Contribution of Gag and protease to variation in susceptibility to protease inhibitors between different strains of HIV-1

Katherine Sutherland

Institution: UCL

Submitted for: Doctor of Philosophy (PhD)

I, Katherine Sutherland, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

Acknowledgements

I would like to thank each of the supervisors involved during the course of my PhD for their time and guidance: Tamyo Mbisa, Chris Parry, Pat Cane and Deenan Pillay. In am particularly indebted to Tamyo for the time he has given me throughout my project and his help producing this thesis. I am grateful to Chris for his time teaching me many of the lab techniques required for this PhD. I would also like to thank Ravi Gupta, UCL, for enthusiasm and support on the MONARK work and for reading drafts of these chapters.

I would like to thank each of the members of the Antiviral Unit (HPA/PHE) past and present for their advice and support, but most importantly for their friendship. I would like to thank Dr Matthew Hannah for his time and expertise in performing the electron microscopy for this project, Dr Stephane Hue for performing the ancestral site reconstruction in chapter 5 and Dr John Gregson (LSHTM) for his assistance with statistical analyses in chapter 6.

Finally, I would like to thank my family and friends for supporting me throughout and at times for keeping me sane! In particular, I want to thank my parents, siblings and partner Rich for your understanding and unwavering support.

Abstract

Recent reports have shown that HIV-1 Gag can directly affect susceptibility to protease inhibitors (PIs) in the absence of known resistance mutations in protease. Inclusion of co-evolved Gag alongside protease in phenotypic drug susceptibility assays can alter PI susceptibility in comparison to protease with a wild-type Gag.

Using a single replication-cycle assay encompassing full-length Gag together with protease, we demonstrate significant variation in PI susceptibility between a number of PI-naïve subtype B viruses. Six publicly available subtype B molecular clones, namely HXB2, NL4-3, SF2, YU2, JRFL and 89.6, displayed up to 9-fold reduction in PI susceptibility. For two molecular clones, YU2 and JRFL, Gag contributed solely to the observed reduction in susceptibility. Gag and protease from treatment-naïve, patient-derived viruses also demonstrated significant variation in susceptibility, with up to a 17-fold reduction to atazanavir. In contrast to the molecular clones, protease was the main determinant of the reduced susceptibility. Common polymorphisms in protease including I13V, L63P and A71T were shown to contribute to this reduction in PI susceptibility, in the absence of major resistance mutations.

The role of variation in PI susceptibility on LPV/r monotherapy treatment failure was investigated. The contribution of suboptimal adherence to treatment failure was shown and the development of reduced PI susceptibility during treatment

observed. In addition, reduced PI susceptibility and single-round infectivity were associated with subsequent treatment failure.

This study demonstrates significant variation in PI susceptibility of treatment-naïve patient viruses and provides further evidence of the independent role of Gag, the protease substrate, and in particular the amino terminus of Gag in PI susceptibility. It also highlights the importance of considering co-evolved Gag and protease when assessing PI susceptibility. These data indicate that reduced PI susceptibility at baseline may contribute to treatment failure on PI monotherapy.

Table of Contents

1 Introduction....21

1.1 HIV and the AIDS epidemic....21

1.1.1 Global and national burden of HIV/AIDS....21

1.1.2 HIV diversity and geographical distribution....23

1.1.3 Natural course of HIV-1 infection....27

1.2 HIV life cycle....29

1.2.1 HIV-1 genome....30

1.2.2 Cell entry....32

1.2.3 Reverse transcription....33

1.2.4 Virus uncoating and nuclear entry....37

1.2.5 Integration....38

1.2.6 Transcription and translation of viral genes....39

1.2.7 Virion assembly and release....41

1.2.8 Virion maturation....44

1.2.9 Cell-to-cell spread....52

1.2.10 Host factors involved in the HIV-1 life cycle....53

1.3 Treatment of HIV-1 infection....55

1.3.1 Entry and fusion inhibitors....56

1.3.2 Reverse transcriptase inhibitors....57

1.3.3 Integrase inhibitors 60
1.3.4 Protease inhibitors 61
1.3.5 Maturation inhibitors 65
1.4 Resistance to antiretroviral therapy 65
1.4.1 Resistance to entry inhibitors 66
1.4.2 Resistance to reverse transcriptase inhibitors 67
1.4.3 Resistance to integrase inhibitors 68
1.4.4 Resistance to protease inhibitors 68
1.4.5 Resistance to maturation inhibitors 85
1.5 Protease inhibitor resistance testing methods 87
1.5.1 Genotypic resistance testing methods 87
1.5.2 Phenotypic resistance assays 87
1.6 Viral fitness assays 89
1.7 Project overview 90
2 Materials and methods 92
2.1 General microbiological techniques 92
2.1.1 Plasmid DNA preparation 92
2.1.2 Polymerase Chain Reaction (PCR) 93
2.1.3 Sequencing 94
2.1.4 Primers 95

2.1.5 Restriction enzyme digestion 95
2.1.6 DNA purification 96
2.1.7 Molecular cloning 96
2.1.8 Site-directed mutagenesis 98
2.1.9 RNA extraction 99
2.1.10 cDNA synthesis 100
2.1.11 Agarose gel electrophoresis 100
2.2 HIV-1 viruses 101
2.2.1 Molecular clones 101
2.2.2 Clinical samples 101
2.2.3 Viral vectors 104
2.3 Tissue culture methods 109
2.3.1 General cell culture 109
2.3.2 Transfection 109
2.3.3 Protease inhibitor susceptibility assay 110
2.3.4 Single-round infectivity assay 112
2.4 Imaging methods 114
2.4.1 Immunofluorescence (IF) 114
2.4.2 Electron microscopy (EM) 115
2.5 Bioinformatic analyses 116

2.5.1 Sequence analysis 116

2.5.2 Genetic distance calculation 116

2.5.3 Positive selection analysis 116

2.5.4 Molecular Modeling 117

2.5.5 Phylogenetics 117

2.5.6 Ancestral site reconstruction 117

2.5.7 Statistical analyses 118

3 Investigation of the role of co-evolved Gag and protease in PI susceptibility of HIV-1 subtype B molecular clones 119

3.1 Introduction 119

3.2 Results 121

3.2.1 Variation in PI susceptibility of full-length Gag-protease observed between six different subtype B molecular clones 121

3.2.2 Confirmation of the reproducibility of variation in PI susceptibility conferred by full-length Gag-protease from subtype B molecular clones . 124

3.2.3 Contribution of Gag to reduction in PI susceptibility observed for full-length Gag-protease of subtype B molecular clones 126

3.2.4 Role of the N terminus of Gag in reduced PI susceptibility exhibited by JRFL and YU2 Gag-protease 129

3.2.5 Genetic determinants in Gag responsible for the reduced susceptibility to PIs in the molecular clones YU2 and JRFL 131

3.2.6 Role of amino acid changes in N terminal region of Gag on PI susceptibility 135

3.2.7 Single-round infectivity of the molecular clone-based vectors.....139

3.2.8 Analysis of the mechanisms underlying the reduced susceptibility to PIs conferred by Gag-protease of molecular clones.....141

3.3 Discussion 159

4 Investigation of the role of co-evolved Gag and protease in susceptibility to PIs of different HIV-1 subtype B viruses derived from treatment-naïve patients 165

4.1 Introduction.....165

Chapter aims:.....166

4.2 Results.....167

4.2.1 Modification of p8.9NSX+ vector for use with viruses from clinical samples.....167

4.2.2 Variation in PI susceptibility of full-length Gag-protease from treatment-naïve HIV-1 infected patients.....174

4.2.3 Contribution of protease to reduction in PI susceptibility observed for full-length Gag-protease in subtype B clinical viruses.....177

4.2.4 Effect of minor resistance mutations and polymorphisms in protease on PI susceptibility of treatment-naïve, patient-derived viruses.179

4.2.5 Effect of amino acid changes in protease on PI susceptibility of assay reference strain.....183

4.2.6 Computer modelling of positions of amino acid changes within the protease homodimer 185

4.3 Discussion 187

5 Investigation of the genotypic and phenotypic correlates of virological failure in treatment-naïve patients receiving LPV/r monotherapy 192

5.1 Introduction 192

5.1.1 Previous PI simplification clinical trials 192

5.1.2 The MONARK trial 195

5.2 Results 198

5.2.1 Clinical and virological information of patients experiencing virological failure on LPV/r monotherapy, selected for this study. 198

5.2.2 Clonal analysis of Gag-protease from patients experiencing treatment failure 202

5.2.3 Positive selection analyses 207

5.2.4 Phylogenetic and genetic distance analysis 210

5.2.5 Phenotypic PI susceptibility and single-round infectivity of Gag-protease derived from patients receiving LPV/r monotherapy 217

5.3 Discussion 224

6 Genotypic and phenotypic comparison between patients achieving a sustained virological response and those experiencing virological failure on LPV/r monotherapy 232

6.1 Introduction 232

6.2 Results 234

6.2.1 Clinical and virological information for controller patients selected for analysis 234

6.2.2 Clonal analysis of Gag-protease from controller patients 235

6.2.3 PI susceptibility and single-round infectivity of viral variants derived from controller patients 237

6.2.4 Association between PI susceptibility at baseline and therapy failure 240

6.2.5 Reduced single-round infectivity was associated with sustained virological response 243

6.2.6 Exploration of genetic correlates of reduced PI susceptibility 245

6.3 Discussion 246

7 Final discussion and future work 250

7.1 Final discussion 250

7.2 Study limitations 256

7.3 Future work 259

7.3.1 Variation in PI susceptibility of subtype B molecular clones 259

7.3.2 Variation in PI susceptibility of subtype B, patient-derived PI-naïve viruses 262

7.3.3 Determinants of treatment failure in LPV/r monotherapy arm of the MONARK trial 264

7.4 Final comments 267

Table of Figures

Chapter 1: Introduction

1.1 The global prevalence of HIV infection……....22

1.2 The global distribution of HIV-1 subtypes…….25

1.3 The natural course of HIV infection in the absence of ART…….28

1.4 The HIV-1 life cycle and involvement of host proteins ……29

1.5 The HIV-1 genome……...31

1.6 Reverse transcription in HIV-1.…….....36

1.7 The HIV-1 protease enzyme……..45

1.8 Ordered cleavage of the Gag polyprotein into its functional subunits by viral protease……...48

1.9 The structure of Gag in the virion……..49

1.10 The structure of the mature virion……..51

1.11 The structure of RT inhibitors; NRTIs……...58

1.12 The structure of protease inhibitors (PIs)……64

1.13 The known PI resistance mutations within protease……..70

1.14 The position of known PI resistance mutations within the HIV protease protein structure……...71

Chapter 2: Materials and methods

2.1 Unique restriction sites within p8.9NSX+ *gag-pol*……....97

2.2 Layout of 96-well plate for PI susceptibility assays……..111

2.3 Layout of 96-well plate for single-round infectivity assays……..113

Chapter 3: Investigation of the role of co-evolved Gag and protease in PI susceptibility of HIV-1 subtype B molecular clones

3.1 Variation in PI susceptibility of full-length Gag-protease derived from six subtype B molecular clones…….……123

3.2 Reproducibility of PI susceptibility data shown testing PI susceptibility of multiple DNA preparations…….……..125

3.3 Reduced PI susceptibility of molecular clone Gag-protease is independently conferred by both the gag and protease genes…….128
3.4 PI susceptibility of chimeric viruses derived from the N terminus and C terminus of molecular clone Gag demonstrates reduced susceptibility is conferred by the N terminus of Gag…….....130
3.5 The amino acid sequences of Gag cleavage sites for the six subtype B molecular clones included in this study demonstrate no correlation between genotype and phenotypic susceptibility …….131
3.6 Amino acid alignment of the N terminus of Gag for subtype B molecular clones……..134
3.7 Role of amino acids 30 and 102 of Gag in PI susceptibility of full-length Gag-protease…….137
3.8 Role of amino acids 30 and 102 of Gag in PI susceptibility of MCgagN derived pseudovirions……..138
3.9 Limited variation between the single-round infectivity of molecular clone derived, p8.9NSX+ based vectors…..…….140
3.10 Electron micrographs of 293T cells transfected using Fugene-6 reagent……..143
3.11 Transfection efficiency of different conditions, measured by GFP expression …….....145
3.12 Pseudovirus yield of different transfection reagents…..…….147
3.13 Immunofluorescence demonstrates successful transfection of 293T cells and Gag expression……..……...150
3.14 Electron micrographs of pseudovirions…….151
3.15 Counting strategy utilised for electron microscopy experiments……...152
3.16 Presence of ATV affects the proportion of mature pseudovirions present for YU2, but not JRFL..…….155
3.17 Electron micrographs of p8.9NSX+ pseudovirions in the presence and absence of the PI ATV……..156

3.18 Electron micrographs of p8YU2gagpro pseudovirions in the presence and absence of the PI ATV……..157
3.19 Electron micrographs of p8JRFLgagpro pseudovirions in the presence and absence of the PI ATV……..158
3.20 Limitations of the counting strategy utilised for electron microscopy …….....164
Chapter 4: Investigation of the role of co-evolved Gag and protease in susceptibility to PIs of different HIV-1 subtype B viruses derived from treatment-naïve patients
4.1 Modifications to p8.9NSX+ vector and creation of p8.9NSXΔpro…….168
4.2 Amino acid alignment of full-length Gag from patients 1-4…….172
4.3 Protease amino acid alignment for viruses derived from patients 1-4..173
4.4 Variation in PI susceptibility of full-length Gag-protease from HIV-1 infected, treatment-naïve patients……...176
4.5 Variation in PI susceptibility of full-length Gag-protease conferred independently by Gag in a single replication-cycle assay…..…….178
4.6 Protease variation of all patient derived viruses subjected to phenotypic PI susceptibility testing……..181
4.7 Protease changes I13V, L63P and A71T affect PI susceptibility…….. 182
4.8 Amino acid changes I13V, L63P and A71T reduce PI susceptibility of reference strain, p8.9NSX+…….....184
4.9 Computer modeling of the location of amino acid positions in protease shown to affect PI susceptibility……..186
Chapter 5: Investigation of genotypic and phenotypic correlates of virological failure in treatment-naïve patients receiving LPV/r monotherapy
5.1 Selection criteria for patients included in this study……..200
5.2 LPV trough concentrations and time to treatment failure for each patient included in this study……..……...201

5.3 Maximum-likelihood phylogeny of Gag-protease viral variants from baseline and failure time points……...213
5.4 Maximum-likelihood phylogeny of full-length Gag from five patients at baseline and tine of treatment failure, derived by clonal analysis..……...215
5.5 Maximum-likelihood phylogeny of the protease gene from five patients at baseline and tine of treatment failure, derived by clonal analysis…..……...216
5.6 Variation in phenotypic PI susceptibility of viruses from five patients...220
5.7 Single-round infectivity of viral variants from five patients……..221
5.8 PI susceptibility of viruses from all patients, derived from EC_{90} data…222
5.9 Location of residues 36 and 64 in the structure of the protease protein …..……230

Chapter 6: Genotypic and phenotypic comparison between patients achieving a sustained virological response and those experiencing virological failure on LPV/r monotherapy

6.1 Limited variation in PI susceptibility of baseline viral variants from controller patients ……...239
6.2 Single-round infectivity of baseline viral variants from controller patients……...240
6.3 Comparison of PI susceptibility of viral variants from non-controller and controller patients shows a trend towards greater reductions in susceptibility for non-controllers……..242
6.4 Reduced single-round infectivity at baseline correlates with treatment failure …….244

Contents of Tables

Chapter 1: Introduction

1.1 Protease Inhibitors approved for clinical use……..61
1.2 Mutations in Gag previously associated with PI resistance or exposure ……76

Chapter 2: Materials and methods

2.1 Primers used for cDNA synthesis, nested PCR and sequence analysis ……95
2.2 Primers designed for site-directed mutagenesis……98
2.3 Viral loads for treatment-naïve patients infected with subtype B viruses ……..101
2.4 Sample details for MONARK clinical trial patients……103
2.5 p8.9NSX+ based vectors derived from subtype B molecular clones, generated during this study……..106
2.6 Subtype B, treatment-naïve, patient-derived, p8.9NSX+ based vectors generated during this study……..107
2.7 Transfection conditions……109
2.8 Top drug concentrations used in phenotypic susceptibility assay…….111
Chapter 3: Investigation of the role of co-evolved Gag and protease in PI susceptibility of HIV-1 subtype B molecular clones
3.1 Mutations in Gag of the molecular clones that have previously been associated with PI exposure or resistance do not correlate with phenotypic susceptibility……...132
3.2 Transfection optimisation conditions tested ...……..142
Chapter 4: Investigation of the role of co-evolved Gag and protease in susceptibility to PIs of different HIV-1 subtype B viruses derived from treatment-naïve patients
4.1 Unique restriction sites present in p8.9NSX+ utilised for cloning……..168
Chapter 5: Investigation of the genotypic and phenotypic correlates of virological failure in treatment-naïve patients receiving LPV/r monotherapy
5.1 Clinical and virological information for the five patients included in this study……199
5.2 Self-reported adherence to trial medication by week……...199
5.3 Drug resistance mutations and polymorphisms present in protease of patient-derived viruses……..206

5.4 Positively selected sites in Gag and protease ……..209
5.5 Summary of clinical and virological information and data for each patient ……223
Chapter 6: Genotypic and phenotypic comparison between patients achieving a sustained virological response and those experiencing virological failure on LPV/r monotherapy
6.1 Clinical and virological information for patients studied……..234
6.2 Protease polymorphisms in viruses derived from non-controller patients ……..237

Abbreviations table

A	Alanine (amino acid)
AIDS	Acquired Immunodeficiency Symdrome
APV	Amprenavir
ATV	Atazanavir
AZT	Zidovudine
bp	Base pairs
CA	Capsid subunit
cm	centimetre
D	Aspartic acid (amino acid)
DMEM	Dulbecco's modified eagle medium
DNA	Deoxyribonucleic Acid
DRV	Darunavir
d4T	Stavudine
E	Glutamic acid (amino acid)
EM	Electron microscopy
Env	Envelope
G	Glycine (amino acid)
GFP	Green fluorescent protein
gp	Glycoprotein
HIV	Human immunodeficiency virus
I	Isoleucine (amino acid)
IF	Immunofluorescence
K	Lysine (amino acid)
KB	Kilobases
L	Leucine (amino acid)
LPV	Lopinavir
MA	Matrix subunit
ml	millilitre
mm	millimetre
N	Asparagine (amino acid)
NC	Nucleocapsid subunit
ng	Nanogram
nM	Nanomolar
nm	Nanometer
nt	Nucleotide
OD	Optical density
P	Proline (amino acid)
PBS	Phosphate-buffered saline
PI	Protease inhibitor
Pol	Polymerase
Pro	Protease

Q	Glutamine (amino acid)
R	Arginine (amino acid)
RC	Replicative capacity
RT	Reverse transcriptase
S	Serine (amino acid)
SIV	Simian immunodeficiency virus
SQV	saquinavir
T	Threonine (amino acid)
TPV	Tipranavir
UV	Ultraviolet
V	Valine (amino acid) OR vaults
X	Times
µg	microgram
µl	microlitre
3TC	Lamivudine

1 Introduction

1.1 HIV and the AIDS epidemic

In 1981, acquired immunodeficiency syndrome (AIDS) was first identified following the presentation of a collection of patients with unusual opportunistic infections such as Kaposi's sarcoma and *Pneumocystis carinii* pneumonia (PCP), and reduced levels of T helper lymphocytes. In 1983, human immunodeficiency virus (HIV) was identified as the causative agent, although the virus was named lymphadenopathyassociated virus (LAV) and Human T-cell Lymphotropic virus type 3 (HTLV-III) until 1986 (Barre-Sinoussi et al., 1983). HIV is a lentivirus from the family *Retroviridae.* HIV predominantly infects the CD4+ T cells of the human immune system and over the course of infection the number of CD4+ T cells falls below the threshold for adequate function of the immune system, leading to the emergence of opportunistic infections that are eventually fatal. HIV first infected humans after a number of zoonosis events of simian immunodeficiency virus (SIV) from various primates into humans. To date there are two HIV 'types' – HIV-1 and HIV-2, both of which occurred from separate transmission events from different primates. In addition, there are four groups of HIV-1 virus, again each of which is predicted to have derived from separate zoonotic transmission events. Molecular phylogenetic methods have suggested that the initial jump into humans occurred in the early 20th century (Korber et al., 2000).

