$1x10^5$	100%	100%

The infectivity of the virus stocks was measured using FACS to detect for the expression of the GFP by infected cells. The amount of green fluorescent events detected was used to measure the number of infected cells. At 24 hours post infection the highest number of infected cells was detected in the 50,000 cells per well plate, at the highest concentration of virus (see Figure 4.8). This was one of the intermediate plating densities used and also produced the highest percentage of infected cells at 48 hours. The lower plating densities suffered from increased cell death due to the virus killing at high virus concentrations, as

a consequence not enough events were counted during FACS and this skewed the data in the 48hrs experiment.

To determine the amount of virus present in the harvested cell-free supernatant stock the number of infected units per ml was calculated as described in Chapter 2 Materials and Methods section 2.5.2. Two independent experiments were used to calculate the titre of the NL4.3 virus stock are shown in Figure 4.9. The bars represent the number of infected cells as determined from the percentage infectivity, the line represents the calculated number of infected cells per ml or the titre/ml. Both the two highest virus dilutions were used to calculate the titre of the virus stock as these gave the most consistent percentage infections over the two experiments. The calculated titre and the standard error of the calculation are shown in Figure 4.9 in the inset. The titres calculated by both MTT assay and GHOST cell titration are almost identical and suggest a $TCID_{50}$/ml of 2000 ±287 or number of infectious particles per ml of 2330 ±373.

Figure 4.4: Titration curves of virus stocks of vector constructs and molecular clones. Error bars indicate Standard error over a range of values obtained from a minimum of 6 independent experiments.



$TCID_{50}$/ml	
C8-PSX	$=9.00 \times 10^{3}$
HXB2	$=2.00 \times 10^{4}$
NL4.3	$=1.13 \times 10^{4}$
X-PSX	$=2.00 \times 10^{4}$

$TCID_{50}$/ml as determined by MTT assay shown above.

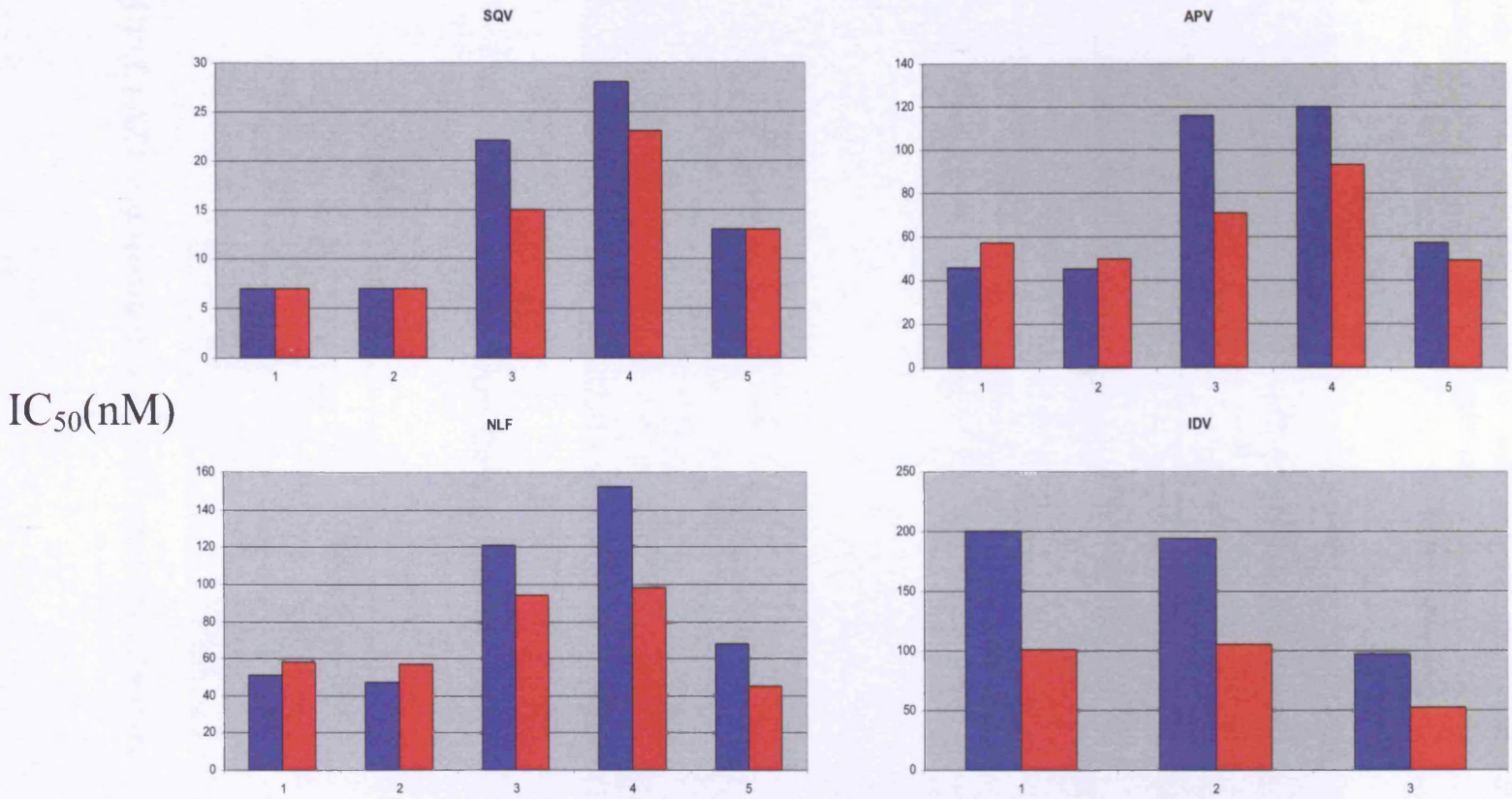
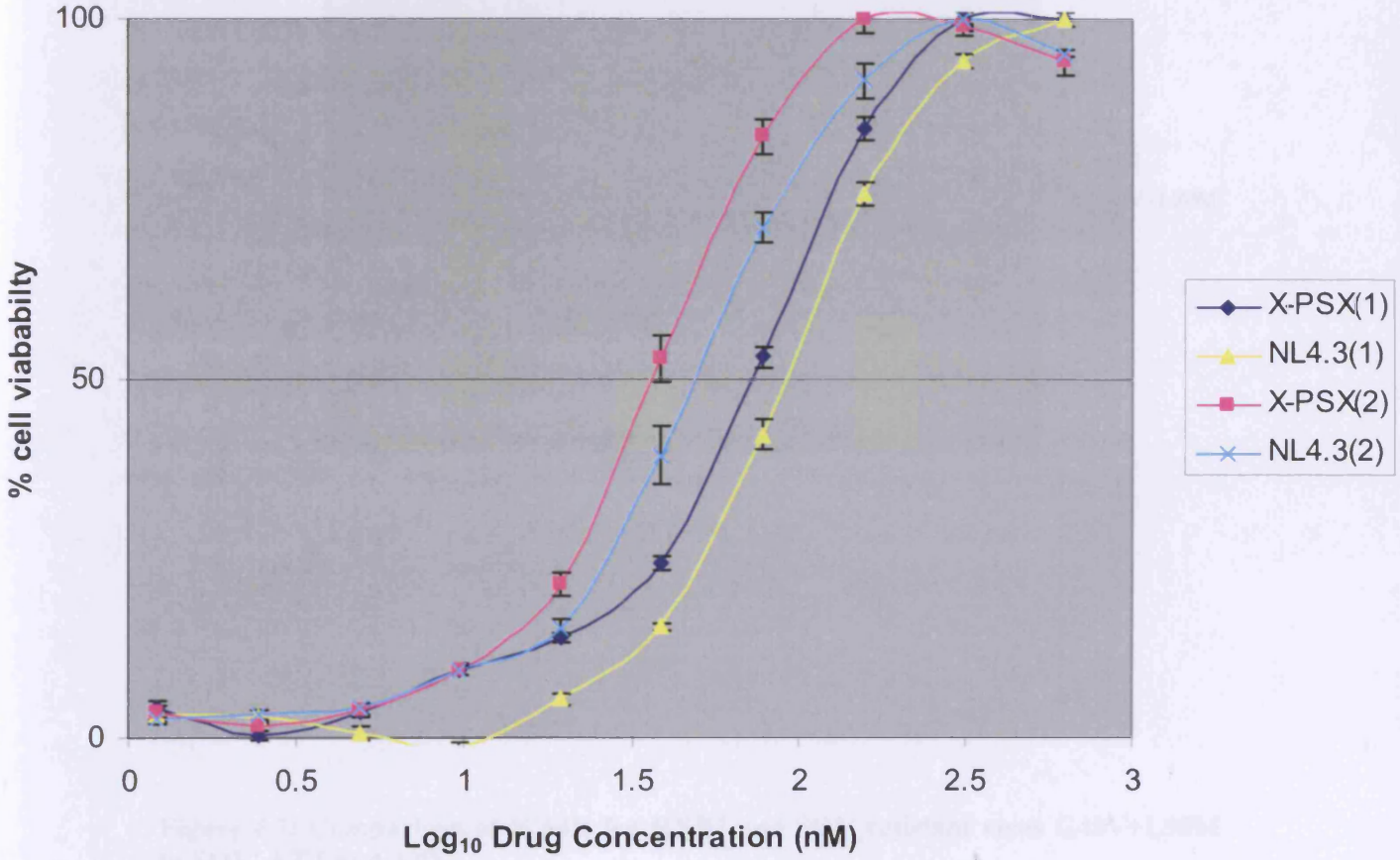


Figure 4.5: Repeat IC50 values for X-PSX and NL4.3. X-PSX (blue) and NL4.3 (red), shown are results for asays against four separate PIs. n=5 (SQV, AMP, and NLF); n=3 for IDV.

Figure 4.6: AZT Drug Sensitivity Assays. Assays carried out on two separate occasions with X-PSX and NL4.3. Error bars indicate the ranges of values obtained from quadruplicate data points in a single experiment.



AZT IC_{50}	
X-PSX (1)	73nM
NL4.3(1)	92nM
X-PSX (2)	37nM
NL4.3(2)	49nM

IC50 for AZT as determined by the MTT assay shown above.

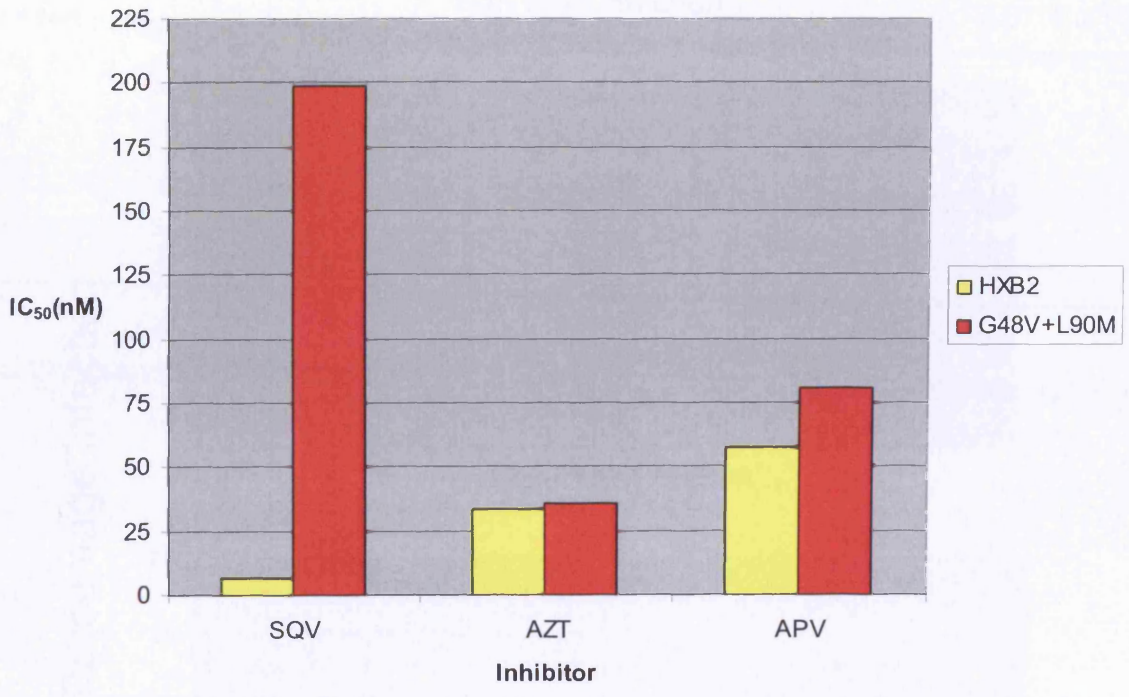


Figure 4.7: Comparison of IC50's for HXB2 and SQV resistant virus G48V+L90M in SQV, AZT and APV.

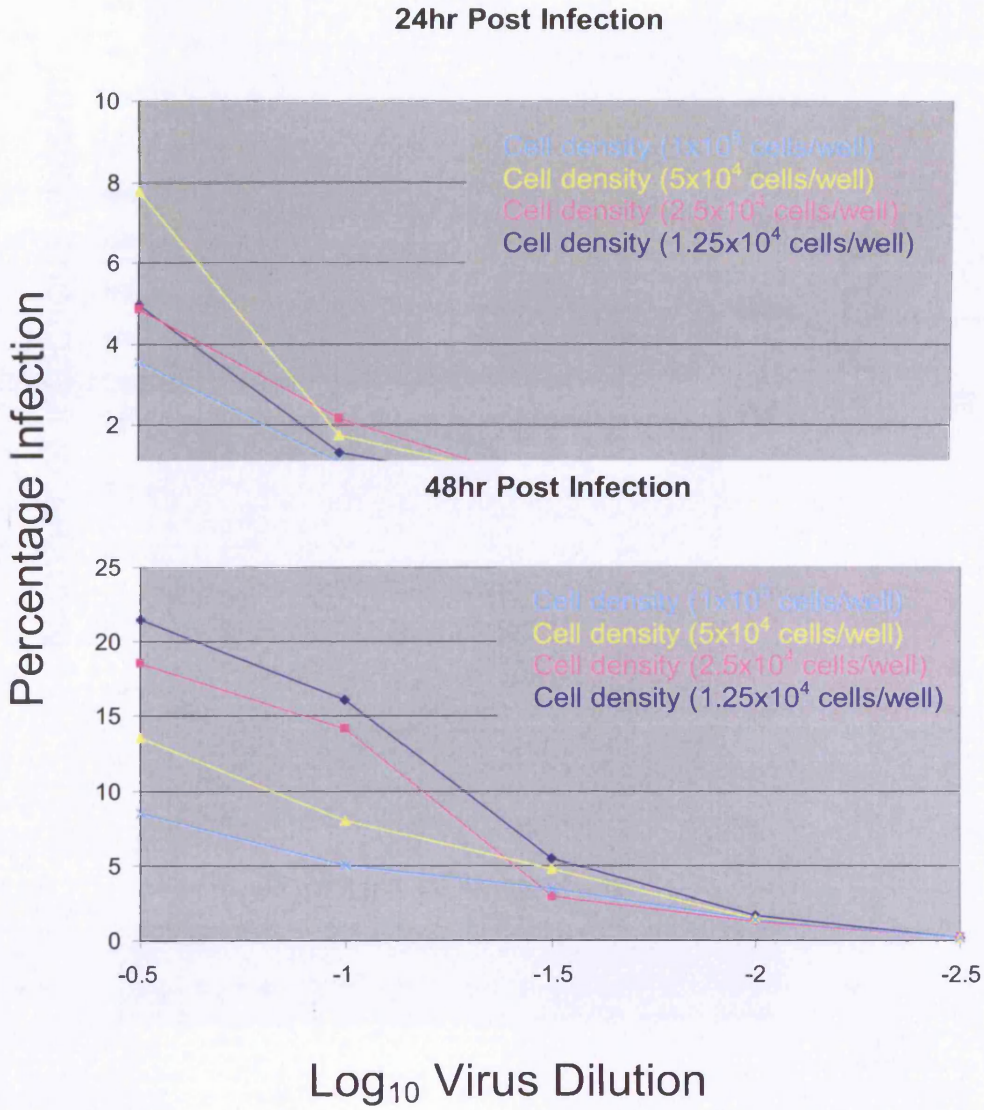
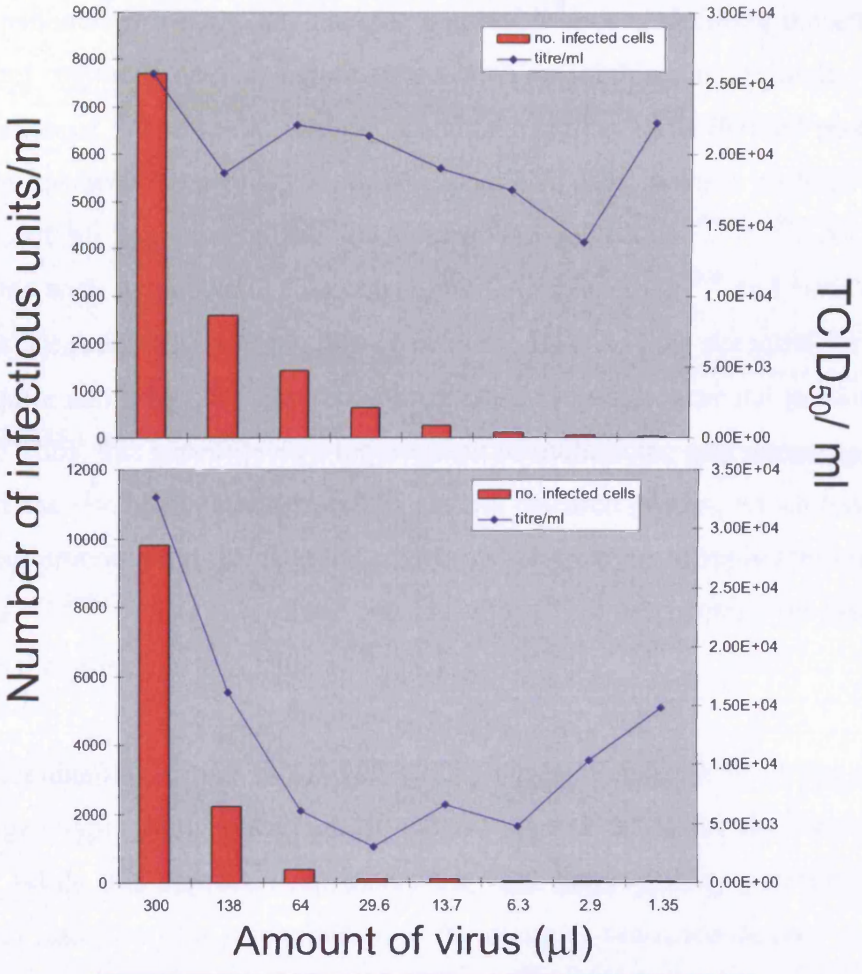


Figure 4.8: Percentage infectivity 24 and 48 hours post infection using NL4.3 and different plating densities on GHOST indicator cells.



Average Titre =2.33 x 10^4 infected cells/ml
Std Error = 3.73 x 10^3

Figure 4.9: Comparison of titrations of NL4.3 using MTT assay and GHOST assay.
Results from two independent assays for MTT assay (line) and GHOST assay (bars)

4.3 Discussion

The use of recombinant virus systems to establish the phenotype of viruses isolated from patients on therapy has become a valuable tool in deciding the effectiveness of antiviral regimens and to decide if a switch in inhibitor is required. [271,293,325]. The generation of recombinant viruses containing patient virus derived protease and RT regions has been the primary focus of research to date, as these proteins are the direct targets of all but one of the FDA approved inhibitors [55,249,329]. As the protease interacts with its substrate Gag, through steric processes [268] and both open reading frames are produced by the highly error prone HIV RT, the potential for the substrate to acquire and select for mutations in response to environmental pressures is highly likely. Thus, this necessitates a requirement to include the *gag* region in RVAs. Such a need has also been demonstrated by several research groups, which have shown that the Gag protein plays an essential role in the restoration of replication in PI resistant viruses [160,281,327]. The C-terminal peptide of the Gag polyprotein, p6 may also play a role in the resistance to NNRTIs [256].

The examination of patient derived virus phenotype through RVA has also used the virus genotypic pool present in the patient plasma, to derive the recombinant virus from. While this approach represents the viral heterogeneity present in individuals with an infection post-seroconversion, the majority sequence derived genotype, may not accurately represent the phenotype due to contributions from minority quasispecies. Such effects may lead to the generation of recombinants and could result in the appearance of low level quasispecies which contribute to the growth characteristics or antiviral resistance of the virus pool, possibly leading to inaccurate phenotype-genotype correlations. These systems are also limited in relating the significance of linkage between gene products, again due to the heterogeneity and recombination potential.

The previous chapter detailed the use of a pseudotyped transfection system that was able to incorporate the *gag* gene from an alternative source and produce a recombinant virus capable only of a single round of replication. This system was only able to detect infectious virus in only one recombinant constructed from a patient-

derived virus isolate and also produced a low level wild-type recombinant and was therefore considered of limited use. A further limitation of this system was that only the first 60 amino acids of protease could be cloned into the Gag-Pol expression vector along with the associated *gag*, as the only available restriction site was located within the protease coding region. Therefore, it was decided to develop a replication competent virus vector in which single or multiple gene products could be replaced and used to generate replication competent virus, with aims of replacing the *gag*, protease and RT with those from HIV positive, antiviral treated patients.

In order to construct this vector and to minimise the amount of sequencing required for confirmation of the vector construction, a cloning strategy was implemented as opposed to direct mutagenesis of an existing vector. If direct mutagenesis were carried out to insert restriction sites, the whole vector would require sequencing to confirm that no other errors had been inadvertently acquired through polymerase error during the mutagenesis reactions. Using the cloning approach would only require sequence confirmation of the *gag* and polymerase regions, (3.8 kb) instead of the entire viral genome.

Production of the molecular clone involved several cloning steps in which the whole 14.8kb vector was manipulated. Each of the cloning steps required meticulous optimization of ligation and transformation conditions in order to construct each of the vectors used in the construction of the molecular clone pNL4.3-PSX. In order to utilise this vector 'entirely' required cloning into the *NarI* site, located 151bp from the *gag* start and present in most HIV clades, except for clade F [176]. Another advantage of using the *NarI* site was that it included the Gag packaging loops, this region of HIV sequence; which in NL4.3 is the 107bp upstream of the *gag* start, may play an important role in evaluating replication capacity of any recombinants, as this region interacts with NC to effectively package the HIV genome [63,131]. However, removing a *NarI* fragment and ligating it back into the pNL4.3-PSX vector was very inefficient, with only two correctly ligated clones selected from >500 colonies screened. Several different approaches were adopted to try and overcome the problem of *NarI* cloning such as different ligation kits, restriction enzyme reaction components, BSA, enzymes suppliers and reaction buffers. An alternative enzyme *KasI*, which is an isoschizomer of *NarI* was also utilised, but proved to be just as inefficient. The addition of the *XbaI*

site 14 bp upstream of the *NarI* site, was positioned such that it included the Gag packing loops. It was hoped this would be easier to clone into, as *NarI* has been shown to demonstrate marked site preferences. For example on pBR322, the *NarI* site at 548 bp is cut very slowly (New England Biolabs product information sheet).
The DNA stocks that were made of the completed cloning vectors were used to produce replication competent virus stocks, by transfecting into the HTLV-I transformed T-cell line, MT-4. The virus stocks were made primarily to test for any deleterious effects of the synonymous mutations inserted during the production of the restriction sites. All the DNA stocks transfected produced CPE, detected by light microscopy, within similar time periods from transfection initiation. The viral infectivity was determined by titration of the cell free supernatant, which showed that all stocks produced similar levels of infectious virus. This would suggest that the mutations inserted do not affect the growth and replication ability of the viruses. This is to be expected, as all the changes introduced were non-coding and would not cause any structural or functional changes to the enzyme. However, it has been recently demonstrated that even synonymous changes can have a substantial impact on virion production. Such an effect has been reported with an A→G change in the *gag* NC coding sequence [128], resulting in a virus that produced normal protein and DNA levels compared to a virus without the A→G change. The drastic effect of the synonymous change was only evident by the lack of any detectable virion production and viral RNA in the cell-free supernatant.

The comparable growth rates seen during production of the viral stocks and their equivalent titres obtained after serial dilution of the stocks, suggests that productive infection rate i.e. the rate of infection, virion production and re-infection, is similar. As each viral clone was transfected with the same number of cells and these were maintained with equivalent amounts of cells and fresh media and the virus stocks were harvested in the same overall volume. Each assay was conducted a minimum of six times and produced $TCID_{50}$ values all within 2-fold of each other (Figure 4.4). This level of variation is to be expected from a cell based assay, especially considering there is a five day incubation period in which any differences in cell plating number between assays can be amplified. The titration value obtained from this assay is primarily used to equalise the amount of virus supernatant used to infect cells for the antiviral assay. This assay system utilises the same cells and plate format

as the titration assay and thus provides a complementary system for the analysis of inhibitor sensitivity. This complementarity was seen in the similar IC_{50} values obtained from the antiviral assays against PIs and AZT (Figures 5 and 6). Although most of the assays with the same virus stock were within a 2-fold range, two sets of assays displayed higher IC_{50}s. This was consistent for all the PIs, as the assays numbered three and four with SQV, NLF and APV were carried out with the IDV assays numbered one and two. This increase in IC_{50}, means that more antiviral was required to inhibit the viruses and could be possible if the antiviral stocks were too dilute e.g. through pipetting error. However this is unlikely as it occurred in all the PIs and over two separate assays, during which fresh stocks were made from the concentrated stock solutions. Another, more likely possibility, is that because the end-point measurement in this system is cell viability; as measured by the conversion of the tetrazolium compound to a formazan compound by viable cells [251], any increased IC_{50} could represent poor cell growth over the course of the assays this may occur after cells have been cultured for long periods. This would be detected as less formazan conversion, which has the same effect as increased cell death. Elevated IC_{50}'s could also be attributed to a low initial cell number in the assays, through inaccurate cell count or through too much virus added initially.

The X-PSX virus appears to show a consistent increase in resistance to IDV, as is shown in Figure 4.5. It is possible that these differences in IC_{50} values may reflect possible errors in the titration at the time of these assays. The increase is nearly two-fold, which is within the limits of error expected for such cell based, replication competent assays [257,272].

The *SexAI* site was introduced in the N-terminal region of RT and involved the insertion of synonymous changes in codons 14 and 15 of this protein. Although neither of these sites have been reported as being involved in resistance to RTI's [40], the activity of this class of antiviral was also checked using AZT. The X-PSX virus and the control virus NL4.3 displayed similar sensitivity to this RTI. Figure 4.6 shows typical sigmoid curves generated when titrating the virus in escalating inhibitor concentrations. Again, the two separate assays are all within a two-fold range of each other, with a less than two-fold difference between the viruses in each assay. Some of the points within these assays appear to display greater than 100% values in terms of cell viability. This occurs in the wells of the assay plates that contain high drug

concentrations; the cells are initially incubated with the virus prior to plating out, the infected cells that are added to high drug concentration wells are afforded complete protection by the compound. The presence of virus has previously been shown to induce growth proliferation in cells [135]. In this system the high drug concentrations prevent further virus growth but do not affect the growth of the cells, which are in a more activated state due to the initial presence of the virus and so, outgrow the uninfected cells in the cell-control, which is used as the 100% cell viability value.

To demonstrate the expected shift in drug sensitivity with a resistant virus genotype, the SQV double mutant, G48V+L90M was grown, titrated and assayed in MT-4 cells. The assays were carried out using the molecular clone HXB2 as the wild-type control, as the mutant virus was constructed by the RVA method, using a protease deleted HXB2 vector [202]. Both viruses showed similar IC_{50}'s for AZT, however in this assay the mutant also seemed to require a greater concentration of APV to inhibit growth of the virus by 50%. Neither of the mutations G48 or L90M is thought to be involved with resistance to APV either *in vitro* or *in vivo* [40,135]. This increase in IC_{50} of 23nM is more likely to reflect variation within the titration values calculated for viruses and results in altered MOI's. The very low MOI of this assay also means that virus input volumes can be low in this system especially with high titre viruses and thus, any errors could mean vastly different virus amounts. However, this error is still below the 2-fold error of such assay systems and was primarily carried out to demonstrate the shift in IC_{50} associated with resistance to SQV.

The MTT assay system determines the titre of virus in the virus stock as the percentage of cell viability i.e. the amount of supernatant required to kill 50% of the cells is the $TCID_{50}$. The MTT assay is a cell based assay that relies on virus induced cell death as means to quantify virus content. In order to maintain consistency between assays, cells are routinely grown to similar cell densities and passaged. A day before assay use, the cell are split again in order to ensure actively dividing cells, which would reduce errors due to cell growth phase differences and ensure that cells are similarly infectable. Each MT-4 cell line is passaged a maximum of 25 times before a new line from frozen cell stocks is used to replace the old cells. As with all cell based assays other variability may occur such cell number between assays, slight volume differences and differing evaporation rates, which could be significant in an

assay incubated for 5 days. This particular problem is reduced through the use of low evaporation plates, which showed no significant difference in cell number in edge wells when plated at the same density and counted one week after plating (data not shown). Each virus grown is titred by serially diluting the virus on a 96 well plate, each well containing a constant number of MT-4 cells. After the 5 day incubation period the cells are treated with MTT, which is metabolised by viable cells and causes a colour change which is measured using a spectrophotometer. The measurement of drug sensitivity is carried out in exactly the same format, with only two differences the first being the incubation of the virus with the cells for 2 hours and ensure the virus amount is constant between wells, the infection is set-up in bulk and aliquoted into each well. The second difference is the addition of the antiviral, which is serially diluted across the 96 well plate. This similarity between the determination of the amount of virus in the stock and the drug sensitivity assay also help to ensure that variability due to assay format differences are reduced.

Each assay, both titration and IC50, are performed with a NL4.3 control, which over time accumulates several separate data determinations. To ensure assay reproducibility, all assays in which the determined values for the NL4.3 control show a greater than two-fold variation, are disregarded. However, it is an indirect measure of the amount of virus present as no viral protein or target is used as the end-point measurement. For this reason an alternative method was explored for virus quantification. The GHOST cell-line contains the GFP gene linked to the HIV-2 LTR which is driven by Tat activation upon HIV infection and thus, is a direct measure of virus production. These cells are derived from Human Osteosarcoma (HOS) cells stably expressing CD4+ receptor and both CXCR4 or CCR5 coreceptor and are therefore able to support infection by HIV. The initial experiments in the development of a titration assay using this cell-line were designed to evaluate the best cell plating density and incubation periods to produce a reproducible infection, proportional to the amount of virus added. The use of this cell-line to measure infectivity and replication capacity of drug-resistant HIV has been previously reported by several groups [19,299]. These studies used this system to quantify the single-cycle infectivity of viruses by measuring the amount of GFP detected after 24hours infection, as the HIV generation time is thought to be between one to two days [5,254]. The single-cycle infectivity of the

NL43 virus was highest with the cell that density that showed approximately 80% confluency on the day of FACS analysis.

The reproducibility of the assay in measuring the amount of infectious particles in the NL4.3 virus stock is shown in Figure 4.9. The two highest virus dilutions were used to calculate the number of infectious units per ml as these values gave more significant percentage infection values than the lower virus dilutions. The NL4.3 virus stock was previously used in the MTT assay system for $TCID_{50}$ determination and this system gave a titre of $2x10^4$ $TCID_{50}$/ml. The GHOST assay also gave a very similar titre for the same virus stock of $2.33x10^4$ infectious units/ml.

Stability of mutations was confirmed by sequence analysis of the areas covering the restriction sites; the presence of the mutations after two to three weeks in cell culture confirmed the stability of these mutations. If these mutations had a detrimental effect on the virus population, then it would be expected that these would be lost overtime and a heterogeneous population would be seen in the chromatogram peaks associated with the restriction sites, however, this was not the case.

The development of this restriction site cloning system has the disadvantage of requiring cloning of the patient derived virus fragment into the vector; this has proved to be difficult and the construction of virus isolated from patients is the subject of the next chapter. The advantage of this system is that the subsequent transfection experiments involve a homogeneous population, ideal for looking at primary infection. Any difference seen from the input genotype must therefore be as a result of evolutionary selection and not due to the presence of any other pre-existing species that is present as a minority. Increasing the diversity of the input genome e.g. to allow for the examination of growth of a more complex heterogeneous system, would simply require the transfection of several different clones. Recombinant clones generated using this system incorporate the whole *gag* region and the packaging signal from the sample source. It is hoped that the development of this vector will allow for the examination of the effects of mutations associated with HAART not only in the protease and RT coding regions but also in the *gag* gene and also any variation that occurs in the packaging loops.

Although the aims of this chapter were achieved, if circumstances were ideal several other approaches would have been explored.

(i) Use of restriction site other than *XbaI* that is less likely to occur in the HIV regions to be studied, although this had previously been used by other researchers as a cloning site in construction of similar replication competent vectors.

(ii) The cloning vector contained the wild-type genes which were to be substituted by the patient virus derived regions of interest, thus gel-purification of the digested X-PSX vector was required in order to remove the wild-type region. The use of small oligonucleotide linkers which could have been ligated into the cut vector would abolish the need for an inefficient and time consuming approach as gel-purification, which also resulted in loss of 30%-50% of vector DNA, (referenced from-Stratgene product insert).

(iii) Had more time been available a pseudotyped vector would have been developed, this approach, similar to that used in Chapter 2 would have been developed to complementary to the live vector used in the remainder of the research. By this it is meant that, the X-PSX vector would have been used as a single cycle vector by either the deletion of the envelope region or insertion of a stop codon to truncate it. This could then be used with either a VSV-G envelope, such as the pMDG vector used in chapter 2, or the patient virus derived envelope could have been used to explore the effects of envelope variation on phenotype. This would also expand the use of this system in that it could be used to determine sensitivity of mutated envelopes to entry inhibitors.

Chapter 5 Use of Live virus vector system for clinical isolates

5.1 Introduction

The assessment of antiviral sensitivity of patient derived viruses through the use of phenotypic assays provides a direct approach and enables a more accurate estimation of the effect of the genotype on antivirals, than genotypic assays. Phenotypic resistance testing using primary isolates has been previously reported and requires the co-cultivation of HIV infected PBMC's with freshly isolated PBMC's [145]. These systems have the advantage of culturing the whole viral genome, but involve long culture times and are labour intensive and so not suitable for routine assay use.

Phenotypic testing is now regularly performed using plasma derived virus PCR products that encompass the protease and RT coding regions. Systems which are based on using homologous recombination to construct the virus genome [136] or more recently single cycle assay systems have been described [257] .
The use of PIs results in the appearance of morphologically aberrant forms of HIV particles, which represent the blockage of cleavage of the Gag polyprotein as determined through EM and analysis of Gag cleavage patterns [52]. The restoration of protease function has been shown to occur thorough the accumulation of resistance and compensatory mutations in both the protease and at the *gag* cleavage sites [84,223,230,292].

The phenotypic systems that have been developed to date are unable to investigate the effects of mutations in patient derived viruses across all three regions, *gag*, protease and RT, although some include the C-terminal portion of *gag* [257,281].
Previously, a live virus vector was constructed which contained restriction sites that could be used to facilitate the introduction of either single or multiple coding regions for *gag*, protease and RT. This chapter details the construction of chimeric constructs which contain the aforementioned coding regions derived from patient virus that had prior exposure to antiretroviral therapy. Several constructs were made and used to generate recombinant virus stocks, which were tested for resistance to a panel of FDA

approved compounds. The ability of some of the clonal HIV recombinant genomes to maintain the introduced sequence in the absence of drug pressure was also studied.