1.1.1 Global and national burden of HIV/AIDS

Since its discovery in 1981, the global burden of HIV has expanded exponentially and the demographic of those infected has also shifted. In 2010, an estimated 35.3 million individuals were infected with HIV across the world, with approximately 25 million of these individuals living in Sub-Saharan Africa (WHO, 2013).

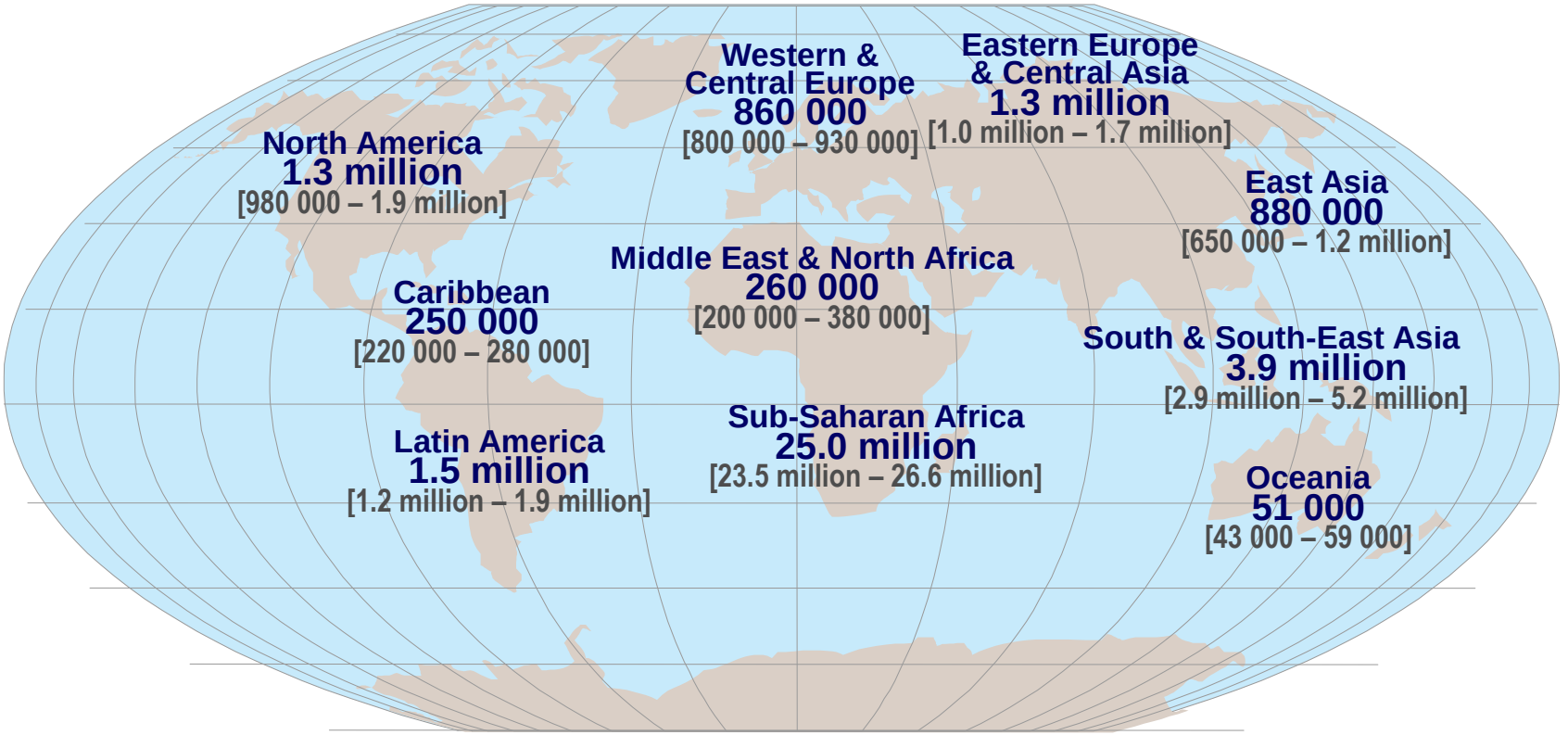


Figure 1.1. The global prevalence of HIV infection. The total estimated number of HIV positive individuals, both adults and children, per region is shown for 2012, with the vast majority of infections occur in Sub-Saharan Africa. This figure is taken from the WHO, 2013.

The global distribution of individuals infected with HIV in 2012 is shown in figure 1.1. Of those infected with HIV in 2010, 3.3 million were children under 15. There are a significant number of AIDS related deaths annually, with 1.6 million reported in 2012 (WHO, 2013). In addition to the direct morbidity and mortality caused by the virus, HIV has a significant economic impact globally and it has been estimated to reduce the GDP in high prevalence countries, with up to an 8% decrease estimated in South Africa (Dixon et al., 2013). The significant financial cost of treatment and research contribute to its impact, with some countries spending up to 20% of their annual health budgets on HIV treatment (Amico et al., 2010).

Within the UK 96,000 individuals are estimated to be HIV positive, but around a quarter of these individuals are unaware of their infection (Health Protection Agency, 2012). The UK epidemic continues to expand, with an estimated 6,280 new infections in 2011. At present, the UK-wide prevalence of HIV is approximately 0.14%, although in some inner city areas this rate rises above 0.2%. It has been estimated that each HIV positive patient will cost £250,000 to treat over their lifetime, providing a significant burden to the National Health Service, NHS (Health Protection Agency, 2012).

1.1.2 HIV diversity and geographical distribution

Globally the majority of HIV infections are with HIV-1, with just a small fraction caused by HIV-2 viruses. Currently, four groups of HIV-1 viruses have been described M ‘major’, N ‘non-M and non-O’, O ‘outlier’ and P with the vast majority of infections worldwide caused by HIV-1 group M viruses. Each of the four HIV-1 groups are thought to have arisen from separate transmission events into humans, with groups M and N originating from SIV infecting chimpanzees (*Pan troglodytes troglodytes)* and O and P from SIV infecting wild gorillas (Plantier et al., 2009; Van Heuverswyn et al., 2006). Viruses from groups N and P, the most recently described group, have been reported in small numbers of

individuals in Cameroon (Plantier et al., 2009; Vallari et al., 2010; Vallari et al., 2011). Group O viruses have also been identified in Cameroon, although at a higher frequency than groups N and P, causing a reported 2% of infections nationwide (Brennan et al., 2008).

Group M viruses can be divided into phylogenetically distinct subtypes: A1, A2, B, C, D, F1, F2, G, H, J and K, although A1, A2, F1 and F2 are sometimes considered sub-subtypes. In addition, in patients dually-infected with two subtypes, recombination between these viruses can occur resulting in the production of unique recombinant forms (URFs). URFs are designated as circulating recombinant forms (CRFs) when reported in at least three epidemiologically distinct individuals, and numbered according to the order of their discovery. Viruses initially identified as subtype E are now widely recognised as CRF01_AE as they comprise predominantly subtype A genomes with a subtype E *env* gene.

The global distribution of HIV-1 subtypes is not uniform, and is shown in figure 1.2. The most predominant subtypes globally are: subtype C (52%), A1 (12%) and B (10%) (Arien et al., 2007). In Southern Africa, the Horn of Africa and Central Asia, subtype C viruses predominate whereas in the Americas and Europe, subtype B viruses cause the majority of infections. CRF01_AE (3.1%) and CRF02_AG (6.7%) are the most predominant CRFs globally, with CRF01_AE being the predominant subtype in South Asia and CRF02_AG causing a significant proportion of infections in West Africa.

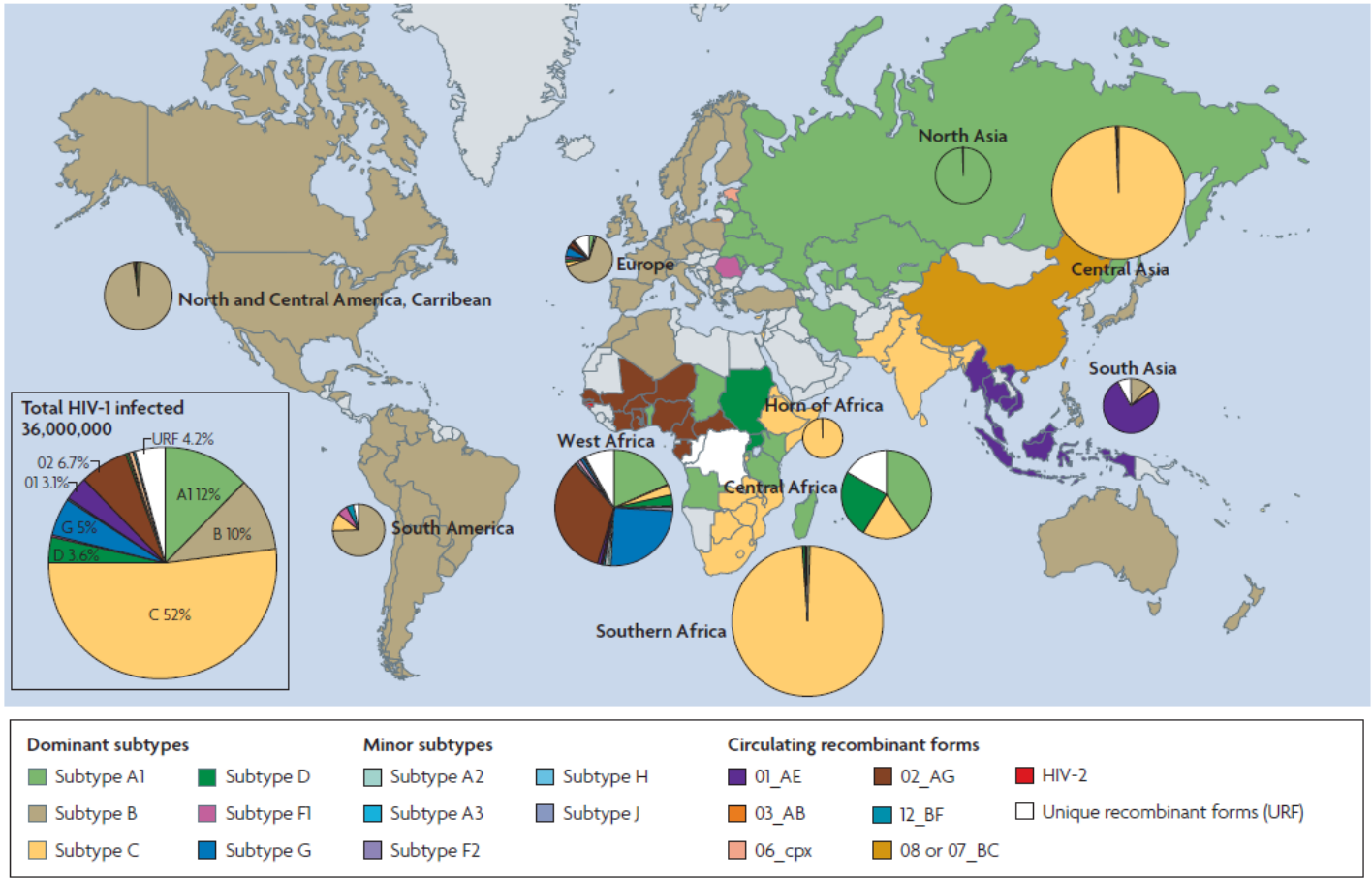


Figure 1.2. The global distribution of HIV-1 subtypes. The frequency of each subtype in each region was estimated and is depicted in the pie-charts, with the colour of the countries representing the predominant subtype in the region. The pie chart on the left represents the global frequency of each subtype and major CRF. Subtype B is the most prevalent in Western Europe and the Americas, whilst subtypes A1, C and CRF02_AG are the most frequent in Africa. Figure taken from Arien et al. 2007.

The HIV-1 subtypes present in the UK are diverse, as a result of a number of separate transmission routes within the UK and migration events into the country. Between 2002 and 2010, the prevalence of subtypes in the UK was as follows: B (39.9%), C (34.3%), A (5%), novel recombinants (9.9%), D (2.5%), G (2.7%) and CRF01_AE (2%) (Dolling et al., 2013a). During this period, men who have sex with men (MSMs) were predominantly infected with subtype B viruses (88.4%), although the percentage of infections with non-B and non-C viruses in MSMs has increased from 4.8% in 2002 to 12.4% in 2010. A different subtype distribution is present in individuals infected through heterosexual contact in the UK, with subtype C viruses the most frequent causing 46.7% of infections in men and 55% in women (Dolling et al. 2013).

HIV-2 was first described in 1986 and has been shown to be most closely related to SIV viruses isolated from sooty mangabeys. HIV-2 infection has primarily been found in West Africa including Guinea Bissau and Senegal, with a few cases reported in western Europe and America (as reviewed by Sharp and Hahn, 2011). HIV-2 viruses are not classified into subtypes, but into groups as the diversity between groups of HIV-2 viruses is similar to that between the HIV-1 virus groups. Unlike HIV-1, HIV-2 recombinants between groups are rare although they have been described (Ibe et al., 2010). HIV-2 has been shown to be less infectious via the mother-to-child transmission route than HIV-1 and most individuals have a lower viral load in comparison with those infected with HIV-1 (as reviewed by Sharp and Hahn, 2011).

1.1.3 Natural course of HIV-1 infection

Transmission of HIV between individuals occurs primarily through direct sexual contact, either heterosexual or homosexual. However, a small proportion of infections occur via direct inoculation of the blood through the use of contaminated needles or blood products. Infection can also occur via the mother-to-child transmission route during gestation, birth or breastfeeding. In the absence of interventions in HIV positive pregnant women, transmission to the baby occurs in approximately 25% of cases (Coutsoudis et al., 1999). However, the use of appropriate prevention strategies has resulted in this rate dropping to around 1% in developed countries such as the UK (Bailey et al., 2011). In the UK in 2009, 95% of pregnant women received an HIV test during pregnancy, which enabled the targeting of appropriate prevention methods to reduce mother-to-child transmission. In 2009, of those HIV positive individuals who were aware of their status and accessing care in the UK, 51% were infected heterosexually, 43% were MSMs, 2% were infected through injecting drug use and 2% by mother-to-child transmission (Health Protection Agency, 2012).

The course of natural HIV-1 infection, without treatment intervention, can be seen in figure 1.3. Following primary infection, approximately 50% of individuals experience flu-like symptoms with the remaining 50% being asymptomatic. During the first phase of infection, the acute phase, high levels of viral replication take place resulting in a viral load reaching 10^6 or 10^7 copies/ml in the blood and cerebral spinal fluid (Cohen et al., 2012). At this stage, infection can be detected reliably using qPCR methods to detect viral RNA. This is followed by the clinical latency phase characterised by a reduced viral load that stabilises at the viral load set point, which has been associated with progression to AIDS (Hodcroft et al., 2014). The clinical latency phase is also characterised by a gradual reduction in CD4+ T cell count, although relatively few clinical symptoms, if any, are observed. The final phase of infection is

characterised by a drop in CD4+ T cell count below levels able to maintain a normal immune response, often at between 250-300 cells/µl. An increase in viral load occurs and clinical symptoms become apparent in the form of opportunistic infections. The time taken for the progression after initial infection to AIDS varies between individuals, but can be a few years to decades (Hodcroft et al., 2014).

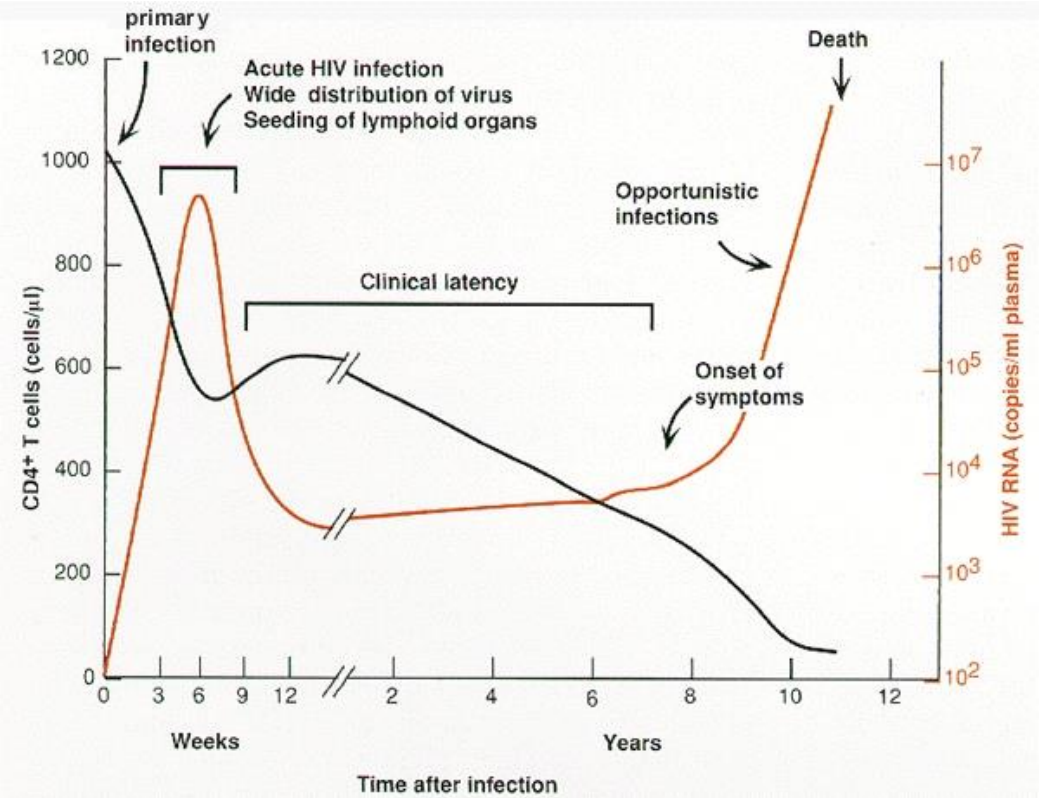


Figure 1.3. The natural course of HIV infection in the absence of ART therapy. Acute infection is associated with a high viral load and rapid decline in CD4+ T cell count. Viral load stabilises at a set point which is maintained during clinical latency. Gradual decline in CD4+ T cell count eventually compromises immune function, leading to increased viral load and onset of AIDS. Figure taken from Coffin et al., 1997.

1.2 HIV life cycle

An overview of the HIV life cycle is displayed in figure 1.4.