Specific Aims and Questions

The specific aim of this chapter was to demonstrate that the addition of *gag* along with the highly resistant protease, affects phenotypic drug sensitivity determinations and therefore, should be routinely considered when such assays are undertaken to determine the effectiveness of a particular drug regimen, especially when there has been extensive prior exposure to PIs . The chosen approach utilises a single genotype, which demonstrates that the mutations shown in the *gag* and protease regions are linked on a single genome and so are more likely to co-exist.

I. Demonstrate the use of the cloning system in the development of recombinants clones, containing single and multiple inserts of *gag*, protease and RT, derived from patient isolated virus.
II. Production of high titre virus stocks from patient virus clones
III. Assessment of phenotypic resistance

5.2 Results

5.2.1 Amplification of *gag*, protease and RT coding regions from Patient isolated virus

Patient samples that were previously identified in Chapter 3 were used to construct the recombinant virus clones. The previous antiviral exposure history, mutations in protease and RT and viral loadof the isolates are also shown in Chapter 3 Table 3.1.

Two approaches were used to generate the replication competent recombinant HIV clones from the patient derived viral RNA. The first was to amplify PCR products with primers that contained the appropriate restriction sites from the patient derived viral cDNA and directly clone into the pNL4.3-PSX vector. The PCR primers conditions were optimised for subtype B, pNL4.3, as most of the samples were of this subtype. However, patient virus 2 was classified as subtype C based on protease and RT sequence and the protease region of patient virus 1, which was classified as subtype D. Also the position of the oligo's was determined by the position of the restriction enzyme sites at the junction of the coding regions of gag, protease and RT,

therefore there was very little possibility for variation, with respect to position. The only variables that could be modified and may have resulted in an increased number of samples being PCR positive were Mg^{2+} concentration and the annealing temperature of the oligos. The use of longer and thus more stringent oligo was not undertaken as amplicons had already been generated in order to demonstrate the cloning required to produce chimeras of the gag, protease and RT coding regions, either separately or in combination.

To generate the *gag* only fragment the primer set 5'Gag-*Nar1* and 3'Gag-*Pvu1* was used, which produced a PCR product of 1612 bp. The *gag* and protease fragment required the primer set 5'Gag-*Nar1* and 3'Pro-*SexA1* and generated a 1976 bp amplicon. The protease only amplicon of 390 bp was produced with the primer set 5'Gag-*Pvu1* and 3' Pro *SexA1*. The protease and RT region was amplified with 5'Gag-*Pvu1* and 3'Integrase-*Xma1*, which produced a 2202 bp amplicon.
As mentioned previously, the lack of success in cloning the *gag* fragment into pNL4.3-PSX was attributable to the *NarI* restriction site and the inability to successfully ligate into it. This obstacle prevented the construction of clones contain the *gag* region derived from alternative sources. Despite this, several clones were produced from ligating into the first molecular construct, pNL4.3-PSX. The eventual production of the pNL4.3X-PSX cloning vector allowed the incorporation of *gag* through the introduction of an *XbaI* site at its N-terminus. To amplify any fragment for cloning the *gag* region was now carried out with the primer 5'Gag-*XbaI*, resulting in amplicons 10bp shorter than those produced with 5'Gag-*Nar1*.

5.2.2 Production of Recombinant Clones

After production of the patient derived virus recombinants, the sequences of clones and the HIV coding region was analyzed by DNA sequencing. The details of these techniques are given in the Materials and Methods section 2.1.13. Clones that contained the desired mutational patterns were assigned the following nomenclature, Pt (Patient) number (1-5), followed by the region (P) protease; (PR) protease and RT), then follwed by the clone number.

5.2.2a Protease and RT clones

Production of these clones was by direct amplification from cDNA the details are given in the Materials and Methods sections 2.1.15 and 2.1.10, using the high fidelity Phusion Polymerase (Finnzymes) and then ligation into the *PvuI* and *XmaI* digested pNL4.3-PSX. Only two patient derived virus samples produced an amplicon with these PCR conditions, virus derived from patients 1 and 5. These two samples that were also those with the highest viral loads, which implies that the PCR with these oligos was inefficient and required further optimisation. Several pNL4.3-PSX patient virus clones were isolated that upon PCR amplification with *gag* and RT primers, located within the vector pNL4.3-PSX produced amplicons. These were gel purified and the protease coding region sequenced as outlined in the Materials and Methods section 2.1.13.

A total of six clones containing the protease and RT coding regions derived from patient virus samples 1 and 5 were produced ; 4 from patient 1 derived virus (Pt1(PR)-C4, C6, C7, C12) and 2 from patient 5 derived virus (Pt5(PR)-C25 and C23). Figure 5.1a shows the protease alignment of the clones from both patient derived virus samples aligned with pNL4.3-PSX. Both virus clones from patient 5 were identical in amino acid sequence and showed 5 changes from pNL4.3-PSX, these were a L10V, T12V, I13V, L63P and I64V.

A total of four recombinant clones were produced from virus derived from patient sample 1 with all but one of the clones showing identical amino acid sequence across the protease region. The only diverse clone contained an additional K20R mutation, but was otherwise identical to the other clones generated. The proteases from patient 1 derived virus all contained a two amino acid insertion at codon 35, composed of a Glutamine and an Asparagine. In all fourteen amino acid changes were seen in these clones compared to the parental clone pNL4.3-PSX. The changes detected were L10V, I13V L24I, L33F, M36L, M46I, I54V, K55R, R57K, I62V, L63P, I64V, L76V and V82A.

5.2.2b Patient Derived Virus Protease clones

As with the protease and RT clones the amplicons for cloning directly into pNL4.3-PSX were generated from the cDNA and cloned into the *PvuI* and *SexAI* digested vector. These ligations proved to be the most efficient, with a 95% of clones selected (19 out of 20) containing the protease region as determined by PCR amplification across the cloning junction, with primers that annealed in *gag* and RT regions (see Materials and Methods section 2.1.9 and Table 2.1). A total of six clones were analysed by DNA sequence analysis and the alignment of the amino acid sequence of the protease is shown in Figure 5.1b. All clones selected except one demonstrated complete sequence homology with each other, as they did with the protease and RT clones, clone 6 of this latter group, contained the K20R mutation.

Figure 5.1a: Alignment of the protease region from protease and RT clones. Clones obtained from both patient derived virus samples 1 and 5, the shaded codons represent changes from pNL4.3-PSX

```
                 ....|....| ....|....| ....|....| ....|--....| ....|....| ....|...
                     5          15         25         35         45         55
Pt1(PR)-C4       PQITLWQRPV VTVKIGGQLK EALIDTGADD TVFEEQNLNLPG RWKPKIIGGI GGFVRVKQ
Pt1(PR)-C6       PQITLWQRPV VTVKIGGQLR EALIDTGADD TVFEEQNLNLPG RWKPKIIGGI GGFVRVKQ
Pt1(PR)-C7       PQITLWQRPV VTVKIGGQLK EALIDTGADD TVFEEQNLNLPG RWKPKIIGGI GGFVRVKQ
Pt1(PR)-C12      PQITLWQRPV VTVKIGGQLK EALIDTGADD TVFEEQNLNLPG RWKPKIIGGI GGFVRVKQ
Pt5(PR)-C23      PQITLWQRPV VVVKIGGQLK EALLDTGADD TVLEE--MNLPG RWKPKMIGGI GGFIKVRQ
Pt5(PR)-C25      PQITLWQRPV VVVKIGGQLK EALLDTGADD TVLEE--MNLPG RWKPKMIGGI GGFIKVRQ
pNL-PSX          PQITLWQRPL VTIKIGGQLK EALLDTGADD TVLEE--MNLPG RWKPKMIGGI GGFIKVRQ
Clustal Co       *********: *::******: ***:****** **:**  :**** *****:**** ***::*:*

                 .| ....|....| ....|....| ....|....| ....|....|
                       65         75         85         95
Pt1(PR)-C4       YD QVPVEICGHK AIGTVVVGPT PANIIGRNLL TQIGCTLNF
Pt1(PR)-C6       YD QVPVEICGHK AIGTVVVGPT PANIIGRNLL TQIGCTLNF
Pt1(PR)-C7       YD QVPVEICGHK AIGTVVVGPT PANIIGRNLL TQIGCTLNF
Pt1(PR)-C12      YD QVPVEICGHK AIGTVVVGPT PANIIGRNLL TQIGCTLNF
Pt5(PR)-C23      YD QIPVEICGHK AIGTVLVGPT PVNIIGRNLL TQIGCTLNF
Pt5(PR)-C25      YD QIPVEICGHK AIGTVLVGPT PVNIIGRNLL TQIGCTLNF
pNL-PSX          YD QILIEICGHK AIGTVLVGPT PVNIIGRNLL TQIGCTLNF
Clustal Co       ** *:::****** *****:**** *:******** *********
```

Figure 5.1b: Alignment of the protease region from protease only clones. Clones obtained from patient derived virus sample 1, the shaded codons represent changes from pNL4.3-PSX

```
                 ....|....| ....|....| ....|....| ....|--....| ....|....| ....|...
                     5          15         25         35         45         55
Pt1(P)-C3        PQITLWQRPV VTVKIGGQLK EALIDTGADD TVFEEQNLNLPG RWKPKIIGGI GGFVRVKQ
Pt1(P)-C8        PQITLWQRPV VTVKIGGQLK EALIDTGADD TVFEEQNLNLPG RWKPKIIGGI GGFVRVKQ
Pt1(P)-C1        PQITLWQRPV VTVKIGGQLK EALIDTGADD TVFEEQNLNLPG RWKPKIIGGI GGFVRVKQ
Pt1(P)-C15       PQITLWQRPV VTVKIGGQLK EALIDTGADD TVFEEQNLNLPG RWKPKIIGGI GGFVRVKQ
Pt1(P)-C18       PQITLWQRPV VTVKIGGQLK EALIDTGADD TVFEEQNLNLPG RWKPKIIGGI GGFVRVKQ
Pt1(P)-C20       PQITLWQRPV VTVKIGGQLK EALIDTGADD TVFEEQNLNLPG RWKPKIIGGI GGFVRVKQ
pNL-PSX          PQITLWQRPL VTIKIGGQLK EALLDTGADD TVLEE--MNLPG RWKPKMIGGI GGFIKVRQ
Clustal Co       *********: **:******* ***:****** **:**  :**** *****:**** ***::*:*

                 .| ....|....| ....|....| ....|....| ....|....
                       65         75         85         95
Pt1(P)-C3        YD QVPVEICGHK AIGTVVVGPT PANIIGRNLL TQIGCTLNF
Pt1(P)-C8        YD QVPVEICGHK AIGTVVVGPT PANIIGRNLL TQIGCTLNF
Pt1(P)-C10       YD QVPVEICGHK AIGTVVVGPT PANIIGRNLL TQIGCTLNF
Pt1(P)-C15       YD QVPVEICGHK AIGTVVVGPT PANIIGRNLL TQIGCTLNF
Pt1(P)-C18       YD QVPVEICGHK AIGTVVVGPT PANIIGRNLL TQIGCTLNF
Pt1(P)-C20       YD QVPVEICGHK AIGTVVVGPT PANIIGRNLL TQIGCTLNF
pNL-PSX          YD QILIEICGHK AIGTVLVGPT PVNIIGRNLL TQIGCTLNF
Clustal Co       ** *:::****** *****:**** *:******** *********
```

The second approach was used to produce *gag*, protease and RT fragment templates for the virus genomes derived from the patients, as this would show linkage of the separate coding regions on a single genome. To carry this out the stored patient viral RNA was reverse transcribed as mentioned in Chapter 3 and the *gag*, protease and RT encoding regions from were amplified by PCR using the primer set 5'Gag-*Nar1* and 3'Integrase-*Xma1*. The resulting 3788 bp amplicons were visualized by EtBr stained agarose gel electrophoresis and were then cloned into the pGEM cloning vector (Promega) as outlined in the Materials and Methods section 2.1.12. The resulting clones were screened by PCR screening of the colonies generated after transformation of the pGEM ligation reaction into *E.coli* HB101's. The screening was performed with oligo's T7 and SP6, resulting in a product of 3.8 kb upon visualization by agarose gel electrophoresis.

Only three samples; virus derived from patient 1, 2 and 5 produced full-length amplicons from the pGEM clones, which contained the whole coding regions for *gag*, protease and RT. Although full-length amplicons were generated from all patient virus derived RNA, suggesting that the 5'Gag-*Nar1* oligo is more efficient, which is to be expected as it was designed to utilise a region which contained an inherent *Nar1* site present in most sub-types.

To minimise sequencing and as the scope of this project was to demonstrate the ability to clone virus derived sequences from alternative sources into the restriction sites, only virus derived from patient samples 1 and 5 were sequenced. The *gag* amino acid sequence data for pGEM clones Pt1pGEM-C2, and Pt5pGEM-C20 is shown in Figure 5.2 and the changes from pNL4.3-PSX for the protease and RT regions are listed in Table 5.1.

Table 5.1: Mutations identified in pGEM clones containing the HIV protease and RT regions isolated from 2 patient derived virus isolates. The mutations in bold are associated with resistance.

	Protease	RT
Patient 1	**L10V,** I13V, **L24I, L33F**, 35QNins, M36L, **M46I, I54V**, K55R, R57K, I62V, L63P, I64V, **L76V, V82A**	**M41L,** K43N, E44D, **L74V, A98G, K103N,** V118I, D123E, I135L, S162C, **Y181C, M184V**, Q207E, **R211K, T215Y,** A272P, I293V, E297R, S322T, Q334D, F346Y, G359S, F389S, K390R, E399D, A400T, D460N, S468P, D471N, K476Q, H483R, L491, A554N
Patient 5	**L10V**, T12V, I13V, L63P, I64V	**D67N**, **K70R**, V90I, A98S, K102N, **Y181I,** S162Y, Q174H, **M184I,** T200A, E203K, Q207E, R211K, **F227L**, A272P, L283I, I293V, L295I, E297K, I329V, R356K, K366R, K390R, E399D, K431T, V467I, L491S, K527N, V531I, A554N

Figure 5.2: Alignment of the gag region contained in the pGEM clones derived from both patient virus samples 1 and 5. The shaded codons represent changes from pNL4.3-PSX

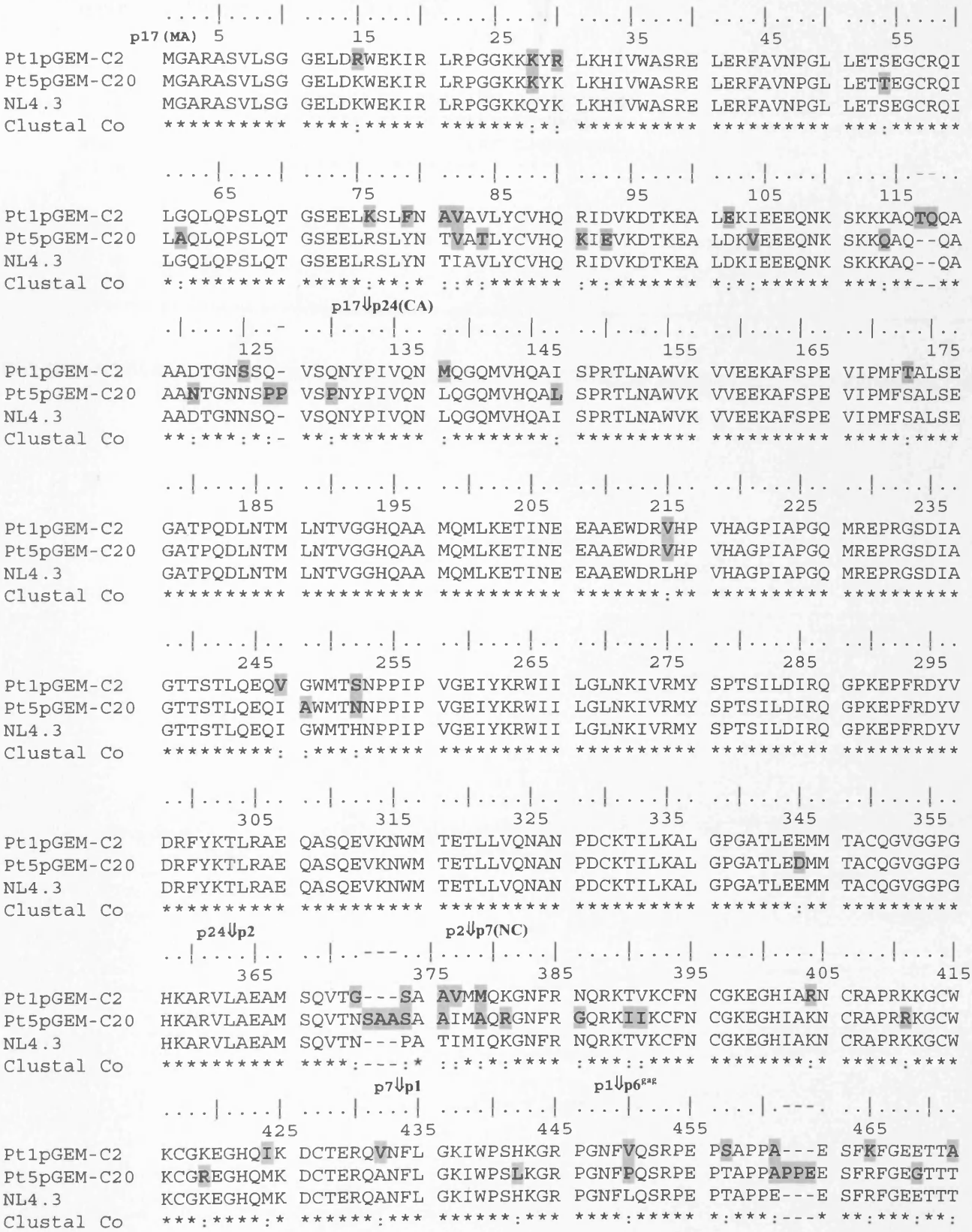


Figure 5.2(continued): Alignment of the gag region contained in the pGEM clones derived from both patient virus samples 1 and 5. The shaded codons represent changes from pNL4.3-PSX

```
                ..|....|.. ..|....|.. ...|....|
                 475        485         495
Pt1pGEM-C2      PPQKQEPIDK ELYPLASLKS LFGNDPSSQ
Pt5pGEM-C20     PSQKQEPIDK ELYPLASLRS LFGNDQ---
NL4.3           PSQKQEPIDK ELYPLASLRS LFGSDPSSQ
Clustal Co      *:******** ********:* ***:*:
```

5.2.2c Patient 1 Derived Virus Gag clones

As the template for this recombinant was the pGEM clone there was only one possible *gag* genome that could have been extracted, therefore only a single clone was constructed, which matched the sequence obtained from the pGEM clone Pt1pGEM-C2. The pGEM clone was used to amplify from, as this was previously shown to have a *gag* that was linked to the protease containing the two amino acid insertion.

5.2.2d Patient 1 Derived Virus Gag and protease clones

The Pt1-(Gag+Pro) clones were constructed from two sources, the first was through PCR from the cDNA and the second was by PCR from the pGEM clone Pt1pGEM-C2. Both gel purified amplicons were ligated into the digested pNL4.3X-PSX vector and a single clone containing the desired coding regions from each ligation was produced. Both clones were sequenced with an oligo that covered the p7/p1 and p1/p6 cleavage sites of *gag*, the whole of protease and up to codon 187 of RT, a fragment length of approximately 1kb. Sequence analysis revealed that both clones contained the A431V and L449V *gag* CSMs and were homogenous in protease with respect to amino acid sequence. The only non-synonymous changes seen between the two clones in this fragment were in the $p6^{Gag}$. The first occurred at codon 453 and was the result of a single base change. Clone 7 contained a Leucine and clone 27 a Proline. The second coding change, also attributed to a single base change was at codon 492. Clone 7 contained a Proline and clone 27 harboured a Leucine.
Two other non-coding changes were also seen between the two Gag+Pro clones; clone 7 contained the AAA at codon 475 in *gag*, while a AAG coded for Lysine in clone 27. The other synonymous change seen occurred in protease at codon 62, where a GTA and a GTG coding for Valine were detected in clones 7 and 27 respectively.

The production of the *gag*, protease RT products was hampered by the presence of existing *XbaI* restriction sites in the pGEM clones from both patients virus. Virus derived from patient 1 contained an *XbaI* site in the RT and virus from patient sample 5 had an *XbaI* site in *gag*. A number of approaches were used to try and generate this *gag*, protease and RT fragment (i) *XbaI* and *XmaI* restriction digest and subsequent ligation of multiple restriction fragments and (ii) production of a new pGEM *gag* to

RT clone of Pt1, in which the *XbaI* site was removed by SDM. (iii) cloning RT fragment of patient 1derived virus into the *gag* and protease clone.

5.2.3 Production of Virus stocks

Clones that contained the desired mutational patterns were used for production of recombinant viruses. The viruses that were generated contained the HIV Protease (termed patient-Pro), *gag* (termed patient-GAG), Protease and Gag (termed patient-Gag+Pro) and protease and RT (termed patient-Pro-RT) from patient derived virus samples in a wild-type HIV-1 molecular clone background.

5.2.3a MT-4 transfections

Large DNA stock preparations (maxi preps, Materials and Methods section 2.1.6) were made of the patient derived virus recombinant molecular clones. These DNA stocks were used to transfect MT-4 cells using DMRIE-C, as outlined in Materials and Methods section 2.4.3 and the resulting viruses were maintained and cultured to produce transfected virus stocks; see Materials and Methods section 2.4.

The transfected molecular clones were monitored for signs of virus production as determined by the presence of CPE in the infected cultures. The first patient virus recombinants to display CPE were the Pt5(PR) clones c23 and c25, after 6 days in culture. This was followed by the Pt1(P)-c8 clone, which showed signs of infection 7 days post-transfection. The Pt1-(PR) clones c6, c7 and c12 all displayed CPE after 10 days in culture. All the recombinant viruses that contained the *gag* region i.e. Pt1-c3(Gag) and the Pt1-(Gag+Pro) clones c7 and c27, required the longest culture times before CPE was visible by light microscopy. Several transfection attempts with the *gag* containing clones were unsuccessful at producing any indication that productive infection was occurring. These viruses required culture periods of 15-21 days before CPE was evident in the cultures.

In order to create high titre stocks, a 500ul aliquot of cell-free supernatant was used to infect $5x10^6$ MT4 cells, infections were cultured until full CPE was observed, see Materials and Methods section 2.4.

5.2.3b 293T transient transfections

Due to the prolonged culture times required for production of the recombinant stocks through transfection in MT-4 cells, an alternative approach was attempted. This involved transfection of the recombinant virus DNA stocks into the human embryonic kidney (HEK) 293T cell-line. This cell-line has been shown to produce high viral titres, due to the high transfection efficiencies [105]. Use of the transfection reagent Fugene (Roche) has been optimized for 293T cells, the details of these techniques are given in the Materials and Methods section 2.4.4. As before, a 500ul aliquot of cell-free supernatants generated from the 293T transfections was used to infect $5x10^6$ MT4 cells, these infections were cultured further until full CPE was observed.

5.2.4 Comparison of Transfection Methods

Viruses generated by both transfection methods were titred using the MTT assay system along with the re-infection stocks, termed 'sub' (sub-master). A comparison of the titres attained by transfecting the viral DNA in either MT-4 or 293T cells is shown in Figure 5.3. Three patient virus derived clones, which included the protease and RT regions were transfected along with the pNL4.3X-PSX and the parental molecular clone pNL4.3. The 293T transient transfections were the least productive method in terms of generating detectable virus. The initial transfection only permits a single round of virus production as the 293T cells lack the receptors to allow infection. The use of these 293T generated viral supernatants for the infection of MT-4 cells, which permitted multiple virus replication cycles, also failed to produce any measurable virus levels in the recombinants derived from the patient 1 virus. Creation of virus stocks with high enough titres that could be used in the assays, required the transfection and a single passage in MT-4 cells to create the patient 1-sub virus stocks.

Both molecular clones pNL4.3X-PSX and pNL4.3, produced high titres when used to transfect either MT-4 or 293T cells and the infection of MT-4 cells to produce the sub stocks, produced the highest recorded titres. The virus clone derived from patient 5 produced lower $TCID_{50}$/ml compared to the wild-type molecular clone produced by the same method.

The highest titres in all the viruses was achieved by MT-4 sub infections, with this method producing 2-fold more virus than through the 293T and MT-4 infection in all the recombinants tested.

5.2.5 Titrations and Antiviral Assays in MT-4

The amount of infectious virus present in the prepared MT-4 virus stock was determined using the MTT Titration Assay as described in Materials and Methods section 2.5.1. A table of $TCID_{50}$/ml values derived from the titration assays is shown in Table 5.2 and examples of the titration curves used to calculate the titres are shown in Figure 5.4.

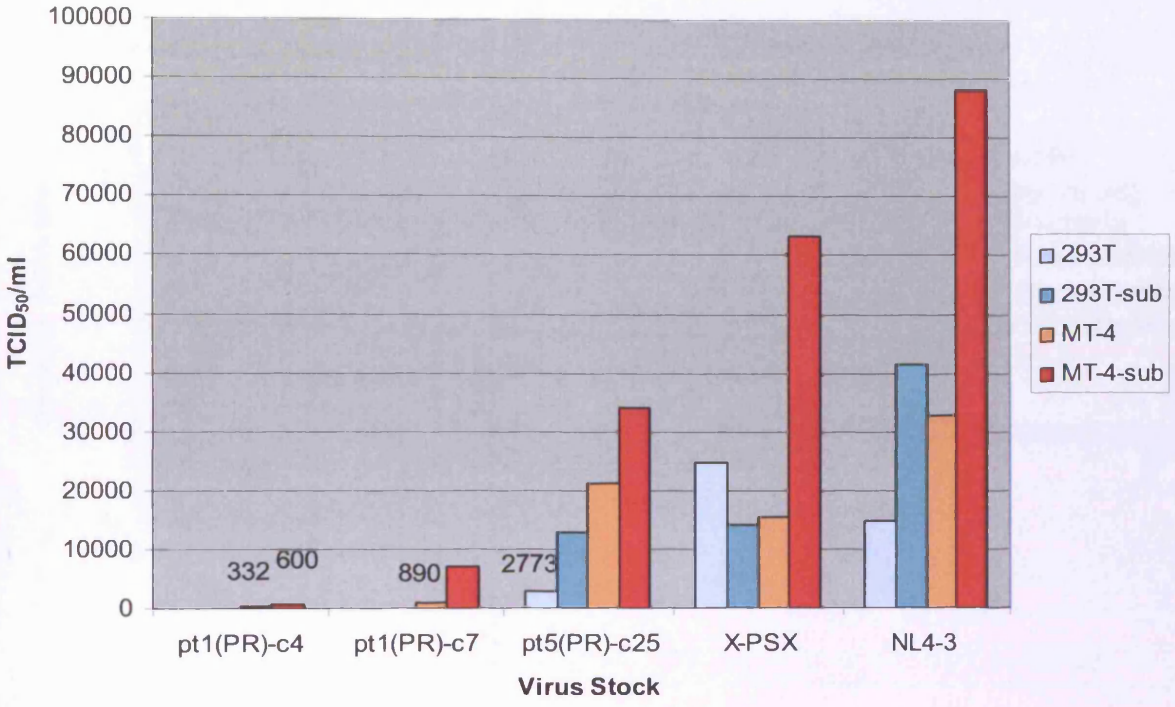


Figure 5.3: A comparison of the titres obtained by transfecting the viral DNA in either MT-4 or 293T cells.

Figure 5.4: MTT Assay derived titration curves, error bars represent the standard error from at least two independent experiments.

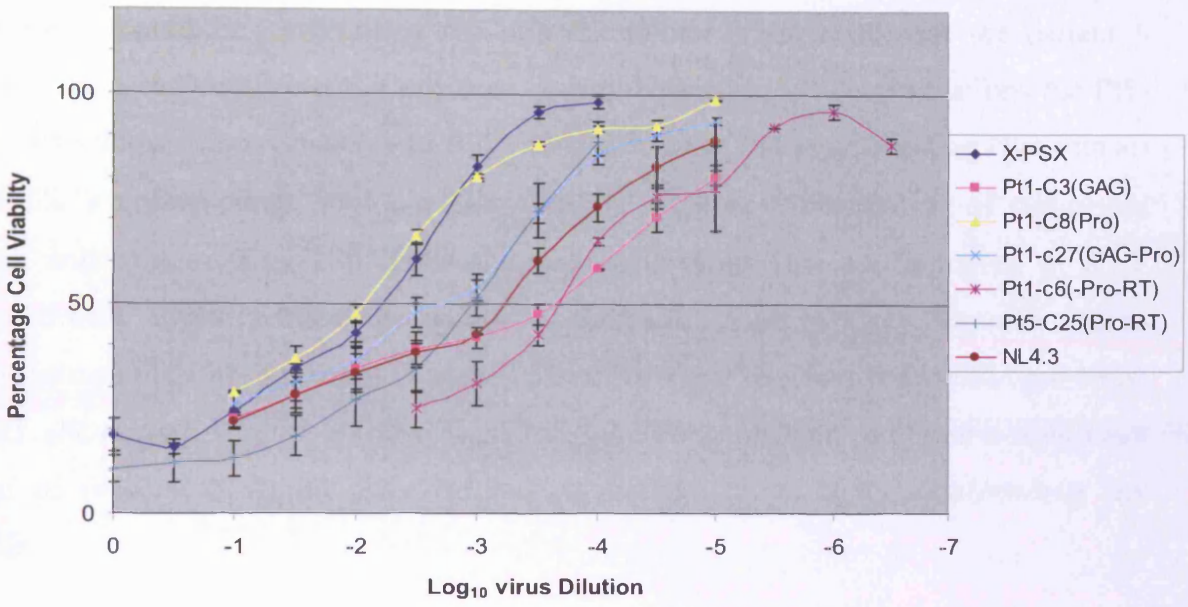


Table 5.2: $TCID_{50}$/ml values derived from the titration curves used to calculate the titres of the recombinant virus clones shown in Figure 5.4.

Virus Stock	$TCID_{50}$/ml
X-PSX	3.20×10^3
Pt1-C3(GAG)	8.27×10^4
Pt1(P)-C8	2.34×10^3
Pt1-c27(GAG+Pro)	9.43×10^3
Pt1(PR)-c6	9.23×10^4
Pt5(PR)-C25	1.80×10^4
NL4.3	3.23×10^4

The $TCID_{50}$/ml values were used to determine the input for the MTT antiviral assay which was used to measure the sensitivity of the recombinant viruses to a panel antiretrovirals. Figure 5.5 shows the comparison of IC_{50}'s obtained for the recombinant viruses. In most cases a minimum of two assays each with four replicates per assay was used to obtain the IC_{50}. The results for the PI RTV are not shown as no IC_{50} value could be determined for this compound when tested for the patient 1 derived viruses containing the protease coding region. In all determinations the Pt1-Gag virus showed no resistance to RTV and both the Pt1-(P) and Pt1-Gag+Pro viruses had IC_{50}'s greater than 20μM, which was the highest concentration of compound tested and greater than 200 times the wild-type IC_{50}. This concentration of RTV showed only slight levels of protection in the MTT viability assay. The aim was to demonstrate that these viruses were significantly more resistant than wild-type virus, which showed an IC_{50} of 100nM. Increasing the concentration to obtain a significant level of protection would have resulted in cellular cytotoxicity, confounding any result.

Figure 5.6 shows the fold changes in IC_{50}'s for the PIs tested of the patient 1 recombinant virus clones compared to NL4.3, the X-PSX virus clone is also included in the comparison.

Both figures show that recombinants which contained the protease derived from patient 1 isolated virus were the only resistant viruses to the PIs tested, an additional increase in IC_{50} was seen with the inclusion of the *gag* region from the patient 1 derived virus. The Pt1-GAG alone virus showed only a slight increase in IC_{50} against SQV and IDV. This approximately 2-fold increase; 1.90 and 2.03 respectively, was less than seen in the X-PSX molecular clone virus for the same PIs; 1.35 and 1.32 respectively, so not considered significantly different.

Pt5(PR)-C25 shows wild-type sensitivity to all the PIs used in the antiviral assays, it was also tested with AZT and was shown to be as susceptible as wild-type as expected as it contains no PI or AZT resistance mutations.

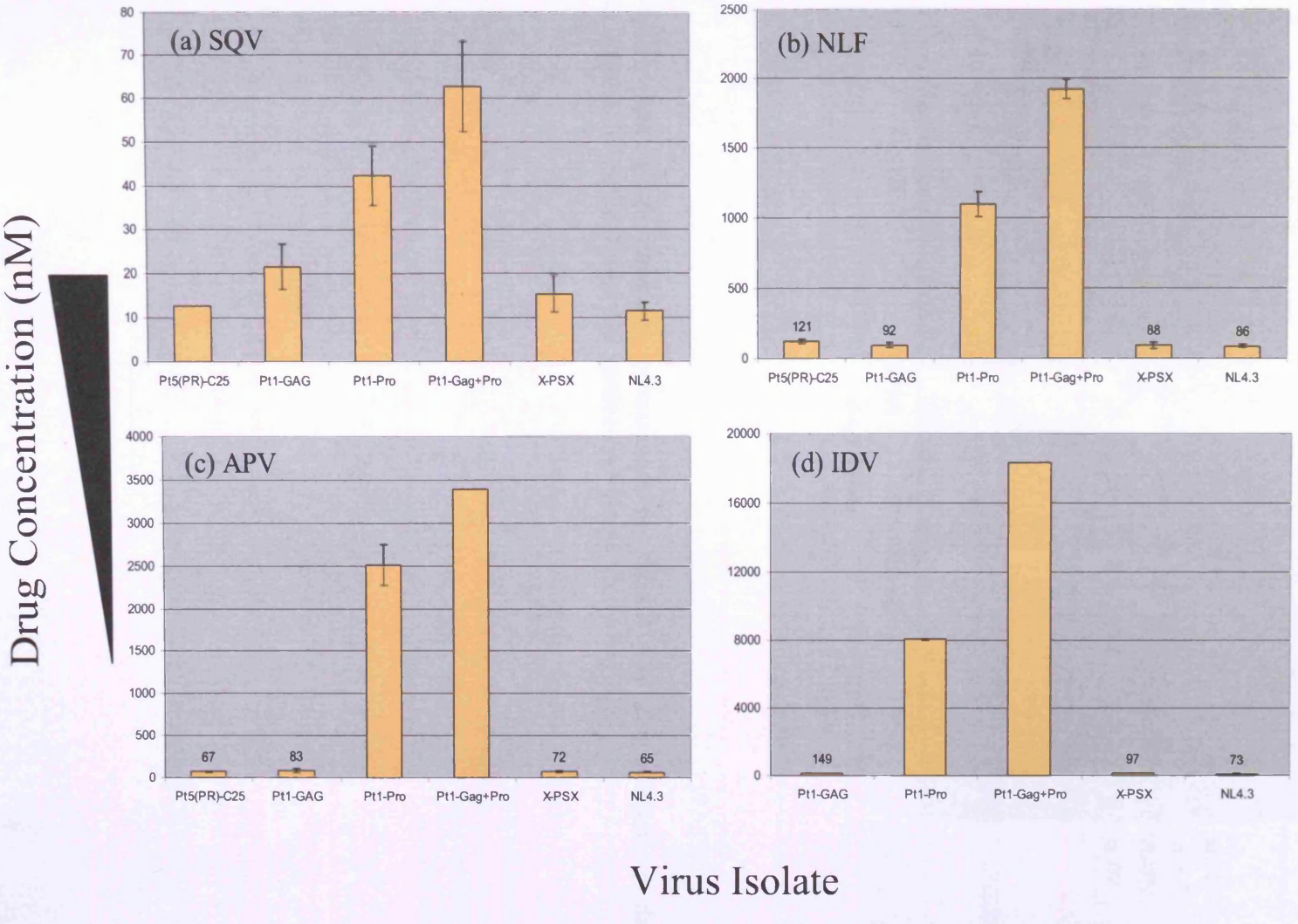


Figure 5.5: Comparison of IC50 values obtained for the recombinant viruses using the MTT Assay system.