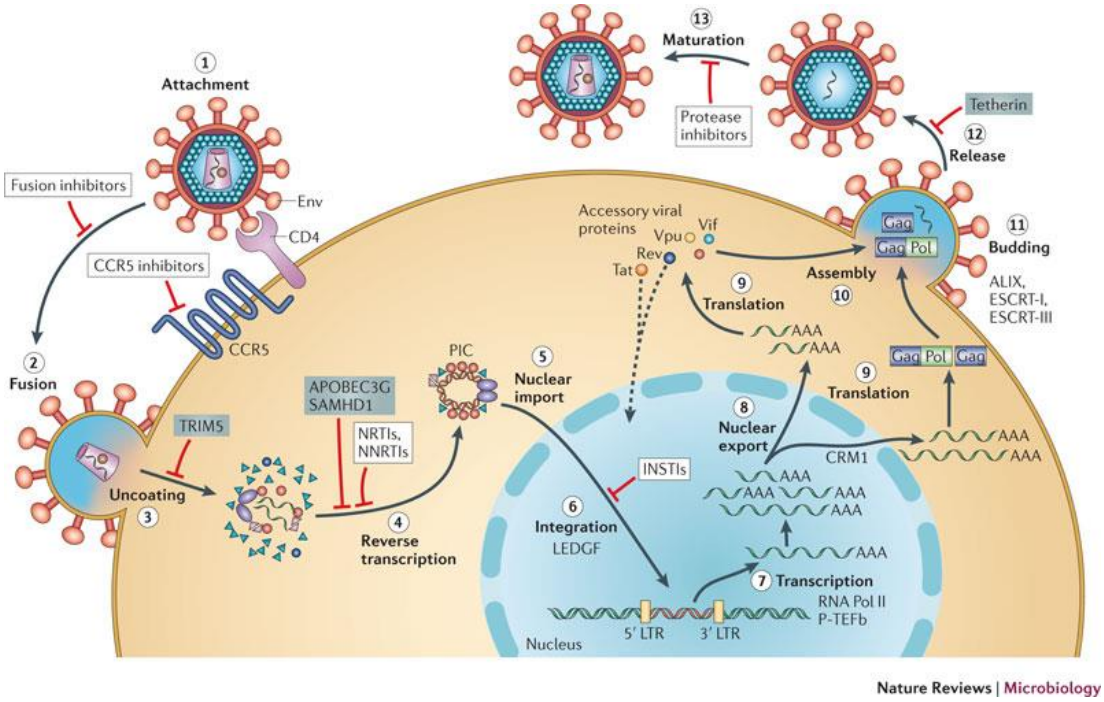


Figure 1.4. The HIV-1 life cycle and involvement of host proteins. The first stage of infection involves the interaction of the Env glycoproteins with CD4 and the CCR5 co-receptor, leading to membrane fusion and entry. Then reverse transcription and uncoating of the viral nucleic acids occurs, although the timing of these two processes in relation to each other remains controversial. The pre-integration complex is imported into the nucleus and cDNA is integrated into the host genome. Transcription of the viral genes takes place and viral mRNAs are exported from the nucleus at different times, depending on the degree of splicing and the action of accessory protein Rev. Assembly of the virus occurs at the plasma membrane and is driven by the Gag polyprotein. The virus buds from the cell and is released, before undergoing protease-dependent maturation. Figure from Engelman and Cherepenov, 2012.

1.2.1 HIV-1 genome

In common with all retroviruses, HIV possesses two copies of a single stranded RNA genome, which must be converted to a double stranded DNA intermediate and integrated into the host genome before viral replication can take place. The HIV-1 genome is 9.8 kb long and encodes 3 essential genes found in all retroviruses (*gag, pol and env*) and a number of accessory and regulatory genes unique to HIV (*vif, vpr, vpu, tat, rev and nef*). A schematic representation of the HIV-1 genome can be seen in figure 1.5. The *gag* gene encodes the structural proteins necessary for viral particle formation, *pol* encodes the enzymes required for the HIV-1 life cycle and *env* encodes envelope glycoproteins necessary for binding of the virus to host cells. The accessory and regulatory genes perform a variety of functions which increase viral transcription and fitness. *Tat* is involved in transcription activation whilst *rev* regulates export of transcribed genomic RNA from the nucleus (Karn and Stoltzfus, 2012). *Vif* has been shown to affect the infectivity of viral particles in part by counteracting APOBEC proteins (apolipoprotein B mRNA-editing enzyme-catalytic peptide-like) (Malim et al., 2012). Nef is involved in infectivity and pathogenicity and has been shown to downregulate the expression of cell surface receptors including CD4 (Landi et al., 2011). Vpr has a number of roles including the induction of cell cycle arrest (Guenzel et al., 2014). Finally, *vpu* is an integral membrane protein that counteracts innate restriction by host tetherin (Neil et al., 2008).

The viral genome is flanked by long terminal repeats (LTRs) at both the 5' and 3' ends which can be divided into 3 elements, U5, U3 and R, as discussed further in sections 1.2.3 and 1.2.6. Following the 5' LTR is the packaging signal (ψ) which labels the RNA for incorporation into virions during particle assembly by the packaging machinery. Ψ is between 80 and 150 nucleotides long and contains four stem-loop structures (SL1-4) (as reviewed by Sundquist and Krausslich, 2012). Additionally, approximately 150 nucleotides from the 5' end

of the RNA genome is a primer binding site (PBS) which is important for the initiation of reverse transcription (as reviewed by Hu and Hughes, 2012).

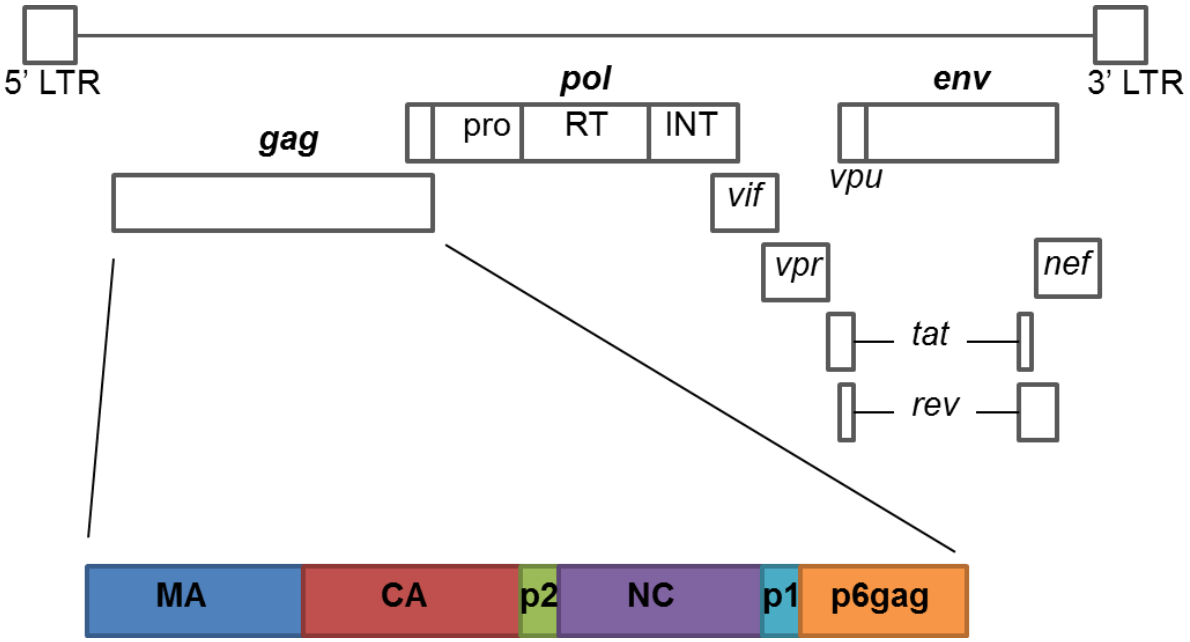


Figure 1.5. The HIV-1 genome. The structure of the 9.8kb, single-stranded RNA HIV-1 genome is shown including the position of *gag, pol, env, t*he accessory and regulatory genes and the LTRs. The three enzymes derived from the *pol* gene are also denoted: protease, reverse transcriptase (RT) and integrase (INT). In addition, the individual subunits of the Gag polyprotein are shown: matrix (MA), capsid (CA), p2, NC (nucleocapsid), p1 and p6.

1.2.2 Cell entry

1.2.2.1 Cellular receptors

HIV-1 binds to one primary cellular receptor (CD4) and one of two co-receptors (CCR5 and CXCR4) to enable entry into host cells, as shown in figure 1.4 (Doms and Peiper, 1997; Bour et al., 1995). CD4 is a member of the immunoglobulin superfamily that interacts, along with the T cell receptor, with MHC class II molecules on antigen presenting cells inducing T cell activation via a signaling cascade (as reviewed by Wilen et al. 2012). CCR5 and CXCR4 have a role in immune cell trafficking and belong to a family of chemokine receptors with a seven-transmembrane domain coupled to a G-protein. CCR5 is expressed on CD+ T cells, whilst CXCR4 is found on a wider range of cells both within and outside the immune system (as reviewed by Didigu and Doms, 2012). Viruses either use CCR5 (R5 tropism), CXCR4 (X4) or both co-receptors (R5X4) for entry. The majority of transmitted viruses use CCR5 for entry, whereas switch to X4-tropism is associated with progression of disease (as reviewed by Wilen et al., 2012).

Prior to the binding of HIV-1 to CD4 and secondary co-receptors, Env-dependent attachment to a number of cell-surface molecules can occur. DC-SIGN, a type II membrane protein with C-type lectin domain, along with several other lectins, have been shown to bind Env and boost *in vitro* infection (Geijtenbeek et al., 2000). DC-SIGN is expressed on dendritic cells (DCs) and captures HIV-1, enabling its presentation to CD4+ T cells, thus enhancing the efficiency of infection (Hijazi et al., 2011).

1.2.2.2 Viral envelope proteins

The HIV-1 *env* gene encodes the Env proteins that mediate attachment and fusion of the virion with the host cell membrane. Env is a type I integral membrane viral glycoprotein and is expressed as a monomeric protein, gp160. Gp160 undergoes post-translational modifications including N-linked glycosylation and is subsequently cleaved by furin cellular proteases, resulting

in the production of two subunits gp120 and gp41 (as reviewed by Wilen et al., 2012). Gp41 is a transmembrane glycoprotein that anchors the Env protein to the virus membrane, whilst gp120 is a surface glycoprotein responsible for receptor binding. Within the viral envelope, Env is trimeric with three gp120s each noncovalently bound to a gp41 subunit. Gp120 is composed of five conserved (C1-C5) and five variable (V1-V5) domains (as reviewed by Didigu and Doms, 2012).

1.2.2.3 Cellular binding and membrane fusion

The viral envelope is derived from the host cell membrane and studies have shown that as few as 7 to 15 Env spikes may be found on each virion. There is evidence that Env spikes are present in a single cluster on the surface of the mature virion and that the clustering is dependent on the cleavage of the Gag MA subunit from the CA subunit during maturation (Chojnacki et al., 2012). Env gp120 initially binds the N terminus of CD4, inducing a conformational change in gp120 that allows the binding of gp120 to either co-receptor. Interactions between the co-receptor binding site and co-receptors result in the exposure of the 15 amino acid fusion peptide within gp41, which is inserted into the plasma membrane and brings about membrane fusion (as reviewed by Wilen et al., 2012).

1.2.3 Reverse transcription

1.2.3.1 HIV reverse transcriptase enzyme

HIV-1 *pol* encodes a reverse transcriptase (RT) enzyme that catalyses the conversion of viral RNA to a double stranded cDNA intermediate, which is later integrated into host DNA forming the proviral template from which transcription of viral genes by host cellular machinery occurs (figure 1.4). RT has two distinct enzymatic activities: RNA polymerase and RNAseH (degrades RNA in an RNA/DNA hybrid). It is composed of two subunits, p66 and p51, that share a common N terminus. The p66 subunit contains two spatially distinct active sites

for the separate activities of RT, whilst the p51 subunit has a structural role (Sarafianos et al., 2009).

1.2.3.2 The process of reverse transcription

Reverse transcription is a complex, multi-stage process, as illustrated in figure 1.6. A $tRNA^{Lys}$ binds to the primer binding site (PBS) 180 nucleotides from the 5' end of RNA, acting as a primer for reverse transcription. DNA synthesis proceeds from the $tRNA^{Lys}$ primer to the 5' end of the RNA generating an RNA/DNA hybrid. The RNA within the RNA/DNA hybrid is degraded by the RNaseH activity of RT leaving a single-stranded DNA fragment, the minus-strand strong stop DNA. Short regions of homology ('R' regions) exist which enable the minus-strand strong stop DNA to 'jump' to the 3' end of the genome, the first DNA strand transfer. Minus-strand DNA synthesis and RNaseH activity proceed (as reviewed by Wu and Hughes, 2012). However, the polypurine tract (PPT) is resistant to degradation and remains to act as a primer for positive strand DNA synthesis. Following positive-strand synthesis, the RNaseH activity removes the $tRNA^{Lys}$ primer from the 3' end of minus-strand DNA, exposing a single stranded DNA sequence complementary to sequences near the 3' end of the plus-strand RNA. The complementary PBS sequences cause the two DNA strands to anneal, the second DNA strand transfer. DNA synthesis is then completed with both the positive and negative strand acting as templates (as reviewed by Wu and Hughes, 2012).

HIV has an additional PPT located in the centre of the genome (cPPT). When positive strand DNA synthesis from the PPT proceeds past the cPPT, the strand produced from the cPPT is displaced for approximately 90-100 nucleotides (Hu et al., 2010). The resulting triple-stranded DNA flap has been reported to be important for nuclear entry, infectivity and protection from APOBEC3 host proteins, although its exact contribution remains controversial (Iglesias et al., 2011).

The RT enzyme has two features that contribute to the high levels of viral variation within an infected individual and between different hosts. Firstly, RT has a high error rate due to the lack of proof-reading capability, with an error occurring every 1 in 2000-7000 base pairs during replication. Secondly, RT has low affinity for RNA and so can switch between the two RNA copies during reverse transcription which will results in a recombinant cDNA molecule if the two RNA molecules are distinct. Studies using different techniques report different frequencies for these recombination events, but the current estimate is between four and five times per genome produced (as reviewed by Onafuwa-Nuga and Telesnitsky, 2009). These features of RT, along with the high rate of viral replication, result in the existence of ‘quasispecies’ within a patient with differing fitness and replication capacities. In the presence of selective pressure, such as that applied by drug therapy, the random distribution of genetic changes present within the quasispecies facilitates the evolution of drug resistance.

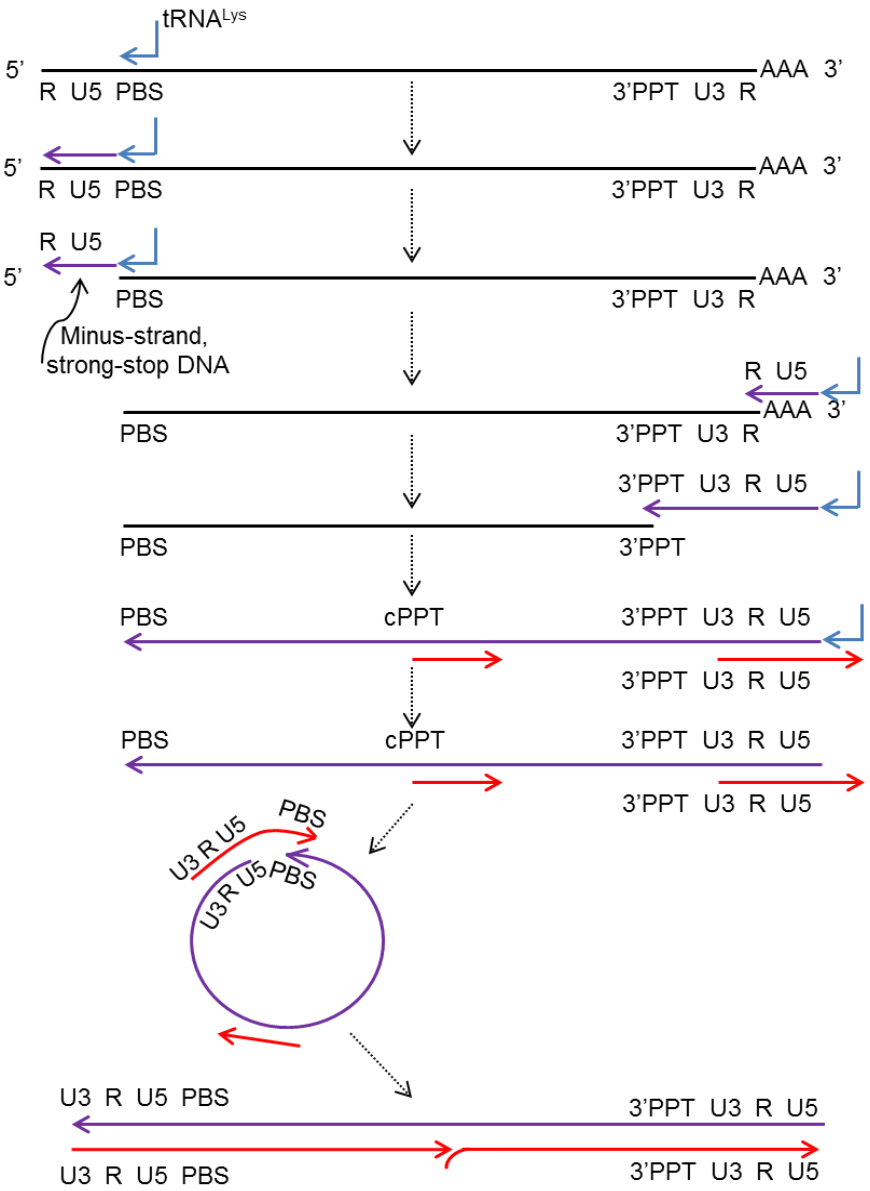


Figure 1.6. Reverse transcription in HIV-1. The conversion of the single-stranded RNA genome to double stranded cDNA is shown. A tRNA primer binds to the PBS at the 5' end of the RNA (light blue) and DNA synthesis proceeds, generating minus-strand DNA (purple). RNAseH activity degrades the RNA and the DNA jumps to homologous R regions at the 3' end of the RNA, enabling minus-strand DNA synthesis to continue accompanied by RNAseH degradation. The PPT is resistant to degradation and acts as a primer for positive-strand DNA synthesis (red line). RNAseH activity results in complementary PBS at each end of the RNA and the second strand transfer occurs, with each DNA strand acting as a template for synthesis. The presence of the cPPT results in the formation of a triple-stranded DNA flap.

1.2.4 Virus uncoating and nuclear entry

Following entry into the cell, the viral capsid moves through the cell by active transport and studies have provided evidence for the role of microtubules and the actin cytoskeleton it this process (as reviewed by Lehmann et al., 2011). HIV-1 is able to enter the nucleus of a non-proliferating cell enabling it to infect important cell types such as terminally differentiated macropahges. However, other retroviruses including murine leukemia virus (MLV) and spleen necrosis virus (SNV) require mitosis for productive infection and cannot integrate in non-diving cells (Yamashita and Emerman, 2006).

The exact mechanisms of the uncoating of the capsid core and the transport of the viral nucleic acid into the nucleus are yet to be fully elucidated. However, the HIV capsid is too large to enter the nucleus through the nuclear pore suggesting that uncoating of the viral cDNA must occur before nuclear entry (Hilditch and Towers, 2014). Despite numerous studies, it is currently unclear when uncoating occurs in relation to the timing of reverse transcription and where in the cell uncoating takes place (as reviewed by Fassati, 2012). Technical difficulties have contributed to this lack of understanding, in particular that some assays measure what occurs in the majority of virions, but these are not necessarily the virions that cause productive infections (as reviewed by Hilditch and Towers, 2014). Studies have provided evidence for the importance of host factors in directing HIV uncoating and nuclear entry, and it has been suggested that the use of host cofactors enables evasion of the innate immune response. When mutations in CA are present that prevent interactions with the host cofactors used for uncoating and nuclear entry, innate sensors are triggered and an antiviral state is induced (Rasaiyaah et al., 2013).

A number of cofactors have been suggested to play a role in the targeting of the viral capsid to the nucleus and in nuclear entry: cyclophilin A (CypA), CPSF6, TNPO3, Nup358 and Nup153. CypA binds CA directly and when binding is

prevented, subsequent interactions with other cofactors and the integration sites are affected (Schaller et al., 2011). CPSF6 also binds CA directly and is has been suggested that this interaction directs HIV to a particular nuclear binding pathway involving TNPO3 and Nup358 and prevents premature reverse transcription. HIV CA interacts directly with nuclear pore protein Nup358 through its cyclophilin-like domain and it has been hypothesised that this interaction tethers the capsid to the nuclear pore enabling the subsequent interactions required for nuclear entry (Schaller et al., 2011). Entry of the viral cDNA is thought to occur through the nuclear basket protein Nup153 and studies have provided evidence for the interaction of Nup153 with integrase and CA (as reviewed by Hilditch and Towers, 2014).

1.2.5 Integration

1.2.5.1 HIV integrase enzyme

Once inside the nucleus the integration of viral DNA into the host genome is catalysed by the viral integrase (IN) enzyme, which enters the nucleus as part of the PIC, as shown in figure 1.4. Integrase has three distinct domains: the N terminal domain, the catalytic core domain and the C terminal domain. The N terminal domain contains an HHCC motif that plays a role in the dimerisation of integrase monomers and the binding of cellular factors. The catalytic domain contains the integrase active site, essential for enzyme function. The C terminal domain contains domains that non-specifically bind to DNA, so has been implicated in stability of the integrase/DNA complex (Delelis et al., 2008). It is thought that integrase functions as a tetramer, composed of two symmetrical dimers of integrase (as reviewed by Mbisa et al. 2011).

1.2.5.2 Integration

Initiation of integration occurs when integrase catalyses the cleavage of two nucleotides from both 3’ ends of the viral DNA, resulting in free 3’ hydroxyl groups. Once in the nucleus, the free 3’ hydroxyl groups attack phosphodiester bonds of the host DNA. This strand transfer reaction results in the ligation of the

viral DNA into the host genome. Host cellular repair enzymes fill the gaps remaining (as reviewed by Craigie and Bushman, 2012). HIV-1 has been shown to preferentially integrate at sites of active genes and integration can lead to latent infection if it occurs at a transcriptionally inactive site. There is evidence that the nuclear entry pathway impacts on integration site selection (Schaller et al., 2011). Host factor LEDGF has been shown to tether the integrase-DNA complex to the host chromatin, enabling correct targeting of integration. In addition, LEDGF stimulates the activity of the integrase enzyme and protects it from degradation (as reviewed by Christ and Debyser, 2013).

In addition to integrated cDNA, other forms of viral DNA are present in the nucleus: linear cDNA, 1-LTR and 2-LTR circles. 1-LTR circles can be formed by homologous recombination of linear cDNA. However, it has been shown that 2-LTR circles are the product of non-homologous end joining (NHEJ) repair mechanisms within the nucleus (as reviewed by Sloan and Wainberg, 2011). 2-LTR circles are used as markers of nuclear entry as PCR can be designed to amplify across the unique LTR-LTR joint. It has also been shown that when integrase is blocked, for example when integrase inhibitors are present, there is an increase in the number of 1-LTR and 2-LTR circles (Buzon et al. 2010).

1.2.6 Transcription and translation of viral genes

Transcription of viral proteins occurs from the proviral template, cDNA integrated into the host genome. The viral LTR contains promoter regions including an initiator and TATA-box, which enable the correct positioning of host RNA polymerase II at the site of initiation of transcription (Berkhout and Jeang, 1992). The cellular NF-κB and viral Tat proteins are involved in increasing the rate of initiation and elongation of viral transcription. Tat stimulates elongation of the viral RNA via its interaction with the Transactivation Response Element (TAR) in the cDNA. Tat is encoded by two exons: the first encodes 72 amino acids and is found upstream of *env* and the second exon varies between 14 and

29 amino acids long and is located at the 3' end of *env,* as shown in figure 1.5. Tat complexes with host factor pTEFb and host proteins CDK9 and CycT1 are also recruited to the elongation complex. Recruitment of CDK9 kinase leads to its activation and the hyperphosphorylation of the RNA polymerase II enzyme, leading to enhanced elongation. In the absence of Tat viral RNA elongation is blocked by the negative elongation factor NELF (as reviewed by Karn and Stoltzfus, 2012). This is different from traditional transcription factors that bind to DNA and promote the recruitment of RNA polymerase.

Studies have shown that large numbers of differentially spliced mRNA species are produced, with up to 40 distinct mRNA species identified (as reviewed by Stoltzfus, 2009). The viral RNA undergoes posttranscriptional processing including the addition of a 5' cap and polyadenylation of the 3' end, as with cellular mRNA to shield it from recognition by the host innate immune system. The early genes *tat, rev* and *nef* are expressed from a fully spliced mRNA species that is expressed early in the HIV life cycle and can be exported from the nucleus without any additional viral mechanisms as it does not contain introns. Other mRNA species that are partially spliced or un-spliced, containing introns, require the presence of Rev that interacts with the rev response element (RRE) for their nuclear export. Rev is encoded by two exons which overlap with those of *tat* and contains a nuclear export signal and an RNA binding region. Rev binds to and mulitmerises on the RRE, and forms an export complex with host proteins Crm1 and RanGTP that is exported from the nucleus through the nuclear pore. Crm1 has been shown to interact with nuclear pore proteins including Nup358 and Nup214 (as reviewed by Kuzembayeva et al., 2014). Once exported, the partially spliced mRNAs express Env, Vif, Vpu, Vpr and Tat whilst the unspliced mRNA expresses full-length Gag-Pol and acts as genomic RNA for incorporation in new virions (as reviewed by Karn and Stoltzfus, 2012).

Translation of *gag* from unspliced mRNA results in two polyproteins of different lengths: Gag and the Gag-Pol polyprotein. The *pol* gene is at a -1 reading frame relative to *gag* so its translation requires a ribosomal frameshift event following the translation of *gag* to express the Gag-Pol polyprotein. This frameshift event occurs in approximately 5% of translations when the ribosome reaches an internal secondary hairpin structure coupled with a slippery sequence, UUUUUU, at the end of the *gag* gene (as reviewed by Karn and Stoltzfus, 2012). The ratio of 1 Gag-Pol polyprotein molecule to 20 Gag polyprotein molecules results in the optimum levels of the polymerase proteins for viral fitness. Translation of the viral proteins occurs in different regions of the cell. Gag and Gag-Pol are translated in the cytosol by free ribosomes before migrating through the cytosol to the plasma membrane, the site of viral particle assembly. *Env* is translated from a spliced transcript as it travels through the rough endoplasmic reticulum to the plasma membrane, via the Golgi apparatus (as reviewed by Sundquist and Krausslich, 2012).

1.2.7 Virion assembly and release

1.2.7.1 Virion assembly

Virion assembly takes place at the host cell plasma membrane from which budding of the virion occurs, as shown in figure 1.4. Particle assembly involves a number of stages: targeting of Gag to and interaction with the plasma membrane, Gag multimerisation, recruitment of genomic RNA, incorporation of Env protein and budding and release of immature particles from the plasma membrane (as reviewed by Ono, 2010). The Gag protein drives particle assembly, in particular it directs virion budding, controls virion size and ensures the inclusion of other vital components. Four Gag subunits are directly involved in different stages of viral assembly: matrix (MA), capsid (CA), nucleocapsid (NC) and p6.

Once synthesized in the cytosol, a highly basic sequence between amino acids 16 to 31 of Gag MA has been identified as the region responsible for targeting

Gag to the plasma membrane. These residues form a cluster on the surface of the MA globular domain that is conserved among retroviruses. A myristic acid is covalently attached to the basic residues in the MA amino terminus which anchors Gag to the plasma membrane lipid bilayer and has been shown to be essential for viral budding (Bukrinskaya, 2004). In particular, amino acids 29 and 31 of MA are necessary for membrane targeting as substitution results in localisation of Gag to intracellular compartments (Ono and Freed, 2004). Gag targets PI(4,5)P2 [phosphatidylinositol-(4,5)-bisphosphate] within the plasma membrane and in its absence Gag localises to internal cellular membranes (Ganser-Pornillos et al., 2008; Ono et al., 2004).

Studies using fluorescence imaging techniques (FRET) have shown that multimerisation of full-length Gag occurs at the plasma membrane (Derdowski et al., 2004). Viral RNA and Gag subunits CA and NC have been implicated in Gag multimerisation. Viral RNA has been shown to promote Gag multimerisation through its interactions with NC, so it is thought that NC does not mediate multimerisation directly (Martin-Serrano et al., 2011). Mutations in the C terminal domain of CA have been shown to affect assembly (as reviewed by Martin-Serrano et al., 2011; Kutluay and Bieniasz, 2010).

Two copies of the viral RNA genome are incorporated into each virion from the full-length, unspliced viral RNA present in the cytoplasm. NC is involved in genomic RNA recruitment and RNA dimerisation is required for its incorporation into virions (as reviewed by Moore and Hu, 2009). Genomic RNA encodes a packaging signal, ψ, in the 5’ region of Gag which labels it for packaging into the virion. Ψ encodes 4 stem loop structures (SL1-4) which interact with two zinc finger motifs of Gag NC to bring about RNA encapsidation. However, most of the 5’ UTR is involved in RNA genome encapsidation (as reviewed by Sundquist and Krausslich, 2012).

The viral envelope proteins, gp120 and gp41, are also sequestered to the host cell membrane for incorporation in the viral membrane. The Env gp160 precursor is synthesised in the rough endoplasmic reticulum and cleaved into two subunits by cellular furin proteases during its transport to the cell surface via the golgi apparatus. The process of Env recruitment is not yet fully understood, but it has been shown to require the MA subunit of Gag (as reviewed by Sundquist and Krausslich, 2012). A mutation at position 49 of MA decrease the incorporation of Env into the virion, but deletion of the gp41 C terminal domain reverses this effect, implying that the interaction occurs between MA and the C terminus of Env (Davis et al., 2006). Recent work has shown that cleavage of Gag, in particular the separation of MA, during maturation is essential for the formation of a single focus of Env, required for viral infectivity (Chojnacki et al., 2012).

A number of viral accessory proteins are also packaged into the assembling virion as they are required in the early stages of the life cycle in the next infected cell. Vpr is incorporated via a direct interaction with the Gag p6 subunit, which has been shown to be sufficient for its inclusion (as reviewed by Sundquist and Krausslich, 2012).

1.2.7.2 Virion budding and release

The Gag lattice that forms during HIV-1 virion assembly is continuous, but consists of a gap thought to cover approximately 30% of the virion, as shown in figure 1.9a. This gap is thought to correlate with the area of the membrane from which scission occurs, required for resolution of the membrane stalk connecting the cell and virion (as reviewed by Briggs and Krausslich, 2011). This scission process requires the p6 subunit of Gag which has late-budding domain (L-domain) activity and mutations in p6 prevent release of budding virus from cells (Gottlinger et al., 1991). The p6 subunit contains two regions required for viral budding: a P(S/T)AP domain and the (L)YPXNL motif. The P(S/T)AP domain interacts with the cellular endosomal sorting (ESCRT) machinery within the cell

to bring about membrane scission. During the assembly and budding stages of HIV-1 virion assembly, ESCRT-I is recruited to the plasma membrane at the site of virion assembly. The P(S/T)AP motif in p6 Gag, required for this recruitment event, binds ESCRT-I via a direct interaction with host protein TSG101 (Martin-Serrano et al., 2001; Garrus et al., 2001). Recruitment of ESCRT-III follows the recruitment of ESCRT-I, although the intermediate proteins responsible have not been conclusively determined (as reviewed by Meng and Lever, 2013). The (L)YPXNL motif binds cellular ALIX (apoptosis-linked gene 2-interacting protein) to promote virion budding (Zhai et al., 2011). It has been shown that in a system where P(S/T)AP is missing and TSG101 mediated budding is blocked, over-expression of ALIX can rescue budding defects (Fisher et al., 2007).

1.2.8 Virion maturation

Once assembled, during or just after budding, the virion undergoes a maturation process which is required for the virion to become infectious, as shown in figure 1.4. This maturation event involves the cleavage of the Gag and Gag-Pol polyproteins into functional units by the viral protease enzyme, resulting in structural changes within the virion.

1.2.8.1 HIV protease enzyme

HIV-1 protease is a 99 amino acid aspartic protease that functions as a homodimer, as shown in figure 1.7 (Fisher et al., 2007). The protease active site is found along the dimer interface and each protease monomer contributes a single catalytic aspartic acid residue. The active site is formed by the triplet Asp-Thr-Gly amino acids from positions 25-27 of each protease monomer. Protease has two flexible flaps, which are dynamic in solution enabling entry of inhibitors and substrate into the binding cleft situated above the active site (Collins et al., 1995).

The mechanism of protease autocleavage from the Gag-Pol polyprotein is not fully understood, but it has been shown that dimerisation of Gag-Pol is required

for protease activation. Following dimerisation, cis cleavage of the protease from within the Gag-Pol polyprotein occurs before protease cleaves the remainder of the Gag-Pol polyprotein in trans at specific cleavage sites, recognised by their secondary structure and not their amino acid sequence (Davis et al., 2012; Prabu-Jeyabalan et al., 2004). It has been shown, *in vitro*, that protease autocleavage can be controlled by the oxidation of cystiene residues 67 and 95 (Daniels et al., 2010).

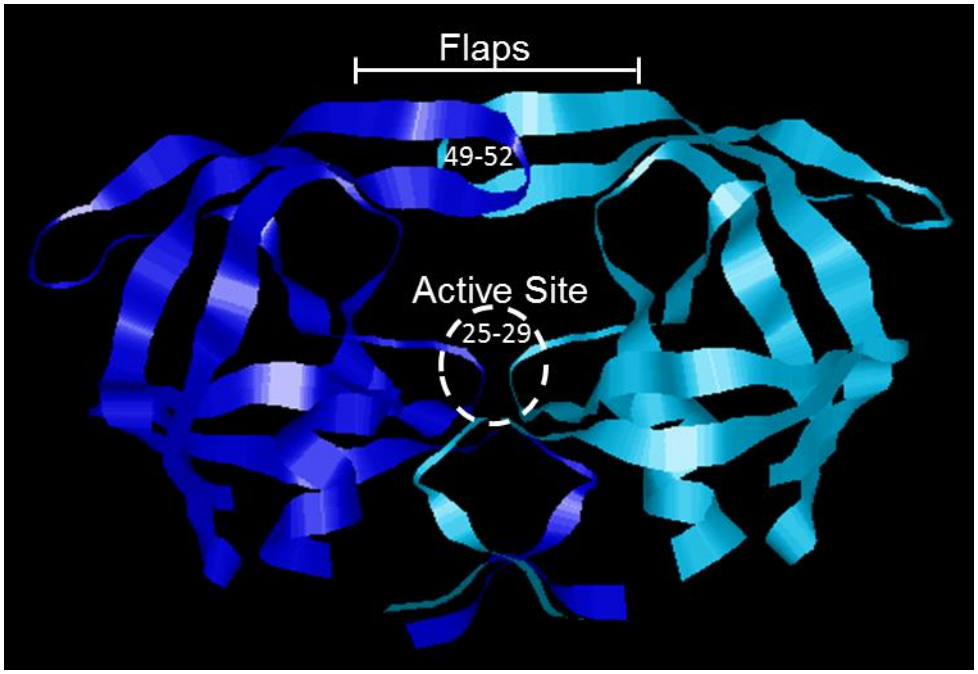


Figure 1.7. The HIV-1 protease enzyme. The structure of the HIV-1 protease homodimer is displayed, with one monomer shown in dark blue and the second in light blue. The active site of the enzyme contains three catalytic amino acids at positions 25-27, Asp-Thr-Gly. The dynamic flaps that enable the entry of substrate and inhibitors into the substrate binding cleft are labelled.

1.2.8.2 Protease cleavage of the Gag polyprotein

Protease cleavage of Gag, during or just after the release of the virion from the plasma membrane, results in the release of the six subunits: matrix (p17, MA), capsid (p24, CA), nucleocapsid (p7, NC), p6 and the spacer peptides p1 and p2. Cleavage of Gag-Pol additionally results in the release of the protease, reverse transcriptase and integrase enzymes. Whilst the exact timing of cleavage remains unknown, it occurs late enough into the budding process to ensure that all cleavage products remain in the budding virion. There are five cleavage sites within Gag that lead to the separation of the six subunits, as shown in figure 1.8. Protease recognition of the cleavage sites is determined by the secondary structure and not the amino acid sequence of each site, hence there is significant variation in amino acid sequences between cleavage sites (Prabu-Jeyabalan et al., 2002). The secondary structures are fairly well conserved between cleavage sites, however there are differences in the protruding side chains which are thought to contribute to the different rates of cleavage at each site.

Whilst amino acids are not conserved across cleavage sites, studies have demonstrated that the amino acids at the cleavage site in direct contact with protease regulate the rate of cleavage (four amino acids before the cleavage site and three after) (Pettit et al., 2002; Pettit et al., 2005). In addition, residues more distant from the cleavage site, in particular the fourth and fifth amino acids following the cleavage site p4' and p5', have been shown to affect rates of cleavage (Dam et al., 2009; Nijhuis et al., 2007b). The amino acid residues are mostly hydrophobic and it has been shown that a hydrophobic residue at the amino acid position before the cleavage site (p1) is required for cleavage to occur (Pettit et al., 1991).

The order of cleavage at the sites within Gag is tightly controlled and, as with the rate of cleavage, is a result of the variation in protruding side chains and amino acid sequences present. The order in which the Gag sites undergo

cleavage is shown in figure 1.8, along with the amino acid positions at each cleavage site (as reviewed by Fun et al., 2012). The first cleavage event occurs at the C terminal of p2, releasing two intermediates: MA-CA-p2 and NC-p1-p6 (Krausslich et al., 1989). This cleavage event permits the condensation of the genomic RNA to the centre of the virion. The second cleavage events occur between MA and CA, leading to release of MA from CA-p2, and between p1 and p6, releasing p6 from NC-p1. These cleavage steps occur at a ten-fold lower rate than the first cleavage step (Pettit et al., 1994). The separation of the spacer peptides (p1 and p2) from NC and CA, respectively, are the final cleavage events and are thought to be the rate limiting steps, occurring at a rate several hundred fold reduced in comparison to that of the first cleavage step (Nijhuis et al., 2007b). The separation of p2 from CA has been shown to be essential for the formation of the mature core surrounding the viral RNA (as discussed in Briggs and Krausslich, 2011). The structural differences between the mature and immature virions can be seen diagrammatically and in electron micrographs in figure 1.9.

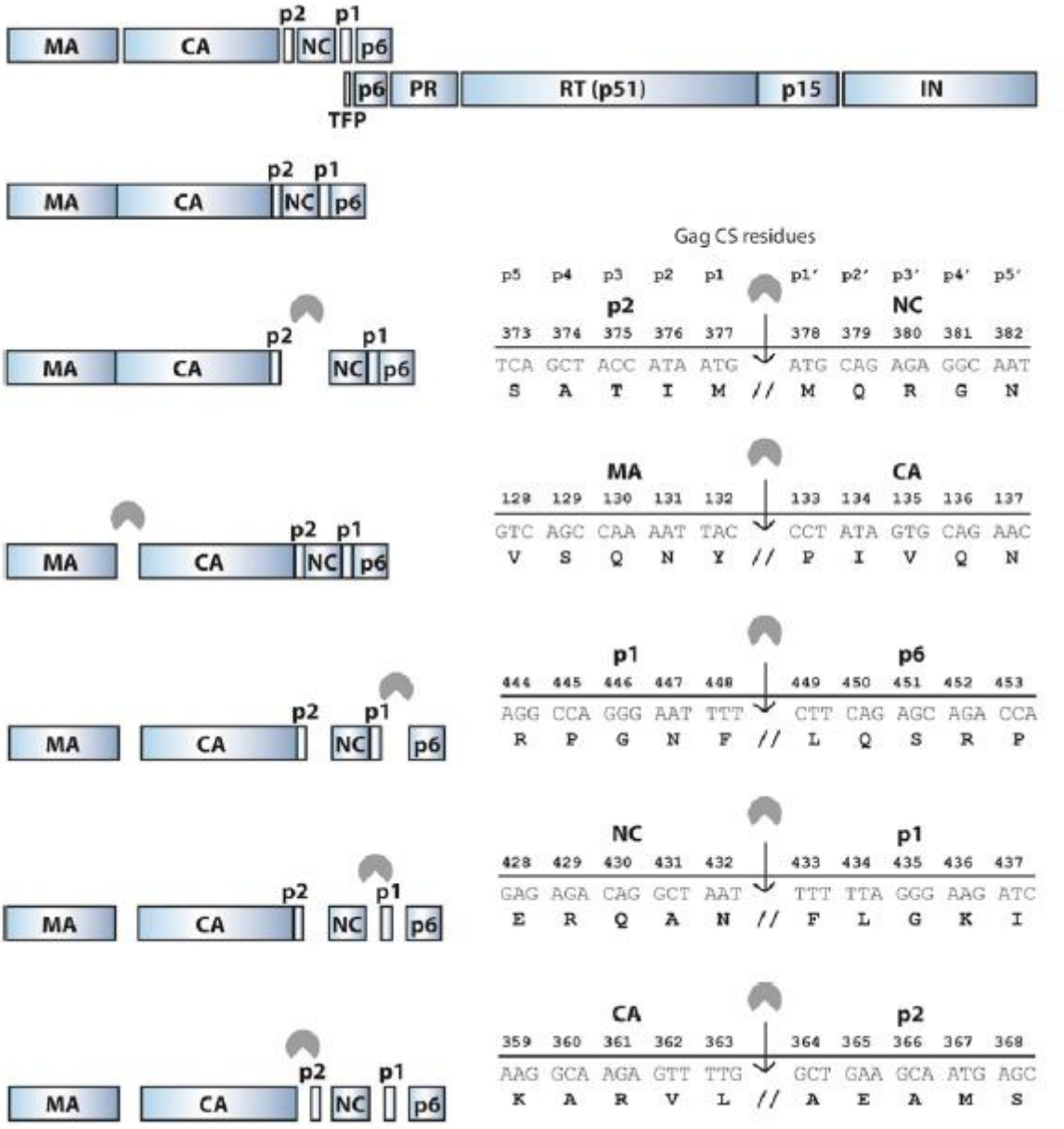


Figure 1.8. Ordered cleavage of the Gag polyprotein into its functional subunits by viral protease. The top panel of the figure shows the subunits of the Gag-Pol polyprotein and highlights the cleavage sites. The order of Gag cleavage and the consensus nucleotide and amino acid sequences at each Gag cleavage site are shown. The first cleavage step results in the release of p2 from NC. The second cleavage steps are the release of MA from CA, and of p6 from p1. The final, rate limiting, steps are the release of CA from p2 and NC from p1. Figure from Nijhuis et al., 2012.