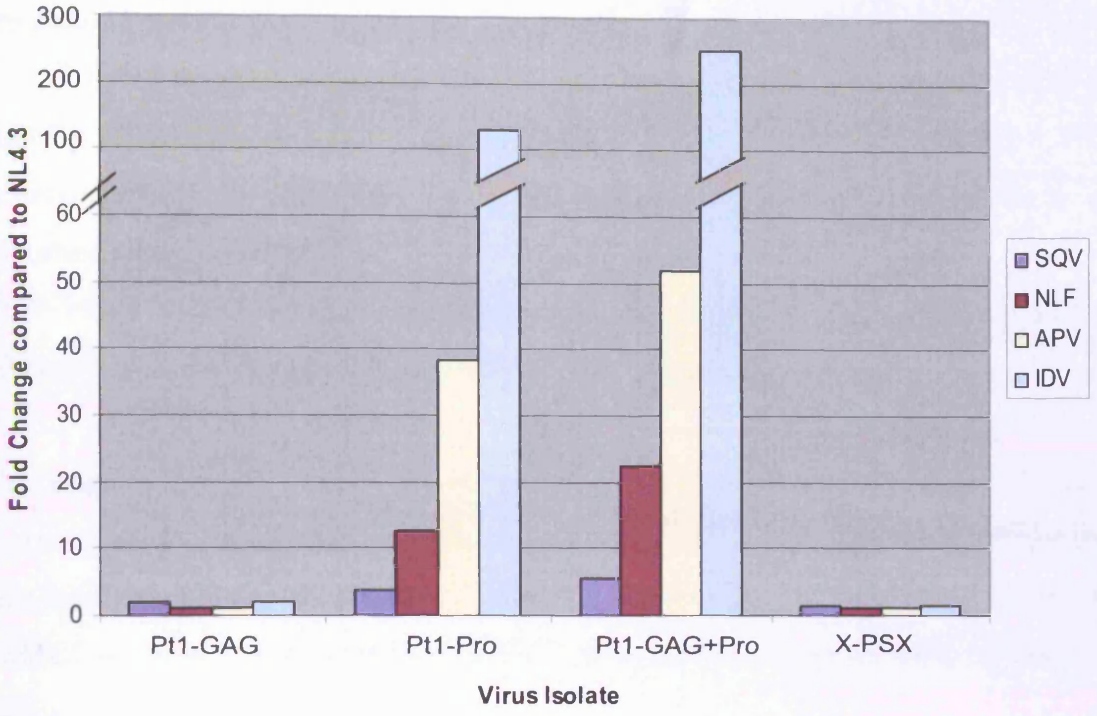


Figure 5.6: Fold changes compared to NL4.3 for the Patient 1 derived recombinant viruses and the molecular clone virus X-PSX.

5.3 Discussion

The prolonged use of antiretrovirals results in the development of resistance mutations to the protein targeted by the compounds. It has been shown that resistance mutations to PIs accumulate in the protease and its substrate, the *gag* polyprotein [12,77,191,358]. Most genotypic and phenotypic monitoring of patients with HIV is carried out by analysis of the two main targeted regions, protease and the RT [292,347]. These interpretations are then used to decide which therapies are most likely to offer the greatest clinical benefit.

The exact nature of the contribution of *gag* mutations that arise during antiviral therapy on the virus phenotype still remains essentially unknown.

In order to demonstrate the utility of the live virus vector, a series of recombinant viruses were constructed, in which all the restriction sites introduced at the termini of the coding regions were utilized. This resulted in the production of chimeric clones which contained *gag*, protease and RT derived from patient virus isolates receiving antiretroviral therapy.

As previously mentioned in Chapter 3, the manipulation of a large DNA vector created several problems during cloning, one of which was the removal of any vector DNA that had not been completely digested before addition into the ligation reaction. To prevent the contamination of digested vector DNA with uncut/partially cut DNA, the restriction digests were resolved on low percentage agarose gels and subjected to very low voltage electrophoresis for in excess of two days, as to prevent any uncut vector which existed in these gel purified DNA preparations would also be transformed into the host *E.coli*. As the fragment from the source sample replaces a fragment of the same size from the original, the presence of wild-type contamination could only be detected after sequence analysis revealed its' existence.

Another problem that occurred was the production of chimeric clones containing the *gag*, protease and RT coding regions. This was hampered in both the patient virus samples chosen, as both contained native *XbaI* restriction sites. In patient 1 derived virus the intrinsic *XbaI* site was located in RT and in patient 5 derived virus, it was located in the *gag* region. Prior to the selection of *XbaI* as the *gag*, N-terminal cloning site, an analysis of several databases was carried to determine the frequency of its

occurrence. The use of this restriction enzyme was also based on its selection by another group which produced a similar cassette system [19], although in this case the *XbaI* site was used at the 5' junction of protease. The *XbaI* site was also used in the construction of the molecular clone pHXB2D; HIV sequence containing segment of this vector is inserted into the backbone vector using *XbaI* sites [91]. The *Xba I* site was only a problem if the gag cloning site were to be utilised on the fragment being cloned in. This only occurred with the Gag to RT fragment from patient 1 and any Gag containing fragment from patient 5. The *Xba I* site was chosen as it was thought to be easier to clone with compared to the original choice of the intrinsic *Nar I* site and a suitable site for its silent insertion was available which also allowed the inclusion of the gag packaging loops. If this vector system were to be used for future high throughput use then it would require modification, as there would likely be a problem with viruses that contained an *Xba I* site in the Gag-Pol coding region; see Table 5.3. Modification of this vector would require a more extensive survey of available sites either to incorporate a less frequently occurring restriction site or to add one further downstream of the packing loops possibly in the 5'LTR.

Table 5.3 The proportion of HIV sequences in the Los Alamos HIV database that contain *Xba I* restriction site in the Gag-Pol region.

Sub-type	Total number	Number containing *Xba I* restriction site	Percentage of total with *Xba I*
CRF's	324	251	77.47
A	80	65	81.25
B	464	131	28.23
C	415	351	84.58
D	21	9	42.86
F	15	12	80.00
G	18	16	88.89
H	3	3	100.00
J	2	2	100.00
K	2	2	100.00
N	5	5	100.00
O	22	18	81.82
U	7	6	85.71
X	13	2	15.38
Total	**1391**	**873**	**62.76**

As mentioned in the results section of this chapter, in order to create the Gag-RT clone, the pGEM clone Pt1pGEM-C2 was mutated to remove the *XbaI* site, to then allow it to be PCR amplified, digested and cloned in to the *XbaI* and *XmaI* cut vector. This step was also hampered by incomplete digestion of the pNL4.3X-PSX vector, using the *XbaI* and *XmaI* restriction enzymes, resulting in partially cut vector contamination as mentioned above. The production of this patient virus derived Gag-RT clone remained unfinished due to time restrictions. However, as all the clones constructed required the use of all the restriction sites at some point, the need to demonstrate any further utilisation was thought less important than continued analysis of the viruses so far generated.

In order to generate stocks of sufficient titre, different methods of transfection were attempted. The highest titres in all the viruses was achieved through MT-4 transfection and culturing followed by further infection using the harvested cell-free supernatant i.e. the MT-4 sub infections. However, this method requires several days in culture before any visible signs of infection are seen, even with a wild-type molecular clone. The alternative method attempted was transient transfection of the patient-derived virus vector constructs into HEK 293T cells, which would then produce virus into the culture medium after 2-3 days. This method was successful for the molecular clones and also for the effectively wild-type virus clone derived from patient 5 isolated virus. The resistant viral constructs derived from patient 1 virus all produced undetectable levels of virus through this transfection method, which would suggest that the differences in virus production are attributable to the protease and RT sequences. Both these enzymes in the patient 1 viruses were heavily mutated, the RT contains several mutations thought to be non-synergistic , such as the AZT resistance mutations M41L and T215Y, along with the NNRTI mutation Y181C [291] and M184V [23], which partially reverse AZT resistance and also are detrimental to RT function [110]. The deficient virus production could also be an effect of the mutant protease. The protease from patient 5 virus contains far fewer mutations which are thought to be detrimental to enzyme function, than are present in virus derived from patient 1 [99,224]. Mutations detrimental to enzyme function are more likely to require additional compensatory changes to restore replication competence [311], thus, *gag* polyprotein would require additional mutations to compensate for the less efficient protease. In Chapter 3, the only isolate to provide a pseudotyped virus was that obtained from

patient 5 derived virus, which also implies it is relatively growth unimpaired. Although the *gag* from patient 1 derived virus failed to amplify with the primers required for construction of the single cycle recombinant, so a direct comparison is not possible.

When transfected into MT-4 cells the two viral clones from patient 1 (Pro-RT) clones c4 and c7 produced different $TCID_{50}$/ml even though they have identical sequences in the protease. Therefore, it would be expected that these viruses would produce similar levels of infectious virus when transfected into the same cells. The differences that were seen in $TCID_{50}$ could reflect the inherent errors in such live virus and cell culture work i.e. the lower $TCID_{50}$/ml detected in the virus stock after the Pt1-c4 transfection could be due to a lower transfection efficiency as a result of experimental variation. This stock was then used to infect more MT-4 cells; unlike the first stocks that were generated by transfection with equal DNA quantities, this infection is equalized by supernatant volume. The differences in the $TCID_{50}$/ml of the input supernatants is then extrapolated further by the multiple rounds of infection and re-infection that occur during the days in culture prior to harvesting. The opposite effect is seen in the X-PSX stock, which produced almost 2-fold higher $TCID_{50}$/ml in the transfection than was seen after the 293T derived virus was used to infect MT-4 cells. This is most likely due to there being too much virus present in the initial inoculum used to infect the MT-4 cells, which could have caused massive cell death and in turn loss of virus titre as there are no cells left to infect.

As expected the patient viruses containing the protease showed resistance to PIs. The highest fold changes were seen in the PIs the patient virus had most exposure to. Pt1-GAG possibly contributes some resistance to SQV and IDV, although less than 2-fold, but it appears to be consistent. However, a similar, increase in fold resistance profile is seen for the X-PSX vector with IDV (Figure 5.6), although the actual difference in IC_{50}'s compared to NL4.3 is much lower than for the Pt1-GAG virus.

The most resistant viruses to the PIs were those that contained the protease isolated from patient 1 derived virus, except for the Pro-RT clones. The lack of any resistance to PIs and RTIs in the Pt1(Pro-RT) clones infers that the genotype does not match the phenotype (Data not shown)e.

To determine if the genotype had reverted, clonal analysis of the RT region amplified from the infected cell pellet was carried out which showed the majority (7/8) clones were carrying the drug resistant mutation at codon 215. This codon position was chosen to determine if reversion had occurred in the RT of the recombinants, as it has been previously demonstrated that this mutation is unstable when transmitted to drug naïve seroconverters [18]. The transfection of a clonal population could be seen as representative of an *in vitro* transmission event, which is also thought to involve a homogeneous population [71,183]. The exact reason for the lack of phenotypic resistance in the Pt-1(Pro-RT) clones is not clearly understood. The Pt-5(Pro-RT) which was cloned and transfected at the same time shows wild-type sensitivity to the antivirals tested, namely, SQV, APV, NFV and AZT. However this should be expected from the sequence data obtained for this sample. Although it would have been expected to show high level resistance to 3TC due to the presence of the M184I mutation, unfortunately no stocks of 3TC were available at the time these assays were carried out.

The sequences derived from the pGEM clones was subjected to the Stanford Genotypic Resistance Interpretation Algorithm, Version 4.2.6 for a predicted phenotype interpretation. As expected the patient 5 virus scored susceptible to all the PIs and showed high level resistance to the NNRTI's DLV and NVP, due the mutation at Y181I and also to 3TC, due to M184V mutation. This patients virus had prior exposure to the 3TC and NVP, therefore the prior treatment explains the genotype and the predicted phenotype. The virus derived from patient 1 scored high level resistance to NFV, IDV, APV, primarily due to the M46I, I54V and V82A mutations. This was also detected in the phenotypic evaluation of the protease containing viruses. The predicted and actual phenotype also correlated with regard to Saquinavir, with low level resistance detected and predicted. The interpretation algorithm also predicted that this virus isolate would show low level resistance only to Tenofovir, due to the absence of K65R, leaving this and SQV as the sole treatment options for this patient.

To evaluate the level of heterogeneity within the patient 1 viral RNA sample, several different clones which contained the protease and RT regions were produced. These clones contained almost identical protease regions, with only one clone from four

displaying a single amino acid variation at codon 20 (K20R). The production of the protease only containing clones failed to generate a single clone with the K20R mutation, from the six clones which were produced by direct PCR from the heterogeneous cDNA, as were the protease and RT clones. This K20R variant may be either required, in which case the virus is evolving to acquire it, or it could be in the process of being discarded. The pGEM clone which contained the Gag-RT region, also derived by the amplification from the cDNA, contained the resistant protease and RT, but not the K20R, further supporting its' low level proportion within the virus quasispecies. The requirement for the K20R could be established by sequencing of the Pt1-(Pro-RT) virus populations, from both the c6 and c7 passage stocks. It would seem that the patient 1 protease only virus does not require the K20R as this was not detected by population sequencing of the infected passage cell pellet. This may suggest that at a low level, the protease containing the K20R, which is thought to be a compensatory mutation [40,154,230], is or was linked to the mutant RT. Bleiber *et al.* noted that the cloning of the cognate RT improved the replication capacity of some protease mutant clones [19].

The homogeneity within the protease region overall would suggest that the major genotype within the sample population contains the highly resistant protease, which also includes a double amino acid insertion at codon position 35. The presence of insertions in the protease coding sequence have previously been reported [24,156,309,314,341]. These insertions are generally duplications of neighbouring sequences caused by slippage of the RT during reverse transcription and have been shown to confer only a modest growth advantage compared to the viruses in which the insertion had been deleted [156]. The growth abilities of the viruses generated is the subject of the next chapter.

The genotypes generated from the patient 2 and 5 viral cDNA used for the cloning into the PSX vector shows almost identical sequence data to that shown in Chapter 3 for the same samples, across the protease and in the case of the virus derived from patient sample 5, also the *gag* region. These results confirm that the PCR amplification selects for the majority genotype. To test that the amplified viral product is representative of the viral population could be done by comparison of sequence data of the whole protease region from that obtained by routine resistance

testing (see Table 3.1) and that of the clones identified in Figure 5.1a and b. This shows only 3 changes with respect to the patient 1 derived virus at L10I, L33Fand L63P from a total of 13. The insertion was also identified in all cases. The pGEM clone sequence (see Table 5.1) for this sample was also identical to the protease clones identified in Figure 5.1a and b. The sequence for the Patient 5 derived virus clones identified in Figure 5.1a, was also almost identical to the routine resistance test sequence in table 3.1, only changes at V3I and S37N were absent from any of the protease clones in Figure 5.1a and in table 5.1.

The RT showed more variation when comparing the sequence obtained by routine resistance testing, and that of the pGEM clone in both patient 1 and 5 derived virus clones. This is primarily due to the larger RT fragment sequenced; 333 and 99 amino acids for RT and protease respectively. Routine resistance testing is also population based and unlike clonal analysis, shows no linkage of mutations. Patient one RT sequence in table 3.1 has 8 additional changes from that of the pGEM clone in table 5.1, from a total of 21 mutations. The variation between the two sequences in RT may also reflect the extensive previous NRTI and NNRTI treatment history, in all this patient had been exposed to 7 different RT inhibitors. The pGEM clone contains more NNRTI resistance associated mutations than the sequence identified in resistance testing and the L74V may be in response to ddI [306].

Patient 5 RT sequence in table 5.1 has 3 additional changes from that of the pGEM clone in table 3.1, from a total of 20 mutations. No further sequence analysis of clones was carried out with RT as this was not primarily involved in the study.
The increased diversity in the RT region could also be attributable to PCR recombination, this could occur in both PCR reactions and during the production of high titre stocks, as the PCR was only optimised for production of good yield amplicons and not to minimise possible recombination events. As all infections originated from a clonal stock any recombination that occurred would represent the out growth of a dominant species.

An alternative, but more labour intensive approach would be limiting dilution PCR, where multiple PCR reactions are set up so that only 30% of the total reactions produce an amplicon. This technique is more likely to result in the production of an

amplicon that is derived from a single template molecule and so is less likely to suffer from recombination.

Real-time PCR may be utilised for the detection of minority species of HIV, sensitivity of such assays are approximately 1% in some studies and quantification of the total viral population was possible as low as 4% [74]. Although another group has shown even greater sensitivity, detecting proportions as low as 0.1 to 21% of total virus population [130]. However this technique does not isolate a product that can be utilised to produce a virus. The only patient virus derived material available was the viral RNA, had PBMCs been available the patient cells could be cultured with stimulated donor cells to derive a population of virus from the infected patient cells. Although, co-culturing may lead to the selection of a minority population that grows well in co-culture in the absence of drug and may not be representative of the starting population. The quasispecies from this population would represent virus produced from integrated proviral DNA in the infected cells. This population may exhibit a different heterogeneity to that derived from plasma viral RNA. The latter is more likely to display less heterogeneity due to the oligo-specific selection during cDNA production i.e. only sequences which hybridise to the oligo used for cDNA production will be selected. To minimise this effect, the oligo used for cDNA synthesis should be located in a highly conserved region.

The interplay between resistance and compensatory mutations is an incredibly complex one, it would require the understanding of the association of each mutation with its effect on several phenotypic markers, such as resistance and replication capacity. What is clear is that the exclusion of the protease substrate *gag*, from phenotypic evaluations of resistance, leads to a underestimation of the level of drug sensitivity, especially to PIs and possibly also to NRTIs [256]. This report demonstrates that inclusion of the *gag* region from a highly PI experienced individual can result in an approximate 2-fold increase in IC_{50}'s.

With more time available this work would have been completed by the production of more patient derived viruses, such as from patient sample 1, the gag, protease and RT clone would have created a 'full-set' of clones from this patient sample. Other patient samples were also selected and viral cDNA from these were amplified, but as

mentioned in the discussion earlier, as all the clones constructed required the use of all the restriction sites, any further utilisation was thought less important than the continued analysis of the viruses generated so far. One further avenue that could also be explored is a greater dissection of the regions of gag involved in PI resistance or compensation of resistance .i.e. are both N and C terminal regions required to bring about an increase in resistance to PIs?

Infected individuals from countries with a healthcare system that provides virological monitoring and access to several classes of antiretrovirals allowing a change of any failing compounds in the regimen, are more likely to be able to control virus load and resistance, which in turn would have an effect on transmission rates [175,324,348]. However, in those countries where limited classes of antivirals are available and little or no virological monitoring exists, such as the poorer and more AIDS burdened countries, the re-cycling of same class compounds is more likely to occur. The standard first line therapy in developing countries consists of two NRTI's plus an NNRTI, or three NRTI's. Second line therapy is likely to be the PIs with at least one NRTI [122]. Research data suggests that RT genotypic diversity in HIV clade C may influence the emergence of drug resistance, especially to NNRTIs, NVP, DLV and EFV and also to the NRTI AZT [182]. The effects of PIs may also be abrogated by naturally occurring mutations in subtype A and C backgrounds, which may exhibit a higher biochemical fitness in the presence of commonly prescribed PIs [326]. Therefore, in order to successfully control viremia and resistance with such a limited supply of antiretrovirals in these developing countries, may require a greater understanding of the complexities of resistance and the associated effects on viral fitness.

The next chapter aims to study the effects of the *gag* region on the growth and replication ability of the recombinant viruses in the context of wild-type and mutant proteases.

Chapter 6 Growth characteristics of Pt1-derived virus constructs

6.1 Introduction

The emergence of highly resistant viruses in individuals on HAART has, in the majority of patients, resulted in the appearance of a less pathogenic strain, which achieves lower viral loads and causes less depletion of CD4+ T-cells [117]. The selection of a particular genotype can result in the fixation of certain mutations [322,327], resulting in the selection of an evolutionary pathway, which in a previous environment, such as under a particular drug selection pressure, produced the most fit species. However, a change in environment, such as the introduction of a new antiviral into a regimen, may require the addition of new mutations for successful viral adaptation. Primary to virus growth, is the requirement for replication, which in this setting necessitates the acquisition of resistance mutations. These new resistance mutations must also maintain functionality of the targeted protein and must also provide synergy with other existing resistance mutations. The first publication relating to this possible lack of synergy was by Chow *et al.* [38] using the combinations of AZT, ddI and NVP. Due to the perceived potential of this regimen to cause HIV to mutate in a stepwise manner to an unviable state, it was termed, convergent combination therapy [278]. Although now it is apparent that the myriad of genomes which make up the quasispecies, especially within a pre-treated individual, are likely to negate the effectiveness of such an approach.

A main factor to be determined is how to measure viral fitness and in what context i.e. which cell line; macrophages vs. T-cells or PBMC's, which virus source; primary isolates vs. cell adapted or recombinant clones. Many caveats exist to each method and a clear consensus on how to determine fitness of different viruses is yet to be achieved.

An understanding of how to use virus fitness information would have several practical applications in designing current regimens, such as, the control of HIV with a reduced number of compounds. This would lead to less toxic and complicated regimens, costing less to maintain. This may also facilitate the design of lifelong antiviral

therapies for HIV infected individuals, in poorer countries. Several studies have compared a single boosted PI with standard triple HAART in patients as a maintenance therapy and found no significant difference in the two treatment arms [108,214].

Previously, a virus had been isolated from a highly antiretroviral experienced individual, which contained multiple resistance mutations in protease and RT coding regions. This virus also contained a novel two amino acid insertion in protease and also displayed mutations in *gag* associated with PI treatment. To investigate the potential effects of the mutated protein coding region on virus growth a series of recombinants were made. The protease containing virus clones demonstrated high levels of resistance to those PIs which the patient had previously been exposed to. The inclusion of the resistant virus *gag*, increased PI resistance, although, *gag* alone afforded no decrease in sensitivity to any of the PIs tested.

The overall aim of this chapter was to determine if the inclusion of the *gag* region results in the modulation of growth characteristics of the recombinant viruses and to see if this was compensated for by the inclusion of the cognate protease. As this vector system includes the region involved in packaging of the viral RNA in new progeny virions, the effects of mutations in this region are also investigated. This region is not included in any recombinant virus systems that have previously been developed to explore the effects of mutations associated with PI therapy and so would provide the first evidence that this region is also under selective pressure from PIs.
The specific aims of this chapter were to:
(a) evaluate different methods to determine replication capacity, this was carried out by

(i) Comparison of growth rates of the recombinant viruses

(ii) Dual Infections with a wild-type virus

(b) determine the difference in reversion of *gag* and protease sequences after culturing in the absence or presence of different antivirals.

6.2Results

6.2.1 Comparison of Virus Titration Methods

To determine which assay system could best be used for the comparison of growth abilities, two infections using MTT titrated stocks were set-up. This involved the growth of a wild-type molecular clone, HXB2 and a Saquinavir resistant double mutant, G48V+L90M. Both infections were set-up at an equivalent MOI of 0.0001, determined by MTT titration, which is the same input used for the antiviral assay determinations. Stock infections were set-up and then aliquoted in equal volumes in a 6-well plate. A single sample time point was collected each day and the cell-free supernatant stored at -80°C, until all time points had been collected. Analysis of all the time points was carried out using either MTT titration or GHOST GFP reporter cell assay, as described in Materials and Methods 2.5. All the time points were titred simultaneously, to reduce any possible inter-assay variation, such as different cell growth rates.

The results obtained for plotting the $TCID_{50}$/ml, as determined by the MTT assay and the number of infectious units per ml, obtained from the GHOST assay, over the period of the growth curve are shown in Figure 6.1. Both methods demonstrate a similar pattern for the level of virus detection across the time points, even though the end-point detection is different in both assay systems. However, the suggested fold difference between the two viruses after six days is greater in the MTT assay than detected by the GHOST assay. A 5-fold difference in $TCID_{50}$/ml was seen by using the MTT assay and a 1.6 fold difference using the GHOST assay.

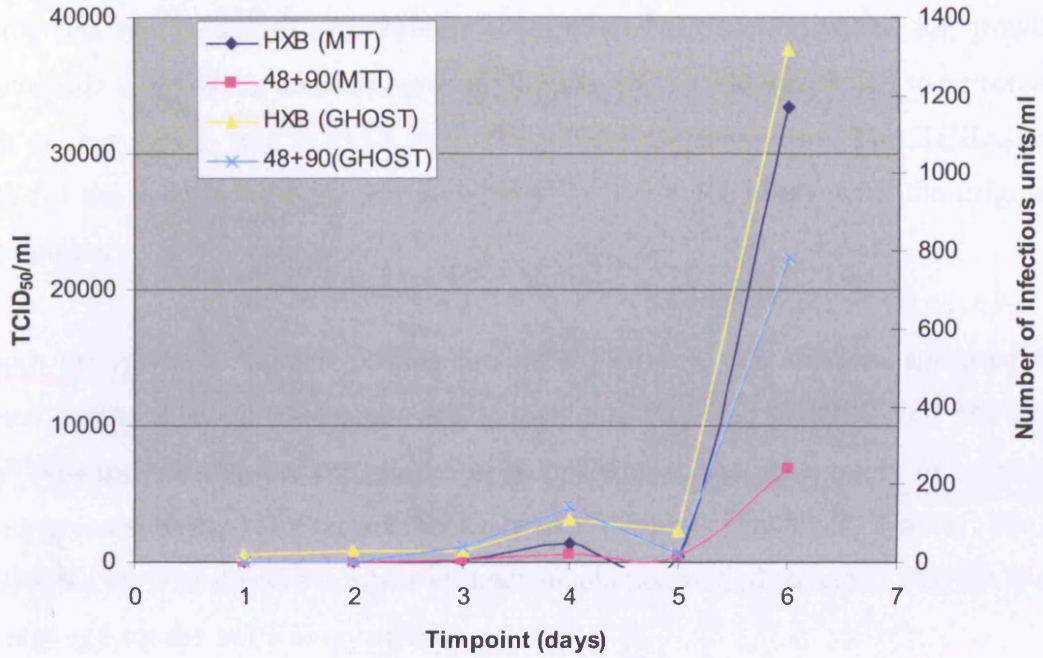


Figure 6.1: Comparison of growth rates of two virus isolates, HXB and Saquinavir double mutant (G48V+L90M), using MTT and GHOST assay systems.

6.2.2 Growth Comparison of Patient 1 Derived Viruses

6.2.2a Titration of Recombinant Viral stocks

As only one representative recombinant virus clone was to be utilised for growth characteristic analysis the nomenclature of the virus stocks was simplified to represent which coding region was derived from the patient isolated virus. The $TCID_{50}$/ml values for the recombinant viruses are shown in Table 5.2 along with the original clone number.

As both the methods for determining amount of virus in the samples appeared to produce similar values, it was decided to continue with the GHOST cell reporter assay. This method had the advantages of quantifying the actual number of infected cells as apposed to the more indirect measure of cell viability in the MTT assay. Also, the GHOST method involved only a 48 hour incubation period, compared to the five days required for the MTT assay system.
The titrations are indicated in Table 6.1

Table 6.1 Titres of Patient 1 derived Recombinant viruses determined from GHOST GFP indicator cell titration.

Virus Stock	Infectious units/ml
XPSX	1.92E+05
Pt1-GAG	3.18E+05
Pt1-Pro	1.56E+05
Pt1-GAG+Pro	1.46E+05
Pt1-Pro+RT	5.61E+05
NL4.3	1.39E+05

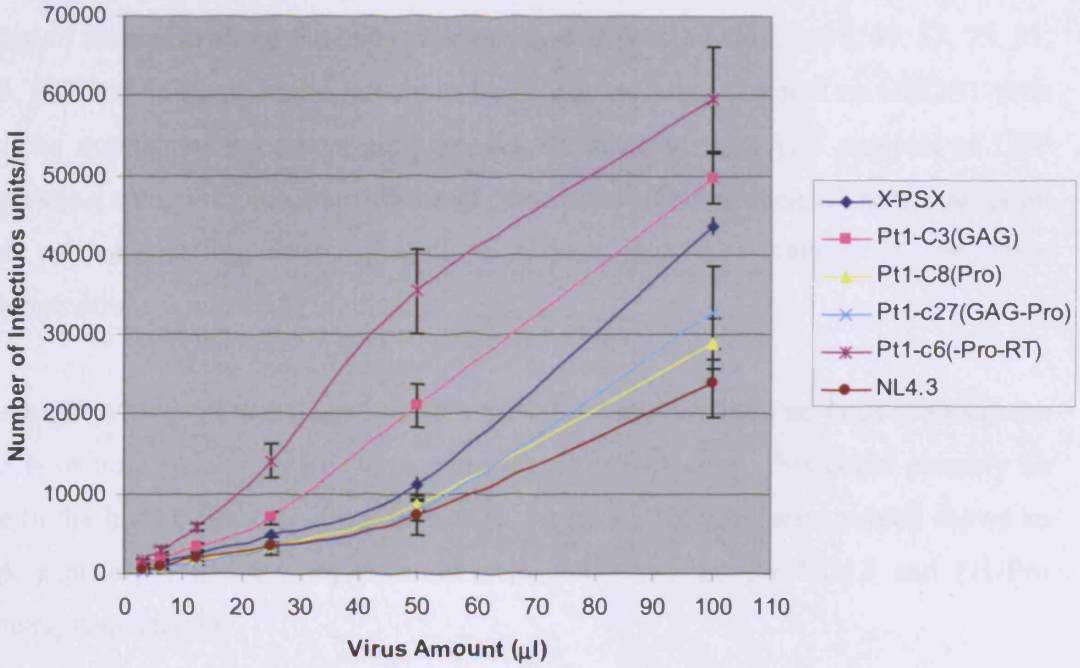


Figure 6.2: Comparison of titres as determined by GHOST assays, error bars indicate standard error over two independent assays.

6.2.2b Comparison of Recombinant Viruses Growth Rates

To asses the rate of growth of the recombinant viruses, each of the virus stocks, previously titred on GHOST cells, was used to infect 5.25 x10^6 SupT1 cells. The amount of cell-free virus supernatant used in the initial infection was that amount which produced 10% infection in the GHOST assay titrations. To minimise any variation in each infection across the time points, the stock infection was aliquoted across 12 wells of a 24-well plate, each well containing 3.75 x 10^5 SupT1 cells in 500µl of media. In all 13 time points were taken at 0, 6, 12, 24, 30, 36, 49, 57, 73, 81, 103, 127 and 150 hours post infection. Each time point was titrated on GHOST cells and the number of infectious units per ml, as detected by FACS analysis of GFP expressing cells, was calculated for each time point. The titration of each time point was carried out on three separate occasions and the average of the three determinations is shown in Figure 6.3.

The highest titres were recorded in both the NL4.3 and the Pt1-Pro virus stocks at the 103 hour time point, which is approximately after four days. This could possibly be due to the higher input, as demonstrated at the zero hour time point, which shows an approximately 1 and 0.5 $\log_{10}$ variation in start titres for the NL4.3 and Pt1-Pro viruses, respectively.

All the viruses showed a continued reduction from the calculated input titres over the first 24 hour period, followed by a gradual increase in detected infection in all but two of the virus stocks. The slowest increase in the number of GFP positive infected cells was observed for both of the viruses derived from patient 1 *gag* containing viruses, with all other viruses demonstrating a similar increase in rate of infection. This slow rate of reproduction for the viruses containing the patient derived *gag* sequences would suggest that these viruses are growth impaired and that this phenotypic trait is related to the *gag* region.

The Pt1-(Pro-RT) virus produced the lowest number of infected cells at peak infection, which occurred approximately 24 hours before the other viruses with similar replication rates.

As the input amount in the growth curve varied and would affect the amount of virus produced, another method of analysis is shown in Table 6.2. This shows a comparison of the amount of virus produced at peak titre over the course of the growth curve for each virus after correction for the start input. Using this method of determination suggests that the X-PSX viral construct replicates best, this is followed by the protease only virus, then by the NL4.3 virus stock. The remaining three virus constructs are ordered Pt1-GAG virus, Pt1-Gag+Pro and finally the virus that replicated most poorly, the protease and RT virus. Using this method of analysis also suggests that there is very little difference between the top 2 viruses.

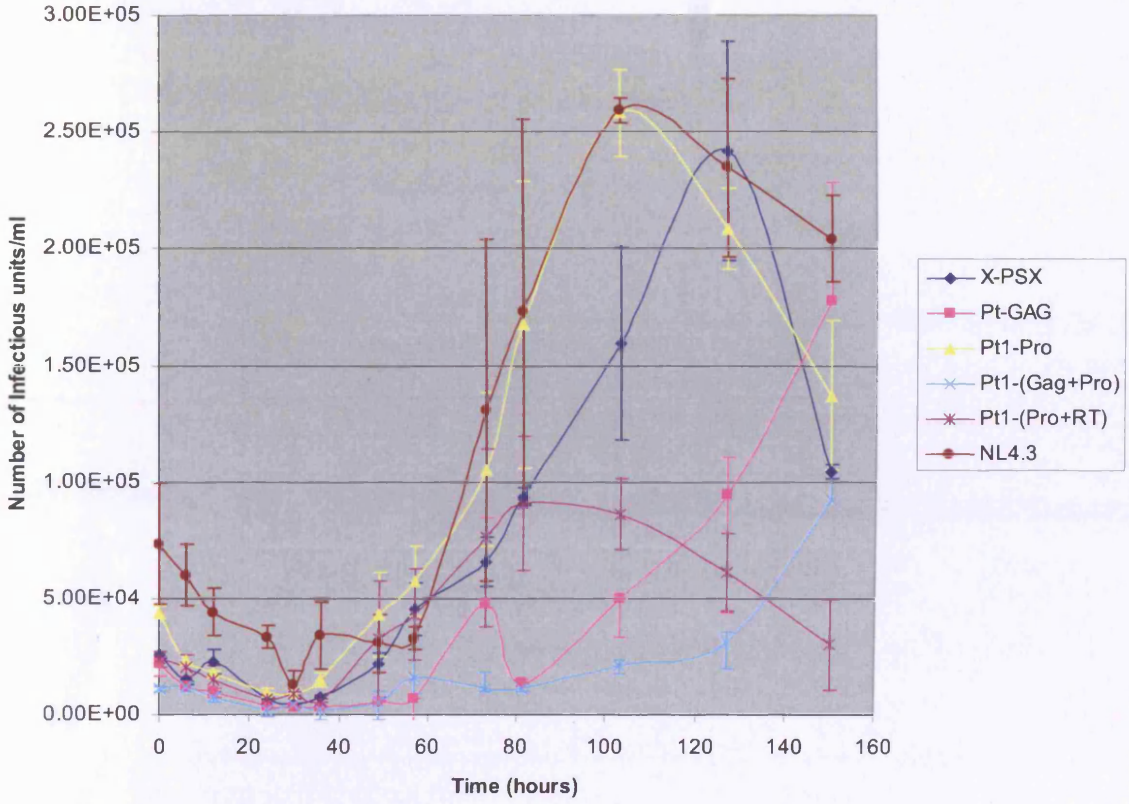


Figure 6.3: Comparison of growth rates of the recombinant viruses and NL4.3 over a 150 hour growth period.
Error bars indicate standard error derived from 3 independent titrations.

Table 6.2 Comparison of the amount of virus produced at peak titre over the course of the growth curve for each of the viruses after adjustment of start input.

Virus Construct	Peak change in titre
X-PSX	2.16E+05
Pt1-Pro	2.14E+05
NL4.3	1.86E+05
Pt1-GAG	1.56E+05
Pt1-(Gag+Pro)	8.04E+04
Pt1-(Pro+RT)	6.71E+04

6.2.2c Competition Growth Study

The aim of this experiment was to assess the growth abilities of the virus stocks using dual infections. Each of the patient 1 derived recombinant viruses was competed against the NL4.3 molecular clone and a comparison of the rate at which each was outgrown by NL4.3, after serial passaging, was to be used as a measure of replication capacity. Initially, two different percentage ratios of recombinant vs. NL4.3 were set-up (50% vs. 50%) for X-PSX vector as it was thought to compete equally with NL4.3 and (80% vs. 20%) for the less fit recombinants, The MOI of each dual infection was 0.001 and the input amounts were derived from the $TCID_{50}$/ml values obtained from a minimum of two MTT titration assays. Each of the infections was harvested after full CPE had been observed and the infected pellets were used for analysis of mixture ratios.

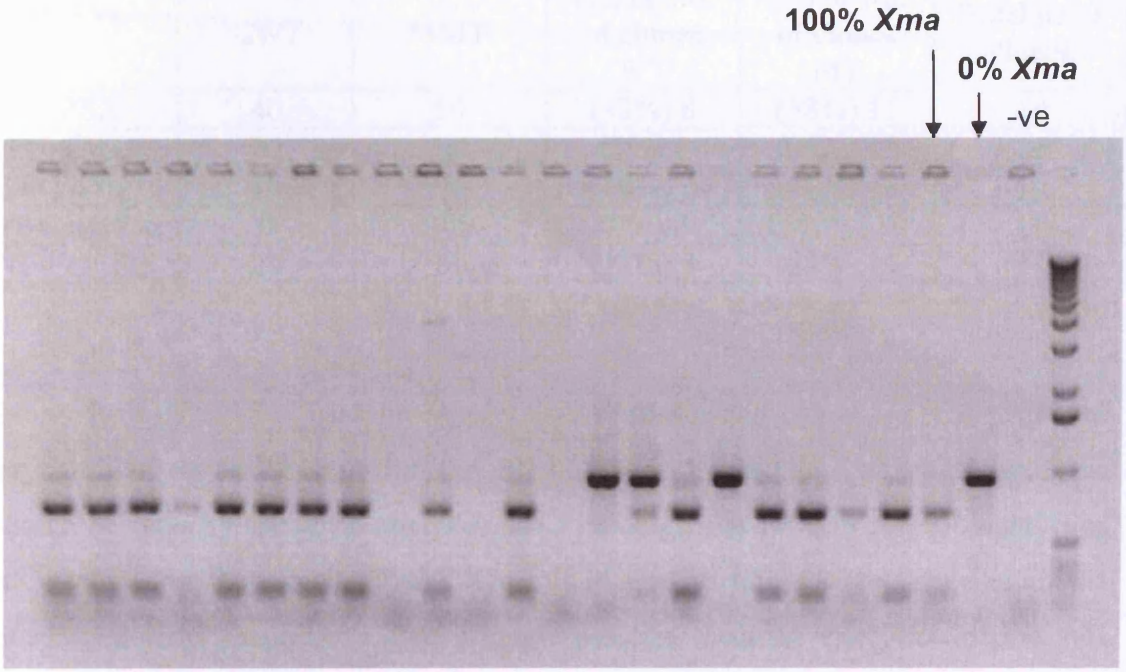
To determine the ratio of NL4.3 and recombinant in the infected cell pellet, two methods of analysis were carried out. As all the recombinants were derived from the X-PSX or PSX vector, but, due to the fact that different sized products were cloned they did not contain all the same restriction sites e.g. the Gag+Pro clones were produced using the *Xba1* and *SexA1* sites and patient derived Gag+Pro sequence, therefore they would not contain the *gag* C-terminal *Pvu1* cloning site. The only common restriction site amongst all the clones was the C-terminal RT cloning site, *Xma1*, which was absent in the parental vector, pNL4.3. Therefore, analysis of mixture ratios in the dual infections was carried out by sequence analysis of a PCR amplicon that encompassed this site. This same amplicon was also used to clone into the pGEM cloning vector and the resulting clones were again PCR amplified across the *Xma1* site. By digesting these PCR products with *Xma1*, the proportion of NL4.3 was determined by the number of clones that did not contain the *Xma1* restriction site. An example of the comparison of the results obtained for one of the dual infections using this approach is shown in Figure 6.4 and the overall results comparison is shown in Table 6.2.

Each of the PCR's were set-up along side positive and negative controls, which were also digested. The 100% *Xma1* represents the PCR amplicon of the X-PSX vector DNA and should only contain products that are sensitive to *Xma1* digestion,

producing fragments of 673bp and 273bp. The 0% *Xma1* amplicon is derived from the PCR of NL4.3, which does not contain the RT cloning site and therefore would remain uncut.

Figure 6.4: Example of a comparison of the clonal analysis and peak height determination. Example shown is for the Pt1-Pro+RT amplicon derived from the NL4.3 vs.Pt1-Pro+RT dual infection, passage 1 infected cell pellet.

(a) A total of 21 amplicons were analysed, 3 clones were uncut, 15 were digested and 3 clones generated negative amplicons. The PCR negative control lane is denoted by the (-ve) symbol.



(b)Chromatogram of bases around the *Xma1* restriction site

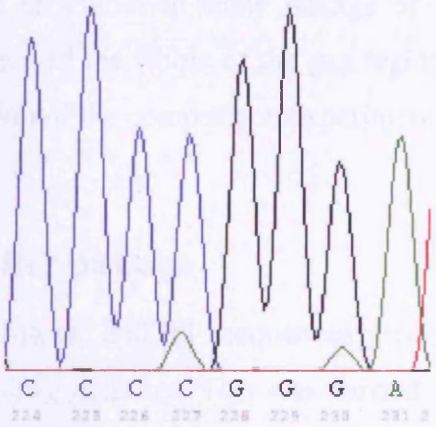


Table 6.2: Comparison of percentage mixtures in competition cultures.
Determined by sequence analysis (peak height comparison at the *Xma1* restriction site) and by clonal restriction enzyme digest analysis after the first passage in the competition growth study.

NL.3 vs.	Sequence Analysis		Clonal Analysis		
	%WT	%MT	(%) or No. of clones WT	(%) or No. of clones MT	Total no of clones
X-PSX	40	60	(42%) 8	(58%) 11	19
Pt1-GAG	20	80	(5%) 1	(95%) 19	20
Pt1-Pro	10	90	(10%) 2	(90%) 18	20
Pt1-Gag+Pro	0	100	(0%) 0	(100%) 19	19
Pt1-Pro+RT	20	80	(17%) 3	(83%) 15	18

The analysis of the first passage data by both sequence and clonal methods produced very similar results, except for the Pt1-GAG dual infection, which showed the predicated ratio by sequence analysis, but a much greater ratio of recombinant virus was seen using the clonal system. Overall, the input ratios were reflected in the determined virus proportions after a single passage, with the exception of the Pt1-Gag+Pro virus, in which no wild-type was detected by either method. Further passing and analysis is needed in this experiment, however, the competition experiment was stopped due to the reversions seen in the high titre stocks. The appearance of reversions meant that the experiment no longer contained two clonal populations competing against each other. This would mean that any subsequent analysis would be need to be carried out on clones to show linkage of mutations and this would require the sequence analysis of the whole of the gag region. It was primarily for this reason that further evaluation of the competition experiments was not carried out.

6.2.3 Sequence data after passage

The stability of the *gag*, protease and RT sequences was assessed by passage of the viruses in the absence of drug pressure. This was carried out by PCR amplification and DNA sequencing of a region covering the *gag* and Pol domains from the infected cell pellet, after harvesting of the cell-free viral supernatant for subsequent passage.

The details of the techniques used to carry out the PCR and DNA sequencing are in the Materials and Methods section 2.1.13.

6.2.3a Patient 1 and 5 viral isolate derived (Pro-RT) Recombinant Virus Sequence Data After Passage

As the Pt1-C6, C7 and Pt5-C25 (Pro-RT) clones contained the protease and RT regions from the patient derived viruses, the sequence data was primarily focused on the stability of the mutations in this fragment. The sequence contig that was assembled and analysed from the infected cell pellet sequence data, spanned the region from the p1/p7 *gag* cleavage site to codon 345 of RT, a region of 1.5kb.
In the Pt5-C25 virus isolate the protease mutations remained stable, the RT region maintained most of the RT mutations that were introduced from the patient derived virus. The sequence after passage showed a reversion of the NNRTI mutation F227L. Two other changes in the RT region from the infected pellet PCR product were also noted. These occurred at codons 277 and 295, which altered the original patient virus amino acid sequence to one that occurs in NL4.3; both of these sites are polymorphic within sub-type B [176].

The PI mutations in Pt1-C6 and C7 remained stable over the growth period and from the infected pellet sequencing, it appeared that the RT mutations were also stable to. The only differences that were observed in the passaged stock from the transfected clone in the RT region were a novel T107I change not seen in other subtypes [40,176], a reversion of the Q207E to the NL4.3 amino acid and a V293I change, also thought to be a consensus B change [176].

Both viruses derived from patient samples 1 and 5,Pro-RT clones contained the majority of the wild-type *gag* region derived from NL4.3. However, the 20 N-terminal amino acids are from the patient virus source, due to the position of the C-terminal cloning site used for the introduction of protease. Therefore, the only changes in the patient 1 and 5 derived virus *gag* fragments from NL4.3 in this region, are due to being introduced during the cloning.

To further investigate the lack of resistance in the Pt1-(Pro-RT) clones, a clonal analysis of the RT region containing the T215Y mutation was carried out to determine the percentage of clones which were still carrying this mutation. A 1.5kb PCR product spanning the RT region was amplified and ligated into the pGEM cloning vector, transformed into E.coli and the DNA isolated from eight clones was sequenced across the region of interest. With the exception of one clone, which reverted to wild-type, all the remaining clones were mutant at codon 215, with a tyrosine at this position.

6.2.3b Patient 1 Derived Virus Protease Recombinant Sequence Data After Passage

The Pt1-(Pro) virus stock showed the least variation from input genotype of any of the recombinant viruses. The genotype derived from the infected cell pellet of the passaged virus stock showed no changes from the wild-type *gag* sequence, or the mutant protease coding regions.

6.2.3c Patient 1 Derived Virus Gag containing Recombinant Viruses Sequence Data after Passage

The Pt1-GAG and the Pt1-(Gag+Pro) virus stocks showed no deviation in the protease sequence from the genotype of the molecular clones used to produce the virus stocks. However, the *gag* regions from the same virus stocks showed an extensive reversion pattern. Figure 6.5 shows the alignment of the *gag* region from the passaged Pt1-GAG and the Pt1-(Gag+Pro) viruses. Also shown in the alignment is the sequence of the *gag* region from the pGEM clone Pt1pGEM-C2, along with NL4.3 reference sequence.

Pt1-GAG virus showed a more extensive loss of mutations from its clonal input DNA stock than the Pt1-(Gag+Pro) virus. This pattern was seen most clearly in the MA and p6 Gag proteins, analysis of which showed that the Pt1-(Gag+Pro) virus maintained all the mutations from the input molecular clone in both these proteins. By contrast the Pt1-GAG virus lost all the analogous mutations in p6 and only four amino acid changes were preserved in p17, three of which are thought to be polymorphic; K15R, K30R and R76K [176]. The Q28K change occurs at a position conserved amongst the different clades, but has been previously shown to appear [311]. The MA insertion

present in the original patient 1 derived virus *gag* was lost by the *gag* only virus, but maintained in the Gag+Pro virus.

The MA of the Pt1-GAG virus was also shown to contain two mutations that occurred at amino acids positions which were originally NL4.3 sequence, namely Q59H and E74V. A database search revealed that these positions are highly conserved amongst different HIV clades, this was also the case for the T81P, which appeared in the *gag* only passage virus. The A82L/V also in p17 occurred in both the Pt1-GAG and the Pt1-(Gag+Pro) viruses respectively. This codon is normally associated with an Isoleucine or Threonine [176].

Both virus clones derived from patient 1 which contained *gag* maintained all the mutations in the CA protein, most of which are polymorphic with the exception of the I247V [176].

The p2 protein showed the highest number of reversions in proportion to its size; 14 amino acids long and four reversions within the space of six amino acids, potentially a 'hot-spot' for adaptation. These occurred in both the Pt1-GAG and the Pt1-(Gag+Pro) viruses and included the mutation at P1' in the p2/p7 cleavage site. The only p7 change that persisted when the *gag* and the corresponding protease were included in the recombinant virus, was the M423I change, which also occurs at a conserved amino acid position across the HIV clades [176].

The A431V change at position P2 in the p7/p1 cleavage site in the original clonal isolate was lost by both the *gag* containing viruses, however the adjacent cleavage site, p1/p6 retained the P1' mutation in the Pt1-Gag+Pro virus.

Figure 6.5: Alignment of gag regions derived from Pt1-GAG and Pt1-Gag+Pro viruses after passage.

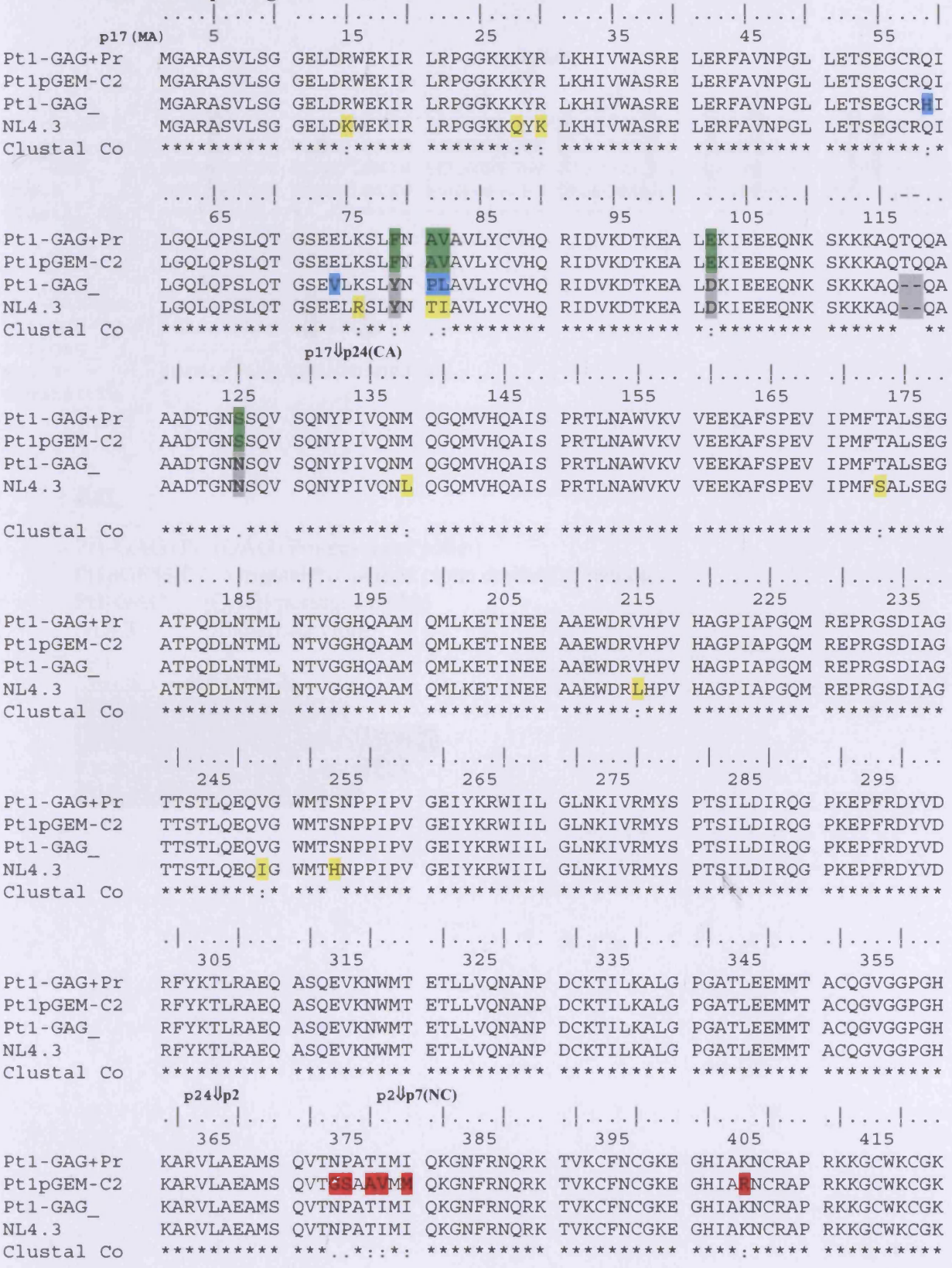
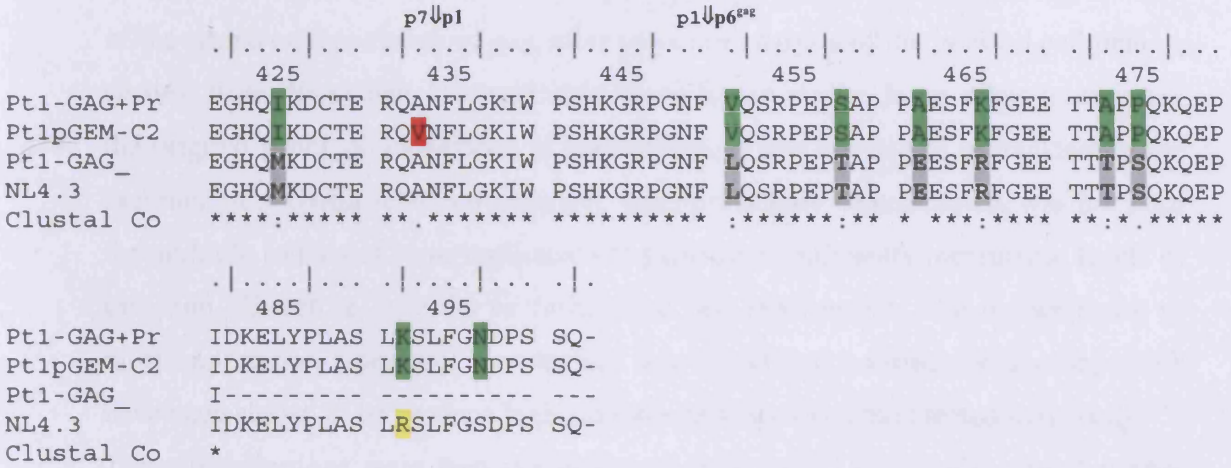


Figure 6.5(continued): Alignment of gag regions derived from Pt1-GAG and Pt1-Gag+Pro viruses after passage.



Key

Pt1-GAG+Pr (GAG+Pro passaged pellet)
Pt1pGEM-C2 Original Pt1 pGEM clone derived sequence
Pt1-GAG_ (GAG) passaged pellet)
NL4.3 molecular clone

Changes in NL4.3 only
Changes in GAG-pellet only
Chhanges in pGem +GAG+Pro pellet
Changes in NL4.3 and GAG pellet
Changes in pGEM clone only

6.2.4 Genotype Stability Determination

In the previous results section it was reported that reversions were occurring at some of the amino acid positions in *gag*, after sequence analysis of the infected cell pellets, derived from the patient 1derived virus recombinant stocks. In an effort to maintain the original genotype introduced at transfection, it was decided to re-transfect in the presence of antiviral selection pressure. Also, previously demonstrated, was the need for multiple rounds of virus replication to generate significantly measurable levels of infection. Therefore, in order to fulfil these two requirements, the former being of most importance, a protocol was devised that would first transfect 293T cells, which have been shown to be produce high amounts of virus when transfected transiently [105]. These transfections were then supplemented with $1x10^7$ MT-4 cells, one day after initial introduction of cloned vector DNA, which, it was hoped would facilitate productive infection of the virus produced through 293T transfection. The infection was allowed to proceed a further 5 days, approximately 2 replication cycles [8], at which point the cell-free supernatant was harvested and used to infect $3x10^6$ SupT1 cells in the presence of 8nM SQV, 40nM IDV, 100nM RTV, 100nM AZT or no inhibitor. The concentrations of antivirals used represents the NL4.3 IC_{50} values determined using the MTT assay system. All infections were allowed to proceed until full CPE was observed, or it was decided that no observable replication was ongoing, by virtue of no visible CPE, throughout the course of infection. 500µl of cell-free supernatant was used to infect a further $2x10^6$ SupT1 cells, this was carried out in the presence or absence of the aforementioned concentrations of antivirals and also constituted the next passage.

The recombinant clone containing the *gag* derived from patient 1 isolated virus, showed no signs of CPE in any of the cultures which contained antivirals, only the 'no drug' infection produced visible CPE. Overall, all the viruses grew poorly, especially during the initial passage in SupT1 cells, with no visible signs of CPE after 8 days in culture.

At passage two, the only virus to display production of CPE was the Pt1-Pro passaged in the absence of any antivirals. During this passage both the protease only clones exposed to RTV and IDV also produced CPE in the SupT1 cultures, before the

cessation of the passaging step. The Pt1-Gag+Pro infections required an additional passage step before CPE was detected in any of these infections, this virus was the only recombinant to produce any visible CPE in the SQV infection.

Only at the end of passage three, after a total amount of time in culture of approximately 30 days, was CPE d.etected in all the Pt1-Pro and the Pt1-Gag+Pro infections. The Pt1-GAG virus only produced detectable infection in the 'no drug' infections. All the infected cell pellets were harvested and amplified as detailed in the Materials and Methods sections 2.4 and 2.1.9.

Figure 6.6: Alignment of gag sequence data generated from the Passage 3 of the Pt1-Gag virus in the absence of inhibitor.

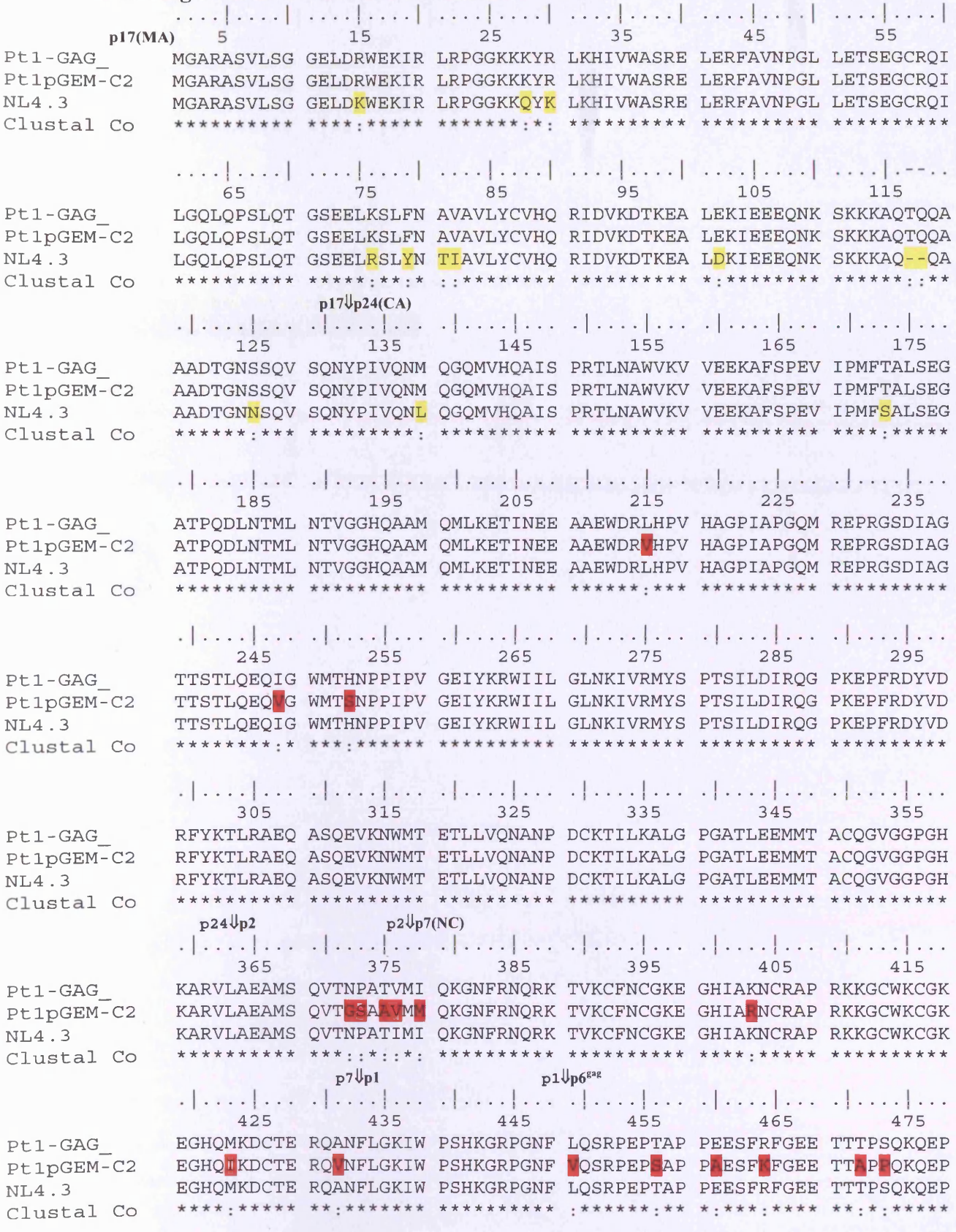


Figure 6.6(continued): Alignment of gag sequence data generated from the Passage 3 of the Pt1-Gag virus in the absence of inhibitor.

```
             .|....|... .|....|... .|.
                     485        495
Pt1-GAG_       IDKELYPLAS LRSLFGSDPS SQ
Pt1pGEM-C2     IDKELYPLAS LKSLFGNDPS SQ
NL4.3          IDKELYPLAS LRSLFGSDPS SQ
Clustal Co     ********** *:****:*** **
```

Key

```
Pt1-GAG_      (GAG) passaged pellet)
pGEM-Pt1      Original Pt1 pGEM clone
NL4.3         wild-type molecular clone
Changes in NL4.3 only
Changes in pGEM clone only
```

The stability of the genomes was assessed by DNA sequencing of the infected cell pellet; this was carried out as detailed in the Materials and Methods section 2.1.13. To investigate the potential of carry-over of the input vector after transfection, a control PCR was set-up which utilised oligonuleotides that annealed to the ampicillin resistance gene located in the recombinant virus vector. Any infected pellets that produced an amplicon from the ampicillin PCR were therefore, likely to contain contaminating input plasmid. All the infected pellet DNA samples were negative for this PCR, therefore, the sequence data obtained from these samples was deemed to be derived from the integrated infectious virus and not as a result of input plasmid contamination.

Analysis of *gag* region sequence data showed that less reversions occurred in the Pt1-GAG virus, than had been previously demonstrated during production and sequencing of the high titre stocks, see Figure 6.6. The main differences occurred in MA, where the all the original mutations in the pGEM clone were maintained by the Pt1-GAG virus. A differential mutational pattern was also observed for CA from the first passage result, with 3 reversions from 5 seen in this experiment, as opposed to complete conservation of the pGEM sequence, demonstrated in the previous passage data set.

The C-terminal region of *gag* reverted to NL4.3 sequence, as was seen previously, although the data obtained from this experiment showed much more heterogeneity in the sequence chromatograms at all the positions where reversions had occurred.
Another disparity from the previous sequence data was seen in the Pt1-Pro virus, here changes were seen in three of the infections; those under the selection pressure of AZT, IDV and SQV, no changes were seen in the 'no drug' infection derived infectious pellet. An identical change was seen for the AZT and IDV infections, an L75R, which was present as a mixture in the IDV infection and as a homogenous codon in the AZT infection. Two other changes were detected in the Pt1-Pro, SQV infection, a synonymous change in MA at codon 10, GGG to GGA, both coding for Glycine and a non-synonymous change, K415I, in NC. Both these non-synonymous changes K75R and K415I, occur at positions which are somewhat polymorphic, although neither of these amino acid substitutions are present in any of the HIV clades [176].

The sequence data obtained for the Pt1-Gag+Pro virus, demonstrated an unchanged genotype from the input sequence in all of the infections.
Overall and in concordance with the previous sequence data, changes were seen primarily in the *gag* regions, no changes were detected in protease or RT in any of the infections.

6.2.4a Analysis of Packaging Sequence

The packaging of viral genomic RNA is aided by the packaging sequence (ψ), located in the 5'UTR [42,215,303,356]. The sequence analysis of the 5'UTR of the study patient derived virus showed several changes in the Ψ, which upon passage showed some reversions, indicating possible selection in this region. These included a 'T' to 'G' change in stem loop 1 (SL1), which may result in mis-pairing of bases in this loop and causing an additional bulge structure, see Figure 6.7. A similar bulge may also occur in SL2, as a result of 'C' to 'A' change, which also causes mis-pairing of nucleotides. Both NL4.3 and the patient virus contained the same nucleotides at these two positions and the mutations acquired after passage are changes from both these viruses. Most of the remaining changes that were detected were reversions back to NL4.3 sequence including the addition of a missing nucleotide at the base of SL3. However, no reversion was detected in the apical loop of SL3, where a 'G' to 'A' change had removed a purine-rich region at the apex of this loop, which others have demonstrated, is a primary site for interaction with Gag [244,356]. This may be as a result of lack of reversion or it may reflect redundancy in the packaging complex.

Perhaps the most important change seen in the original virus genotype was a two bp deletion in the palindromic, 'kissing-hairpin loop' region. Abolishment of this palindromic sequence has been shown to cause dramatic loss in virus infectivity [173,319]. In this study it was shown that after passage this sequence showed partial reversion, with the insertion of a 'GG' sequence. Although this did not restore the palindromic sequence it did result in the appearance of a double guanine motif at the loop apex, which is thought to increase affinity to NC [198,244].

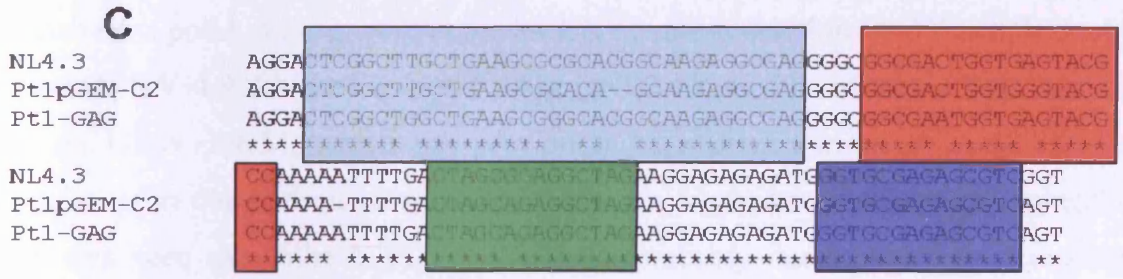


Figure 6.7:Amino acid and nucleotide sequences for HIV-1 NL4.3 NC and Ψ.

A) Amino acid sequence of the NL4.3 NC protein showing the zinc-binding knuckles. Residues that contact the RNA in the NC-SL3 complex are denoted by open letters. Asterisks denotes residues that contained mutations in the Pt1-GAG recombinant that reverted after passage.

B) Nucleotide sequence and secondary structure of the NL4.3 Ψ sequence. The dimer initiation (DIS), major splice donor sites (SD) and the *gag* initiation codon (AUG) shown in open letters.

C) Alignment of Ψ of Pt1-GAG virus after passage, compared against original pGEM and NL4.3. The colour shaded boxes represent the stem loops (SL1-4).

Figures A and B adapted from De Guzman *et al.* [63].

6.3 Discussion

The development of resistance to PIs has been shown to reduce the activity of the viral protease, requiring restoration of enzyme function through the accumulation of compensatory mutations [230,358]. These mutations have been identified in both the protease and the *gag* coding regions of viruses isolated from highly antiretroviral experienced individuals. One such isolate was identified in this study and the highly mutated protease was also shown to contain a two amino acid insertion. To assess the impact of the *gag* and protease regions, derived from the patient isolate on virus growth ability, synchronous infections using the recombinants were set-up. All viruses were titred on GHOST indicator cells, as prior to this growth curve, a smaller scale experiment was used to determine which method would be used for the determination of virus quantity.

Both the MT-4 and GHOST titrations gave very similar results for the amount of virus at each time point in the growth curve for HXB2 molecular clone and the SQV double mutant G48V+L90M (see Figure 6.1). The HXB2 virus achieved higher titres than the mutant G48V+L90M, which has previously been demonstrated to show reduced replication in comparison to wild-type virus [283]. A greater difference in replication rate was seen using the MTT assay, this is probably due to the low percentage infections detected using the GHOST assay system. An equivalent amount of virus was added to the GHOST titration as was added to the MTT assay, the latter system had several more days in culture, during which any differences in replication rate could be magnified. The selection of the GHOST assay to determine virus quantity was due to the reduced incubation time and the more direct end-point measurement. Other methods of determining titres of cell free supernatant could also have been investigated, such as, p24 ELISA assays assess the amount of protein present, therefore this must have been produced by viral replication and so is a measure of active replication. However, it relies on binding of anti-p24 antibody to the capsid region of Gag. As this was one of the regions that might vary in clinical isolates due to drug induced or replication restoration required mutations, this could result in variable binding to antibody and therefore under estimates. Assays to measure p24

also only measure protein levels and may not distinguish between uncleaved proteins which could accumulate due to gag and protease mutations in the virus clones. The chosen method of analysis required infection of GHOST cells in order to produce a GFP signal. Analysis of infections where only 30% of cells are infected reduced the possibility of underestimation of virus quantity due to re-infection of an already infected cell. Viral load measurements determine the amount of viral RNA in a sample and do not elude to the amount of live/infectious virus present. All assays show some level of variability when assaying for the amount of p24/RNA viral load and $TCID_{50}$. The latter is a measure of virus induced death in the MTT assay and a measure of the number of infectious particles produced in the GHOST cell assay. Both these measures were compared to each other using the same virus stock and gave similar results (see Figure 4.9).

In the growth rate comparison experiments the recombinants were used to infect T-cells and the amount of virus produced over the course of infection was monitored at 15 time points. The variation in the start titres may have been caused by inaccuracies in the titre determinations, as the Pt1-Pro virus demonstrated a larger variation in the GHOST cell titration curves shown in Figure 6.2. These two viruses, NL4.3 and Pt1-Pro also represent the viruses with the lowest and highest titres, so it is possible that the growth curve inputs are subject to more variation, as it is more likely to under or over estimate the amount of virus added. Therefore, any inaccuracies in the titre determinations are more likely to affect the samples at the extremes, when calculating the input for the growth curve. For example, with low titre viruses, it is conceivable that an under estimation of titre can occur as not enough virus infection occurs through cell to cell spreading infection. Thus, when larger volumes are used, this effect is reduced, leading to an under estimation of the titre. With high titre viruses, a similar under estimation of titre can be achieved if cell killing occurs too quickly, causing a reduction in the overall amount of virus produced.