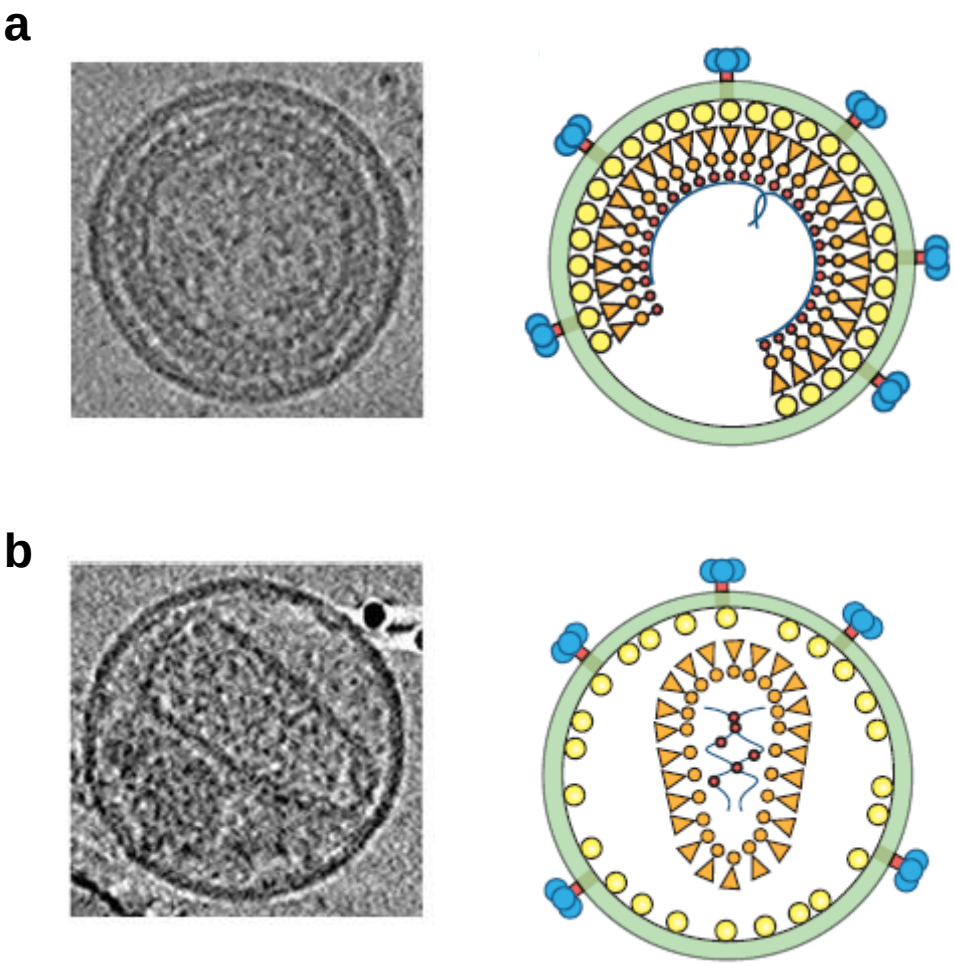


Figure 1.9. The structure of Gag in the virion. (a) An electron micrograph and schematic representation of the structure of immature HIV-1 virions. The Env protein spikes (blue), viral envelope (green), Gag matrix (yellow), capsid (orange) and nucleocapsid subunits (red) are shown. In the immature virion all Gag subunits are attached to the outer viral envelope. (b) An electron micrograph and schematic representation of the structure of mature HIV-1 virions. Gag cleavage releases the capsid and nucleocapsid subunits. Nucleocapsid complexes with RNA and capsid forms the conical capsid surrounding the RNA. Figure adapted from Sundquist and Krausslich, 2012.

1.2.8.3 Structure of the mature virion

The structure of the HIV-1 mature virion is shown in figures 1.9 and 1.10. The host cell membrane derived envelope surrounds the virion, studded with viral Env proteins. The MA protein is found attached to the virus envelope, although this lattice is not continuous containing a gap correlating with the area where scission from the host cell occurred. Viral protease is found between the envelope and the core as it is necessary for the maturation process. The CA protein forms the capsid which surrounds the genome-nucleocapsid complex. It comprises multiple CA hexamers which form a tube and a number of CA pentamers which result in the formation of a conical shape, as shown in figure 1.10 (Pornillos et al., 2011). The capsid contains two copies of the single stranded RNA genome complexed with NC and p6, protecting it from digestion by host nucleases. The viral enzymes reverse transcriptase and integrase are also packaged within the capsid alongside the viral genome.

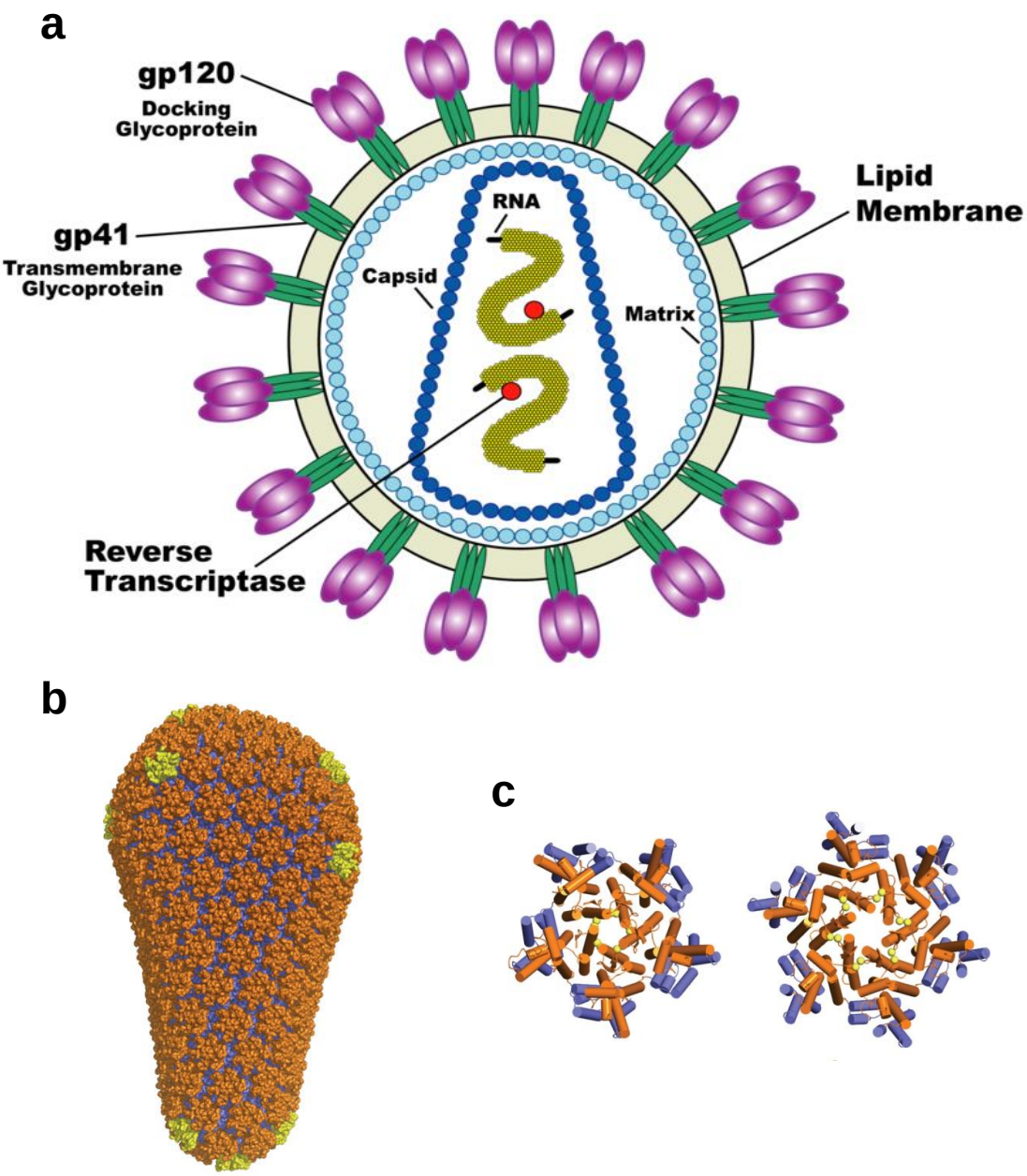


Figure 1.10. The structure of the mature virion. (a) A schematic showing the structure of the mature HIV-1 virion. Env spikes comprise trimeric gp41 and gp120 proteins. Gag MA subunit is stuck to the viral envelope, CA has formed the conical capsid and NC is in complex with the viral RNA. Reverse transcriptase is packaged along with the RNA. (b) The structure of the mature capsid surrounding the viral RNA is shown, comprised of CA hexamers and pentamers, shown in (c). Figure taken from NIH website (a) and Pornillos et al. 2011 (b,c).

1.2.9 Cell-to-cell spread

In addition to the above mechanism of infection involving release of the virion into the extracellular environment (cell-free spread), direct cell-to-cell transmission of HIV can occur in which the virus is never released into the extra-cellular environment. This occurs via the formation of the virological synapse between infected (donor) and uninfected cells (Chen et al., 2007). Polarisation of viral proteins and cellular organelles occurs at the site of contact between cells. Cell-to-cell spread has been shown to be mediated by the viral Env protein in the infected cell and the CD4 receptor of the uninfected cell (Jolly et al., 2004).

The importance of cell-to-cell spread *in vivo* is still under investigation, although it has been suggested to play a role in the spread of virus in areas densely populated by CD4+ cells such as lymph nodes. *In vitro*, cell-to-cell spread is more efficient than cell free spread with both reverse transcription and integration occurring more rapidly, possibly as rate limiting steps in cell-free infection such as virion attachment are not required. Different studies have reported varying significance of cell-to-cell spread *in vitro* with some studies reporting cell-to-cell spread as the main mechanism (Chen et al., 2007) and others reporting an equal contribution of cell-free and cell-to-cell spread (Komarova et al., 2013).

Cell-to-cell spread has been implicated in the persistence of low levels of viral replication in the presence of ART and as such there are questions over the effectiveness of ART in preventing this mode of transmission (Sigal et al., 2011). Studies have shown that entry inhibitors and protease inhibitors are effective in preventing cell-to-cell spread (Martin et al., 2010; Sigal et al., 2011). However, a recent study has shown that whilst protease inhibitors are equally effective at blocking cell-to-cell spread, certain reverse transcriptase inhibitors are between 4 and 20 fold less effective (Titanji et al. 2013). In addition, cell-to-

cell spread is thought to enable immune evasion from both antibodies and the complement system.

1.2.10 Host factors involved in the HIV-1 life cycle

A number of host factors interact with HIV-1 during the course of its life cycle to affect its infectivity, as shown in figure 1.4. HIV uses cellular factors and pathways for replication and conversely cells also produce proteins that suppress replication (restriction factors). During the early stages of the life cycle, cyclophilin A (CypA) is recruited into virions as they assemble following interaction with Gag (Franke et al., 1994). It binds to a loop on the CA subunit and this binding can be blocked by mutations at positions 89 and 90 of CA (as reviewed by Hilditch and Towers. 2014). It has been hypothesised that in binding CA, CypA inhibits the CA-CA interactions necessary for the stability of the viral core leading to progressive uncoating of the core as increasing amounts of CypA bind (Gamble et al., 1996). Recent data has shown that blocking the binding of CypA and CA influences the pathway of nuclear entry and alters integration site targeting (Schaller et al., 2011)

Trim5α is a component of the innate immune system and its expression is induced by interferon (Asaoka et al., 2005). Trim5α contains a tripartite motif (an RBCC) domain which comprises a RING domain, a B Box 2 domain and a coiled coil. The exact mechanism by which Trim5α inhibits viral infection has yet to be determined. However, it has been suggested that Trim5α may interfere with uncoating of the viral core or the trafficking of the viral core to the nucleus (as reviewed by Towers, 2007). Trim5α binds HIV-1 CA directly capsids and is heavily ubiquitinylated in cells (Diaz-Griffero et al., 2006). It is rapidly degraded by the proteosome in a RING domain dependent manner, therefore it has been suggested that once bound to virions it targets them for proteosomal degradation, leading to a block in early infection. However, when the proteosome is inhibited the virus remains un-infectious and proteosome

independent mechanisms of degradation have also been described (Chatterji et al., 2006; as reviewed by Malim and Bieniasz, 2012). Whilst Trim5α from other animals successfully restricts HIV-1 infection, HIV-1 is resistant to restriction by human Trim5α.

Another important restriction factor that acts in the early stage of the HIV-1 life cycle is APOBEC3G, one of a family of cytidine deaminase enzymes of the APOBEC3 family. APOBEC3G causes steric inhibition of reverse transcription by causing the conversion of deoxycytidine to deoxyuridine in the negative DNA strand, resulting in G to A hypermutations and eventually leading to the production of uninfectious virions (Bishop et al., 2008). The viral protein Vif prevents APOBEC3G action in human cells by targeting it for proteosomal degradation (Sheehy et al., 2003). In the absence of Vif, APOBEC3G is packaged into the assembling virion through interactions with RNA and Gag NC (as reviewed by Malim and Bieniasz, 2012).

SAMHD1 is a more recently described restriction factor shown to block the early stage of viral replication in dendritic and resting CD4+ T cells (Baldauf et al., 2012). SAMHD1 functions by hydrolysing cellular dNTPs which inhibits reverse transcription and prevents cDNA synthesis (Goldstone et al., 2011; Lahouassa et al., 2012). SAMHD1 restriction can be overcome by the Vpx accessory protein, which is thought to sequester SAMHD1 and target it for proteaosomal degradation (Sharova et al., 2008). However, Vpx is not present in HIV-1 although it is present in HIV-2 and SIV viruses.

Tetherin restricts infectivity of HIV-1 at the budding stage of the virus life cycle. In the absence of the antagonist Vpu the virus assemble normally, but tetherin anchors the virions to the cell membrane and virions are internalised, leading to their accumulation in endosomes (Neil et al., 2008). Tetherin contains a coiled coil domain that enables dimerization, thought to result in a protein tether which links virions to the cell (as reviewed by Neil and Bieniasz, 2009). The HIV-1

accessory protein Vpu is the tetherin antagonists, whereas SIVs use Nef.or Env (Gupta et al., 2009a; Sauter et al., 2009). The viral proteins Vpu, Env and Nef reduce the surface expression of tetherin by a number of mechanisms including interference with its transport to the membrane, internalisation and degradation (as reviewed by Sauter, 2014).

1.3 Treatment of HIV-1 infection

Currently there is no vaccine or cure for HIV, the only treatment available is antiretroviral therapy (ART) which halts the progression of infection to AIDS. There are a variety of classes of inhibitors that act at different stages of the HIV life cycle to prevent virus replication: cellular entry, reverse transcription, integration, protease cleavage and virion maturation (as reviewed by Arts and Hazuda, 2012). Current guidelines recommend the use of multiple ART drugs, known as highly active antiretroviral therapy (HAART). HAART normally comprises three drugs; either two nucleoside reverse transcription inhibitors (NRTIs) and one protease inhibitor (PI) or two NRTIs and one non-nucleoside reverse transcriptase inhibitor (NNRTI) (World Health Organisation, 2013). Currently, first line therapy of two NRTIs and one NNRTI, PI or INI is recommended as the starting point of therapy (Williams et al., 2012). If treatment failure occurs, other drugs in the same classes of ART can be considered.

HAART is used with the aim of suppressing viral replication to levels below the limit of detection in commercial viral load assays, approximately 50 RNA copies/ml. When viral suppression is achieved, the patients CD4+ T cell count rises again towards that of normal levels, resulting in improved immune function (as reviewed by Arts and Hazuda, 2012). As well as improving the prognosis of the individual patient, HAART reduces transmission rates between sero-discordant couples and mothers and their babies (Cohen et al., 2011). Until recently WHO guidelines recommended commencement of HAART in patients with a CD4+ T count below 350 cells/ml but in the summer of 2013 this was

raised to 500 cells/ml given the benefits of earlier treatment (Kitahata et al., 2009; World Health Organisation, 2013).

1.3.1 Entry and fusion inhibitors

Entry inhibitors are a relatively new class of drug which target the cell entry steps of the HIV life cycle. There are two classes of entry inhibitors that have compounds that are FDA approved: fusion inhibitors and CCR5 co-receptor antagonists.

Fusion inhibitors act to prevent the fusion of the virion envelope with the host cell membrane. Since they act outside the host cell, they have minimal side effects in comparison to other ARTs. Enfuviritide (T20) is a fusion inhibitor which binds to viral gp41 and disrupts the interaction between the N and C terminal heptad repeats (NHR and CHR, respectively) of the gp41 ectodomain by mimicking the CHR. Hence it prevents the subsequent interaction between gp41 and cellular receptors required for virus entry. It has been shown to be effective again viruses of all tropisms, R5, X4 and R5X4 (Fletcher, 2003). It was approved by the FDA in 2003, however its short half-life of less than four hours has limited its use, as it must be administered twice daily at fairly high doses. Additionally, it is administered as a subcutaneous injection which is expensive and can result in side effects at the injection site. A second fusion inhibitor, sifuviritide, has been developed and is currently in Phase II clinical trials. It mimics the gp41 CHR more closely than T20 resulting in a higher affinity for the NHR peptide of gp41. It has improved potency (20-fold higher) and half-life in comparison to T20 and a different mechanism of action means cross resistance between the two fusion inhibitors does not appear to occur (He et al., 2008).

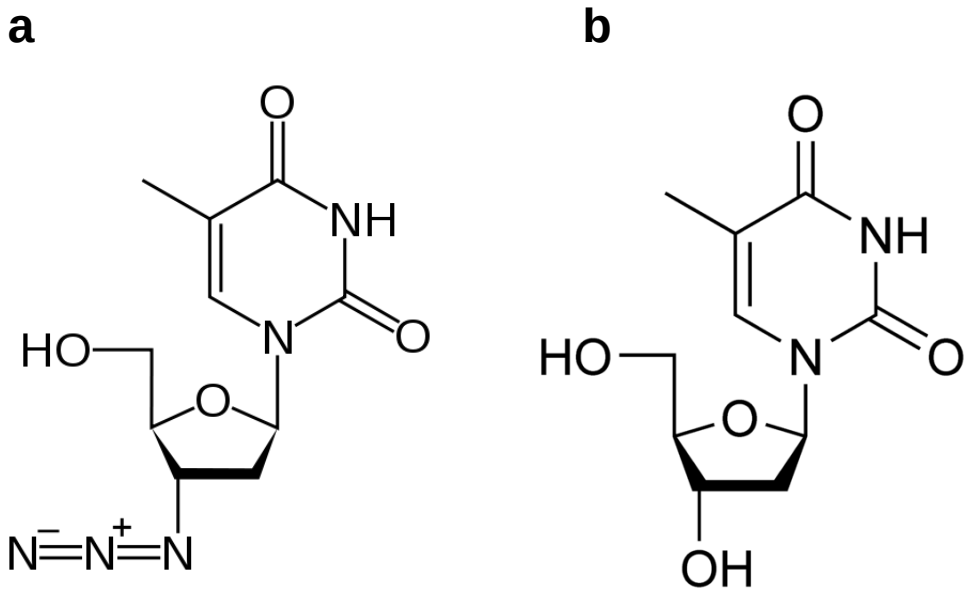
Co-receptor antagonists have been developed, particularly to CCR5 as this is the co-receptor regularly involved in early infection. Maraviroc is the only FDA approved CCR5 receptor inhibitor (in 2007), and the only FDA approved drug that functions by targeting a host protein (CCR5) and not the virus (Latinovic et

al., 2009). Co-receptor antagonists are allosteric inhibitors of CCR5, binding hydrophobic pockets within the transmembrane helices of CCR5. This causes a conformational change in the second extracellular loop of CCR5 which the V3 stem loop of gp120 HIV-1 cannot recognise (Dragic et al., 2000). In the clinical setting it has limited use as it only works against R5 viruses, not X4 or dual-tropic, so each patient must be tested to ensure they are only infected with R5 virus before maraviroc can be introduced to their treatment regimen. A second CCR5 co-receptor antagonist, vicriviroc, has been developed however its development was abandoned in Phase III clinical trials.

1.3.2 Reverse transcriptase inhibitors

1.3.2.1 Nucleoside/Nucleotide reverse transcriptase inhibitors (NRTIs)

Nucleoside/nucleotide RT Inhibitors (NRTIs) were the first class of drug to be approved by the FDA, with Zidovudine (AZT) receiving approval in 1986. NRTIs are alternative substrate inhibitors of RT that compete with the host dNTPs for incorporation into cDNA. NRTIs lack the 3' OH group on the ribose ring, preventing phosphodiester bond formation to the next nucleotide. This leads to a block in DNA elongation and results in chain termination (as reviewed by Tsibris and Hirsch, 2010). A total of eight NRTIs have now been approved by the FDA for treatment of HIV: zidovudine (AZT/ZDV), didanosine (ddI, second to be approved), zalcitabine (ddC, now discontinued), stavudine (d4T), lamivudine (3TC, approved for both HIV-1 and Hepatitis B treatment), abacavir (ABC), emtricitabine (FTC, approved for HIV and in trials for Hepatitis B treatment) and tenofovir (TDF) (as reviewed by Arts and Hazuda, 2012). Each of these competes with one of the nucleotides for incorporation into the elongating DNA strand. Figure 1.11 demonstrates the structural similarity of the NRTI AZT and the thymidine nucleotide.



Zidovudine (drug) Deoxy-Thymidine (nucleotide)

Figure 1.11. The structure of RT inhibitors; NRTIs. The structure of the NRTI Zidovudine (AZT) (a) and dNPT deoxy-thymidine (b) are shown, demonstrating the structural similarity of the NRTI to the dNTPs they competitively inhibit. Figure adapted from Arts and Hazuda, 2012.

Nucleoside RTIs are administered in pro-drug form and require phosphorylation by intracellular kinases within the body before they function and can be incorporated into the elongating DNA strand. Conversely, nucleotide RTIs do not require phosphorylation before they become biologically active (as reviewed by Cihlar and Ray, 2010). NRTIs form the backbone of most HAART regimens, partly because the long half-life of the metabolised form enables once daily dosing. In addition, clinical trials are ongoing to test the effectiveness of TDF based microbicide gels in the prevention of HIV infection. The CAPRISA 004 trial reported up to 54% reduction in the rate of new infections in individuals with good adherence, and the VOICE MTN 003 and MDP 302 trials with the same gel and different dosing are ongoing (Abdool et al., 2010). NRTIs given orally as part of HAART have side effects that are caused by the inhibition of cellular DNA polymerase activity, including mitochondrial DNA polymerase which is particularly susceptible.

1.3.2.2 Non-nucleoside reverse transcriptase inhibitors

Non-nucleoside RT Inhibitors (NNRTIs) are allosteric, non-competitive inhibitors of RT. NNRTIs bind to a hydrophobic pocket approximately 10Å from the active site of RT, inducing a conformational change that reduces the rate of nucleotide incorporation into DNA (Bacheler et al., 2001; Kohlstaedt et al., 1992). The affinity of RT to the dNTPs is not affected by the binding of NNRTIs. NNRTIs differ from NRTIs in that they do not require intracellular processing for drug activity.

The first NNRTI, nevirapine (NVP), received FDA approval in 1996 followed by delavirdine (DLV) in 1997 and efavirenz (EFV) in 1998. NVP and EFV are regularly used as the third agent in HAART regimens, with EFV the most frequently used NNRTI in the developed world. However, the low genetic barrier to resistance means there is a requirement for two fully-active additional non-NNRTIs (as reviewed by Arts and Hazuda, 2012). In addition NVP has been shown to be safe in pregnant women and superior to AZT in the

prevention of mother-to-child transmission as a single dose in the developing world, although resistance can occur in the mother (as reviewed by de Bethune, 2010). The fourth drug in the class, and the first of the second generation drugs, etravirine (ETV) was the first NNRTI shown to be efficacious in patients harbouring resistance to other first generation NNRTIs in the DUET trials (Katlama et al., 2010a). Rilpivirine (RPV) was approved in 2011 for use in treatment-naive patients, following evidence of efficacy and safety from the ECHO and THRIVE trials. ETV and RPV have molecular flexibility which enables binding to RT containing certain resistance mutations (as reviewed by Jayaweera and Dilanchian, 2012).

1.3.3 Integrase inhibitors

IN inhibitors are the most recent class of antiretroviral drugs to be developed that target viral enzymes. The first inhibitor, raltegravir (RAL), was approved by the FDA in 2007 and other integrase inhibitors are currently in late clinical trials including elvitegravir (EVG) and dolutegravir (DTG), which is approved for use in the US (Arts and Hazuda, 2012; Sato et al., 2006). However, studies have shown extensive cross-resistance between RAL and EVG (as reviewed by Shimura and Kodama, 2009). Second generation DTG has shown good efficacy in extensively PI-experienced patients including those failing on RAL regimens in the VIKING trial (as reviewed by Katlama et al., 2010b).

IN inhibitors competitively inhibit the DNA strand transfer reaction of DNA by binding in the IN active site (Espeseth et al., 2000). This DNA strand transfer stage is targeted as IN inhibitors only bind to the complex formed between viral DNA and the integrase enzyme. Binding of two metal ion co-factors in the integrase active site is required for enzyme function *in vitro*. Hence, IN inhibitors are designed with two functional components – a metal binding core which enables the sequestering of magnesium ions from the active site and a hydrophobic group which interacts with both the viral DNA and the IN enzyme

(Grobler et al., 2002). Structural methods have shown that integase is a tetramer and that the binding of IN inhibitors displaces the reactive viral DNA from the active site, preventing integration (Hare et al., 2010).

1.3.4 Protease inhibitors

Protease inhibitors (PIs) are a potent class of inhibitors which are also used against other viruses including Hepatitis C. Nine PIs in total have received FDA approval, of which 6 are currently in regular clinical use, as detailed in Table 1.1.

Table 1.1. Protease inhibitors approved for clinical use.

Protease Inhibitor	Abbreviation
(fos) Amprenavir	APV
Atazanavir	ATV
Darunavir	DRV
Indinavir	IDV
Lopinavir	LPV
Nelfinavir	NFV
Ritonavir	RTV
Saquinavir	SQV
Tipranavir	TPV

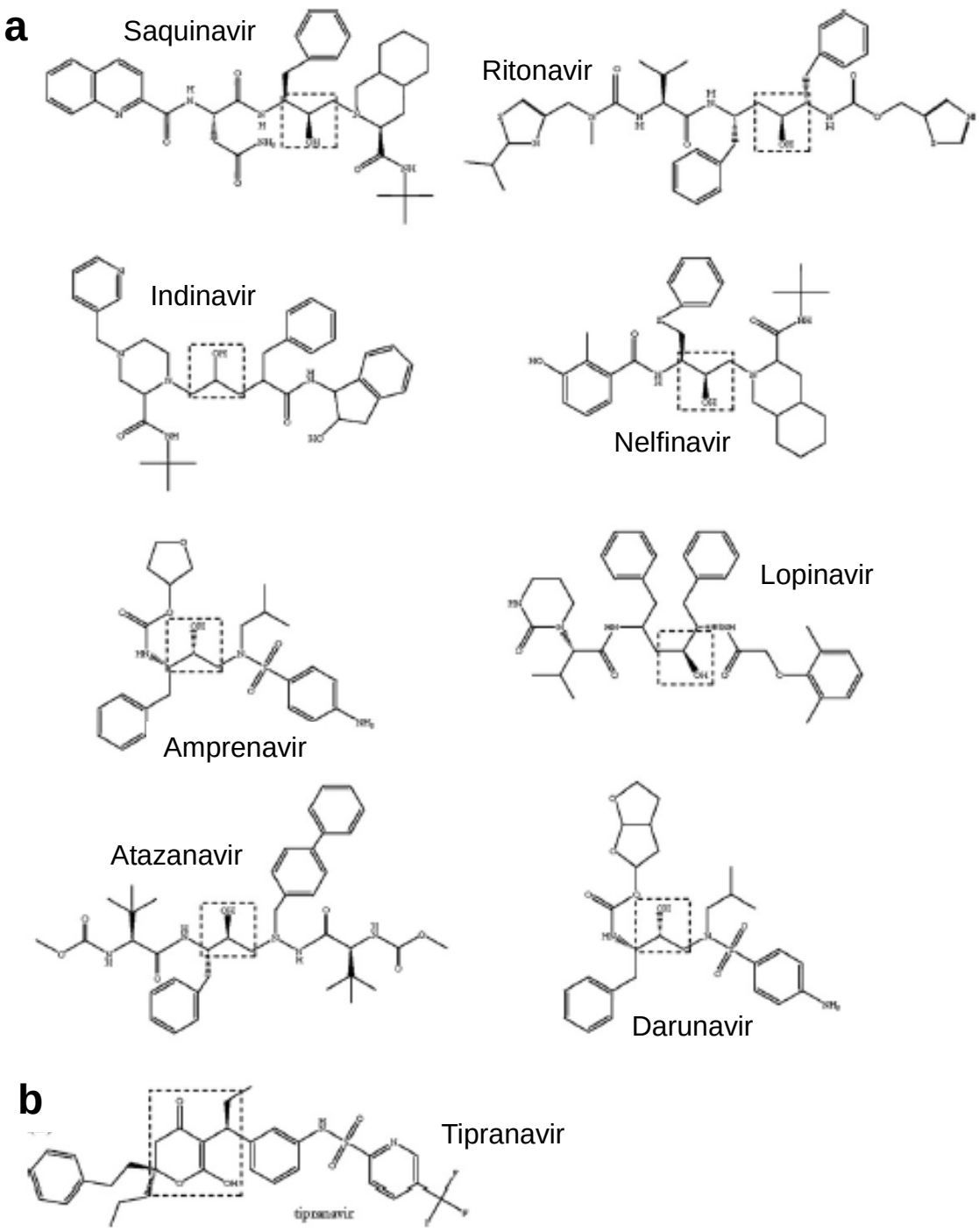
PIs are competitive inhibitors of the protease enzyme, which enter and block the active site of protease, preventing cleavage of the natural substrates the Gag and Gag-Pol polyproteins (Wensing et al., 2010). Without this cleavage step, immature virions are still produced but maturation does not occur, so the virions remain immature and thus un-infectious. A recent study identified that using a PI as the third drug in HAART led to significantly fewer patients experiencing virological failure than HAART using NNRTI as the third agent (Gupta et al., 2008).

PIs were rationally designed using the known crystal structure of HIV protease resulting in competitive peptidomimetic inhibitors which mimic Gag, the protease natural substrate. All PIs, except Tipranavir (TPV), contain a hydroxyethylene core which cannot be cleaved by the viral protease, as shown in figure 1.12 (Craig et al., 1991; Kempf et al., 1995). In its place, TPV contains a dihydropyrone ring as a central scaffold (Turner et al., 1998). Saquinavir (SQV) was the first PI to receive FDA approval in 1995. When used as monotherapy it proved relatively unsuccessful with 50% of patients harbouring viruses with reduced susceptibility after one year, but it was efficacious when used as part of combination therapy once bioavailability issues had been addressed (Jacobsen et al., 1996; Schapiro et al., 1996). The next two PIs were approved in 1996 – ritonavir (RTV) and indinavir (IDV). RTV was shown to be efficacious, but poorly tolerated with side effects including severe gastrointestinal symptoms and as such it is no longer used as a protease inhibitor on its own (Notermans et al., 1998). IDV, as with SQV, proved to have limited efficacy when given as monotherapy but was efficacious when given as part of triple therapy (Gulick et al., 1997; Stein et al., 1996). Nelfinavir (NFV) was FDA approved in 1997 and showed good efficacy as part of triple therapy in clinical trials (Saag et al., 2001).

Although SQV, RTV, IDV and NFV proved efficacious, poor bioavailability and short half-lives meant frequent dosing was required. 'Boosting' of PIs by administering with low doses of RTV was introduced to increase the therapeutic levels without having to increase the dose – the second generation of PI therapy (van Heeswijk et al., 2001). RTV is an inhibitor of CYP3A4 – part of the cytochrome p450 pathway which degrades PIs in the liver. Boosting with ritonavir reduces the rate that the body metabolises PIs thereby raising the therapeutic levels without having to increase the dose and toxicity (as reviewed by Wensing et al., 2010). Boosting strategies cannot be used with IDV, as it increases nephrotoxicity, or NFV, as it did not increase its bioavailability (Boyd

et al., 2005). As a result, the use of IDV and NFV is now limited, with NFV being used only during pregnancy as it has been shown to be relatively safe.

Amprenavir (APV) received FDA approval in 1999, but suboptimal concentrations in the blood lead to weak activity and frequent development of resistance (Sadler et al., 2001). Introduction of a pro-drug fosamprenavir and RTV boosting have resulted in increased efficacy when given as part of a HAART regimen (Rodriguez-French et al., 2004). PIs lopinavir (LPV), atazanavir (ATV), TPV and darunavir (DRV) were FDA approved in 2000, 2003, 2005 and 2006, respectively (Arts and Hazuda, 2012). All are co-administered with RTV and are available in formulations that can be taken once daily, reducing pill burden for patients and hence increasing drug adherence. LPV is available co-formulated with RTV (LPV/r) and marketed under the trade name Kaletra. LPV was shown to be efficacious with 75% of patients achieving viral suppression after 48 weeks on LPV based HAART (Walmsley et al., 2002). LPV/r was the first PI shown to be efficacious in treatment-naïve patients, thus demonstrating efficacy of PIs as first line therapy in HAART (Riddler et al., 2008). Both LPV and ATV have been shown to be efficacious in PI-experienced individuals, with viral suppression achieved in 50% of individuals in the CASTLE study (Molina et al., 2008). TPV was approved for use in patients with resistance to multiple PIs (Hicks et al., 2006). Although shown to be superior in salvage therapy in the RESIST trial, TPV is not widely used as it does not demonstrate superiority over other PIs in PI-naïve patients and has significant side effects. The most recently developed PI, DRV, was specifically designed to inhibit viruses that are resistant to other PIs for use in treatment experienced individuals. The POWER studies demonstrated superiority of DRV in comparison to other PIs in treatment-experienced patients (Katlama et al., 2007). In addition, the ARTEMIS trial showed the superiority of DRV in comparison to LPV as part of a HAART regimen in treatment-naïve patients (Mills et al., 2009).



1.12. The structures of protease inhibitors (PIs). The chemical structures of all FDA approved PIs are shown. (a) PIs with a hydroxyethylene core and with (b) a dihydropyrone ring are shown. The scaffolds are highlighted with a dashed box. Figure taken from Wensing at al. 2010

1.3.5 Maturation inhibitors

Maturation inhibitors are the only class of drug, along with entry inhibitors, that do not target a viral enzyme, but instead they target the maturation of new virions during or after budding from the cell (as reviewed by Arts and Hazuda, 2012; as reviewed by Tilton and Doms, 2010). There are only two drugs in this class that has been evaluated in clinical trials. Bevirimat inhibits the late step of Gag processing, preventing the conversion of CA-SP1 to mature CA protein by protease cleavage (Ghosn et al., 2011). It has been demonstrated that this block in processing occurs because bevirimat stabilises the immature Gag lattice, thus preventing the subsequent maturation steps (Keller et al., 2011). This results in the production of immature and aberrant virions.

The development of bevirimat is currently on hold because of its performance in trials to date. A significant number of patients do not respond, which is thought to be caused by polymorphisms in SP1 residues 6 to 8, which have been found to prevent its function (Wainberg and Albert, 2010). It has been shown that the sequence of the SP1 can be used to predict the function of bevirimat *in vivo* (Heider et al., 2010). The development of a second maturation inhibitor, vivecon MP-9055, has also been discontinued after it reached phase II clinical trials.

1.4 Resistance to antiretroviral therapy

Even when HAART is given, HIV-1 is still able to develop resistance. A recent study estimated that after 10 years of HAART treatment, 9.2% of patients will have developed resistance to all three classes of drug included in the treatment regimen, and that the risk of death within 5 years of the development of extensive triple-class resistance is 10% (Phillips et al., 2007). As well as the risk to the individual patient, drug resistant viruses can be transmitted and can persist in the absence of the selective pressure of drugs complicating the treatment of patients with newly acquired drug resistant infections in the future

(Brenner et al., 2002; Pao et al., 2004). Of particular concern are viruses that develop resistance to multiple drug classes or cross-resistance to multiple drugs within one class (Condra et al., 1995). As a result, it is recommended that all newly diagnosed individuals in the UK are subjected to resistance testing before starting therapy. Several factors enable the development of resistance to ART including the nature of RT (its high error rate and low RNA affinity, as discussed in 1.2.3.2), the high level of replication and sub-optimal drug levels in patients, caused by non-compliance or by sub-optimal treatment regimens.

1.4.1 Resistance to entry inhibitors

Resistance to fusion inhibitor, T20, is conferred via mutations at the T20 binding site within gp41 including G36D, I37T, V38A/M, N42T/D and N43K (Wei et al., 2002). These mutations reduce the affinity of T20 to gp41, however they also confer reduced fitness as a result of decreased fusion efficiency. Compensatory mutations in the CHR domain of gp41occur that can restore the viral fusion kinetics and thus viral fitness (Ray et al., 2009).

The mechanisms of resistance to the CCR5 co-receptor antagonist maraviroc are not yet fully understood. Lack of response can be conferred by a switch in viral tropism to the use of the CXCR4 co-receptor. This has been observed in patients, although some studies indicate that this mechanism of resistance only occurs in patients with pre-existing CXCR4 tropic viruses (Westby et al., 2006). Secondly, the virus can evolve to use the maraviroc bound CCR5 receptor for cellular entry, noncompetitive resistance, which has been observed in patients failing therapy in the absence of a switch in co-receptor usage (Westby et al., 2007). Resistance has been shown to map to multiple regions of the gp120 and gp41 proteins (Westby et al., 2007). There are two additional theoretical mechanisms of resistance although their role *in vivo* is not yet known: an increase in affinity of the virus for the CCR5 receptor and an increase in efficiency of the viral entry process (as reviewed by Arts and Hazuda, 2012).

1.4.2 Resistance to reverse transcriptase inhibitors

Resistance to NRTIs is conferred by two main mechanisms: first, ATP dependent pyrophosphorolysis resulting in the removal of the 3' end of the elongating DNA chain and reversal of chain termination and second, increased discrimination between native dNTPs and the NRTI by RT. Resistance via ATP dependent pyrophosphorolysis is conferred by the development of thymidine analogue mutations (TAMs) in RT in response to AZT and d4T exposure. There are two distinct TAM mutation pathways that confer resistance: TAM1 (M41L, L210W, T215Y and occasionally D67N) and TAM2 (D67N, K70R, T215F and 219E/Q). The second mechanism of resistance, the prevention of incorporation of NRTIs into the elongating DNA chain, is conferred by mutations at two amino acid positions M184V/I and K65R (as reviewed by Arts and Hazuda, 2012). K65R develops on exposure to TDF, ddC, ddI, d4T and ABC (Wainberg et al., 1999). Conversely, M184V/I emerges following 3TC or FTC therapy (Schinazi et al., 1993). Each of these resistance mutations confer a reduction in viral fitness, and secondary mutations accumulate in RT that partially compensate for this reduced fitness.