All the recombinant viruses gradually reduced in titre from the input amount over the first 24hr period, the reduction reflecting the entry of viruses during the replication cycle (eclipse phase). The subsequent increases from this nadir level, which suggest the production of new virions were seen in most viruses at the next time point at 30 hours. In the case of both Pt1-GAG viruses a recovery in titre to that of the initial

input amount, took over 80 hours. Recovery of the Pt1-GAG appears to be better than that of the Pt1-(Gag+Pro) virus, this would suggest that protease compromises the virus which contains it; however, the Pt1-Pro virus grew to similar rates to NL4.3. This anomaly could be explained by the sequence data generated from the infected pellets of the high titre stocks, used as the input viruses in this growth rate comparison study. As reported in Results section 6.2.3c of this chapter, the reversion pattern observed for the Pt1-GAG was greater than that seen in the Pt1-Gag+Pro, this was especially true for the p6 protein in *gag*, from both passage data sets. Sequences within this region are thought to be involved with the late stages of the viral Life cycle, such as budding, through interactions of the PTAPP motif and host proteins [30,94,334,335]. This sequence data also showed the retention of the PTAPP to PSAPP by the *gag* and protease virus, but reversion in the *gag* only, even though this difference is thought to be polymorphic [241]. This co-evolution with the mutant protease would appear to force *gag* to maintain the mutation(s) which may account for the slower replication seen compared to the virus that originally contained only the same *gag* region. Another point to note is that the level of reversion in the *gag* only clone was not apparent at the start of this growth curve and to what extent further reversions have occurred during the course of the experiment is also unknown.

The protease only virus replicated to high levels despite containing multiple mutations known to reduce enzyme efficiency [230]. This high replication also occurred without the acquisition of any *gag* mutations, as demonstrated by the passage sequence data. Although any changes may have been undetectable by population sequencing, the majority of the virus genotype would appear to be the same as was introduced by transfection. The presence of the two amino acid insertion, which occurs in the flap region, encompassing residues 33-62 [268,290] of the protease protein, may play a role in maintaining the replication capacity of the drug resistant phenotype. Molecular modelling experiments on SIV proteases have shown that the insertions in the protease gene between codons 34 to 37 cause conformational changes in the geometry of the flap region and contribute to structural alterations in more distant regions of the molecule [156]. The appearance of insertions at this position in protease have been previously reported in PI naïve, although RT inhibitor experienced and PI treated individuals [245,309]. The emergence of such duplications in PI naïve, but RT inhibitor treated individuals, suggests that these insertions, which have been shown to represent

duplications of the preceding sequence [156], may be associated with RT inhibitors and the inhibition of RT processivity. It is has also been demonstrated that these insertions do not decrease sensitivity to PIs, but do normally confer a growth advantage in the presence of other PI resistance mutations [156,245,309]. Thus their emergence in PI naïve individuals, without other PI mutations, also implies that these proteases confer a selective advantage even in the absence of PI therapy.

The Pt1-(Pro-RT) virus, produced the lowest amount of infection in the growth curve, this may imply that the cloning of the cognate highly resistant RT has a detrimental effect on virus growth. As the protease only virus grew to high levels, the production of an RT-only clone would help to clarify this, however due to time constraints this was not completed.

The lack of growth associated with the *gag* containing virus clones, would appear to be solely attributable to the *gag* coding region, as the Pt1-GAG and the Pt1-Gag+Pro viruses displayed poorer growth kinetics compared to the protease only virus stock.
One method of interpretation of the data generated from the growth rate comparison is to compare the start titre with the peak titre detected over the course of the growth curve for each of the viruses; shown in table 6.2. This method of evaluation removes the error associated with the differences in the start input for each of the viruses and orders the virus constructs based on the amount of virus produced in the 'static' conditions of the experiment i.e. no new cells were added to the assay. The molecular clone NL4.3 had the largest input and this may explain the lower peak change in titre. The Pt1-GAG construct also would seem to out perform the Pt1-Gag+Pro virus using this method of analysis. This may be explained by the more extensive reversion seen in the sequence analysis of this virus, i.e. the reversions that occurred during the production of this virus stock may have made the virus more adapted to the culture conditions.

The large error bars represent the difficulties in using a replication competent system to asses growth rates, as variation in titres, not only in the determination of the input amounts, but also in the final determination of virus quantity, undoubtedly affect the values at each time point. Other factors such as different plating densities and growth rates of the reporter cells also affect the final result. However, the overall positions of

the recombinants, relative to each other remained almost constant throughout the three separate determinations.

A more direct measure of virus growth comparisons is achieved by competing two viruses in dual culture experiments. To assess the growth rates of the patient derived virus recombinants, each was added in excess to a mixed infection containing NL4.3. The amount of mutant to wild-type virus was measured as the percentage of clones that were isolated which contained the *Xmal* restriction site, present only in the mutants. This experiment was only continued for a single passage, although variation was detected from the assumed input, this is more likely to reflect errors in the determined titres, rather than any immediate outgrowth by the presumed less fit recombinants.

The assessment of the virus proportions was shown to be almost identical when using a clonal analysis versus a population sequence analysis. Only one dual infection produced disparate results; NL4.3 vs. Pt1-GAG. This could be due to the sequence analysis being less sensitive than the clonal method, which by analysing 20 clones per infection, allows a sensitivity of 5%. The sensitivity of population sequencing has been shown to be much higher than this [151]. It is also possible that an underestimation of wild-type clones occurred, due to the low frequency at which they would be distributed, therefore, selection of these colonies would be less likely to occur.
The similar results obtained for determination of ratios would imply that further analysis using just the sequence method of determination, which is a much less labour intensive approach, would provide an adequate interpretation of the virus proportions within the dual infections, after further passaging.

Continued culturing of the virus clones in the absence of drug pressure is thought to select for the fittest virus in the quasispecies produced by the error prone RT, even in the low passage numbers used here. Those mutations which are present purely to prevent the antiviral compound from interfering/binding, are no longer required and so will be lost. This is because these mutations have been shown to be detrimental to the activity of the enzyme targeted by the antivirals [20,230,253]. It is also assumed that the original wild-type virus before any exposure to antivirals and before the acquisition of resistance was the most genetically fit species. The persistence of the resistance

mutations in both the protease and RT sequences of the patient 1 viruses would seem to suggest that they may not be detrimental to virus replication capacity, or reversion may require an intermediate that is even less fit. However, as this system was clonal, the acquisition of mutations or reversions could only be due to their appearance following an error during replication, followed by their selection as a fitter variant. Thus, the persistence of these mutations is solely due to the time taken for virus to acquire and select them. Such a similar delay in reversions was seen after longitudinal analysis of plasma virus from an individual who acquired HIV-1 that contained multiple RT and protease mutations [99]. The resistance mutations in protease were maintained for longer than 300 days after diagnosis of HIV-1 infection. Another hypothesis could be that the mutational pathway required to select for these reversions necessitates a reduction in overall replication capacity levels and thus, this pathway would initially be selected against, until other reversions or mutations which work co-operatively, allow for their selection [88,355]. The only reversion of a resistance mutation amongst the recombinant viruses generated in this study, was the AZT, T215Y mutation, resulting in a heterogeneous population similar to that seen by Bezemer *et al.* [18].

The viruses which displayed the highest level of variability in sequence were those recombinants which contained the patient 1 derived virus *gag* region. The *gag* region also proved to be the most variable after transfection and subsequent passage. Such a high level of heterogeneity in *gag*, after recent transmissions, was shown to occur by Zhang *et al.*, the MA also demonstrated greater selection pressure than the V3 loop region of the *env* gene [357]. The selection pressure on *gag* has also been shown to occur *in vitro*, after serial bottle-neck experiments starting from a clonal population [353].
The reversion pattern seen in both the viruses derived from patient 1 containing *gag* from the first passage data set, also suggests that there is a sizeable selection pressure placed on this polyprotein. However, the inclusion of the mutant protease seems to have applied another form of selection on the complementary cloned *gag*, this is most evident in the MA and p6 *gag* regions. As both *gag* containing viruses initially contained the same mutant *gag* sequences, with the exception of the terminal 21 amino acids of p6; here the only differences were due to the position of the *PvuI* cloning site. Therefore, any sequence differences detected between the two mutant *gag* clones after passage, could be attributable to their associated protease.

The changes in blue shown in Figure 6.5, represent those changes which occur only in the Pt1-GAG virus after passage, all of these mutations appear in MA. Two of these changes also occur at amino acid positions where originally the sequence in the pGEM clone was that of NL4.3, at codons 59 and 74. Both of these changes are thought to be non-polymorphic [176], the Q59H change occurs in a region in which mutations have been shown to cause a reduction in the efficiency of particle production [98] and this codon has also been shown to be involved in the formation of MA tetramers [137]. Mutations in MA, at codons 79, 81 and 82, which are seen in the pGEM clone, reverted when the wild-type protease was included, mutations at these positions have been shown to cause virus assembly defects [94,97]. Although, changes at these positions occurred in the *gag* only clone, the resulting amino acid sequence was different to that of NL4.3.

The TQ insertion which also occurs in MA is similar in position to an insertion previously reported by Tamiya *et al.*, where it was demonstrated that an AQQA insertion restored the replication capacity of the recombinant virus containing a mutant protease [311]. The Gag cleavage pattern of the virus containing both the mutant *gag* and protease was also shown to be comparable to that of the wild-type control in both the infected cell and cell-free virion preparations. It was also shown that the corresponding mutations in the p6 protein were required in order for optimum Gag polyprotein processing. The Pt1-(Gag+Pro) virus generated in this report, like that described by Tamiya *et al.*, co–evolved with the insertion and also maintained the p6 mutations, including those mutations around the PTAP, late domain budding motif. Whereas, the *gag* only virus did not retain the insertion and p6 mutations, this would agree with the findings of Tamiya *et al.*.

Muller *et al.* have described the construction of a virus containing a 12 amino acid epitope tag at the C-terminus of MA. The insert did not interfere with viral replication in tissue culture, and the inserted sequence could be detected after multiple rounds of infection [219]. This insert was located in a region that forms a random coil conformation and was therefore thought not to interfere with the overall functional arrangement of Gag. The insert described in this report and that by Tamiya *et al.*, occurs in an extended α-helix region of MA projecting away from the globular

domain of the protein [137,219] and may serve to connect MA and CA domains in $Pr55^{Gag}$ [29]. Therefore, as these inserts were generated by evolution and not through *in vitro* introduction, they are more likely to be involved in one of the many roles of MA or $Pr55^{Gag}$. The multifunctional role played by MA in the HIV Life cycle, may be the reason that this protein is under such selection pressure from the onset. It plays a key role in regulation of the early and late steps of viral morphogenesis and is required to interact with cellular proteins such as, nuclear import of the PIC by interactions with cellular importins [29,94,96]

Both the viruses derived from patient 1 containing *gag* maintained all the p24 mutations from the original input clones. This is thought to be because this region is more dependant on the conservation of overall protein structure, as opposed to preservation of nucleotide or amino acid sequence [350].

The first Gag cleavage site to be processed by the viral protease, p2/NC [60,261,262], reverted to NL4.3 sequence in both Pt1-GAG and Pt1-(Gag+Pro) viruses. These changes included all the mutations at the p2/p7 cleavage site, which have previously been shown to be highly polymorphic, both *in vitro* and *in vivo* [191,194,299,299,350]. Interestingly, the P5 position of the original patient 1derived virus pGEM clone, contains the same amino acid found in the HIV subtype B clone HXB2, however, this mutation also reverted back to the analogous amino acid found in the NL4.3 subtype B clone. This suggests that this codon position is highly restricted even within HIV B sub-type.

The K403R change seen in the original pGEM clone also reverts in both viruses, this mutation occurs in one of the zinc finger domains located in NC [94] and retains the basic charge of the amino acid. Basic residues within NC have been shown to be required to interact electrostatically with the phosphate backbone in nucleic acid, which is thought to be required to promote the assembly of the mature virions [29,94,114]. This Arginine to Lysine reversion, retained the basic nature of the amino group and may be better at performing the RNA/DNA binding function of this region, through the acquisition of a less bulky side chain, which could lead to less steric hindrance.
The A431V mutation occurs at the P2 position of the NC/p1 cleavage site and is therefore likely to be in direct contact with the enzyme substrate-binding pocket [268]. It

has previously been demonstrated that changing the Alanine to a Valine, results in enhanced cleavage at the NC/p1 cleavage site by both the wild-type protease [231] and also by mutant proteases [85,188,224,311,327]. Viruses carrying protease only mutations showed reduced infectivity which correlated significantly with reduced processing at the NC/p1 cleavage site [277]. The reversion of the A431V codon in the Pt1-(Gag+Pro) virus, suggests that there is selection against this mutation, in the absence of PI pressure and in the context of cloning with the associated mutant protease. The A431V mutation has also been associated with proteases containing multiple PI resistance mutations, therefore, its loss after sequential passage in the Pt1-(Gag+Pro) virus is contradictory to the evidence presented by some research groups [51,191,327]. . However all these references refer to viruses found within patients who maintained PI therapy and thus continued the pressure to maintain the A431V mutation. The A431 mutation may also have a role as primary PI mutation and thus be dependant on PI pressure to maintain its presence. The introduction of the A431into a wild-type reference strain was shown to confer 3.8-fold reduced susceptibility to ritonavir [231]. This same research group also found that 1% of samples from a cohort of 28,241HIV patients harbored the A431V change without any primary PI mutations. Peptide cleavage analysis has demonstrated that Valine at this position provides a better substrate [60], however, Alanine is present as the wild-type amino acid at this codon. This is thought to be because this cleavage site is among the slowest processing events in Gag [60,268]. The inclusion of the less efficient Alanine as the P2 residue maintains the order of Gag cleavage and functions as a rate-limiting site. This prevents the premature cleavage of this site, which is thought to be detrimental to virion maturation [260,262,340]. However, since resistance mutations contained within protease are thought to be detrimental to cleavage, replacement of the Alanine with Valine increases the processivity at this site and compensates for reduced enzymatic activity of the mutant protease [191,224,231]. The lack of linkage of the A431V with the mutant protease demonstrated in this study could be attributable to these viruses not being produced in the presence of selective pressure i.e. PIs. Also the association of this A431V CSM and PI mutations has only been demonstrated for clinical isolates derived from heterogeneous populations, which do not show direct linkage on the same genome. Therefore, it is possible that these mutations do not co-exist on the same genome, which would explain its removal after transfection of the clone as seen in the Pt1-(Gag+Pro) viruses. Maguire *et al.*, demonstrated that population

sequencing can potentially over emphasise the linkage seen with *gag* and protease. Population sequencing of passage experiments with APV, suggested the appearance of a double mutant, L449F and I50V. Clonal sequence data from the same selection experiment showed that only one out of three I50V containing sequences actually showed linkage with the L449F mutation [188].

The only NC mutation to be retained by either of the viruses derived from patient 1 which contained *gag,* was the M423I change seen in the Pt1-(Gag+Pro) virus. This change is rarely seen across the different sub-types and other variations at this codon are also seen, usually a Leucine. The retention of this mutation implies it may also co-evolve with the mutant protease.

The results also show that the p1/p6 cleavage site mutation, L449V appears to co-evolve with the mutant protease, this association has also been shown to occur *in vivo* [327]. This codon change normally associated with protease resistance at this site is usually an L449F, it has been proposed that the L-F change develops a sequence that may facilitate the ribosomal frame-shifting required to produce the Gag-Pol encoded proteins, protease, RT and integrase. As this mutation generates a more homopolymeric or 'slippery' sequence at the nucleotide sequence level [109], thus compensating for reduced Gag processing by the inefficient mutant protease by increasing the amount of protease [60,79,109,295]. The L449F mutation would change the F448 and L449F sequence to UUUUUU from UUUCUU, the L449V change seen at this codon in Pt1-(Gag+Pro) virus contains the less slippery UUUGUU sequence. The loss of this mutation in the Pt1-GAG virus suggests that there is selection pressure for co-evolution at this site with other *gag* mutations or with the co-expressed protease. However, the less slippery ribosomal frame-shifting sequence is maintained by mutation to the NL4.3 codon, Leucine at 449. In many of the circulating recombinant forms that are seen today such as AE, AG and AB, this codon is L449P [176], which is also a less slippery sequence motif. These viruses are produced *in vivo* by recombination with different sub-types and are thought more likely to generate the fittest phenotype from a heterogeneous population [234]. As yet most research which has examined viruses that contain the homopolymeric 'U' stretch have been unable to demonstrate that this results in an increase in viral protease production, through increased ribosomal shifting [109,231].

Cleavage at the NC/p1 and p1/p6 sites appear to be rate limiting steps in polyprotein processing [77,358] and these sites also show a high level of polymorphism, in virus isolated from treated and untreated patients [51,85,191,327]. It is thought that mutations at the rate limiting cleavage sites are more likely to occur as these sites are not optimised, such as the A431V [85]. Other sites that cleave efficiently are less able to mutate, since they are already optimised and thus, any deviation in sequence will result in a less hydrolysable substrate.

The passage of the recombinant containing only the mutant protease derived from the patient 1 virus, showed no deviation in both the *gag* and protease from the input genotype. This may be because the NL4.3 *gag* is more T-cell adapted, therefore it requires less adaptation to occur as a result of the new transfection. It has also been shown that differences in replication capacity can exist between T-cell lines when infected with the same virus [324]. The virus isolated from this patient was probably exposed to a different host cell environment, than the one that is created during *in vitro* culture. There may be some selection pressure on the *gag* as a function of adapting to a 'new' protease, which may require some level of variation or complementation as one is the substrate for the other. However, no changes were detected in the *gag* from the Pt1-Pro virus, this could be due to any adaptation being at a low level, probably below the limit of detection of the heterogeneous PCR product sequencing of the infected cell pellet. Such techniques are known to have a sensitivity of 35-50% [93,151] i.e. any population that exists at a percentage lower than 35% will not be routinely detected.

Another possibility for the lack of *gag* adaptation seen, could be that the virus derived from patient 1 which contained only the protease region also had less time in tissue to culture to adapt, as this recombinant produced detectable CPE several days earlier after transfection compared to those viruses which contained the patient derived virus *gag* region.

The changes in the recombinant virus genotypes seen over the passage period can only be attributed to deviations from the original, input genotype. In other RVA approaches, the use of heterogeneous populations means that any other sequence variations could have always been present and as a result could lead to the formation

of artificial genotypes through recombination, which are then selected for. *In vivo*, recombination and minority populations are thought to play a major role in the development of heterogeneity and resistance.[235,352]. Therefore, these systems are able to provide valuable information on the contribution of quasispecies to appearance of a fitter, resistant phenotype. The advantage of this clonal system is that it provides a means to assess the impact of a single virus genotype on the replication phenotype.

Several caveats exist to this research.

(i) All the passage data was subject to heterogeneous analysis, which may not detect or underestimate the presence of minority variants.

(ii) This also provides no information on the linkage of mutations, not only in the same protein coding region, but also on those regions or proteins more distal.

(iii) It is also unclear how much the recombinants have evolved, given more time in culture further adaptations would probably occur, resulting in the appearance of fitter genotypes.

The sequence data obtained from the first passage experiment infers that without PI selection pressure the mutant *gag* would be likely to revert to a more wild-type sequence. To abrogate this, transfections were repeated and then transferred into media containing antivirals, each infection was passaged under constant selection pressure. The transfection in the presence of antiretrovirals would potentially allow for the maintenance of the resistance mutations. However, it may also result in no virus production as the sample chosen (patient 1 derived virus) contained highly mutated gag, protease and RT coding regions. Due to the lack of PI resistance mutations and the apparent lack of replication capacity of the *gag* only virus, no growth was detected in any of the antiviral containing infections (see section 6.2.4). All the infections required multiple rounds of passage before a sustainable infection could be attained. As demonstrated by the growth rate comparisons, the Pt1-Pro virus again would appear to be the least impaired of the recombinants, this is also reflected in the lack of sequence reversions in the Pt1-Pro virus isolate, in both the protease and *gag* regions.

It may also be possible that Pt1-Gag+Pro virus was eventually able to grow better than the Pt1-Pro virus in the presence of SQV, as was demonstrated by the appearance

of CPE in this infection at passage three. Although the only comparative analysis that was carried out was by a subjective method, such as CPE scoring i.e. assigning a relative score to each culture, based on the level of CPE visible.

The sequence data derived from the transfection and passage of the viruses in the presence of antiretrovirals has shown less reversions than those from the high titre passaged stocks, this could be attributable to the different experimental conditions used. The main difference between the two methods was that the high titred stocks were generated to maximise virus production, which involved culturing with much larger numbers of cells and the cultures were maintained until full CPE was observed. The latter transfections were carried out primarily to maintain the genotype of the input recombinant, although a 'no drug' control was included, most of the cultures failed to reach the stage where extensive CPE was observed. This would imply that the transfections and subsequent culturing in the prensence of antivirals would result in a reduction in the amount of virus produced in these infections. Therefore, each infection would have a lower number of overall replication cycles in a given volume, than occurred in development of the high titre stocks.

The most reversion was again seen in the Pt1-GAG virus, however, on this occasion, all the MA mutations and the insertion were preserved. This data also showed that the CA region reverted to NL4.3 sequence, these changes may be adaptive to the MA insertion, as both these regions interact in the uncleaved Pr55 ^{Gag} polyprotein [29]. The much more extensive heterogeneity seen in the latter sequence data, i.e. the mixtures in the sequence chromatograms also implies that reversions may have been occurring, as less time had been allowed for changes to take place. The same reversion pattern was seen in the C-terminal region of *gag*, in both passage experiments, from the Pt1-GAG viruses, although with more extensive heterogeneity at the reverted sites.
The previous sequence data demonstrated less reversion occurred in the Pt1-Gag+Pro virus than in the *gag* alone, this also appears to have transpired in the second sequence analyses, as here, no changes were detected when the patient virus *gag* was cloned with the cognate protease. This again could be due to the lower number of virus replication cycles and the low detection limit of population sequencing. As already mentioned, the possibility of contamination exists, however, the inclusion of the

ampicillin detection PCR to verify that the input plasmid was not contaminating the virus infected cell pellet amplicons, provided a control for this.

Also in agreement with the initial passage data, is the lack of reversion in the Pt1-Pro virus, this maintenance of the original input vector sequence, was seen across both the *gag* and protease coding regions. However, mutations were detected in the cultures under antiviral selection. Both the AZT and IDV infections produced an L75R change in *gag*, which has been previously seen with APV passage experiments. This change also conferred replication advantages, especially when an additional p24 mutation, H219Q was also present [107]. The H219Q mutation is thought to increase replication of HIV in CyPA rich cells, by decreasing the amount of CyPA incorporated into progeny virions [106]. This mutation is polymorphic, occurring in clade F and G isolates and was also spontaneously generated in *in vitro* passage of HIV NL4.3, when L75R was also detected [107]. The appearance of the L75R mutation in the AZT infection and its more apparent homogeneity than in the IDV infection, as seen in the sequence chromatograms, could suggest that this mutation arises through an overall response to low viral replication and not as a result of PI selection. The Pt1-Pro virus displayed high level resistance to IDV and would not be resistant to any RT inhibitors, as the cognate RT was missing. Thus, AZT would have a much greater inhibitory effect on virus replication than IDV, which would apply only minimal selection pressure, therefore, if the existence of the L75R is low viral replication driven, then it would be selected for more in the AZT containing infection, thus, explaining the homogeneity seen in the sequence chromatograms.

HIV like most other retroviruses, packages two unspliced copies of genomic RNA in each progeny virion, through the ψ, located in the 5'UTR. This region has been predicted to form a highly ordered secondary structure, comprising of four RNA stem-loops that each contribute to the overall viral RNA packaging efficiency (see Figure 6.7b). Encapsidation of viral genomic RNA is a highly specific process that must distinguish unspliced viral RNA, against a high background of cellular and spliced viral RNA [42,150]. The NC protein is thought to be one of the determinants for this specificity and may also have a role in stabilisation of Ψ secondary structure [42,218,303,319].

The two encapsidated RNA copies are held together in an apparent parallel orientation, as a non-covalent dimer at the dimer inititiation site (DIS), located in SL1 of Ψ. This contains a pallindromic, GC-rich loop, that promotes Watson-Crick base pairing between the two pallindromic viral RNA sequences by forming stable, 'kissing-hairpin loop' dimers [82,215,218,310]. Incompatibilities between the pallindromic sequences have been shown to result in a 9-fold reduction in RNA packaging [215]. Although in the patient virus the DIS may be less functional through deletion of the pallindomic sequence, it is thought that secondary sites can also act as packaging sites, but function at a much lower efficiency [356]. This may explain the slow growth initially of both the recombinants containing the patient virus derived Ψ and the heterogeneous changes detected in the sequence chromatograms for the Pt1-GAG virus in the second passage data set.

Other supporting evidence that suggests the Ψ is under selection pressure during the passage experiments, are the mutations in the C-terminal region of *gag*, in both NC and p2. These two proteins have been shown to be involved with viral RNA dimerisation and encapsidation [150,216,218,284,319]. The NC contains two zinc knuckle motifs, these both bind zinc ions, in a process that facilitates the folding and stabilisation of the viral RNA-NC packaging complex and is also thought to be one of the primary events involved in the selectivity of the encapsidation process [205]. Both the zinc knuckle motifs contained a mutation from NL4.3 sequence in the original study virus clone, a K403R and M423I change. After passage, reversions back to NL4.3 sequence occurred at both these mutated sites, the K403R change occurs at a codon thought make contact with the viral RNA in the SL3-NC complex [63,244], see Figure 6.7.

It has been previously demonstrated that *trans*-packaging, (as opposed to *cis*-packaging; encapsidation from the RNA from which the Gag was translated) of Gag is a major mechanism of HIV-1 RNA encapsidation and may contribute to the rescue of *gag* mutants by wild-type virus [232]. This has also been shown to occur with restoration of replication after deletions in the SL1 loop and compensation by mutations after 18 passages in both p1 and NC [180,181]. Deletions in SL3 were also shown to be compensated for by mutations in both p2 and NC [282].

It is unclear if the mutations in the *gag* p2/NC were in response to the mutations in the Ψ and may have lead to increased encapsidation or dimerisation of viral RNA. Or if the mutations in the Ψ were as a result other mutations in *gag* p2/NC associated with compensation of PI resistant viruses. In which case this would demonstrate that PIs have pleiotropic effects even beyond those in *gag* and that they may affect packaging of viral RNA. The distal effects of PI therapy have previously been shown to influence the genetic variability of *env* and cause strong bottleneck events [229].

This chapter highlights the identification of highly mutated protease which appeared to grow efficiently despite the presence of a number of debilitating PI resistance mutations. The presence of a two amino acid insertion in the protease gene; not previously reported, may provide a means of compensation, through extension of the flap region [156] providing increased substrate recognition [161,290]. It was also shown that cloning of the cognate *gag*, reduced the replication capacity of the resistant protease when included in the recombinant virus. This observation is not in agreement with the commonly held theory of *gag* restoration of aberrant protease cleavage. Indeed the addition of mutations in *gag* to compensate for the deleterious effects of PI resistance mutations would be an elegant means of affording a small protein like protease, a greater potential for genetic variability, to produce a resistant and replication competent phenotype. However, the selection of this genotype may not have primarily been through a pathway that was able to select for the fittest species, it is conceivable that this genotype has been forced to exist as a result of fixation of certain mutations. This anomaly may also be explained by the level of reversion that is occurring in the gag, i.e. the slow replication is a result of the time taken for the clonal population to discard the growth deficient mutations and acquire more fit ones. Reduced replication of constructs with the patient 1 virus derived gag may also be due to the altered packaging signal region present in both viruses containing patient 1 virus derived gag region. This region was shown to mutate back to a sequence similar to that of wild-type in the second passage experiment. The mutated packaging signal present in the original patient derived virus contained a deletion in the palindromic, 'kissing-hairpin loop' region, which others have shown causes a substantial loss in virus infectivity [131,173,310,319]. Such maintenance of an unfit virus with multiple PI mutations and containing a two amino acid insertion at codon 35 in protease, has previously been

reported [245]. The aforementioned study demonstrated that the protease insertion when present in a PI sensitive background, grew as well as the wild-type control virus. However, virus containing the resistance mutations and the insertion, demonstrated much poorer growth than the analogous virus minus the insertion. The virus that was isolated from the study patient had been exposed to extensive therapy and although the *gag* would appear to be detrimental to the growth ability of the virus, its existence may be due to the enhanced PI resistance, which in some cases results in IC_{50} shifts that may exceed the nadir drug levels present in the patient plasma.

The results presented here raise more questions and avenues of research which are discussed further in the next chapter. Some relevant points relating to this chapter are also discussed below.

(i) Other methods to determine the start input of the growth curve, such as p24, RNA viral load and RT activity could also have been used. Although cell viability gives a measure of infectious virus. Others have shown that p24 antigen concentration is not always reflective of the level of infectious virus contained in the stock [28].

(ii) Although the reversions that were seen in the virus stocks prevented any further continuation of the competition or dual infection study, these types of experiment are thought more likely to determine small fitness differences between viruses. This would have allowed a greater dissection of the differences between the Pt1-Gag+Pro and the Pt1-Gag viruses, which showed similar profiles in the comparison of growth rates experiment. Such an experiment would also have helped to determine if the Pt1-Pro virus was also growth impaired compared to the wild-type stocks, NL4.3 and X-PSX.

(iii) The analysis of the sequence data after passage, in both experiments was carried out using the infected cell pellet as the source for the viral integrated DNA. This would represent the viral genome that has successfully completed entry, reverse transcription and integration. The analysis used a heterogeneous PCR product to determine the viral sequence, a more complete and relevant analysis would have been achieved if the sequence data were generated from clonal analysis, although, this would have required much more time to complete.

Chapter 7 Conclusions and Future Work

7.1 Conclusions

The initial aim of this study was to develop a cloning vector that could be used to determine the antiviral sensitivity of recombinant viruses containing the *gag*, protease and RT coding regions derived from antiviral therapy treated individuals. This was initially undertaken through the use of a single cycle replication system. Due to the lack of infectious particles generated using the pseudotyped system, when recombinants containing *gag* derived from antiviral therapy treated individuals were generated, a live virus system was constructed. The chosen study patient virus sample was shown to contain a novel two amino acid insertion in the protease gene, which demonstrated high resistance to several PIs tested and was also shown to replicate efficiently. The recombinants containing the *gag* from the patient isolated virus, were shown to cause a reduction in growth capacity of the viruses, even when cloned with the cognate protease. However, the virus containing the patient virus derived *gag* and protease, demonstrated a 2-fold increase in resistance to some of the PIs tested, compared to the virus in which only the mutant protease was cloned. Recombinants which included the *gag* region showed multiple reversions in *gag* and the reversion pattern was shown to be dependant on whether or not the recombinant contained the highly resistant protease. Therefore, the data suggests the emergence and eventual dominance of an unfit virus, with respect to *gag*, under drug selection pressure. The *gag* persists even though alone it provides no resistance and replacement of the patient virus derived *gag* region with that of wild-type, results in a virus that is more fit but slightly less resistant.

These observations raise the question, how does such an unfit virus come to exist as the majority species? The restoration of growth impairment may occur through recombination with other genomes present as minority species. Recombination is common to all retroviruses, primarily due to persistent breaks during reverse transcription, which can lead to template switching between the two genomic RNA strands [235,286]. Heterozygous virions can also be produced by cells that are co-infected by two separate proviruses [58].

The role of recombination in repair of a dysfunctional genome could possibly be more likely to occur in *gag*, due to its multi-functional role within the HIV Life cycle. Gag performs functions either as correctly cleaved proteins or as the uncleaved polyprotein [29,94,95]. Muller *et al.* constructed a virus containing the GFP gene located in *gag*, at the C-terminus of MA. The GFP viruses displayed greatly reduced infectivity and demonstrated an accumulation of budding structures at the plasma membrane when visualised by EM [219]. Co-transfection of the GFP virus and wild-type resulted in viruses displaying wild-type protein composition, morphology, and infectivity. However, serial passage of the co-transfected population resulted in the rapid loss of the modified variant. This dominant interference mechanism has been proposed for the inhibition of wild-type virus by HIV *gag* mutants by several groups [174,315] and rescue of *gag* deletion mutants by wild-type *gag* has also been demonstrated [33].

Another possibility is the restoration of growth by the patient derived virus *gag*, may occur through pleiotropic means, this was shown with *gag* and *Env*, where virus replication capacity of MA mutants was restored by truncations in the envelope glycoprotein gp41 [193]. The association of $p6^{Gag}$ with Vpr is also another example of the pleiotropic effect of *gag*, efficient incorporation of Vpr into virions has been shown to require the C-terminal *gag* protein, p6 [162,252]. The transframe protein (p6*), has also been reported to associate with Nef in progeny virions in producer cells. This Gag-Pol and Nef interaction may allow Nef to target HIV-1 budding to lipid rafts [50]. It has also been shown that mutation of Nef results in a loss of infectivity in the molecular clone virus NL4.3, but not in a closely related virus, LAI. These clones share 93.2% and 96.5% sequence homology in *gag* and *Pol* respectively and only four amino acid changes in the *Nef* gene. Nonetheless, the C-terminal region of *gag* was shown to be responsible for the diversity seen in infectivity of the *Nef* mutants. This suggests that a mechanism exists by which the Gag region regulates the viral infectivity in combination with Nef [239].

The virus genotype isolated from the study patient was selected for under HAART and thus, was either already adapted to this environment or was still in the process of adapting. This may explain the low percentage of the K20R containing protease, which was found in only 1 clone. This mutation has been shown to compensate for

growth restoration of PI resistant viruses [48,213,230] and may represent a low percentage, emerging population isolated during a transition state to a fit species.

The addition of antivirals results in the selection of an evolutionary pathway that would previously not have been selected for due to the detrimental effects on virus growth of the resistance mutations. This pathway now allows for the selection of a combination of mutations that results in a species fitter than the original wild-type strain [230]. This mechanism is also proposed for illustrating the molecular basis for the emergence of the AZT resistance associated mutations, T215Y and Q151M. Dual infection experiments determined that the Q151M demonstrated greater replicative capacity than the T215Y in both the presence and absence of dideoxynucleosides. Both mutations require double nucleotide substitutions, yet the Q151M is found only in patients after long-term therapy with multiple RT inhibitors. Researchers demonstrated that the basis for lack of selection of the Q151M was the poor growth abilities of this viruses intermediates, whereas those for T215Y were fitter [164].

The vector system developed in this study is not high throughput and so would not be ideal for routine evaluation of clinical isolates. This is primarily due to the difficulties during cloning and the other difficulties a associated with virus production. However, the latter problem was in part due to the difficult sample selected for evaluation i.e. the patient 1 derived virus had extensive prior exposure to multiple antivirals and was thus highly mutated and resulted in low level replicating virus amounts, if the gag region was included in the recombinant. However, the general aim was to develop a vector system that could be used to evaluate the effects of the gag, protease and RT coding regions either separately or in combination. This has successfully been demonstrated, albeit with a very small number of samples.

This study demonstrates that the inclusion of *gag* in determining phenotypic sensitivity to antivirals should be routinely undertaken, as an increased resistance profile an be demonstrated when this region is considered. Although, p2/NC is included in some of the antiviral assay systems [257], these may require the addition of the N-terminal portion of *gag* to fully determine the resistance capacity of viruses in the presence PIs. This is also of paramount importance if fitness is also to be

examined and this study also suggests that regions upstream of the *gag* initiation codon, which are involved in packaging of viral RNA, should also be considered.