Resistance to NNRTIs is conferred by mutations in the NNRTI binding pocket of RT including those at amino acid positions L100, K101, K103, E138, V179, Y181, Y188 and G190A (Tantillo et al., 1994). Resistance mutations K103N and Y181C are most common, and notably single NNRTI resistance mutations do not convey significant reductions in fitness, as observed for NRTI resistance mutations (Dykes et al., 2001). However, ETV is active against HIV viruses harbouring the K103N mutation and the DUET trials identified amino acid substitutions Y181I/V as conveying the highest levels of resistance to ETV. Following RPV exposure the E138K mutation is most frequently observed and it often appears alongside M184I (as reviewed by Jayaweera and Dilanchian, 2012).

1.4.3 Resistance to integrase inhibitors

Resistance to the IN inhibitor RAL is conferred by one of three independent pathways of mutations at positions Y143, N155 and Q148 (Fransen et al., 2009). These mutations are thought to be mutually exclusive in that they don't occur on the same viral genome, although viruses with resistance via different pathways can be harboured within the quasispecies of a single patient. These mutations are located in the catalytic core domain of IN which is located near the IN inhibitor binding site. The N155H pathway is associated with secondary resistance mutations V151I, T97A, G163R and L74M. Primary resistance mutations Q148 K/R/H are associated with secondary mutations G140S/A and E138K (as reviewed by Arts and Hazuda, 2012). The mechanisms of resistance to DTG are yet to be fully elucidated. However, studies suggest that mutations at position Q148 confer resistance to DTG, but changes at positions Y143 and N155 do not (as reviewed by Katlama et al., 2010b).

1.4.4 Resistance to protease inhibitors

Resistance to PIs develops via the accumulation of resistance mutations in protease and Gag which fall into two categories: major and minor resistance mutations (Molla et al., 1996). Major mutations confer resistance to PIs by reducing the affinity of the protease to the PI, preventing the PI from blocking the protease active site and leading to protease activity in the presence of PI (Condra et al., 1995). Major mutations also reduce the affinity of the protease to the natural substrate, Gag, thus reducing the fitness of the resistant virus (Zennou et al., 1998). Minor mutations may partially compensate for the reduced fitness of major mutations in resistant protease (Croteau et al., 1997). These are found in protease and within Gag, both at cleavage sites and other distant sites, as discussed in 1.4.4.4 (Clavel and Mammano, 2010; Gatanaga et al., 2002).

1.4.4.1 Resistance mutations in protease

To date, the Stanford Resistance Database describes PI resistance mutations at 33 of the 99 amino acids positions in protease (Liu and Shafer, 2006). These are collated every year by the International Aids Society, as shown in figure 1.13. Mutations at twelve of these positions are described as major mutations, conferring high levels of phenotypic resistance or evidence of a reduced response to treatment *in vivo* for at least one PI. These 12 positions are: 30, 32, 46, 47, 48, 50, 54, 76, 82, 84, 88 and 90 (Johnson et al., 2011). Of these 30, 32, 48, 50, 82 and 84 are located in the substrate binding cleft of protease and 46, 47 and 54 near the cleft (Kagan et al., 2005; Wu et al., 2003). These mutations confer resistance by enlarging the binding cleft, thus reducing the affinity of the PI (as reviewed by Nijhuis et al., 2007a). Mutations at positions 76, 88 and 90 affect PI susceptibility indirectly and are not located at or near the substrate binding cleft (Louis et al., 2011). The position of each of the PI resistance mutations within protease is shown in figure 1.14.

In addition, there are also a number of minor mutations found in protease which affect PI susceptibility primarily through compensatory mechanisms which are further discussed in 1.4.4.4. These cause reduced PI susceptibility when found in combination with other PI resistance mutations. Non-polymorphic mutations are those that are not found in treatment-naïve patients and only observed as a result of PI exposure. Conversely, polymorphic mutations are found in a significant proportion of treatment-naïve viruses, especially in non-B subtypes. Of interest, figure 1.13 demonstrates that the PI resistance mutations that occur are different for each PI, indicating different pathways to the development of resistance for each drug.

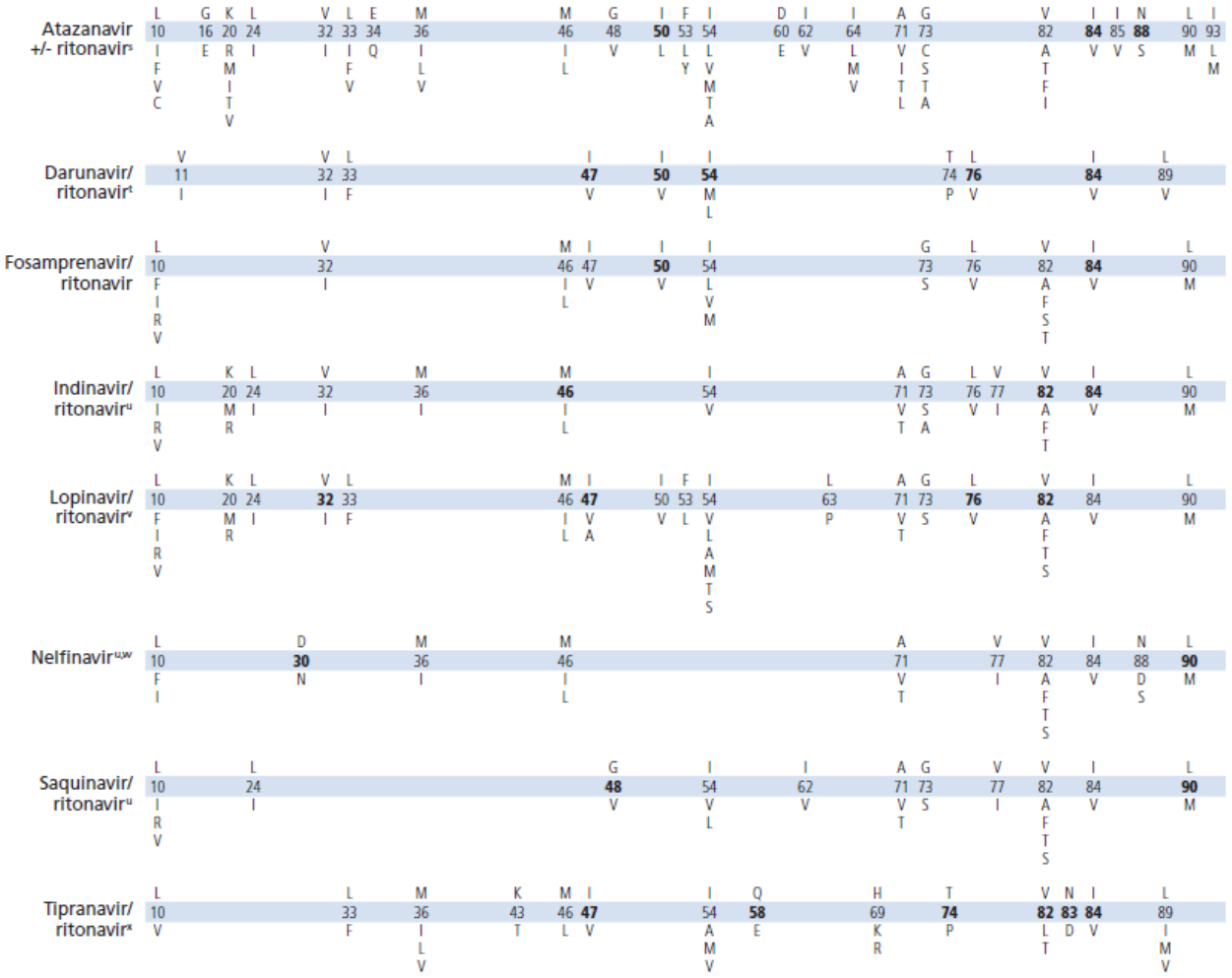


Figure 1.13. The known PI resistance mutations within protease. Diagram showing the PI resistance mutations within protease that have been documented to date. The mutations for each of the nine FDA approved PIs are shown, with major mutations shown in **bold**. Figure produced by the International AIDS Society (IAS), taken from Johnson et al., 2013.

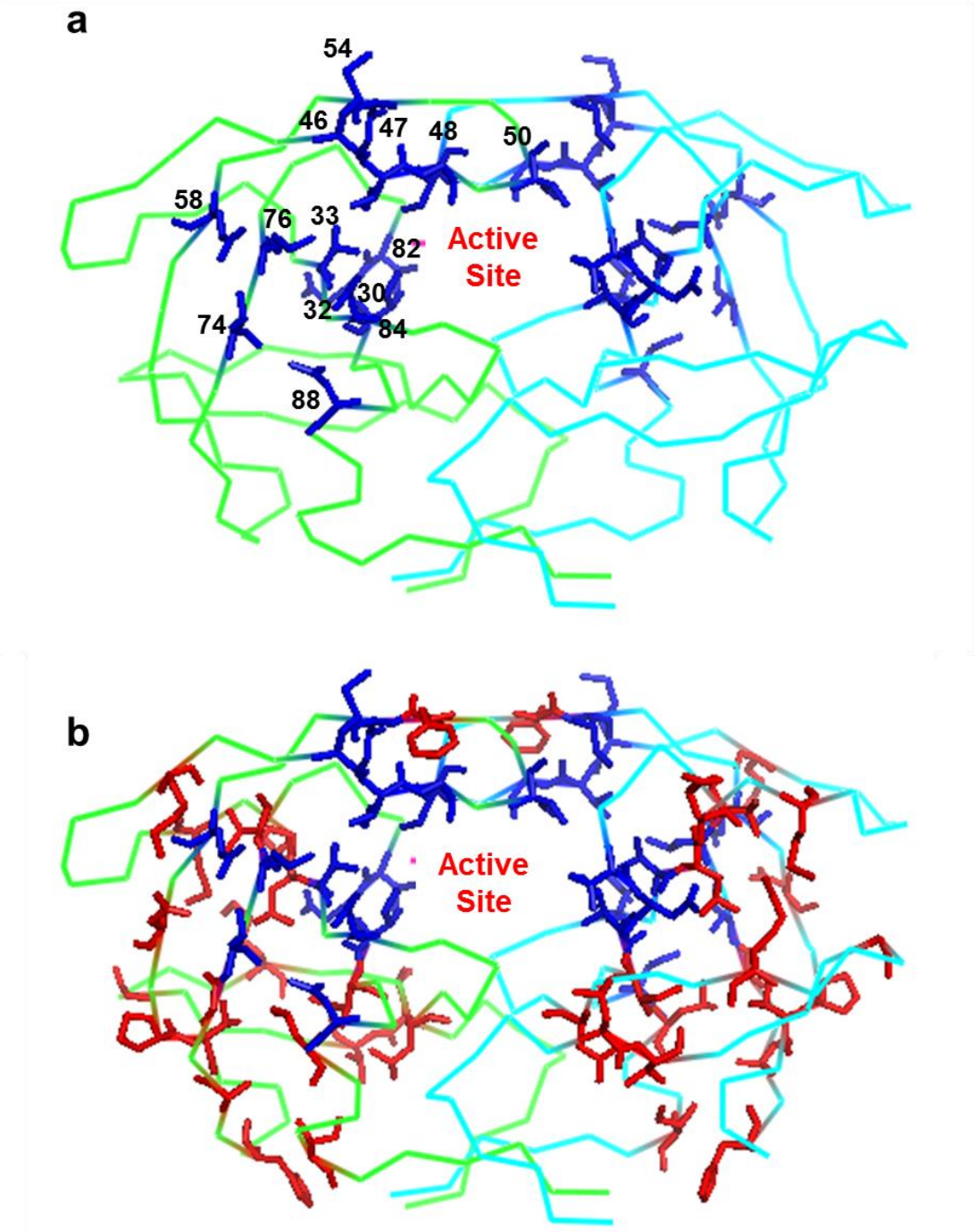


Figure 1.14. The position of known PI resistance mutations within the HIV protease protein structure. (a) The position of known major resistance mutations to the PI LPV are shown in blue, as defined by the IAS. (b) In addition to the position of major resistance mutations, the location of minor resistance mutations to the PI LPV are also shown, highlighted in red. Diagrams were produced using PyMol software from the 3ELI model of protease (RBSC PDB).

For APV/r the following major resistance mutations have been described to date: V32I, M46I/L, I47V/A, I50V, I54A/T/A/L/M, L76V, V82A/T/S/F, I84V, L90M (Rhee et al., 2003). In addition minor mutations L10F/I/R/V and G73S have been observed. Four main pathways to resistance to APV/r were identified in previously PI-naïve patients, either I50V, I54L/M, I84V or V32I with I47V (Maguire et al., 2002). These main pathways to resistance were also observed in PI-experienced patients (Ait-Khaled et al., 2003). Conversely, the I50V mutation, accompanied by L10F, M46I/L or I47V, is the key resistance pathway observed following *in vitro* passage with APV/r (Partaledis et al., 1995; Tisdale et al., 1995). It is thought that this difference in *in vitro* and *in vivo* observations occurs because I50V confers the greatest reduction in susceptibility, but it also has the most significant impact on viral fitness.

To date the following major resistance mutations to ATV/r have been described: V32I, M46I/L, I47V, G48V/M, I50L, I54V/T/A/L/M, V82A/T/F/S, I84V, N88S and L90M. In addition changes L10I/F/V/C, G16E, K20R/M/I/T/V, L24I, L33I/F/V, E34Q, F53L/Y, D60E, I62V, I64L/M/V, A71V/I/T/L, G73C/S/T/A, I85V and I93L/M have been described as minor resistance mutations (Rhee et al., 2003). I50L is the most commonly described major resistance mutation *in vivo* and is recognized as the signature resistance mutation to ATV (Colonno et al., 2004). Following *in vitro* passage with ATV the emergence of V32I, M46I, I50L, I84V and N88S was observed in resistant strains (Gong et al., 2000). Phenotypic analysis of a large panel of PI-resistant patient viruses demonstrated an association between L10I/V/F, K20R/M/I, L24I, L33I/F/V, M46I/L, G48V, A71V/I/T/L, G73C/S/T/A, V82A/F/S/T, I84V and L90M with reduced susceptibility to ATV (Colonno et al., 2003).

The following major resistance mutations to DRV/r have been described mainly identified during the POWER clinical trials: V32I, I47V/A, I50V, I54L/M, L76V, V82F, I84V (De et al., 2008). In addition minor resistance mutations V11I, L33F, T74P and L89V have been described (Rhee et al., 2003). Identification

of major mutation in previously PI-naïve patients has been difficult as many experiencing virological failure on DRV/r based therapy harbour no PI resistance mutations. As a result, most DRV/r resistance mutations have been described in previously PI-experienced patients. Attempts to create resistance viruses by *in vitro* passage with DRV have also proven to be challenging due to its high genetic barrier (Koh et al., 2010).

For IDV/r the following major resistance mutations have been reported: V32I, M46I/L, I47V, I54A/T/A/L/M, L76V, V82A/T/S/F, I84V, N88S, L90M. Minor resistance mutations L10I/R/V, K20M/R, L24I, M36I, A71V/T, G73S/A and V77I have also been described (Rhee et al., 2003). Clinical trials identified that the M46I/L and V82A/F/T mutations account for most treatment failures with IDV, but that these mutations singly do not convey reduced susceptibility phenotypically (Condra et al., 1996). Following therapy *in vivo*, changes at positions 46, 54, 71, 82, 89 and 90 have been observed (Zhang et al., 1997).

Currently, the following major resistance mutations have been described for LPV/r: V32I, M46I/L, I47V/A, G48V/M, I50V, I54A/T/A/L/M, L76V, V82A/T/S/F, I84V, L90M (Rhee et al., 2003). Mutations V32I, M46I and I47V/A have been observed following *in vitro* passage of virus with LPV, as was I50L at a lesser frequency (Carrillo et al., 1998). In LPV/r treated patients M46I/L, I54V/T/A/M/S and V82A/F/S have been described (Maguire et al., 2002). A recent UK study showed that of 291 patients failing LPV/r based therapy, PI resistance mutations were detected in only 11%. The most frequent mutations were I54V, M46I, V82A and L76V (Barber et al., 2012). In a French study, L10V, K20R, L33F, M36I, I47V, I54V, A71V and I85V were selected by LPV/r based therapy (Lambert-Niclot et al., 2012).

At present the following major resistance mutations have been described for NFV: D30N, M46I/L, I47V, G48V/M, I54A/T/A/L/M, V82A/T/S/F, I84V, N88D/S, L90M (Rhee et al., 2003). D30N has been described in patients following

treatment with NFV (Patrick 1998). Mutations at positions 48, 82, 84 and 90 have been associated with reduced rates of virological response in patients (Walmsley et al., 2001). In addition, the K20I mutation has been shown to reduced susceptibility to NFV when present in combination with other mutations (Perno et al., 2001). L23I is a rare mutation positioned in the substrate binding cleft which has been shown to cause low-level resistance to NFV (Johnston et al., 2004). The VIRAPHAR study reported the development of L10I, D30N, M36I, V77I, N88S/D and L90M in response to NFV/r based therapy (Pellegrin et al., 2002).

The following major resistance mutations for SQV/r have been described to date: G48V/M, I54A/T/A/L/M, V82A/T, I84V, N88S and L90M (Rhee et al., 2003). Mutations G48V, V82A and L90M were observed following *in vitro* passage of virus with SQV (Jacobsen et al., 1996). Mutations at position 73 – G73S/T/C/A – are important accessory mutations for SQV/r, particularly in combination with the L90M major resistance mutation. L90M is the most commonly occurring major resistance mutation on failure with SQV and has been reported to occur with I84V, conferring a significant reduction in susceptibility (Rhee et al., 2006)

For TPV/R, the following major resistance mutations have been described: V32I, M46I/L, I47V/A, I54V/A/M, V82T/L, I84V. The major mutations V32I, V82A and I84V were most commonly observed following *in vitro* passage of virus with TPV/r (Doyon et al., 2005). As TPV/r is approved for use in previously PI-exposed patients, V82T develops only in the presence of the major mutation V82A at baseline. In patients who had wild-type virus at baseline, V82L was observed (Baxter et al., 2006). The RESIST trial demonstrated mutations T74P, I47V, V82L/T, Q58E and N83D were the best predictors of viral failure to TPV/r. The most commonly observed minor mutation is L33F, shown to be associated with reduced virological response in combination with other mutations (Vermeiren et al., 2007).

1.4.4.2 Mutations at Gag cleavage sites

It is now well established that Gag cleavage site mutations contribute to PI resistance by partially compensating for the reduction in replicative capacity (RC) caused by major mutations in protease (Mammano et al., 1998). However, several recent studies have shown that Gag mutations can directly influence PI resistance, independent of compensating for reduced single-round infectivity and in the absence of mutations in protease (Dam et al., 2009; Jinnopat et al., 2009; Nijhuis et al., 2007b; Parry et al., 2009). Mutations in Gag have mainly been reported to occur at the protease cleavage sites (cleavage site mutations, CSMs) and at sites distant from cleavage sites, which will be discussed later (Clavel and Mammano, 2010; Gatanaga et al., 2002). All mutations in Gag reported to be associated with PI exposure or resistance to date are displayed in table 1.2.