7.2 Further Work

The study has focused primarily on the development of a vector, cloning of patient virus derived *gag*, protease and RT sequences and assessment of the contribution of these coding regions to virus growth capacity and resistance phenotype. The identification of a highly mutated genotype with multiple insertions and an apparent dysfunctional Ψ raises other avenues of research that deserve further and more detailed examination, such as:

i. to facilitate the cloning using the pNL4.3X-PSX vector, a series of small oligonucleotide linkers were constructed, but due to time constraints this was not completed. These linkers would allow the pre-digested vector i.e. minus the *gag*, protease and RT regions, to be used directly, thus, negating the need to gel purify the vector prior to cloning and preventing uncut contamination. Linkers would also enable the use of a PCR screen to differentiate between colonies containing the region of interest and those not containing any insert.

ii. the requirement to determine titres or viral input more accurately, as these affect any determinations of replication capacity, although a reasonable number of replicates may abrogate this. Both systems used in this study to calculate titres were measures of live replication and or infectivity. An alternative would be p24. However, this is also an indirect measure and may be affected by Gag processing defects. Discordances between measures such as p24 and viral infectivity have previously been reported [8,248]. Other studies using viral RNA level have shown that the ribonucleoside analog, ribavirin inhibits virus production without substantially affecting RNA synthesis [53]. Another alternative would be the development of a single cycle system with the pNL4.3X-PSX vector, by first removing the *Env* and then pseudotyping with the pMDG vector, encoding the VSV-G envelope protein. This would produce a more accurate measure of virus production due to the limited number of replication cycles [272] than the live systems used in this study.

Although, a potential drawback to this approach could be very low titres if the *gag* region is included in the recombinant, as was seen in this study.

iii. the dissection of the effects of the two amino acid insertion in protease. This would involve the construction of a series of site directed mutants with and without PI resistance mutations and *gag* mutations. A similar insertion to 35QN, namely 35TN, was shown to be detrimental in the presence of PI resistance mutations. Although the authors suggest the affect of insertions is dependant on the amino acid insertion and the drug resistance mutations present [245].

iv. continued passing of the recombinants in the presence or absence of PIs. This would determine the divergence of the genotypes under antiviral selection and establish which mutations in *gag* are drug resistance associated. In the second passage experiment at passage three, the Pt1-Pro virus produced higher $TCID_{50}$s/ml in the IDV and RTV infection than in the 'no drug' control (data not shown). This may represent increased replication in the presence of the these two PIs to which the virus had prolonged previous exposure. A virus containing highly mutated protease and an insertion in *gag*, also in helix 5 of MA, was shown to confer NFV-dependent replication enhancement phenotype on NL4-3, this phenotype only required protease and MA. Also, at passage three the Pt1-Gag+Pro virus produced higher $TCID_{50}$/ml than the Pt1-Pro virus, in the infections passaged in the absence of antivirals. This virus may have become more fit that the Pt1-Pro virus, due to cloning with the cognate protease.

v. continuation of the competition growth studies and initiation of dual infections between the Pt1-(GAG) only virus vs. Pt1-(Gag+Pro) virus to determine the effect of cloning the cognate protease on the order of replication capacity.

vi. the implications of RNA packaging restrictions due to compensatory mutations in *gag*, primarily in p2 and NC, as a result of the acquisition of PI resistance mutations. Although the frequency of mutations in the Ψ in viruses from patients on PIs is not known, as this region is very rarely, if ever considered.

Another aim of this study is to submit this work for publication. This work could be divided into several areas depending on the focus of research for publication.

(i) A report primarily on the development of the vector system, this would require more patient samples to demonstrate the versatility of the vector system. The amplification conditions may also require optimisation to include more samples. The remaining downstream steps, such as cloning, production of the high titre virus stocks and drug susceptibility assays do not require any further improvement, other than those previously highlighted, such as an alternative to the *Xba I* cloning site. Another alteration could be the adaptation of this system to one similar to a single cycle assay, as mentioned in section 7.2 (ii). This would mean the analysis of replication capacity would be similar to that used by several diagnostic laboratories e.g. Monogram Biosciences, ViroLogic and Virco.

(ii) One of the most interesting aspects of the patient 1 derived virus is the identification of a novel two amino acid insertion in protease. Further evaluation of this patient derived virus would require the production of mutants containing combinations of the 35QN insert, with and with other PI mutations and also clones with and without the cognate gag. In order to determine which cloned viruses grow best would require more growth studies such as competition growth experiments and also drug sensitivity determinations to ascertain if growth is compromised for decreased drug susceptibility. The evaluation of mutant and wild-type gag cleavage by the 35QN containing protease clones would also be worth investigating using western blots analysis.

(iii) Possibly the most interesting finding was the changes in the ψ, at the 'kissing-hairpin loop' region in the Pt1-GAG clone after only three passages. To explore this further would require examining other heavily PI and RTI treated patients. A control group for this study would be patient derived virus samples which had only been exposed to NRTIs and NNRTIs, this is because others have shown that insertions and deletions are associated with RTI treatment.

The observation in this study of the mutated packaging signal is thought to be related primarily to PI treatment and the development of compensatory mutations in NC, which require changes in the ψ in order to facilitate both resistance and RNA genome packaging. Although this latter point has not been demonstrated, it could be shown by cloning the ψ from the patient derived virus into the pCSGW vector and comparing it to one conating a wild-type ψ. The vector pCSGW was used in this report to package the HIV Gag-Pol region in order to produce the VSV-G pseudotyped viruses. Differences in ψ function could be measured using FACS analysis to measure GFP production in the pseudotyped viruses.

The development of antiretrovirals to several HIV protein targets, RT, protease, Env, integrase [172,201], RNase H [159] and *gag* [179,307], provides the future potential for a chemotherapeutic arsenal that could possibly force HIV into an unfavourable mutational pathway. However, this " reduced fitness approach" may only be feasible for relatively newly acquired infections, as within untreated infected individuals there are virions that encode every possible single, double, and triple amino acid substitution [8]. The use of a more targeted approach to treatment at a very early stage of the infection to prevent the appearance of such highly resistant strains which are fully compensated could also be enhanced by fitness reductions through Muller's Ratchet [8], the stochastic accumulation of deleterious mutations. This occurs under strong selective pressure or in populations undergoing repetitive bottlenecks. High mutation frequencies will result in the accumulation of deleterious mutations, which can become fixed, reducing overall fitness. As the population size is low [355] and relatively homogenous [16,71,328], this reduces the chance of escape from Muller's Ratchet through recombination [8,353]. Development of such 'fitness targeted' regimens could also be improved by selection of compounds leading to non-synergistic resistance pathways, such as AZT and 3TC. Also the PI, Atazanavir, which selects for an I50L mutation in protease, which has been shown to cause hypersusceptibility to other currently approved PIs [346]. One caveat to this approach at present is the lack of viral targets for antiretrovirals. It would require more new compounds to become available that can be directed at different HIV proteins or regions. and also an increase in our understanding of resistance and fitness.

However, the evidence would suggest that HIV is be able to correct any deleterious phenotypes through the myriad of compensatory mutations and mechanisms alluded to in this study.

Chapter 8 Reference List

1. **Adachi, A., H. E. Gendelman, S. Koenig, T. Folks, R. Willey, A. Rabson, and M. A. Martin**. 1986. Production of acquired immunodeficiency syndrome-associated retrovirus in human and nonhuman cells transfected with an infectious molecular clone. J. Virol. **59**:284-291.

2. **Adamson, C. S., S. D. Ablan, I. Boeras, R. Goila-Gaur, F. Soheilian, K. Nagashima, F. Li, K. Salzwedel, M. Sakalian, C. T. Wild, and E. O. Freed**. 2006. In vitro resistance to the human immunodeficiency virus type 1 maturation inhibitor PA-457 (Bevirimat). J. Virol. **80**:10957-10971.

3. **Adamson, C. S. and I. M. Jones**. 2004. The molecular basis of HIV capsid assembly--five years of progress. Rev. Med. Virol. **14**:107-121.

4. **Ali, A. and O. O. Yang**. 2006. A novel small reporter gene and HIV-1 fitness assay. J. Virol. Methods **133**:41-47.

5. **Althaus, C. L. and S. Bonhoeffer**. 2005. Stochastic interplay between mutation and recombination during the acquisition of drug resistance mutations in human immunodeficiency virus type 1. J. Virol. **79**:13572-13578.

6. **Ambrose, Z., V. Boltz, S. Palmer, J. M. Coffin, S. H. Hughes, and V. N. KewalRamani**. 2004. In vitro characterization of a simian immunodeficiency virus-human immunodeficiency virus (HIV) chimera expressing HIV type 1 reverse transcriptase to study antiviral resistance in pigtail macaques. J. Virol. **78**:13553-13561.

7. **Ambrose, Z., J. G. Julias, P. L. Boyer, V. N. KewalRamani, and S. H. Hughes**. 2006. The level of reverse transcriptase (RT) in human immunodeficiency virus type 1 particles affects susceptibility to nonnucleoside RT inhibitors but not to lamivudine. J. Virol. **80**:2578-2581.

8. **Anderson, J. P., R. Daifuku, and L. A. Loeb**. 2004. Viral error catastrophe by mutagenic nucleosides. Annu. Rev. Microbiol. **58**:183-205.

9. **Arts, E. J., M. E. Quinones-Mateu, J. L. Albright, J. P. Marois, C. Hough, Z. Gu, and M. A. Wainberg**. 1998. 3'-Azido-3'-deoxythymidine (AZT) mediates cross-resistance to nucleoside analogs in the case of AZT-resistant human immunodeficiency virus type 1 variants. J. Virol. **72**:4858-4865.

10. **Auewarakul, P., P. Wacharapornin, S. Srichatrapimuk, S. Chutipongtanate, and P. Puthavathana**. 2005. Uncoating of HIV-1 requires cellular activation.. Virology **337**:93-101.

11. **Back, N. K., M. Nijhuis, W. Keulen, C. A. Boucher, B. O. Oude Essink, A. B. van Kuilenburg, A. H. van Gennip, and B. Berkhout**. 1996. Reduced replication of 3TC-resistant HIV-1 variants in primary cells due to a processivity defect of the reverse transcriptase enzyme. EMBO J. **15**:4040-4049.

12. **Bally, F., R. Martinez, S. Peters, P. Sudre, and A. Telenti**. 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**:1209-1213.

13. **Barbour, J. D., F. M. Hecht, T. Wrin, M. R. Segal, C. A. Ramstead, T. J. Liegler, M. P. Busch, C. J. Petropoulos, N. S. Hellmann, J. O. Kahn, and R. M. Grant**. 2004. Higher CD4+ T cell counts associated with low viral pol replication capacity among treatment-naive adults in early HIV-1 infection. J. Infect. Dis. **190**:251-256.

14. **Barré-Sinoussi F, Chermann JC, Rey F, Nugeyre MT, Chamaret S, Gruest J, Dauguet C, Axler-Blin C, Vézinet-Brun F, Rouzioux C, Rozenbaum W, Montagnier L.** 1983. Isolation of a T-lymphotropic retrovirus from a patient at risk for acquired immune deficiency syndrome (AIDS). Science **220**:868-71.

15. **Beerenwinkel, N., M. Daumer, T. Sing, J. Rahnenfuhrer, T. Lengauer, J. Selbig, D. Hoffmann, and R. Kaiser**. 2005. Estimating HIV evolutionary pathways and the genetic barrier to drug resistance. J. Infect. Dis. **191**:1953-1960.

16. **Bergstrom, C. T., P. McElhany, and L. A. Real**. 1999. Transmission bottlenecks as determinants of virulence in rapidly evolving pathogens. Proc. Natl. Acad. Sci. U. S. A **96**:5095-5100.

17. **Besnier, C., Y. Takeuchi, and G. Towers**. 2002. Restriction of lentivirus in monkeys. Proc. Natl. Acad. Sci. U. S. A **99**:11920-11925.

18. **Bezemer, D., A. de Ronde, M. Prins, K. Porter, R. Gifford, D. Pillay, B. Masquelier, H. Fleury, F. Dabis, N. Back, S. Jurriaans, and H. L. van der**. 2006. Evolution of transmitted HIV-1 with drug-resistance mutations in the absence of therapy: effects on CD4+ T-cell count and HIV-1 RNA load. Antivir. Ther. **11**:173-178.

19. **Bleiber, G., M. Munoz, A. Ciuffi, P. Meylan, and A. Telenti**. 2001. Individual contributions of mutant protease and reverse transcriptase to viral infectivity, replication, and protein maturation of antiretroviral drug-resistant human immunodeficiency virus type 1. J. Virol. **75**:3291-3300.

20. **Boden, D. and M. Markowitz**. 1998. Resistance to human immunodeficiency virus type 1 protease inhibitors. Antimicrob. Agents Chemother. **42**:2775-2783.

21. **Boucher, C. A., E. O'Sullivan, J. W. Mulder, C. Ramautarsing, P. Kellam, G. Darby, J. M. Lange, J. Goudsmit, and B. A. Larder**. 1992. Ordered appearance of zidovudine resistance mutations during treatment of 18 human immunodeficiency virus-positive subjects. J. Infect. Dis. **165**:105-110.

22. **Boyer, P. L., T. Imamichi, S. G. Sarafianos, E. Arnold, and S. H. Hughes**. 2004. Effects of the Delta67 complex of mutations in human immunodeficiency virus type 1 reverse transcriptase on nucleoside analog excision. J. Virol. **78**:9987-9997.

23. **Boyer, P. L., S. G. Sarafianos, E. Arnold, and S. H. Hughes**. 2002. The M184V mutation reduces the selective excision of zidovudine 5'-monophosphate (AZTMP) by the reverse transcriptase of human immunodeficiency virus type 1. J. Virol. **76**:3248-3256.

24. **Brann, T. W., R. L. Dewar, M. K. Jiang, A. Shah, K. Nagashima, J. A. Metcalf, J. Falloon, H. C. Lane, and T. Imamichi**. 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**:6136-6145.

25. **Brik, A. and C. H. Wong**. 2003. HIV-1 protease: mechanism and drug discovery. Org. Biomol. Chem. **1**:5-14.

26. **Briz, V., E. Poveda, and V. Soriano**. 2006. HIV entry inhibitors: mechanisms of action and resistance pathways. J. Antimicrob. Chemother. **57**:619-627.

27. **Brown, A. J. and D. D. Richman**. 1997. HIV-1: gambling on the evolution of drug resistance? Nat. Med. **3**:268-271.

28. **Brown, B. K., J. M. Darden, S. Tovanabutra, T. Oblander, J. Frost, E. Sanders-Buell, M. S. de Souza, D. L. Birx, F. E. McCutchan, and V. R. Polonis**. 2005. Biologic and genetic characterization of a panel of 60 human immunodeficiency virus type 1 isolates, representing

clades A, B, C, D, CRF01_AE, and CRF02_AG, for the development and assessment of candidate vaccines. J. Virol. **79**:6089-6101.

29. **Bukrinskaya, A.** 2007. HIV-1 matrix protein: A mysterious regulator of the viral life cycle. Virus Res. **124**:1-11.

30. **Bukrinskaya, A. G.** 2004. HIV-1 assembly and maturation. Arch. Virol. **149**:1067-1082.

31. **Campbell, T. B., K. Schneider, T. Wrin, C. J. Petropoulos, and E. Connick**. 2003. Relationship between in vitro human immunodeficiency virus type 1 replication rate and virus load in plasma. J. Virol. **77**:12105-12112.

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

33. **Carriere, C., B. Gay, N. Chazal, N. Morin, and P. Boulanger**. 1995. Sequence requirements for encapsidation of deletion mutants and chimeras of human immunodeficiency virus type 1 Gag precursor into retrovirus-like particles. J. Virol. **69**:2366-2377.

34. **Carteau, S., S. C. Batson, L. Poljak, J. F. Mouscadet, H. de Rocquigny, J. L. Darlix, B. P. Roques, E. Kas, and C. Auclair**. 1997. Human immunodeficiency virus type 1 nucleocapsid protein specifically stimulates Mg2+-dependent DNA integration in vitro. J. Virol. **71**:6225-6229.

35. **CDC**. 1982. Current Trends Update on Acquired Immune Deficiency Syndrome (AIDS) -- United States. Morbidity and Mortality Weekly **31**:507-508.

36. **Centlivre, M., P. Sommer, M. Michel, F. R. Ho Tsong, S. Gofflo, J. Valladeau, N. Schmitt, S. Wain-Hobson, and M. Sala**. 2006. The HIV-1 clade C promoter is particularly well adapted to replication in the gut in primary infection. AIDS **20**:657-666.

37. **Chin, M. P., T. D. Rhodes, J. Chen, W. Fu, and W. S. Hu**. 2005. Identification of a major restriction in HIV-1 intersubtype recombination. Proc. Natl. Acad. Sci. U. S. A **102**:9002-9007.

38. **Chow, Y. K., M. S. Hirsch, D. P. Merrill, L. J. Bechtel, J. J. Eron, J. C. Kaplan, and R. T. D'Aquila**. 1993. Use of evolutionary limitations of HIV-1 multidrug resistance to optimize therapy. Nature **361**:650-654.

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

40. **Clark S, Calef C, and Mellors J**. 2005. Mutations in Retroviral Genes Associated with Drug Resistance. HIV Sequence Compendium 2005., p. 80-175. HIV Sequence Compendium 2005. Theoretical Biology and Biophysics Group, Los Alamos National Laboratory, Los Alamos, NM. LA-UR 06-0680.

41. **Clavel, F., D. Guetard, F. Brun-Vezinet, S. Chamaret, M. A. Rey, M. O. Santos-Ferreira, A. G. Laurent, C. Dauguet, C. Katlama, C. Rouzioux, and .** 1986. Isolation of a new human retrovirus from West African patients with AIDS. Science **233**:343-346.

42. **Clever, J. L., D. Miranda, Jr., and T. G. Parslow**. 2002. RNA structure and packaging signals in the 5' leader region of the human immunodeficiency virus type 1 genome. J. Virol. **76**:12381-12387.

43. **Cocchi, F., A. L. DeVico, A. Garzino-Demo, S. K. Arya, R. C. Gallo, and P. Lusso**. 1995. Identification of RANTES, MIP-1 alpha, and MIP-1 beta as the major HIV-suppressive factors produced by CD8+ T cells. Science **270**:1811-1815.

44. **Coffin J.,Haase A.,Levy J.A.,Montagnier L.,Oroszlan S.,Teich N.,Temin H.,Toyoshima K.,Varmus H.,Vogt P.,Weiss R.A.** 1986. What to call the AIDS virus? Nature **321**:10.

45. **Coffin, J. M.** 1995. HIV population dynamics in vivo: implications for genetic variation, pathogenesis, and therapy. Science **267**:483-489.

46. **Coffin, J. M., S. H. Hughes, H. Varmus, National Center for Biotechnology Information (, and National Institutes of Health (**. 1997. Retroviruses. Cold Spring Harbor, NY : Cold Spring Harbor Laboratory Press .

47. **Colman, P. M. and M. C. Lawrence**. 2003. The structural biology of type I viral membrane fusion.. Nat. Rev. Mol. Cell Biol. **4**:309-319.

48. **Condra, J. H., D. J. Holder, W. A. Schleif, O. M. Blahy, R. M. Danovich, L. J. Gabryelski, D. J. Graham, D. Laird, J. C. Quintero, A. Rhodes, H. L. Robbins, E. Roth, M. Shivaprakash, T. Yang, J. A. Chodakewitz, P. J. Deutsch, R. Y. Leavitt, F. E. Massari, J. W. Mellors, K. E. Squires, R. T. Steigbigel, H. Teppler, and E. A. Emini**. 1996. Genetic correlates of in vivo viral resistance to indinavir, a human immunodeficiency virus type 1 protease inhibitor. J. Virol. **70**:8270-8276.

49. **Condra, J. H., W. A. Schleif, O. M. Blahy, L. J. Gabryelski, D. J. Graham, J. C. Quintero, A. Rhodes, H. L. Robbins, E. Roth, M. Shivaprakash, and .** 1995. In vivo emergence of HIV-1 variants resistant to multiple protease inhibitors. Nature **374**:569-571.

50. **Costa, L. J., Y. H. Zheng, J. Sabotic, J. Mak, O. T. Fackler, and B. M. Peterlin**. 2004. Nef binds p6* in GagPol during replication of human immunodeficiency virus type 1. J. Virol. **78**:5311-5323.

51. **Cote, H. C., Z. L. Brumme, and P. R. Harrigan**. 2001. Human immunodeficiency virus type 1 protease cleavage site mutations associated with protease inhibitor cross-resistance selected by indinavir, ritonavir, and/or saquinavir. J. Virol. **75**:589-594.

52. **Craig, J. C., I. B. Duncan, D. Hockley, C. Grief, N. A. Roberts, and J. S. Mills**. 1991. Antiviral properties of Ro 31-8959, an inhibitor of human immunodeficiency virus (HIV) proteinase. Antiviral Res. **16**:295-305.

53. **Crotty, S., D. Maag, J. J. Arnold, W. Zhong, J. Y. Lau, Z. Hong, R. Andino, and C. E. Cameron**. 2000. The broad-spectrum antiviral ribonucleoside ribavirin is an RNA virus mutagen. Nat. Med. **6**:1375-1379.

54. **Cullen, B. R.** 1991. Regulation of HIV-1 gene expression. FASEB J. **5**:2361-2368.

55. **D'Cruz, O. J. and F. M. Uckun**. 2006. Dawn of non-nucleoside inhibitor-based anti-HIV microbicides. J. Antimicrob. Chemother. **57**:411-423.

56. **Daar, E. S. and D. D. Richman**. 2005. Confronting the emergence of drug-resistant HIV type 1: impact of antiretroviral therapy on individual and population resistance. AIDS Res. Hum. Retroviruses **21**:343-357.

57. **Dalgleish, A. G., P. C. Beverley, P. R. Clapham, D. H. Crawford, M. F. Greaves, and R. A. Weiss**. 1984. The CD4 (T4) antigen is an essential component of the receptor for the AIDS retrovirus. Nature **312**:763-767.

58. **Dang, Q., J. Chen, D. Unutmaz, J. M. Coffin, V. K. Pathak, D. Powell, V. N. KewalRamani, F. Maldarelli, and W. S. Hu**. 2004. Nonrandom HIV-1 infection and double infection via direct and cell-mediated pathways. Proc. Natl. Acad. Sci. U. S. A **101**:632-637.

59. **Darwin, C.** 1860. On the origin of the species by means of natural selection : or, The preservation of favoured races in the struggle for life. London : John Murray.

60. **Dauber, D. S., R. Ziermann, N. Parkin, D. J. Maly, S. Mahrus, J. L. Harris, J. A. Ellman, C. Petropoulos, and C. S. Craik**. 2002. Altered substrate specificity of drug-resistant human immunodeficiency virus type 1 protease. J. Virol. **76**:1359-1368.

61. **De Clercq, E.** 1998. The role of non-nucleoside reverse transcriptase inhibitors (NNRTIs) in the therapy of HIV-1 infection. Antiviral Res. **38**:153-179.

62. **De Clercq, E.** 2004. Non-nucleoside reverse transcriptase inhibitors (NNRTIs): past, present, and future. Chem. Biodivers. **1**:44-64.

63. **De Guzman, R. N., Z. R. Wu, C. C. Stalling, L. Pappalardo, P. N. Borer, and M. F. Summers**. 1998. Structure of the HIV-1 nucleocapsid protein bound to the SL3 psi-RNA recognition element. Science **279**:384-388.

64. **de Mendoza, C., O. Gallego, and V. Soriano**. 2002. Mechanisms of resistance to antiretroviral drugs--clinical implications. AIDS Rev. **4**:64-82.

65. **De Meyer, S., H. Azijn, D. Surleraux, D. Jochmans, A. Tahri, R. Pauwels, P. Wigerinck, and M. P. de Bethune**. 2005. TMC114, a novel human immunodeficiency virus type 1 protease inhibitor active against protease inhibitor-resistant viruses, including a broad range of clinical isolates. Antimicrob. Agents Chemother. **49**:2314-2321.

66. **de Oliveira, T., S. Engelbrecht, v. R. Janse, M. Gordon, K. Bishop, M. J. zur, S. W. Barnett, and S. Cassol**. 2003. Variability at human immunodeficiency virus type 1 subtype C protease cleavage sites: an indication of viral fitness? J. Virol. **77**:9422-9430.

67. **Deeks, S. G.** 2003. Treatment of antiretroviral-drug-resistant HIV-1 infection. Lancet **362**:2002-2011.

68. **Deeks, S. G., J. Lu, R. Hoh, T. B. Neilands, G. Beatty, W. Huang, T. Liegler, P. Hunt, J. N. Martin, and D. R. Kuritzkes**. 2007. Interruption of enfuvirtide in HIV-1 infected adults with incomplete viral suppression on an enfuvirtide-based regimen. J. Infect. Dis. **195**:387-391.

69. **DeJesus, E., D. Berger, M. Markowitz, C. Cohen, T. Hawkins, P. Ruane, R. Elion, C. Farthing, L. Zhong, A. K. Cheng, D. McColl, and B. P. Kearney**. 2006. Antiviral activity, pharmacokinetics, and dose response of the HIV-1 integrase inhibitor GS-9137 (JTK-303) in treatment-naive and treatment-experienced patients. J. Acquir. Immune. Defic. Syndr. **43**:1-5.

70. **Delaugerre, C., L. Morand-Joubert, M. L. Chaix, O. Picard, A. G. Marcelin, V. Schneider, A. Krivine, A. Compagnucci, C. Katlama, P. M. Girard, and V. Calvez**. 2004. Persistence of multidrug-resistant HIV-1 without antiretroviral treatment 2 years after sexual transmission. Antivir. Ther. **9**:415-421.

71. **Delwart, E., M. Magierowska, M. Royz, B. Foley, L. Peddada, R. Smith, C. Heldebrant, A. Conrad, and M. Busch**. 2002. Homogeneous quasispecies in 16 out of 17 individuals during very early HIV-1 primary infection. AIDS **16**:189-195.

72. **Delwart, E. L., H. W. Sheppard, B. D. Walker, J. Goudsmit, and J. I. Mullins**. 1994. Human immunodeficiency virus type 1 evolution in vivo tracked by DNA heteroduplex mobility assays. J. Virol. **68**:6672-6683.

73. **Delwart, E. L., E. G. Shpaer, J. Louwagie, F. E. McCutchan, M. Grez, H. Rubsamen-Waigmann, and J. I. Mullins**. 1993. Genetic relationships determined by a DNA heteroduplex mobility assay: analysis of HIV-1 env genes. Science **262**:1257-1261.

74. **Detsika, M. G., B. Chandler, S. H. Khoo, C. Winstanley, P. Cane, D. J. Back, and A. Owen**. 2007. Detection and quantification of minority HIV isolates harbouring the D30N mutation by real-time PCR amplification. J. Antimicrob. Chemother. **60**:881-884.

75. **Dismuke, D. J. and C. Aiken**. 2006. Evidence for a functional link between uncoating of the human immunodeficiency virus type 1 core and nuclear import of the viral preintegration complex. J. Virol. **80**:3712-3720.

76. **Dorr, P., M. Westby, S. Dobbs, P. Griffin, B. Irvine, M. Macartney, J. Mori, G. Rickett, C. Smith-Burchnell, C. Napier, R. Webster, D. Armour, D. Price, B. Stammen, A. Wood, and M. Perros**. 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**:4721-4732.

77. **Doyon, L., G. Croteau, D. Thibeault, F. Poulin, L. Pilote, and D. Lamarre**. 1996. Second locus involved in human immunodeficiency virus type 1 resistance to protease inhibitors. J. Virol. **70**:3763-3769.

78. **Dull, T., R. Zufferey, M. Kelly, R. J. Mandel, M. Nguyen, D. Trono, and L. Naldini**. 1998. A third-generation lentivirus vector with a conditional packaging system. J. Virol. **72**:8463-8471.

79. **Dulude, D., M. Baril, and L. Brakier-Gingras**. 2002. Characterization of the frameshift stimulatory signal controlling a programmed -1 ribosomal frameshift in the human immunodeficiency virus type 1. Nucleic Acids Res. **30**:5094-5102.

80. **El Safadi, Y., V. Vivet-Boudou, and R. Marquet**. 2007. HIV-1 reverse transcriptase inhibitors. Appl. Microbiol. Biotechnol.

81. **Emerman, M.** 2006. How TRIM5alpha defends against retroviral invasions. Proc. Natl. Acad. Sci. U. S. A **103**:5249-5250.

82. **Ennifar, E. and P. Dumas**. 2006. Polymorphism of bulged-out residues in HIV-1 RNA DIS kissing complex and structure comparison with solution studies. J. Mol. Biol **356**:771-782.

83. **Erice, A., D. L. Mayers, D. G. Strike, K. J. Sannerud, F. E. McCutchan, K. Henry, and H. H. Balfour, Jr.** 1993. Brief report: primary infection with zidovudine-resistant human immunodeficiency virus type 1. N. Engl. J. Med. **328**:1163-1165.

84. **Erickson, J. W. and S. K. Burt**. 1996. Structural mechanisms of HIV drug resistance. Annu. Rev. Pharmacol. Toxicol. **36**:545-571.

85. **Feher, A., I. T. Weber, P. Bagossi, P. Boross, B. Mahalingam, J. M. Louis, T. D. Copeland, I. Y. Torshin, R. W. Harrison, and J. Tozser**. 2002. Effect of sequence polymorphism and drug resistance on two HIV-1 Gag processing sites. Eur. J. Biochem. **269**:4114-4120.

86. **Fellay, J., K. V. Shianna, D. Ge, S. Colombo, B. Ledergerber, M. Weale, K. Zhang, C. Gumbs, A. Castagna, A. Cossarizza, A. Cozzi-Lepri, A. De Luca, P. Easterbrook, P. Francioli, S. Mallal, J. Martinez-Picado, J. M. Miro, N. Obel, J. P. Smith, J. Wyniger, P. Descombes, S. E. Antonarakis, N. L. Letvin, A. J. McMichael, B. F. Haynes, A. Telenti, and D. B. Goldstein**. 2007. A whole-genome association study of major determinants for host control of HIV-1.. Science **317**:944-947.

87. **Feng, Y., C. C. Broder, P. E. Kennedy, and E. A. Berger**. 1996. HIV-1 entry cofactor: functional cDNA cloning of a seven-transmembrane, G protein-coupled receptor. Science **272**:872-877.

88. **Fernandez, G., B. Clotet, and M. A. Martinez**. 2007. Fitness landscape of human immunodeficiency virus type 1 protease quasispecies. J. Virol. **81**:2485-2496.

89. **Fields, B. N., D. M. Knipe, P. M. Howley, and R. M. Chanock**. 1996. Fields virology. Philadelphia : Lippincott-Raven.

90. **Finney, D. J.** 1978. Statistical Method in Biological Assay. Griffin, London.

91. **Fisher, A. G., E. Collalti, L. Ratner, R. C. Gallo, and F. Wong-Staal**. 1985. A molecular clone of HTLV-III with biological activity. Nature **316**:262-265.

92. **Forns, X., J. Bukh, R. H. Purcell, and S. U. Emerson**. 1997. How Escherichia coli can bias the results of molecular cloning: preferential selection of defective genomes of hepatitis C virus during the cloning procedure. Proc. Natl. Acad. Sci. U. S. A **94**:13909-13914.

93. **Frater, A. J., C. C. Chaput, J. N. Weber, and M. O. McClure**. 2000. HIV-1 resistance genotyping by sequencing produces inconsistent results for mixed viral populations. AIDS **14**:1473-1475.

94. **Freed, E. O.** 1998. HIV-1 gag proteins: diverse functions in the virus life cycle. Virology **251**:1-15.

95. **Freed, E. O.** 2001. HIV-1 replication. Somat. Cell Mol. Genet. **26**:13-33.

96. **Freed, E. O.** 2004. HIV-1 and the host cell: an intimate association. Trends Microbiol. **12**:170-177.

97. **Freed, E. O. and M. A. Martin**. 1994. HIV-1 infection of non-dividing cells. Nature **369**:107-108.

98. **Freed, E. O., J. M. Orenstein, A. J. Buckler-White, and M. A. Martin**. 1994. Single amino acid changes in the human immunodeficiency virus type 1 matrix protein block virus particle production. J. Virol. **68**:5311-5320.

99. **Gandhi, R. T., A. Wurcel, E. S. Rosenberg, M. N. Johnston, N. Hellmann, M. Bates, M. S. Hirsch, and B. D. Walker**. 2003. Progressive reversion of human immunodeficiency virus type 1 resistance mutations in vivo after transmission of a multiply drug-resistant virus. Clin. Infect. Dis. **37**:1693-1698.

100. **Gao, F., S. G. Morrison, D. L. Robertson, C. L. Thornton, S. Craig, G. Karlsson, J. Sodroski, M. Morgado, B. Galvao-Castro, H. von Briesen, S. Beddows, J. Weber, P. M. Sharp, G. M. Shaw, and B. H. Hahn**. 1996. Molecular cloning and analysis of functional envelope genes from human immunodeficiency virus type 1 sequence subtypes A through G. The WHO and NIAID Networks for HIV Isolation and Characterization. J. Virol. **70**:1651-1667.

101. **Gao, F., D. L. Robertson, C. D. Carruthers, S. G. Morrison, B. Jian, Y. Chen, F. Barre-Sinoussi, M. Girard, A. Srinivasan, A. G. Abimiku, G. M. Shaw, P. M. Sharp, and B. H. Hahn**. 1998. A comprehensive panel of near-full-length clones and reference sequences for non-subtype B isolates of human immunodeficiency virus type 1. J. Virol. **72**:5680-5698.

102. **Gao, F., E. Bailes, D. L. Robertson, Y. Chen, C. M. Rodenburg, S. F. Michael, L. B. Cummins, L. O. Arthur, M. Peeters, G. M. Shaw, P. M. Sharp, and B. H. Hahn**. 1999. Origin of HIV-1 in the chimpanzee Pan troglodytes troglodytes. **397**:436-441.

103. **Gao, Q., M. A. Parniak, Z. Gu, and M. A. Wainberg**. 1992. Generation of nucleoside-resistant variants of HIV-1 by in vitro selection in the presence of AZT or DDI but no by combinations. Leukemia **6 Suppl 3**:192S-195S.

104. **Garcia-Lerma, J. G., H. MacInnes, D. Bennett, H. Weinstock, and W. Heneine**. 2004. Transmitted human immunodeficiency virus type 1 carrying the D67N or K219Q/E mutation evolves rapidly to zidovudine resistance in vitro and shows a high replicative fitness in the presence of zidovudine. J. Virol. **78**:7545-7552.

105. **Gasmi, M., J. Glynn, M. J. Jin, D. J. Jolly, J. K. Yee, and S. T. Chen**. 1999. Requirements for efficient production and transduction of human immunodeficiency virus type 1-based vectors. J. Virol. **73**:1828-1834.

106. **Gatanaga, H., D. Das, Y. Suzuki, D. D. Yeh, K. A. Hussain, A. K. Ghosh, and H. Mitsuya**. 2006. Altered HIV-1 Gag protein interactions with cyclophilin A (CypA) on the acquisition of H219Q and H219P substitutions in the CypA binding loop. J. Biol Chem. **281**:1241-1250.

107. **Gatanaga, H., Y. Suzuki, H. Tsang, K. Yoshimura, M. F. Kavlick, K. Nagashima, R. J. Gorelick, S. Mardy, C. Tang, M. F. Summers, and H. Mitsuya**. 2002. Amino acid substitutions in Gag protein at non-cleavage sites are indispensable for the development of a high multitude of HIV-1 resistance against protease inhibitors. J. Biol. Chem. **277**:5952-5961.

108. **Ghosn, J., I. Pellegrin, C. Goujard, C. Deveau, J. P. Viard, J. Galimand, M. Harzic, C. Tamalet, L. Meyer, C. Rouzioux, and M. L. Chaix**. 2006. HIV-1 resistant strains acquired at the time of primary infection massively fuel the cellular reservoir and persist for lengthy periods of time. AIDS **20**:159-170.

109. **Girnary, R., L. King, L. Robinson, R. Elston, and I. Brierley**. 2007. Structure-function analysis of the ribosomal frameshifting signal of two human immunodeficiency virus type 1 isolates with increased resistance to viral protease inhibitors. J. Gen. Virol. **88**:226-235.

110. **Goldschmidt, V. and R. Marquet**. 2004. Primer unblocking by HIV-1 reverse transcriptase and resistance to nucleoside RT inhibitors (NRTIs). Int. J. Biochem. Cell Biol **36**:1687-1705.

111. **Gomez, C. and T. J. Hope**. 2005. The ins and outs of HIV replication. Cell Microbiol. **7**:621-626.

112. **Goodenow, M. M., S. L. Rose, D. L. Tuttle, and J. W. Sleasman**. 2003. HIV-1 fitness and macrophages. J. Leukoc. Biol **74**:657-666.

113. **Gotte, M., X. Li, and M. A. Wainberg**. 1999. HIV-1 reverse transcription: a brief overview focused on structure-function relationships among molecules involved in initiation of the reaction. Arch. Biochem. Biophys. **365**:199-210.

114. **Gottlinger, H. G.** 2001. The HIV-1 assembly machine. AIDS **15 Suppl 5**:S13-S20.

115. **Goudsmit, J., A. de Ronde, E. de Rooij, and R. de Boer**. 1997. Broad spectrum of in vivo fitness of human immunodeficiency virus type 1 subpopulations differing at reverse transcriptase codons 41 and 215. J. Virol. **71**:4479-4484.

116. **Goudsmit, J., A. de Ronde, D. D. Ho, and A. S. Perelson**. 1996. Human immunodeficiency virus fitness in vivo: calculations based on a single zidovudine resistance mutation at codon 215 of reverse transcriptase. J. Virol. **70**:5662-5664.

117. **Grant, R. M., F. M. Hecht, M. Warmerdam, L. Liu, T. Liegler, C. J. Petropoulos, N. S. Hellmann, M. Chesney, M. P. Busch, and J. O. Kahn**. 2002. Time trends in primary HIV-1 drug resistance among recently infected persons. JAMA **288**:181-188.

118. **Greenberg, M., N. Cammack, M. Salgo, and L. Smiley**. 2004. HIV fusion and its inhibition in antiretroviral therapy. Rev. Med. Virol. **14**:321-337.

119. **Greenberg, M. L. and N. Cammack**. 2004. Resistance to enfuvirtide, the first HIV fusion inhibitor. J. Antimicrob. Chemother. **54**:333-340.

120. **Grobler, J. A., G. Dornadula, M. R. Rice, A. L. Simcoe, D. J. Hazuda, and M. D. Miller**. 2007. HIV-1 reverse transcriptase plus-strand initiation exhibits preferential sensitivity to non-nucleoside reverse transcriptase inhibitors in vitro. J. Biol Chem. **282**:8005-8010.

121. **Gummuluru, S., V. N. KewalRamani, and M. Emerman**. 2002. Dendritic cell-mediated viral transfer to T cells is required for human immunodeficiency virus type 1 persistence in the face of rapid cell turnover.. J. Virol. **76**:10692-10701.

122. **Gupta, R. K. and D. Pillay**. 2007. HIV resistance and the developing world. Int. J. Antimicrob. Agents.

123. **Gurer, C., A. Hoglund, S. Hoglund, and J. Luban**. 2005. ATPgammaS disrupts human immunodeficiency virus type 1 virion core integrity.. J. Virol. **79**:5557-5567.

124. **Gurney, K. B., J. Elliott, H. Nassanian, C. Song, E. Soilleux, I. McGowan, P. A. Anton, and B. Lee**. 2005. Binding and transfer of human immunodeficiency virus by DC-SIGN+ cells in human rectal mucosa. J. Virol. **79**:5762-5773.

125. **Hachiya, A., S. Aizawa-Matsuoka, M. Tanaka, Y. Takahashi, S. Ida, H. Gatanaga, Y. Hirabayashi, A. Kojima, M. Tatsumi, and S. Oka**. 2001. Rapid and simple phenotypic assay for drug susceptibility of human immunodeficiency virus type 1 using CCR5-expressing HeLa/CD4(+) cell clone 1-10 (MAGIC-5). Antimicrob. Agents Chemother. **45**:495-501.

126. **Hahn, B. H., G. M. Shaw, K. M. De Cock, and P. M. Sharp**. 2000. AIDS as a zoonosis: scientific and public health implications. Science **287**:607-614.

127. **Hallenberger, S., M. Moulard, M. Sordel, H. D. Klenk, and W. Garten**. 1997. The role of eukaryotic subtilisin-like endoproteases for the activation of human immunodeficiency virus glycoproteins in natural host cells. J. Virol. **71**:1036-1045.

128. **Hamano, T., K. Matsuo, Y. Hibi, A. F. Victoriano, N. Takahashi, Y. Mabuchi, T. Soji, S. Irie, P. Sawanpanyalert, H. Yanai, T. Hara, S. Yamazaki, N. Yamamoto, and T. Okamoto**. 2007. A single-nucleotide synonymous mutation in the gag gene controlling human immunodeficiency virus type 1 virion production. J. Virol. **81**:1528-1533.

129. **Hanahan, D.** 1983. Studies on transformation of Escherichia coli with plasmids. J. Mol. Biol **166**:557-580.

130. **Hance, A. J., V. Lemiale, J. Izopet, D. Lecossier, V. Joly, P. Massip, F. Mammano, D. Descamps, F. Brun-Vezinet, and F. Clavel**. 2001. Changes in human immunodeficiency virus type 1 populations after treatment interruption in patients failing antiretroviral therapy. J. Virol. **75**:6410-6417.

131. **Harrison, G. P., G. Miele, E. Hunter, and A. M. Lever**. 1998. Functional analysis of the core human immunodeficiency virus type 1 packaging signal in a permissive cell line. J. Virol. **72**:5886-5896.

132. **Hatziioannou, T., D. Perez-Caballero, S. Cowan, and P. D. Bieniasz**. 2005. Cyclophilin interactions with incoming human immunodeficiency virus type 1 capsids with opposing effects on infectivity in human cells. J. Virol. **79**:176-183.

133. **Haubrich, R.** 2007. Defining HIV susceptibility to new antiretroviral agents--darunavir. J. Infect. Dis. **196**:1125-1127.

134. **Havlir, D. V. and D. D. Richman**. 1996. Viral dynamics of HIV: implications for drug development and therapeutic strategies. Ann. Intern. Med. **124**:984-994.

135. **Hayes, M. J., A. Koundouris, N. Gruis, W. Bergman, G. G. Peters, and A. J. Sinclair**. 2004. p16(INK4A)-independence of Epstein-Barr virus-induced cell proliferation and virus latency. J. Gen. Virol. **85**:1381-1386.

136. **Hertogs, K., M. P. de Bethune, V. Miller, T. Ivens, P. Schel, A. Van Cauwenberge, E. C. Van Den, G. Van, V, H. Azijn, M. Van Houtte, F. Peeters, S. Staszewski, M. Conant, S. Bloor, S. Kemp, B. Larder, and R. Pauwels**. 1998. A rapid method for simultaneous

detection of phenotypic resistance to inhibitors of protease and reverse transcriptase in recombinant human immunodeficiency virus type 1 isolates from patients treated with antiretroviral drugs. Antimicrob. Agents Chemother. **42**:269-276.

137. **Hill, C. P., D. Worthylake, D. P. Bancroft, A. M. Christensen, and W. I. Sundquist**. 1996. Crystal structures of the trimeric human immunodeficiency virus type 1 matrix protein: implications for membrane association and assembly. Proc. Natl. Acad. Sci. U. S. A **93**:3099-3104.

138. **Hirsch, M. S., F. Brun-Vezinet, B. Clotet, B. Conway, D. R. Kuritzkes, R. T. D'Aquila, L. M. Demeter, S. M. Hammer, V. A. Johnson, C. Loveday, J. W. Mellors, D. M. Jacobsen, and D. D. Richman**. 2003. Antiretroviral drug resistance testing in adults infected with human immunodeficiency virus type 1: 2003 recommendations of an International AIDS Society-USA Panel. Clin. Infect. Dis. **37**:113-128.

139. **Hladik, F., P. Sakchalathorn, L. Ballweber, G. Lentz, M. Fialkow, D. Eschenbach, and M. J. McElrath**. 2007. Initial Events in Establishing Vaginal Entry and Infection by Human Immunodeficiency Virus Type-1. Immunity. **26**:257-270.

140. **Ho, D. D., A. U. Neumann, A. S. Perelson, W. Chen, J. M. Leonard, and M. Markowitz**. 1995. Rapid turnover of plasma virions and CD4 lymphocytes in HIV-1 infection. Nature **373**:123-126.

141. **Hornak, V., A. Okur, R. C. Rizzo, and C. Simmerling**. 2006. HIV-1 protease flaps spontaneously open and reclose in molecular dynamics simulations. Proc. Natl. Acad. Sci. U. S. A **103**:915-920.

142. **Hsu, M., J. M. Harouse, A. Gettie, C. Buckner, J. Blanchard, and C. Cheng-Mayer**. 2003. Increased mucosal transmission but not enhanced pathogenicity of the CCR5-tropic, simian AIDS-inducing simian/human immunodeficiency virus SHIV(SF162P3) maps to envelope gp120. J. Virol. **77**:989-998.

143. **Isaka, Y., S. Miki, S. Kawauchi, A. Suyama, H. Sugimoto, A. Adachi, T. Miura, M. Hayami, O. Yoshie, T. Fujiwara, and A. Sato**. 2001. A single amino acid change at Leu-188 in the reverse transcriptase of HIV-2 and SIV renders them sensitive to non-nucleoside reverse transcriptase inhibitors. Arch. Virol. **146**:743-755.

144. **Jacobsen, H., K. Yasargil, D. L. Winslow, J. C. Craig, A. Krohn, I. B. Duncan, and J. Mous**. 1995. Characterization of human immunodeficiency virus type 1 mutants with decreased sensitivity to proteinase inhibitor Ro 31-8959. Virology **206**:527-534.

145. **Japour, A. J., D. L. Mayers, V. A. Johnson, D. R. Kuritzkes, L. A. Beckett, J. M. Arduino, J. Lane, R. J. Black, P. S. Reichelderfer, R. T. D'Aquila, and .** 1993. Standardized peripheral blood mononuclear cell culture assay for determination of drug susceptibilities of clinical human immunodeficiency virus type 1 isolates. The RV-43 Study Group, the AIDS Clinical Trials Group Virology Committee Resistance Working Group. Antimicrob. Agents Chemother. **37**:1095-1101.

146. **Johnson, V. A., F. Brun-Vezinet, B. Clotet, H. F. Gunthard, D. R. Kuritzkes, D. Pillay, J. M. Schapiro, and D. D. Richman**. 2007. Update of the Drug Resistance Mutations in HIV-1: 2007. Top. HIV. Med. **15**:119-125.

147. **Jolly, C., K. Kashefi, M. Hollinshead, and Q. J. Sattentau**. 2004. HIV-1 cell to cell transfer across an Env-induced, actin-dependent synapse. J. Exp. Med. **199**:283-293.

148. **Kantor, R., D. A. Katzenstein, B. Efron, A. P. Carvalho, B. Wynhoven, P. Cane, J. Clarke, S. Sirivichayakul, M. A. Soares, J. Snoeck, C. Pillay, H. Rudich, R. Rodrigues, A. Holguin, K. Ariyoshi, M. B. Bouzas, P. Cahn, W. Sugiura, V. Soriano, L. F. Brigido, Z. Grossman, L. Morris, A. M. Vandamme, A. Tanuri, P. Phanuphak, J. N. Weber, D. Pillay, P. R. Harrigan, R. Camacho, J. M. Schapiro, and R. W. Shafer**. 2005. Impact of

HIV-1 subtype and antiretroviral therapy on protease and reverse transcriptase genotype: results of a global collaboration. PLoS. Med. **2**:e112.

149. **Kaushik, R. and L. Ratner**. 2004. Role of human immunodeficiency virus type 1 matrix phosphorylation in an early postentry step of virus replication. J. Virol. **78**:2319-2326.

150. **Kaye, J. F. and A. M. Lever**. 1998. Nonreciprocal packaging of human immunodeficiency virus type 1 and type 2 RNA: a possible role for the p2 domain of Gag in RNA encapsidation. J. Virol. **72**:5877-5885.

151. **Kearney M, Palmer S, Maldarelli F, Bixby C, Bazmi H, Rock D, Falloon J, Davey R, Dewar R, Metcalf J, Mellors J, and Coffin J.** Single-genome Sequencing Is More Sensitive than Standard Genotype Analysis for Detection of HIV-1 Drug-resistance Mutations. 11th Conf Retrovir Opportunistic Infect Febr 8 11 2004 San Franc CA Conf Retrovir Opportunistic Infect 11th 2004 San Franc Calif. 2004 Feb 8-11; 11: abstract no. 695. 2004. 8-2-2004.

152. **Kellam, P., C. A. Boucher, and B. A. Larder**. 1992. Fifth mutation in human immunodeficiency virus type 1 reverse transcriptase contributes to the development of high-level resistance to zidovudine. Proc. Natl. Acad. Sci. U. S. A **89**:1934-1938.

153. **Kellam, P. and B. A. Larder**. 1994. Recombinant virus assay: a rapid, phenotypic assay for assessment of drug susceptibility of human immunodeficiency virus type 1 isolates. Antimicrob. Agents Chemother. **38**:23-30.

154. **Kempf, D. J., J. D. Isaacson, M. S. King, S. C. Brun, Y. Xu, K. Real, B. M. Bernstein, A. J. Japour, E. Sun, and R. A. Rode**. 2001. Identification of genotypic changes in human immunodeficiency virus protease that correlate with reduced susceptibility to the protease inhibitor lopinavir among viral isolates from protease inhibitor-experienced patients. J. Virol. **75**:7462-7469.

155. **Kiernan, R. E., A. Ono, G. Englund, and E. O. Freed**. 1998. Role of matrix in an early postentry step in the human immunodeficiency virus type 1 life cycle. J. Virol. **72**:4116-4126.

156. **Kim, E. Y., M. A. Winters, R. M. Kagan, and T. C. Merigan**. 2001. Functional correlates of insertion mutations in the protease gene of human immunodeficiency virus type 1 isolates from patients. J. Virol. **75**:11227-11233.

157. **Kinter, A., J. Arthos, C. Cicala, and A. S. Fauci**. 2000. Chemokines, cytokines and HIV: a complex network of interactions that influence HIV pathogenesis. Immunol. Rev. **177**:88-98.

158. **Kleiman, L. and S. Cen**. 2004. The tRNALys packaging complex in HIV-1. Int. J. Biochem. Cell Biol. **36**:1776-1786.

159. **Klumpp, K., J. Q. Hang, S. Rajendran, Y. Yang, A. Derosier, K. Wong, I, H. Overton, K. E. Parkes, N. Cammack, and J. A. Martin**. 2003. Two-metal ion mechanism of RNA cleavage by HIV RNase H and mechanism-based design of selective HIV RNase H inhibitors. Nucleic Acids Res. **31**:6852-6859.

160. **Koch, N., N. Yahi, J. Fantini, and C. Tamalet**. 2001. Mutations in HIV-1 gag cleavage sites and their association with protease mutations. AIDS **15**:526-528.

161. **Kolli, M., S. Lastere, and C. A. Schiffer**. 2006. Co-evolution of nelfinavir-resistant HIV-1 protease and the p1-p6 substrate. Virology **347**:405-409.

162. **Kondo, E., F. Mammano, E. A. Cohen, and H. G. Gottlinger**. 1995. The p6gag domain of human immunodeficiency virus type 1 is sufficient for the incorporation of Vpr into heterologous viral particles. J. Virol. **69**:2759-2764.

163. **Korber, B., M. Muldoon, J. Theiler, F. Gao, R. Gupta, A. Lapedes, B. H. Hahn, S. Wolinsky, and T. Bhattacharya**. 2000. Timing the ancestor of the HIV-1 pandemic strains. Science **288**:1789-1796.

164. **Kosalaraksa, P., M. F. Kavlick, V. Maroun, R. Le, and H. Mitsuya**. 1999. Comparative fitness of multi-dideoxynucleoside-resistant human immunodeficiency virus type 1 (HIV-1) in an In vitro competitive HIV-1 replication assay. J. Virol. **73**:5356-5363.

165. **Kovalevsky, A. Y., Y. Tie, F. Liu, P. I. Boross, Y. F. Wang, S. Leshchenko, A. K. Ghosh, R. W. Harrison, and I. T. Weber**. 2006. Effectiveness of nonpeptide clinical inhibitor TMC-114 on HIV-1 protease with highly drug resistant mutations D30N, I50V, and L90M. J. Med. Chem. **49**:1379-1387.

166. **Kukhanova, M. K., N. F. Zakirova, A. V. Ivanov, L. A. Alexandrova, M. V. Jasco, and A. R. Khomutov**. 2005. Hypophosphoric acid is a unique substrate of pyrophosphorolysis catalyzed by HIV-1 reverse transcriptase. Biochem. Biophys. Res. Commun. **338**:1335-1341.

167. **Lacey, S. F. and B. A. Larder**. 1994. Mutagenic study of codons 74 and 215 of the human immunodeficiency virus type 1 reverse transcriptase, which are significant in nucleoside analog resistance. J. Virol. **68**:3421-3424.

168. **Lacey, S. F., J. E. Reardon, E. S. Furfine, T. A. Kunkel, K. Bebenek, K. A. Eckert, S. D. Kemp, and B. A. Larder**. 1992. Biochemical studies on the reverse transcriptase and RNase H activities from human immunodeficiency virus strains resistant to 3'-azido-3'-deoxythymidine. J. Biol Chem. **267**:15789-15794.

169. **Lafeuillade, A., C. Tamalet, P. Pellegrino, P. de Micco, C. Vignoli, and R. Quilichini**. 1994. Correlation between surrogate markers, viral load, and disease progression in HIV-1 infection. J. Acquir. Immune. Defic. Syndr. **7**:1028-1033.

170. **Lambert-Niclot, S., P. Flandre, A. Canestri, G. Peytavin, C. Blanc, R. Agher, C. Soulie, M. Wirden, C. Katlama, V. Calvez, and A. G. Marcelin**. 2008. Factors associated with the selection of mutations conferring resistance to protease inhibitors (PIs) in PI-experienced patients displaying treatment failure on darunavir. Antimicrob. Agents Chemother. **52**:491-496.

171. **Lastere, S., C. Dalban, G. Collin, D. Descamps, P. M. Girard, F. Clavel, D. Costagliola, and F. Brun-Vezinet**. 2004. Impact of insertions in the HIV-1 p6 PTAPP region on the virological response to amprenavir. Antivir. Ther. **9**:221-227.

172. **Lataillade, M. and M. J. Kozal**. 2006. The hunt for HIV-1 integrase inhibitors. AIDS Patient. Care STDS. **20**:489-501.

173. **Laughrea, M., L. Jette, J. Mak, L. Kleiman, C. Liang, and M. A. Wainberg**. 1997. Mutations in the kissing-loop hairpin of human immunodeficiency virus type 1 reduce viral infectivity as well as genomic RNA packaging and dimerization. J. Virol. **71**:3397-3406.

174. **Lee, P. P. and M. L. Linial**. 1995. Inhibition of wild-type HIV-1 virus production by a matrix deficient Gag mutant. Virology **208**:808-811.

175. **Leigh Brown, A. J., S. D. Frost, W. C. Mathews, K. Dawson, N. S. Hellmann, E. S. Daar, D. D. Richman, and S. J. Little**. 2003. Transmission fitness of drug-resistant human immunodeficiency virus and the prevalence of resistance in the antiretroviral-treated population. J. Infect. Dis. **187**:683-686.

176. **Leitner T, Foley B, Hahn B, Marx P, McCutchan F, Mellors J, Wolinsky S, and Korber B**. 2005. HIV Sequence Compendium 2005.

177. **Levy JA, Hoffman AD, Kramer SM, Landis JA, Shimabukuro JM, Oshiro LS.**. 1984. Isolation of lymphocytopathic retroviruses from San Francisco patients with AIDS. Science **225**:840-2.

178. **Levy, D. N., G. M. Aldrovandi, O. Kutsch, and G. M. Shaw**. 2004. Dynamics of HIV-1 recombination in its natural target cells. Proc. Natl. Acad. Sci. U. S. A **101**:4204-4209.

179. **Li, F., R. Goila-Gaur, K. Salzwedel, N. R. Kilgore, M. Reddick, C. Matallana, A. Castillo, D. Zoumplis, D. E. Martin, J. M. Orenstein, G. P. Allaway, E. O. Freed, and C. T. Wild**. 2003. PA-457: a potent HIV inhibitor that disrupts core condensation by targeting a late step in Gag processing. Proc. Natl. Acad. Sci. U. S. A **100**:13555-13560.

180. **Liang, C., L. Rong, M. Laughrea, L. Kleiman, and M. A. Wainberg**. 1998. Compensatory point mutations in the human immunodeficiency virus type 1 Gag region that are distal from deletion mutations in the dimerization initiation site can restore viral replication. J. Virol. **72**:6629-6636.

181. **Liang, C., L. Rong, Y. Quan, M. Laughrea, L. Kleiman, and M. A. Wainberg**. 1999. Mutations within four distinct gag proteins are required to restore replication of human immunodeficiency virus type 1 after deletion mutagenesis within the dimerization initiation site. J. Virol. **73**:7014-7020.

182. **Loemba, H., B. Brenner, M. A. Parniak, S. Ma'ayan, B. Spira, D. Moisi, M. Oliveira, M. Detorio, and M. A. Wainberg**. 2002. Genetic divergence of human immunodeficiency virus type 1 Ethiopian clade C reverse transcriptase (RT) and rapid development of resistance against nonnucleoside inhibitors of RT. Antimicrob. Agents Chemother. **46**:2087-2094.

183. **Long, E. M., H. L. Martin, Jr., J. K. Kreiss, S. M. Rainwater, L. Lavreys, D. J. Jackson, J. Rakwar, K. Mandaliya, and J. Overbaugh**. 2000. Gender differences in HIV-1 diversity at time of infection. Nat. Med. **6**:71-75.

184. **Louie, M. and M. Markowitz**. 2002. Goals and milestones during treatment of HIV-1 infection with antiretroviral therapy: a pathogenesis-based perspective. Antiviral Res. **55**:15-25.

185. **Lu, J. and D. R. Kuritzkes**. 2001. A novel recombinant marker virus assay for comparing the relative fitness of hiv-1 reverse transcriptase variants. J. Acquir. Immune. Defic. Syndr. **27**:7-13.

186. **Luster, A. D.** 1998. Chemokines--chemotactic cytokines that mediate inflammation. N. Engl. J. Med. **338**:436-445.

187. **Madani, N., A. L. Perdigoto, K. Srinivasan, J. M. Cox, J. J. Chruma, J. LaLonde, M. Head, A. B. Smith, III, and J. G. Sodroski**. 2004. Localized changes in the gp120 envelope glycoprotein confer resistance to human immunodeficiency virus entry inhibitors BMS-806 and #155. J. Virol. **78**:3742-3752.

188. **Maguire, M. F., R. Guinea, P. Griffin, S. Macmanus, R. C. Elston, J. Wolfram, N. Richards, M. H. Hanlon, D. J. Porter, T. Wrin, N. Parkin, M. Tisdale, E. Furfine, C. Petropoulos, B. W. Snowden, and J. P. Kleim**. 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**:7398-7406.

189. **Makhija, M. T.** 2006. Designing HIV integrase inhibitors--shooting the last arrow. Curr. Med. Chem. **13**:2429-2441.

190. **Malet, I., O. Delelis, M. A. Valantin, B. Montes, C. Soulie, M. Wirden, L. Tchertanov, G. Peytavin, J. Reynes, J. F. Mouscadet, C. Katlama, V. Calvez, and A. G. Marcelin**. 2008.

Mutations associated with failure of raltegravir treatment affect integrase sensitivity to the inhibitor in vitro. Antimicrob. Agents Chemother. **52**:1351-1358.

191. **Malet, I., B. Roquebert, C. Dalban, M. Wirden, B. Amellal, R. Agher, A. Simon, C. Katlama, D. Costagliola, V. Calvez, and A. G. Marcelin**. 2006. Association of Gag cleavage sites to protease mutations and to virological response in HIV-1 treated patients. J. Infect.

192. **Malet, I., M. Wirden, A. Derache, A. Simon, C. Katlama, V. Calvez, and A. G. Marcelin**. 2007. Primary genotypic resistance of HIV-1 to the maturation inhibitor PA-457 in protease inhibitor-experienced patients. AIDS **21**:871-873.

193. **Mammano, F., E. Kondo, J. Sodroski, A. Bukovsky, and H. G. Gottlinger**. 1995. Rescue of human immunodeficiency virus type 1 matrix protein mutants by envelope glycoproteins with short cytoplasmic domains. J. Virol. **69**:3824-3830.

194. **Mammano, F., C. Petit, and F. Clavel**. 1998. Resistance-associated loss of viral fitness in human immunodeficiency virus type 1: phenotypic analysis of protease and gag coevolution in protease inhibitor-treated patients. J. Virol. **72**:7632-7637.

195. **Mangasarian, A., M. Foti, C. Aiken, D. Chin, J. L. Carpentier, and D. Trono**. 1997. The HIV-1 Nef protein acts as a connector with sorting pathways in the Golgi and at the plasma membrane. Immunity. **6**:67-77.

196. **Mangeat, B., P. Turelli, G. Caron, M. Friedli, L. Perrin, and D. Trono**. 2003. Broad antiretroviral defence by human APOBEC3G through lethal editing of nascent reverse transcripts. Nature **424**:99-103.

197. **Maniatis, T., E. F. Fritsch, and J. Sambrook**. 1982. Molecular cloning a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.

198. **Mark-Danieli, M., N. Laham, M. Kenan-Eichler, A. Castiel, D. Melamed, M. Landau, N. M. Bouvier, M. J. Evans, and E. Bacharach**. 2005. Single point mutations in the zinc finger motifs of the human immunodeficiency virus type 1 nucleocapsid alter RNA binding specificities of the gag protein and enhance packaging and infectivity. J. Virol. **79**:7756-7767.

199. **Markosyan, R. M., F. S. Cohen, and G. B. Melikyan**. 2003. HIV-1 envelope proteins complete their folding into six-helix bundles immediately after fusion pore formation. Mol. Biol Cell **14**:926-938.

200. **Markowitz, M., H. Mohri, S. Mehandru, A. Shet, L. Berry, R. Kalyanaraman, A. Kim, C. Chung, P. Jean-Pierre, A. Horowitz, M. La Mar, T. Wrin, N. Parkin, M. Poles, C. Petropoulos, M. Mullen, D. Boden, and D. D. Ho**. 2005. Infection with multidrug resistant, dual-tropic HIV-1 and rapid progression to AIDS: a case report. Lancet **365**:1031-1038.

201. **Markowitz, M., J. O. Morales-Ramirez, B. Y. Nguyen, C. M. Kovacs, R. T. Steigbigel, D. A. Cooper, R. Liporace, R. Schwartz, R. Isaacs, L. R. Gilde, L. Wenning, J. Zhao, and H. Teppler**. 2006. Antiretroviral activity, pharmacokinetics, and tolerability of MK-0518, a novel inhibitor of HIV-1 integrase, dosed as monotherapy for 10 days in treatment-naive HIV-1-infected individuals. J. Acquir. Immune. Defic. Syndr. **43**:509-515.

202. **Maschera, B., E. Furfine, and E. D. Blair**. 1995. Analysis of resistance to human immunodeficiency virus type 1 protease inhibitors by using matched bacterial expression and proviral infection vectors. J. Virol. **69**:5431-5436.

203. **McDonald, D., M. A. Vodicka, G. Lucero, T. M. Svitkina, G. G. Borisy, M. Emerman, and T. J. Hope**. 2002. Visualization of the intracellular behavior of HIV in living cells. J. Cell Biol **159**:441-452.

204. **McMichael, A. J. and R. E. Phillips**. 1997. Escape of human immunodeficiency virus from immune control. Annu. Rev. Immunol **15**:271-296.

205. **Mely, Y., H. de Rocquigny, N. Morellet, B. P. Roques, and D. Gerad**. 1996. Zinc binding to the HIV-1 nucleocapsid protein: a thermodynamic investigation by fluorescence spectroscopy. Biochemistry **35**:5175-5182.

206. **Michael, N. L., T. Mo, A. Merzouki, M. O'Shaughnessy, C. Oster, D. S. Burke, R. R. Redfield, D. L. Birx, and S. A. Cassol**. 1995. Human immunodeficiency virus type 1 cellular RNA load and splicing patterns predict disease progression in a longitudinally studied cohort. J. Virol. **69**:1868-1877.

207. **Mikhail, M., B. Wang, P. Lemey, B. Beckholdt, A. M. Vandamme, M. J. Gill, and N. K. Saksena**. 2005. Full-length HIV type 1 genome analysis showing evidence for HIV type 1 transmission from a nonprogressor to two recipients who progressed to AIDS. AIDS Res. Hum. Retroviruses **21**:575-579.

208. **Mild, M., J. Esbjornsson, E. M. Fenyo, and P. Medstrand**. 2007. Frequent intrapatient recombination between HIV-1 R5 and X4 envelopes: Implications for coreceptor switch. J. Virol.

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

210. **Mitsuya, H., M. Popovic, R. Yarchoan, S. Matsushita, R. C. Gallo, and S. Broder**. 1984. Suramin protection of T cells in vitro against infectivity and cytopathic effect of HTLV-III. Science **226**:172-174.

211. **Mitsuya, H., K. J. Weinhold, P. A. Furman, M. H. St Clair, S. N. Lehrman, R. C. Gallo, D. Bolognesi, D. W. Barry, and S. Broder**. 1985. 3'-Azido-3'-deoxythymidine (BW A509U): an antiviral agent that inhibits the infectivity and cytopathic effect of human T-lymphotropic virus type III/lymphadenopathy-associated virus in vitro. Proc. Natl. Acad. Sci. U. S. A **82**:7096-7100.

212. **Mitsuya, H., R. Yarchoan, and S. Broder**. 1990. Molecular targets for AIDS therapy. Science **249**:1533-1544.

213. **Molla, A., M. Korneyeva, Q. Gao, S. Vasavanonda, P. J. Schipper, H. M. Mo, M. Markowitz, T. Chernyavskiy, P. Niu, N. Lyons, A. Hsu, G. R. Granneman, D. D. Ho, C. A. Boucher, J. M. Leonard, D. W. Norbeck, and D. J. Kempf**. 1996. Ordered accumulation of mutations in HIV protease confers resistance to ritonavir. Nat. Med. **2**:760-766.

214. **Molto, J., J. R. Santos, E. Negredo, C. Miranda, S. Videla, and B. Clotet**. 2007. Lopinavir/ritonavir monotherapy as a simplification strategy in routine clinical practice. J. Antimicrob. Chemother. **60**:436-439.

215. **Moore, M. D., W. Fu, O. Nikolaitchik, J. Chen, R. G. Ptak, and W. S. Hu**. 2007. Dimer initiation signal of human immunodeficiency virus type 1: its role in partner selection during RNA copackaging and its effects on recombination. J. Virol. **81**:4002-4011.

216. **Morellet, N., S. Druillennec, C. Lenoir, S. Bouaziz, and B. P. Roques**. 2005. Helical structure determined by NMR of the HIV-1 (345-392)Gag sequence, surrounding p2: implications for particle assembly and RNA packaging. Protein Sci. **14**:375-386.

217. **Morner, A., A. Bjorndal, J. Albert, V. N. KewalRamani, D. R. Littman, R. Inoue, R. Thorstensson, E. M. Fenyo, and E. Bjorling**. 1999. Primary human immunodeficiency virus type 2 (HIV-2) isolates, like HIV-1 isolates, frequently use CCR5 but show promiscuity in coreceptor usage. J. Virol. **73**:2343-2349.

218. **Mujeeb, A., N. B. Ulyanov, S. Georgantis, I. Smirnov, J. Chung, T. G. Parslow, and T. L. James**. 2007. Nucleocapsid protein-mediated maturation of dimer initiation complex of full-length SL1 stemloop of HIV-1: sequence effects and mechanism of RNA refolding. Nucleic Acids Res. **35**:2026-2034.

219. **Muller, B., J. Daecke, O. T. Fackler, M. T. Dittmar, H. Zentgraf, and H. G. Krausslich**. 2004. Construction and characterization of a fluorescently labeled infectious human immunodeficiency virus type 1 derivative. J. Virol. **78**:10803-10813.

220. **Murdoch, C.** 2000. CXCR4: chemokine receptor extraordinaire. Immunol. Rev. **177**:175-184.

221. **Murphy, E. L., A. C. Collier, L. A. Kalish, S. F. Assmann, M. F. Para, T. P. Flanigan, P. N. Kumar, L. Mintz, F. R. Wallach, and G. J. Nemo**. 2001. Highly active antiretroviral therapy decreases mortality and morbidity in patients with advanced HIV disease. Ann. Intern. Med. **135**:17-26.

222. **Muzammil, S., A. A. Armstrong, L. W. Kang, A. Jakalian, P. R. Bonneau, V. Schmelmer, L. M. Amzel, and E. Freire**. 2007. Unique Thermodynamic Response of Tipranavir to HIV-1 Protease Drug Resistance Mutations. J. Virol.

223. **Muzammil, S., P. Ross, and E. Freire**. 2003. A major role for a set of non-active site mutations in the development of HIV-1 protease drug resistance. Biochemistry **42**:631-638.

224. **Myint, L., M. Matsuda, Z. Matsuda, Y. Yokomaku, T. Chiba, A. Okano, K. Yamada, and W. Sugiura**. 2004. Gag non-cleavage site mutations contribute to full recovery of viral fitness in protease inhibitor-resistant human immunodeficiency virus type 1. Antimicrob. Agents Chemother. **48**:444-452.

225. **Naldini, L., U. Blomer, F. H. Gage, D. Trono, and I. M. Verma**. 1996. Efficient transfer, integration, and sustained long-term expression of the transgene in adult rat brains injected with a lentiviral vector. Proc. Natl. Acad. Sci. U. S. A **93**:11382-11388.

226. **Naldini, L., U. Blomer, P. Gallay, D. Ory, R. Mulligan, F. H. Gage, I. M. Verma, and D. Trono**. 1996. In vivo gene delivery and stable transduction of nondividing cells by a lentiviral vector. Science **272**:263-267.

227. **Narayan, S. and J. A. Young**. 2004. Reconstitution of retroviral fusion and uncoating in a cell-free system. Proc. Natl. Acad. Sci. U. S. A **101**:7721-7726.

228. **Nicastri, E., L. Sarmati, G. d'Ettorre, L. Palmisano, S. G. Parisi, I. Uccella, A. Rianda, E. Concia, V. Vullo, S. Vella, and M. Andreoni**. 2003. Replication capacity, biological phenotype, and drug resistance of HIV strains isolated from patients failing antiretroviral therapy. J. Med. Virol. **69**:1-6.

229. **Nijhuis, M., C. A. Boucher, P. Schipper, T. Leitner, R. Schuurman, and J. Albert**. 1998. Stochastic processes strongly influence HIV-1 evolution during suboptimal protease-inhibitor therapy. Proc. Natl. Acad. Sci. U. S. A **95**:14441-14446.

230. **Nijhuis, M., R. Schuurman, D. de Jong, J. Erickson, E. Gustchina, J. Albert, P. Schipper, S. Gulnik, and C. A. Boucher**. 1999. Increased fitness of drug resistant HIV-1 protease as a result of acquisition of compensatory mutations during suboptimal therapy. AIDS **13**:2349-2359.

231. **Nijhuis, M., N. M. van Maarseveen, S. Lastere, P. Schipper, E. Coakley, B. Glass, M. Rovenska, D. de Jong, C. Chappey, I. W. Goedegebuure, G. Heilek-Snyder, D. Dulude, N. Cammack, L. Brakier-Gingras, J. Konvalinka, N. Parkin, H. G. Krausslich, F. Brun-Vezinet, and C. A. Boucher**. 2007. A Novel Substrate-Based HIV-1 Protease Inhibitor Drug Resistance Mechanism. PLoS. Med. **4**:e36.

232. **Nikolaitchik, O., T. D. Rhodes, D. Ott, and W. S. Hu**. 2006. Effects of mutations in the human immunodeficiency virus type 1 Gag gene on RNA packaging and recombination. J. Virol. **80**:4691-4697.

233. **Nisole, S. and A. Saib**. 2004. Early steps of retrovirus replicative cycle.. Retrovirology. **1**:9.

234. **Njai, H. F., Y. Gali, G. Vanham, C. Clybergh, W. Jennes, N. Vidal, C. Butel, E. Mpoudi-Ngolle, M. Peeters, and K. K. Arien**. 2006. The predominance of Human Immunodeficiency Virus type 1 (HIV-1) circulating recombinant form 02 (CRF02_AG) in West Central Africa may be related to its replicative fitness. Retrovirology. **3**:40.

235. **Nora, T., C. Charpentier, O. Tenaillon, C. Hoede, F. Clavel, and A. J. Hance**. 2007. Contribution of Recombination to the Evolution of Human Immunodeficiency Viruses Expressing Resistance to Antiretroviral Treatment. J. Virol.

236. **Oberg, B.** 2006. Rational design of polymerase inhibitors as antiviral drugs. Antiviral Res. **71**:90-95.

237. **Ohi, Y. and J. L. Clever**. 2000. Sequences in the 5' and 3' R elements of human immunodeficiency virus type 1 critical for efficient reverse transcription. J. Virol. **74**:8324-8334.

238. **Ono, A., A. A. Waheed, and E. O. Freed**. 2007. Depletion of cellular cholesterol inhibits membrane binding and higher-order multimerization of human immunodeficiency virus type 1 Gag. Virology **360**:27-35.

239. **Ono, T., Y. Iwatani, A. Nishimura, A. Ishimoto, and H. Sakai**. 2000. Functional association between the nef gene product and gag-pol region of HIV-1. FEBS Lett. **466**:233-238.

240. **Ostrowski, M. A., S. J. Justement, A. Catanzaro, C. A. Hallahan, L. A. Ehler, S. B. Mizell, P. N. Kumar, J. A. Mican, T. W. Chun, and A. S. Fauci**. 1998. Expression of chemokine receptors CXCR4 and CCR5 in HIV-1-infected and uninfected individuals. J. Immunol **161**:3195-3201.

241. **Ott, D. E., L. V. Coren, T. D. Gagliardi, and K. Nagashima**. 2005. Heterologous late-domain sequences have various abilities to promote budding of human immunodeficiency virus type 1. J. Virol. **79**:9038-9045.

242. **Pantaleo, G. and A. S. Fauci**. 1996. Immunopathogenesis of HIV infection. Annu. Rev. Microbiol. **50**:825-854.

243. **Pantophlet, R. and D. R. Burton**. 2006. GP120: target for neutralizing HIV-1 antibodies. Annu. Rev. Immunol **24**:739-769.

244. **Paoletti, A. C., M. F. Shubsda, B. S. Hudson, and P. N. Borer**. 2002. Affinities of the nucleocapsid protein for variants of SL3 RNA in HIV-1. Biochemistry **41**:15423-15428.

245. **Paolucci, S., F. Baldanti, L. Dossena, and G. Gerna**. 2006. Amino acid insertions at position 35 of HIV-1 protease interfere with virus replication without modifying antiviral drug susceptibility. Antiviral Res. **69**:181-185.

246. **Partin, K., G. Zybarth, L. Ehrlich, M. DeCrombrugghe, E. Wimmer, and C. Carter**. 1991. Deletion of sequences upstream of the proteinase improves the proteolytic processing of human immunodeficiency virus type 1. Proc. Natl. Acad. Sci. U. S. A **88**:4776-4780.

247. **Paterson, D. L., S. Swindells, J. Mohr, M. Brester, E. N. Vergis, C. Squier, M. M. Wagener, and N. Singh**. 2000. Adherence to protease inhibitor therapy and outcomes in patients with HIV infection. Ann. Intern. Med. **133**:21-30.

248. **Paulus, C., C. Ludwig, and R. Wagner**. 2004. Contribution of the Gag-Pol transframe domain p6* and its coding sequence to morphogenesis and replication of human immunodeficiency virus type 1. Virology **330**:271-283.

249. **Pauwels, R.** 2006. Aspects of successful drug discovery and development. Antiviral Res. **71**:77-89.

250. **Pauwels, R., J. Balzarini, M. Baba, R. Snoeck, D. Schols, P. Herdewijn, J. Desmyter, and E. De Clercq**. 1988. Rapid and automated tetrazolium-based colorimetric assay for the detection of anti-HIV compounds. J. Virol. Methods **20**:309-321.

251. **Pauwels, R., E. De Clercq, J. Desmyter, J. Balzarini, P. Goubau, P. Herdewijn, H. Vanderhaeghe, and M. Vandeputte**. 1987. Sensitive and rapid assay on MT-4 cells for detection of antiviral compounds against the AIDS virus. J. Virol. Methods **16**:171-185.

252. **Paxton, W., R. I. Connor, and N. R. Landau**. 1993. Incorporation of Vpr into human immunodeficiency virus type 1 virions: requirement for the p6 region of gag and mutational analysis. J. Virol. **67**:7229-7237.

253. **Pazhanisamy, S., C. M. Stuver, A. B. Cullinan, N. Margolin, B. G. Rao, and D. J. Livingston**. 1996. Kinetic characterization of human immunodeficiency virus type-1 protease-resistant variants. J. Biol Chem. **271**:17979-17985.

254. **Perelson, A. S., A. U. Neumann, M. Markowitz, J. M. Leonard, and D. D. Ho**. 1996. HIV-1 dynamics in vivo: virion clearance rate, infected cell life-span, and viral generation time. Science **271**:1582-1586.

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

256. **Peters, S., M. Munoz, S. Yerly, V. Sanchez-Merino, C. Lopez-Galindez, L. Perrin, B. Larder, D. Cmarko, S. Fakan, P. Meylan, and A. Telenti**. 2001. Resistance to nucleoside analog reverse transcriptase inhibitors mediated by human immunodeficiency virus type 1 p6 protein. J. Virol. **75**:9644-9653.

257. **Petropoulos, C. J., N. T. Parkin, K. L. Limoli, Y. S. Lie, T. Wrin, W. Huang, H. Tian, D. Smith, G. A. Winslow, D. J. Capon, and J. M. Whitcomb**. 2000. A novel phenotypic drug susceptibility assay for human immunodeficiency virus type 1. Antimicrob. Agents Chemother. **44**:920-928.

258. **Pettit, S. C., J. C. Clemente, J. A. Jeung, B. M. Dunn, and A. H. Kaplan**. 2005. Ordered processing of the human immunodeficiency virus type 1 GagPol precursor is influenced by the context of the embedded viral protease. J. Virol. **79**:10601-10607.

259. **Pettit, S. C., L. E. Everitt, S. Choudhury, B. M. Dunn, and A. H. Kaplan**. 2004. Initial cleavage of the human immunodeficiency virus type 1 GagPol precursor by its activated protease occurs by an intramolecular mechanism. J. Virol. **78**:8477-8485.

260. **Pettit, S. C., S. Gulnik, L. Everitt, and A. H. Kaplan**. 2003. The dimer interfaces of protease and extra-protease domains influence the activation of protease and the specificity of GagPol cleavage. J. Virol. **77**:366-374.

261. **Pettit, S. C., J. N. Lindquist, A. H. Kaplan, and R. Swanstrom**. 2005. Processing sites in the human immunodeficiency virus type 1 (HIV-1) Gag-Pro-Pol precursor are cleaved by the viral protease at different rates. Retrovirology. **2**:66.

262. **Pettit, S. C., M. D. Moody, R. S. Wehbie, A. H. Kaplan, P. V. Nantermet, C. A. Klein, and R. Swanstrom**. 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**:8017-8027.

263. **Piatak, M., Jr., M. S. Saag, L. C. Yang, S. J. Clark, J. C. Kappes, K. C. Luk, B. H. Hahn, G. M. Shaw, and J. D. Lifson**. 1993. High levels of HIV-1 in plasma during all stages of infection determined by competitive PCR. Science **259**:1749-1754.

264. **Pilcher, C. D., H. C. Tien, J. J. Eron, Jr., P. L. Vernazza, S. Y. Leu, P. W. Stewart, L. E. Goh, and M. S. Cohen**. 2004. Brief but efficient: acute HIV infection and the sexual transmission of HIV. J. Infect. Dis. **189**:1785-1792.

265. **Pillai, S. K., B. Good, S. K. Pond, J. K. Wong, M. C. Strain, D. D. Richman, and D. M. Smith**. 2005. Semen-specific genetic characteristics of human immunodeficiency virus type 1 env. J. Virol. **79**:1734-1742.

266. **Pommier, Y., A. A. Johnson, and C. Marchand**. 2005. Integrase inhibitors to treat HIV/AIDS. Nat. Rev. Drug Discov. **4**:236-248.

267. **Poon, A. F., F. I. Lewis, S. L. Pond, and S. D. Frost**. 2007. Evolutionary Interactions between N-Linked Glycosylation Sites in the HIV-1 Envelope. PLoS. Comput. Biol **3**:e11.

268. **Prabu-Jeyabalan, M., E. A. Nalivaika, K. Romano, and C. A. Schiffer**. 2006. Mechanism of substrate recognition by drug-resistant human immunodeficiency virus type 1 protease variants revealed by a novel structural intermediate. J. Virol. **80**:3607-3616.

269. **Prado, J. G., N. T. Parkin, B. Clotet, L. Ruiz, and J. Martinez-Picado**. 2005. HIV type 1 fitness evolution in antiretroviral-experienced patients with sustained CD4+ T cell counts but persistent virologic failure. Clin. Infect. Dis. **41**:729-737.

270. **Prado, J. G., T. Wrin, J. Beauchaine, L. Ruiz, C. J. Petropoulos, S. D. Frost, B. Clotet, R. T. D'Aquila, and J. Martinez-Picado**. 2002. Amprenavir-resistant HIV-1 exhibits lopinavir cross-resistance and reduced replication capacity. AIDS **16**:1009-1017.

271. **Qari, S. H., R. Respess, H. Weinstock, E. M. Beltrami, K. Hertogs, B. A. Larder, C. J. Petropoulos, N. Hellmann, and W. Heneine**. 2002. Comparative analysis of two commercial phenotypic assays for drug susceptibility testing of human immunodeficiency virus type 1. J. Clin. Microbiol. **40**:31-35.

272. **Quinones-Mateu, M. E. and E. J. Arts**. 2002. Fitness of drug resistant HIV-1: methodology and clinical implications. Drug Resist. Updat. **5**:224-233.

273. **Ray, N., J. E. Harrison, L. A. Blackburn, J. N. Martin, S. G. Deeks, and R. W. Doms**. 2007. Clinical resistance to enfuvirtide does not affect susceptibility of human immunodeficiency virus type 1 to other classes of entry inhibitors. J. Virol. **81**:3240-3250.

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

275. **Reeves, J. D., S. A. Gallo, N. Ahmad, J. L. Miamidian, P. E. Harvey, M. Sharron, S. Pohlmann, J. N. Sfakianos, C. A. Derdeyn, R. Blumenthal, E. Hunter, and R. W. Doms**. 2002. Sensitivity of HIV-1 to entry inhibitors correlates with envelope/coreceptor affinity, receptor density, and fusion kinetics. Proc. Natl. Acad. Sci. U. S. A **99**:16249-16254.

276. **Reeves, J. D., F. H. Lee, J. L. Miamidian, C. B. Jabara, M. M. Juntilla, and R. W. Doms**. 2005. Enfuvirtide resistance mutations: impact on human immunodeficiency virus envelope function, entry inhibitor sensitivity, and virus neutralization. J. Virol. **79**:4991-4999.

277. **Resch, W., N. Parkin, T. Watkins, J. Harris, and R. Swanstrom**. 2005. Evolution of human immunodeficiency virus type 1 protease genotypes and phenotypes in vivo under selective pressure of the protease inhibitor ritonavir. J. Virol. **79**:10638-10649.

278. **Richman, D. D.** 1993. Human immunodeficiency virus. Playing chess with reverse transcriptase. Nature **361**:588-589.

279. **Rigourd, M., G. Bec, P. Benas, S. F. Le Grice, B. Ehresmann, C. Ehresmann, and R. Marquet**. 2003. Effects of tRNA 3 Lys aminoacylation on the initiation of HIV-1 reverse transcription. Biochimie **85**:521-525.

280. **Roberts, N. A., J. A. Martin, D. Kinchington, A. V. Broadhurst, J. C. Craig, I. B. Duncan, S. A. Galpin, B. K. Handa, J. Kay, A. Krohn, and .** 1990. Rational design of peptide-based HIV proteinase inhibitors. Science **248**:358-361.

281. **Robinson, L. H., R. E. Myers, B. W. Snowden, M. Tisdale, and E. D. Blair**. 2000. HIV type 1 protease cleavage site mutations and viral fitness: implications for drug susceptibility phenotyping assays. AIDS Res. Hum. Retroviruses **16**:1149-1156.

282. **Rong, L., R. S. Russell, J. Hu, M. Laughrea, M. A. Wainberg, and C. Liang**. 2003. Deletion of stem-loop 3 is compensated by second-site mutations within the Gag protein of human immunodeficiency virus type 1. Virology **314**:221-228.

283. **Rose, R. E., Y. F. Gong, J. A. Greytok, C. M. Bechtold, B. J. Terry, B. S. Robinson, M. Alam, R. J. Colonno, and P. F. Lin**. 1996. Human immunodeficiency virus type 1 viral background plays a major role in development of resistance to protease inhibitors. Proc. Natl. Acad. Sci. U. S. A **93**:1648-1653.

284. **Russell, R. S., A. Roldan, M. Detorio, J. Hu, M. A. Wainberg, and C. Liang**. 2003. Effects of a single amino acid substitution within the p2 region of human immunodeficiency virus type 1 on packaging of spliced viral RNA. J. Virol. **77**:12986-12995.

285. **Sagar, M., L. Lavreys, J. M. Baeten, B. A. Richardson, K. Mandaliya, B. H. Chohan, J. K. Kreiss, and J. Overbaugh**. 2003. Infection with multiple human immunodeficiency virus type 1 variants is associated with faster disease progression. J. Virol. **77**:12921-12926.

286. **Saint Vincent, B. R. and G. M. Wahl**. 1983. Homologous recombination in mammalian cells mediates formation of a functional gene from two overlapping gene fragments. Proc. Natl. Acad. Sci. U. S. A **80**:2002-2006.

287. **Sanders, R. W., E. C. de Jong, C. E. Baldwin, J. H. Schuitemaker, M. L. Kapsenberg, and B. Berkhout**. 2002. Differential transmission of human immunodeficiency virus type 1 by distinct subsets of effector dendritic cells. J. Virol. **76**:7812-7821.

288. **Schmidtmayerova, H., M. Alfano, G. Nuovo, and M. Bukrinsky**. 1998. Human immunodeficiency virus type 1 T-lymphotropic strains enter macrophages via a CD4- and CXCR4-mediated pathway: replication is restricted at a postentry level.. J. Virol. **72**:4633-4642.

289. **Schmit, J. C., K. Van Laethem, L. Ruiz, P. Hermans, S. Sprecher, A. Sonnerborg, M. Leal, T. Harrer, B. Clotet, V. Arendt, E. Lissen, M. Witvrouw, J. Desmyter, E. De Clercq, and A. M. Vandamme**. 1998. Multiple dideoxynucleoside analogue-resistant (MddNR) HIV-1 strains isolated from patients from different European countries. AIDS **12**:2007-2015.

290. **Scott, W. R. and C. A. Schiffer**. 2000. Curling of flap tips in HIV-1 protease as a mechanism for substrate entry and tolerance of drug resistance. Structure. **8**:1259-1265.

291. **Selmi, B., J. Deval, K. Alvarez, J. Boretto, S. Sarfati, C. Guerreiro, and B. Canard**. 2003. The Y181C substitution in 3'-azido-3'-deoxythymidine-resistant human immunodeficiency virus, type 1, reverse transcriptase suppresses the ATP-mediated repair of the 3'-azido-3'-deoxythymidine 5'-monophosphate-terminated primer. J. Biol Chem. **278**:40464-40472.

292. **Shafer, R. W., S. Y. Rhee, D. Pillay, V. Miller, P. Sandstrom, J. M. Schapiro, D. R. Kuritzkes, and D. Bennett**. 2007. HIV-1 protease and reverse transcriptase mutations for drug resistance surveillance. AIDS **21**:215-223.

293. **Shafer, R. W. and J. M. Schapiro**. 2005. Drug resistance and antiretroviral drug development. J. Antimicrob. Chemother. **55**:817-820.

294. **Shattock, R. J. and J. P. Moore**. 2003. Inhibiting sexual transmission of HIV-1 infection. Nat. Rev. Microbiol. **1**:25-34.

295. **Shehu-Xhilaga, M., S. M. Crowe, and J. Mak**. 2001. Maintenance of the Gag/Gag-Pol ratio is important for human immunodeficiency virus type 1 RNA dimerization and viral infectivity. J. Virol. **75**:1834-1841.

296. **Shi, C. and J. W. Mellors**. 1997. A recombinant retroviral system for rapid in vivo analysis of human immunodeficiency virus type 1 susceptibility to reverse transcriptase inhibitors. Antimicrob. Agents Chemother. **41**:2781-2785.

297. **Shimura, K., E. Kodama, Y. Sakagami, Y. Matsuzaki, W. Watanabe, K. Yamataka, Y. Watanabe, Y. Ohata, S. Doi, M. Sato, M. Kano, S. Ikeda, and M. Matsuoka**. 2008. Broad antiretroviral activity and resistance profile of the novel human immunodeficiency virus integrase inhibitor elvitegravir (JTK-303/GS-9137). J. Virol. **82**:764-774.

298. **Sierra, S., B. Kupfer, and R. Kaiser**. 2005. Basics of the virology of HIV-1 and its replication. J. Clin. Virol. **34**:233-244.

299. **Simon, V., N. Padte, D. Murray, J. Vanderhoeven, T. Wrin, N. Parkin, M. Di Mascio, and M. Markowitz**. 2003. Infectivity and replication capacity of drug-resistant human immunodeficiency virus type 1 variants isolated during primary infection. J. Virol. **77**:7736-7745.

300. **Smith, D. M., D. D. Richman, and S. J. Little**. 2005. HIV superinfection. J. Infect. Dis. **192**:438-444.

301. **Snyder, G. A., J. Ford, P. Torabi-Parizi, J. A. Arthos, P. Schuck, M. Colonna, and P. D. Sun**. 2005. Characterization of DC-SIGN/R Interaction with Human Immunodeficiency Virus Type 1 gp120 and ICAM Molecules Favors the Receptor's Role as an Antigen-Capturing Rather than an Adhesion Receptor. The Journal of Virology **79**:4589-4598.

302. **Sokolskaja, E., L. Berthoux, and J. Luban**. 2006. Cyclophilin A and TRIM5alpha independently regulate human immunodeficiency virus type 1 infectivity in human cells. J. Virol. **80**:2855-2862.

303. **Song, R., J. Kafaie, L. Yang, and M. Laughrea**. 2007. HIV-1 Viral RNA Is Selected in the Form of Monomers that Dimerize in a Three-step Protease-dependent Process; the DIS of Stem-Loop 1 Initiates Viral RNA Dimerization. J. Mol. Biol **371**:1084-1098.

304. **Spence, R. A., W. M. Kati, K. S. Anderson, and K. A. Johnson**. 1995. Mechanism of inhibition of HIV-1 reverse transcriptase by nonnucleoside inhibitors. Science **267**:988-993.

305. **Srinivas, R. V., D. Middlemas, P. Flynn, and A. Fridland**. 1998. Human immunodeficiency virus protease inhibitors serve as substrates for multidrug transporter proteins MDR1 and MRP1 but retain antiviral efficacy in cell lines expressing these transporters. Antimicrob. Agents Chemother. **42**:3157-3162.

306. **St Clair, M. H., J. L. Martin, G. Tudor-Williams, M. C. Bach, C. L. Vavro, D. M. King, P. Kellam, S. D. Kemp, and B. A. Larder**. 1991. Resistance to ddI and sensitivity to AZT induced by a mutation in HIV-1 reverse transcriptase. Science **253**:1557-1559.

307. **Sticht, J., M. Humbert, S. Findlow, J. Bodem, B. Muller, U. Dietrich, J. Werner, and H. G. Krausslich**. 2005. A peptide inhibitor of HIV-1 assembly in vitro. Nat. Struct. Mol. Biol **12**:671-677.

308. **Strebel, K.** 2003. Virus-host interactions: role of HIV proteins Vif, Tat, and Rev. AIDS **17 Suppl 4**:S25-S34.

309. **Sturmer, M., S. Staszewski, H. W. Doerr, and K. Hertogs**. 2003. A 6-base pair insertion in the protease gene of HIV type 1 detected in a protease inhibitor-naive patient is not associated with indinavir treatment failure. AIDS Res. Hum. Retroviruses **19**:967-968.

310. **Sun, X., Q. Zhang, and H. M. Al Hashimi**. 2007. Resolving fast and slow motions in the internal loop containing stem-loop 1 of HIV-1 that are modulated by Mg2+ binding: role in the kissing-duplex structural transition. Nucleic Acids Res. **35**:1698-1713.

311. **Tamiya, S., S. Mardy, M. F. Kavlick, K. Yoshimura, and H. Mistuya**. 2004. Amino acid insertions near Gag cleavage sites restore the otherwise compromised replication of human immunodeficiency virus type 1 variants resistant to protease inhibitors. J. Virol. **78**:12030-12040.

312. **Taylor, S., R. P. van Heeswijk, R. M. Hoetelmans, J. Workman, S. M. Drake, D. J. White, and D. Pillay**. 2000. Concentrations of nevirapine, lamivudine and stavudine in semen of HIV-1-infected men. AIDS **14**:1979-1984.

313. **Tessmer, U. and H. G. Krausslich**. 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**:3459-3463.

314. **Tramuto, F., F. Bonura, S. Mancuso, N. Romano, and F. Vitale**. 2005. Detection of a new 3-base pair insertion mutation in the protease gene of human immunodeficiency virus type 1 during highly active antiretroviral therapy (HAART). AIDS Res. Hum. Retroviruses **21**:420-423.

315. **Trono, D., M. B. Feinberg, and D. Baltimore**. 1989. HIV-1 Gag mutants can dominantly interfere with the replication of the wild-type virus. Cell **59**:113-120.

316. **Troyer, R. M., K. R. Collins, A. Abraha, E. Fraundorf, D. M. Moore, R. W. Krizan, Z. Toossi, R. L. Colebunders, M. A. Jensen, J. I. Mullins, G. Vanham, and E. J. Arts**. 2005. Changes in human immunodeficiency virus type 1 fitness and genetic diversity during disease progression. J. Virol. **79**:9006-9018.

317. **Tsibris, A. M. and D. R. Kuritzkes**. 2007. Chemokine antagonists as therapeutics: focus on HIV-1. Annu. Rev. Med. **58**:445-459.

318. **Turner, B. G. and M. F. Summers**. 1999. Structural biology of HIV. J. Mol. Biol. **285**:1-32.

319. **Turner, K. B., N. A. Hagan, and D. Fabris**. 2007. Understanding the isomerization of the HIV-1 dimerization initiation domain by the nucleocapsid protein. J. Mol. Biol **369**:812-828.

320. **UNAIDS/WHO**. 2006. AIDS epidemic update: December 2006.

321. **Upcroft, P., B. Carter, and C. Kidson**. 1980. Analysis of recombination in mammalian cells using SV40 genome segments having homologous overlapping termini. Nucleic Acids Res. **8**:2725-2736.

322. **van Maarseveen, N. M., A. M. Wensing, D. de Jong, M. Taconis, J. C. Borleffs, C. A. Boucher, and M. Nijhuis**. 2007. Persistence of HIV-1 variants with multiple protease inhibitor (PI)-resistance mutations in the absence of PI therapy can be explained by compensatory fixation. J. Infect. Dis. **195**:399-409.

323. **van Opijnen, T. and B. Berkhout**. 2005. The host environment drives HIV-1 fitness. Rev. Med. Virol. **15**:219-233.

324. **van Opijnen, T., A. de Ronde, M. C. Boerlijst, and B. Berkhout**. 2007. Adaptation of HIV-1 Depends on the Host-Cell Environment. PLoS. ONE. **2**:e271.

325. **Vandamme, A. M., F. Houyez, D. Banhegyi, B. Clotet, G. De Schrijver, K. A. De Smet, W. W. Hall, R. Harrigan, N. Hellmann, K. Hertogs, C. Holtzer, B. Larder, D. Pillay, E. Race, J. C. Schmit, R. Schuurman, E. Schulse, A. Sonnerborg, and V. Miller**. 2001. Laboratory guidelines for the practical use of HIV drug resistance tests in patient follow-up. Antivir. Ther. **6**:21-39.

326. **Velazquez-Campoy, A., S. Vega, and E. Freire**. 2002. Amplification of the effects of drug resistance mutations by background polymorphisms in HIV-1 protease from African subtypes. Biochemistry **41**:8613-8619.

327. **Verheyen, J., E. Litau, T. Sing, M. Daumer, M. Balduin, M. Oette, G. Fatkenheuer, J. K. Rockstroh, U. Schuldenzucker, D. Hoffmann, H. Pfister, and R. Kaiser**. 2006. Compensatory mutations at the HIV cleavage sites p7/p1 and p1/p6-gag in therapy-naive and therapy-experienced patients. Antivir. Ther. **11**:879-887.

328. **Verhofstede, C., E. Demecheleer, N. De Cabooter, P. Gaillard, F. Mwanyumba, P. Claeys, V. Chohan, K. Mandaliya, M. Temmerman, and J. Plum**. 2003. Diversity of the human immunodeficiency virus type 1 (HIV-1) env sequence after vertical transmission in mother-child pairs infected with HIV-1 subtype A. J. Virol. **77**:3050-3057.

329. **Vivet-Boudou, V., J. Didierjean, C. Isel, and R. Marquet**. 2006. Nucleoside and nucleotide inhibitors of HIV-1 replication. Cell Mol. Life Sci. **63**:163-186.

330. **Wainberg, M. A.** 2004. HIV-1 subtype distribution and the problem of drug resistance. AIDS **18 Suppl 3**:S63-S68.

331. **Wakefield, J. K., S. A. Jablonski, and C. D. Morrow**. 1992. In vitro enzymatic activity of human immunodeficiency virus type 1 reverse transcriptase mutants in the highly conserved YMDD amino acid motif correlates with the infectious potential of the proviral genome. The Journal of Virology **66**:6806-6812.

332. **Wang, J. H., A. M. Janas, W. J. Olson, V. N. KewalRamani, and L. Wu**. 2007. CD4 coexpression regulates DC-SIGN-mediated transmission of human immunodeficiency virus type 1. J. Virol. **81**:2497-2507.

333. **Wei, X., S. K. Ghosh, M. E. Taylor, V. A. Johnson, E. A. Emini, P. Deutsch, J. D. Lifson, S. Bonhoeffer, M. A. Nowak, B. H. Hahn, and .** 1995. Viral dynamics in human immunodeficiency virus type 1 infection. Nature **373**:117-122.

334. **Welsch, S., O. T. Keppler, A. Habermann, I. Allespach, J. Krijnse-Locker, and H. G. Krausslich**. 2007. HIV-1 buds predominantly at the plasma membrane of primary human macrophages. PLoS. Pathog. **3**:e36.

335. **Welsch, S., B. Muller, and H. G. Krausslich**. 2007. More than one door - Budding of enveloped viruses through cellular membranes. FEBS Lett. **581**:2089-2097.

336. **Westby, M., C. Smith-Burchnell, J. Mori, M. Lewis, M. Mosley, M. Stockdale, P. Dorr, G. Ciaramella, and M. Perros**. 2007. 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**:2359-2371.

337. **Whitehurst, N., C. Chappey, C. Petropoulos, N. Parkin, and A. Gamarnik**. 2003. Polymorphisms in p1-p6/p6* of HIV type 1 can delay protease autoprocessing and increase drug susceptibility. AIDS Res. Hum. Retroviruses **19**:779-784.

338. **Whittaker, G. R. and A. Helenius**. 1998. Nuclear import and export of viruses and virus genomes. Virology **246**:1-23.

339. **Wiegers, K. and H. G. Krausslich**. 2002. Differential dependence of the infectivity of HIV-1 group O isolates on the cellular protein cyclophilin A. Virology **294**:289-295.

340. **Wiegers, K., G. Rutter, H. Kottler, U. Tessmer, H. Hohenberg, and H. G. Krausslich**. 1998. Sequential steps in human immunodeficiency virus particle maturation revealed by alterations of individual Gag polyprotein cleavage sites. J. Virol. **72**:2846-2854.

341. **Winters, M. A., R. M. Kagan, P. N. Heseltine, and T. C. Merigan**. 2005. New two-amino acid insertion near codon 70 of the HIV type 1 protease gene. AIDS Res. Hum. Retroviruses **21**:311-313.

342. **Winters, M. A. and T. C. Merigan**. 2005. Insertions in the human immunodeficiency virus type 1 protease and reverse transcriptase genes: clinical impact and molecular mechanisms. Antimicrob. Agents Chemother. **49**:2575-2582.

343. **Woolfson, D. N. and D. H. Williams**. 1990. The influence of proline residues on alpha-helical structure. FEBS Lett. **277**:185-188.

344. **Wu, T. D., C. A. Schiffer, M. J. Gonzales, J. Taylor, R. Kantor, S. Chou, D. Israelski, A. R. Zolopa, W. J. Fessel, and R. W. Shafer**. 2003. Mutation patterns and structural correlates in human immunodeficiency virus type 1 protease following different protease inhibitor treatments. J. Virol. **77**:4836-4847.

345. **Wyatt, R. and J. Sodroski**. 1998. The HIV-1 envelope glycoproteins: fusogens, antigens, and immunogens. Science **280**:1884-1888.

346. **Yanchunas, J., Jr., D. R. Langley, L. Tao, R. E. Rose, J. Friborg, R. J. Colonno, and M. L. Doyle**. 2005. Molecular basis for increased susceptibility of isolates with atazanavir resistance-conferring substitution I50L to other protease inhibitors. Antimicrob. Agents Chemother. **49**:3825-3832.

347. **Yeni, P.** 2006. Update on HAART in HIV. J. Hepatol. **44**:S100-S103.

348. **Yerly, S., S. Jost, A. Telenti, M. Flepp, L. Kaiser, J. P. Chave, P. Vernazza, M. Battegay, H. Furrer, B. Chanzy, P. Burgisser, M. Rickenbach, M. Gebhardt, M. C. Bernard, T. Perneger, B. Hirschel, and L. Perrin**. 2004. Infrequent transmission of HIV-1 drug-resistant variants. Antivir. Ther. **9**:375-384.

349. **Yin, P. D., D. Das, and H. Mitsuya**. 2006. Overcoming HIV drug resistance through rational drug design based on molecular, biochemical, and structural profiles of HIV resistance. Cell Mol. Life Sci. **63**:1706-1724.

350. **Yoshimura, F. K., K. Diem, G. H. Learn, Jr., S. Riddell, and L. Corey**. 1996. Intrapatient sequence variation of the gag gene of human immunodeficiency virus type 1 plasma virions. J. Virol. **70**:8879-8887.

351. **Yoshimura, K., R. Kato, M. F. Kavlick, A. Nguyen, V. Maroun, K. Maeda, K. A. Hussain, A. K. Ghosh, S. V. Gulnik, J. W. Erickson, and H. Mitsuya**. 2002. A potent human immunodeficiency virus type 1 protease inhibitor, UIC-94003 (TMC-126), and selection of a novel (A28S) mutation in the protease active site. J. Virol. **76**:1349-1358.

352. **Yusa, K., M. F. Kavlick, P. Kosalaraksa, and H. Mitsuya**. 1997. HIV-1 acquires resistance to two classes of antiviral drugs through homologous recombination. Antiviral Res. **36**:179-189.

353. **Yuste, E., C. Lopez-Galindez, and E. Domingo**. 2000. Unusual distribution of mutations associated with serial bottleneck passages of human immunodeficiency virus type 1. J. Virol. **74**:9546-9552.

354. **Yuste, E., J. D. Reeves, R. W. Doms, and R. C. Desrosiers**. 2004. Modulation of Env content in virions of simian immunodeficiency virus: correlation with cell surface expression and virion infectivity. J. Virol. **78**:6775-6785.

355. **Yuste, E., S. Sanchez-Palomino, C. Casado, E. Domingo, and C. Lopez-Galindez**. 1999. Drastic fitness loss in human immunodeficiency virus type 1 upon serial bottleneck events. J. Virol. **73**:2745-2751.

356. **Zeffman, A., S. Hassard, G. Varani, and A. Lever**. 2000. The major HIV-1 packaging signal is an extended bulged stem loop whose structure is altered on interaction with the Gag polyprotein. J. Mol. Biol **297**:877-893.

357. **Zhang, L. Q., P. MacKenzie, A. Cleland, E. C. Holmes, A. J. Brown, and P. Simmonds**. 1993. Selection for specific sequences in the external envelope protein of human immunodeficiency virus type 1 upon primary infection. J. Virol. **67**:3345-3356.

358. **Zhang, Y. M., H. Imamichi, T. Imamichi, H. C. Lane, J. Falloon, M. B. Vasudevachari, and N. P. Salzman**. 1997. Drug resistance during indinavir therapy is caused by mutations in the protease gene and in its Gag substrate cleavage sites. J. Virol. **71**:6662-6670.

359. **Zhou, J., X. Yuan, D. Dismuke, B. M. Forshey, C. Lundquist, K. H. Lee, C. Aiken, and C. H. Chen**. 2004. Small-molecule inhibition of human immunodeficiency virus type 1 replication by specific targeting of the final step of virion maturation. J. Virol. **78**:922-929.

360. **Zhu, P., J. Liu, J. Bess, Jr., E. Chertova, J. D. Lifson, H. Grise, G. A. Ofek, K. A. Taylor, and K. H. Roux**. 2006. Distribution and three-dimensional structure of AIDS virus envelope spikes. Nature **441**:847-852.

361. **Zufferey, R., D. Nagy, R. J. Mandel, L. Naldini, and D. Trono**. 1997. Multiply attenuated lentiviral vector achieves efficient gene delivery in vivo. Nat. Biotechnol. **15**:871-875.