Table 1.2. Mutations in Gag previously associated with PI resistance or exposure

Gag Mutations associated with PI exposure or resistance	Position within Gag	References
E12K	MA	Gatanaga et al. 2002, Aoki et al. 2009
G62R	MA	Koh et al. 2009
L75R	MA	Gatanaga et al. 2002, Aoki et al. 2009
R76K	MA	Parry et al. 2011
Y79F	MA	Parry et al. 2011
T81A	MA	Parry et al. 2011
K112E	MA	Callebaut et al. 2007
V128A/I/T/del		Callebaut et al. 2007, Dierynck et al. 2007, Larrouy et al. 2010, Knops et al. 2010
165	CA	Kameoka et al. , Jinnopat et al.
M200I	CA	Callebaut et al. 2007
H219Q	CA	Gatanaga et al. 2002, Aoki et al. 2009
A360V	CA	Mammano 1998
V362I	CA	Doyon et al. 2005, Margot et al. 2010
L363M/F/C/N/Y	p2	Koh et al. 2009
S368C/N	p2	Cote et al. 2001, Verheyen et al. 2010
Q369H	p2	Verheyen et al. 2010
T371del	p2	Verheyen et al. 2010
S373P	p2	Cote et al. 2001, Malet et al. 2007, Verheyen et al. 2010
A374P/S	p2	Malet et al. 2007
T375N/S	p2	Mammano et al. 1998, Malet et al. 2007
I376V	p2	Cote et al. 2001, Mammano et al. 1998

R380K	NC	Myint et al. 2004
G381S	NC	Mammano et al. 1998
I389T	NC	Myint et al. 2004
V390D/A	NC	Gatanaga et al. 2002, Aoki et al. 2009
I401T/V	NC	Myint et al. 2004, Callebaut et al. 2007
R409K	NC	Gatanaga et al. 2002, Aoki et al. 2009, Callebaut et al. 2007, Koh et al. 2009
E428G	NC	Dierynck et al. 2007
Q430R	NC	Doyon et al 1996
A431V	NC	Zhang et al. 1997, Mammano et al 1998, Kolli et al. 2009 and others as reviewed in Fun et al. 2012
K436E/R	p1	Verheyen et al. 2006, Nijhuis et al. 2007, Kolli et al. 2009, van Maarseveen et al. 2012
I437V/T	p1	Zhang et al. 1997, Mammano et al 1998, Nijhuis et al. 2007, Dam et al 2009 and others as reviewed in Fun et al. 2012
L449F/P	p6	Doyon et al. 1996, Zhang et al. 1997, Maguire et al. 2002, Myint et al. 2004, Kolli et al. 2009 and others as reviewed in Fun et al. 2012
S451N/T/G/R	p6	Dierynck et al. 2007, Kolli et al. 2009
P452S/K	p6	Yates et al. 2006, Verheyen et al. 2006, Dierynck et al. 2007, Mo et al. 2007, Kolli et al. 2009
T456S	p6	Breuer et al. 2011
E468K	p6	Gatanaga et al. 2002
Q474L	p6	Myint et al. 2007
R479	p6	McKinnon et al. 2010
A487S	p6	Myint et al. 2007
P497L	p6	Myint et al. 2007

Most CSMs are found in the C terminus of Gag - the NC/p1/p6 region – and have been identified *in vitro*, following passage of virus with PIs (Carrillo et al., 1998; Doyon et al., 1996) and *in vivo* in patients undergoing PI treatment, some of whom were failing therapy (Cote et al., 2001; Zhang et al., 1997). The A431V mutation in the NC/p1 CS is the most frequently observed CSM following exposure to PIs and has been observed both *in vitro,* following passage with PIs, (Carrillo et al., 1998; Doyon et al., 1996) and *in vivo* in patients undergoing therapy with RTV, IDV, NFV and SQV (Cote et al., 2001; Nijhuis et al., 2007b). A431V has been shown to convey reduced susceptibility to all PIs except DRV in phenotypic drug susceptibility assays (Dam et al., 2009; Nijhuis et al., 2007b).

Changes at positions 436 and 437 in the NC/p1 cleavage site have also been observed following exposure to PIs both *in vitro* and *in vivo* (Mammano et al., 1998; Zhang et al., 1997; Cote et al., 2001). A recent study demonstrated that these two mutations can directly affect susceptibility in the absence of mutations in protease (Nijhuis et al., 2007b). *In vitro* passage with a PI produced a number of mutants exhibiting substitutions in the Gag p2 spacer peptide – K436E and I437T/V – and none in protease. Synthesis of recombinant virus with wild-type protease showed that the presence of the mutant Gag alone significantly increased the EC_{50} of the recombinant virus to various PIs. A second study examining A431V and I437V, showed that the reversion of the Gag CSM to wild type had a significant effect on PI resistance, but a far more modest effect on RC, indicating a direct effect on resistance of these mutations independent of replication capacity (RC) compensation (Dam et al., 2009; van Maarseveen et al., 2012).

Mutations at the p1/p6 CS, in particular at positions 449, 452 and 453 have been observed following exposure to PIs both *in vitro* and *in vivo*. In particular L449F/V/P is commonly associated with therapy and observed following exposure *in vivo* to a number of PIs: APV, ATV, IDV, NFV, RTV and SQV (Fun

et al., 2012; Larrouy et al., 2010; Maguire et al., 2002; Mammano et al., 1998; Zhang et al., 1997). L449F has been selected following *in vitro* passage with both APV and LPV, as well as a number of experimental PIs (Carrillo et al., 1998; Doyon et al., 1996). Mutations at positions 452 and 453 of Gag have been observed *in vivo* following exposure to a number of PIs (Bally et al., 2000; Maguire et al., 2002). In addition changes at these positions have also been observed *in vitro* following PI exposure, P452K, P453L and P453T, although these have been shown to only convey resistance when found with resistance mutations in protease (Yates et al., 2006).

The only CSM in the N terminus of Gag reported to affect susceptibility is near the MA/CA CS at position 128. It has been associated with PI exposure *in vivo* and *in vitro*, and associated with an increased risk of virological failure following administration of DRV-containing regimens (Callebaut et al., 2007; Larrouy et al., 2010).

There is evidence regarding the effect of Gag CSMs present at baseline on the outcome of subsequent PI-based therapy. Mutations at position 128 near the MA/CA CS have been negatively associated with viral response in the ANRS 127 trial and V128I was observed in >10% patients experiencing virological rebound in the POWER 1, 2 and 3 clinical trials (Dierynck et al., 2007; Larrouy et al., 2010). The ANRS 127 trial also showed that the presence of L449P was negatively associated with virological response (Larrouy et al., 2010). Analyses of data from the POWER trials showed a number of other mutations were also associated with reduced response to DRV-based regimens: E428G, S451T and R452S (Dierynck et al., 2007). Surprisingly, given its frequency and proven role in PI resistance *in vitro*, A431V was not associated with poor outcomes in several clinical trials and was actually associated with positive outcomes with DRV-based treatments (Larrouy et al., 2011; Malet et al., 2007). I437V was associated with reduced virological responses to a number of PIs in the

NARVAL trial and with virological failure in DRV-based regimens (Meynard et al., 2002).

Mutations in Gag cleavage sites are thought to affect PI resistance by enhancing the efficiency of Gag cleavage, as discussed in 1.4.4.4 (Nijhuis et al., 2007b). Several mechanisms by which Gag cleavage site mutations may cause resistance independent of compensating for reduced fitness have been suggested. Gag mutations may improve the kinetics of cleavage, beyond the level of compensating for fitness, and even in the absence of PIs. Alternatively, Gag CSMs may result in a substrate that competes more effectively with PIs for the protease active site (as reviewed by Clavel and Mammano, 2010).

1.4.4.3 Non-Cleavage site mutations in Gag

A number of mutations occurring at sites distant from the cleavage sites have been shown to be involved in protease inhibitor resistance, as shown in table 1.2. Since these are found in different regions of Gag, it is thought that the mechanisms by which they convey resistance are distinct. Previous work in our lab has identified changes within MA which directly affect PI susceptibility. In a single cycle *in vitro* assay, the N terminal portion of Gag from a treatment-experienced patient exhibiting high-level resistance was shown to convey resistance when separated from its cognate protease (Parry et al., 2009). Specifically, three amino acids in MA were shown to be responsible for decreased susceptibility these being: R76K, Y79F and T81A (Parry et al., 2011). These mutations are found in a predicted alpha helix region of MA, and it is thought that the changes may cause the loss of a hydrogen bond resulting in increased flexibility of the region. It has been suggested that this alteration in the conformation of the Gag protein may enhance the efficiency of cleavage (Parry et al., 2011). Other changes in MA have been identified following *in vitro* exposure to APV, E12K and L75R, and an experimental PI, G62R and K112E (Aoki et al., 2009; Callebaut et al., 2007; Gatanaga et al., 2002). It is possible

that these mutations affect Gag multimerisation during virion assembly, explaining their involvement in PI resistance.

Several mutations in CA have been observed following exposure to PIs *in vitro*, at positions M200I and H219Q (Aoki et al., 2009; Callebaut et al., 2007; Gatanaga et al., 2002; Gatanaga et al., 2002; Callebaut et al., 2007; Aoki et al., 2009), although mutations in CA have not been reported to be associated with therapy failure *in vivo* (Fun et al., 2012). The H219Q/R mutation affects the binding of CypA to Gag, and it is thought that this mutation may reduce the requirement of CypA for replication thus increasing replication efficiency. Another study has shown that the presence of a lysine residue at position 165 of Gag in subtype CRF01_AE viruses can directly reduce PI susceptibility (Kameoka et al., 2010).

To date, the role of changes in other regions of Gag in PI resistance remains unclear. Whilst mutations at positions 369-371 within the p2 spacer peptide accumulate during PI therapy *in vivo*, they are not associated with treatment failure nor observed following *in vitro* PI exposure (Myint et al., 2004). In addition, mutations at these positions have not been shown to convey changes in susceptibility *in vitro*. Similarly, the role of changes within NC remains unknown. Several mutations have been associated with treatment failure, I389T and I401V, however the mechanism by which they affect treatment outcome is unknown (Myint et al., 2004). A number of mutations have been observed following PI exposure *in vitro,* V390A/D, I401T and R409K; however they have not been shown to confer changes in susceptibility *in vitro* (Aoki et al., 2009; Callebaut et al., 2007; Gatanaga et al., 2002).

1.4.4.4 Effect of PI resistance mutations on viral fitness

PIs were designed to have the maximum affinity for the protease active site, and as a result PIs occupy more space within the active site than the natural substrates (Prabu-Jeyabalan et al., 2003). Most major PI resistance mutations

confer resistance to PIs by enlarging the protease active site, thus reducing the affinity of the PI to the protease enzyme. It has been suggested that the enlarging of the active site will have more of an effect on the binding of PIs than the substrate to protease (Nijhuis et al., 2007a). Despite this, major PI resistance mutations reduce the rate of Gag and Gag-Pol processing, resulting in a reduction of viral fitness (Zennou et al., 1998). Major resistance mutations are usually observed first and the accumulation of secondary mutations follows with continuing PI-exposure.

Different PI resistance mutations affect viral fitness to different extents and it can be difficult to measure the effect of single mutations on fitness as resistance mutations are often found in combination (methods used to measure fitness are described in 1.5.3). When present in clinical isolates or molecular clones, the following major PI resistance mutations have been associated with a significant reduction in viral fitness: D30N, M46I/L, G48V, I50L, I54V, V82A/T, I84V, N88D/S and L90M (Luca, 2006). A number of studies have examined the exact role of resistance mutations on viral fitness *in vitro*, both independently and when present in combination. For example, major mutations V82T, I84V, M36I and I54V were observed following *in vitro* passage of virus with RTV and a significant reduction in fitness was observed. However, continued passage led to the appearance of A71V and K20R, which increased the fitness of the virus towards that of wild-type (Nijhuis et al., 1999). Mutations I84V and V82A independently confer a reduction in fitness to around 80-90% that of wild-type, but the mutation L10I has been shown to compensate for this reduction in fitness when present with either change (Devereux et al., 2001; Menzo et al., 2003). L10I, alongside L63P, has also been shown to compensate for the reduced fitness conferred by the major mutation L90M (Martinez-Picado et al., 1999).

A second mechanism of compensation for reduced fitness is via mutations in Gag; both at the cleavage sites and at other sites distant from the cleavage

sites. Cleavage site mutations found in concert with protease mutations are thought to alter the tertiary structure of the Gag to make it fit more tightly within the active site of mutant protease, thus compensating for the reduced RC caused by mutant protease. The protease resistance mutation V82A creates a gap between the protease active site and the PI, simultaneously reducing the affinity of Gag to protease and to the PI. A413V in Gag has been shown to create a protrusion into the substrate binding pocket, compensating at least partially for the reduction in affinity to protease (Prabu-Jeyabalan et al., 2004). In addition, L449F is associated with increased fitness when present with major PI resistance mutations (Maguire et al., 2002).

1.4.4.5 Association between mutations in protease and Gag

The emergence of mutations in Gag in response to PI-exposure provides compelling evidence as to the co-evolution of Gag-protease in response to selective pressure. As mentioned above, mutations in Gag can develop to compensate for reduced fitness conferred by mutations in protease. In addition, a number of studies have reported strong associations between specific mutations in Gag and protease that develop together following therapy (as reviewed by Clavel and Mammano, 2010). The exact mechanisms and fitness advantages conferred by this co-evolution are yet to be fully elucidated.

Gag mutations A431V and I437V are associated with V82A in protease, although A431V is also associated with M46I and changes at position 54 of protease (Bally et al., 2000; Verheyen et al., 2006). Mutations at Gag positions L449 and P453 have been linked to the I50V resistance mutation in protease, and shown to confer reduced susceptibility to APV and improve viral fitness *in vitro* (Maguire et al., 2002). The Gag mutation L449F is associated with D30N and I84V, it has even been suggested that the emergence of I84V is protease in favoured by the pre-existence of L449F. P453L has also been demonstrated to emerge with protease mutation I84V (Bally et al., 2000). In addition P453L has been associated with D30N and N88D in protease and it has been

hypothesised that the presence of P453L restores the hydrogen bond that occurs between D30 of protease and position 452 of Gag that is disrupted by the D30N mutation (Shibata et al., 2011).

1.4.4.6 Transmission and persistence of PI resistance

After the development of resistance whilst a patient is on therapy, it is possible that these resistant viruses could be transmitted to other individuals. Studies have shown that patients infected with drug resistant viruses still engage in high-risk sexual behavior, with up to 27% reporting engaging in unprotected sexual intercourse with a partner of either negative or unknown HIV status (Chin-Hong et al., 2005). The acquisition of resistant viruses at new diagnoses is an important consideration in countries where ART is widely used. Infection with a resistant virus can limit treatment options and it has been shown to negatively impact treatment response and decrease time to first virological failure (Little et al., 2002). Transmission of resistant viruses has been well documented throughout geographical areas where ART is in use. Rates of transmitted drug resistance (TDR) are normally monitored by looking at the prevalence of drug resistance in treatment-naïve, recently infected individuals. Reported rates of TDR vary depending on location, time and the methods of the individual study, but have been reported to be more than 10% in every continent by at least one study (http://hivdb.stanford.edu/surveillance/map/). Rates of up to 24.9% have been reported in the USA (Smith et al., 2009), although other reports have found rates of 12.6% and 8.3% (Rhee 2012, CROI poster). Within Europe, the SPREAD study reported TDR rates of 8.4% (Vercauteren et al., 2009).

Since resistant viruses have lower replicative capacities, it has previously been assumed that transmitted resistant viruses would revert to wild-type in the absence of the selective pressure of drug therapy in the newly infected individual. The disappearance of PI DRMs over time has been demonstrated in several reports (Brenner et al., 2002). Conversely, a number of studies have

reported the persistence of resistant viruses over time, despite the absence of drug pressure and the reduced replicative capacity. A study followed 3 patients with PI resistance mutations which persisted throughout the observation period, median time of one year (Barbour et al., 2004). For this reason, all newly diagnosed patients in the UK are now subjected to a genotypic resistance test before therapy begins.

As well as the role for persistence of drug resistance in transmission, it is also an important consideration in the context of treatment interruptions. If a patient develops resistance to PIs during therapy, then the persistence or reversion of the resistance mutations once PI treatment has been stopped has important implications for future treatment options. If resistance mutations persist over time then PI therapy cannot be re-introduced successfully, however, if the mutations disappear and revert to wild-type it may be possible to use the same PI successfully. A study of 25 patients observed a reduction in the number of drug resistance mutations in 50% of patients, over a median of 53 weeks follow up (Gianotti and Lazzarin, 2005). However, in the remaining 50% no reduction in the number of DRMs was observed. A second study reported that after 4 years observation, PI DRMs did not revert to wild type in two patients. The authors suggest that PI DRMs do not revert to wild type due to compensatory fixation - the reversion of one PI DRM to wild type results in a loss of replicative capacity, effectively blocking the route of evolution back to wild type (van Maarseveen et al., 2006).

1.4.5 Resistance to maturation inhibitors

Resistance to the maturation inhibitor, bevirimat, can be predicted by the amino acid sequence of the p2 spacer peptide SP1. Gag positions 369 to 376 have the highest impact, with positions 370 and 372 shown to be particularly important (Heider et al., 2010). *In vitro* studies have shown six amino acids that independently confer resistance to bevirimat: 3 at or near the C terminus of CA

(CA-H226Y, -231F and -231M) and the first three amino acids of SP1 (SP-A1V, -A3T and –A3V). Of these, CA-H226Y, CA-L231F/M and SP1-A1V have no significant effect on the fitness of the virus (Adamson et al., 2006). The prevalence of these mutations in patient viruses is high, with 30% of treatment-naïve patient-derived viruses and 45% of PI-resistant patient-derived viruses containing at least one mutation conferring resistance to bevirimat therapy. These changes increase in frequency in patient whose viruses contain more than three PI associated resistance mutations (Verheyen et al., 2010). Of particular interest, a study showed that PI resistance mutations may delay the development of resistance to bevirimat (Adamson et al., 2009).

1.5 Protease inhibitor resistance testing methods

1.5.1 Genotypic resistance testing methods

Genotypic assays are generally used in clinical practice and involve the prediction of the resistance of a virus based on the presence or absence of known resistance mutations. PCR amplification of *pol* sequences from the patient viruses using RT-PCR is carried out, often using a cDNA synthesis step and nested PCRs. This is followed by DNA sequencing and analysis of sequence data. Commercial kits are available or some centres have developed in-house assays, such as the one used in the Antiviral Unit at PHE Colindale (as described in the materials and methods section). Genotypic analysis involves a prediction of the susceptibility of a virus, based on the presence of known resistance mutations within the sequence. This prediction can be complicated as the patterns of resistance mutations are complex. Algorithms which predict the susceptibility of a virus based on DNA sequences are available, such as the Stanford HIVdb Genotypic Resistance Interpretation Algorithm (http://hivdb.stanford.edu/) (Rhee et al., 2003).

1.5.2 Phenotypic resistance assays

Conversely, phenotypic assays involve a direct measure of the susceptibility of a virus or enzyme *in vitro*, as opposed to a prediction. These generally include replication based assays, such as the one which will be utilised in this project, or enzyme susceptibility assays. These assays are used to determine the EC_{50} of each virus, the concentration of drug that inhibits 50% of replication or enzyme activity, and this value is compared to that of a designated wild-type virus.

There is currently only one commercially available phenotypic drug susceptibility assay – Phenosense, provided by Monogram – following the discontinuation of the Antivirogram assay which was provided by Virco. Both assays are replication-based cell assays, not enzyme based, and use

recombinant viruses. Use of recombinant viruses avoids the need to isolate live virus from patient samples which can be technically challenging and means the tropism of the patient virus does not affect the performance of the assay.

The Phenosense assay utilises a resistance test vector with part of the *pol* gene (N-termial region) that contains known PI and RTI resistance mutations derived from patient virus, but with a luciferase reporter gene in the place of the HIV *env* gene. The vectors are co-transfected into cells with a plasmid encoding a MLV-*env*, resulting in the production of pseudovirions. The infectivity in the presence of drug is tested using a single-round of replication assay, since the nature of viruses limits them to a single-round of replication. Infectivity is measured by recording the amount of luciferase activity (Petropoulos et al., 2000). A number of studies have examined the degree of concordance between the Monogram and Virco commercial assays, and in general the results are good. One study showed that 92% of results were in agreement between the two assays, and most of the viruses for which different results were observed fell just under the respective clinical cut-offs (Qari et al., 2002).

Thus, it follows that one of the complications associated with using phenotypic resistance assays is the determination of clinical cut-offs, the fold difference EC_{50} values at which a difference in clinical outcome would be observed. To determine these values, large numbers of paired clinical outcome and phenotypic susceptibility data are required. For example the clinical cut-offs defined for the Virco assay required >6500 data sets for their determination (Winters et al., 2009).

1.6 Viral fitness assays

A number of different methods to measure viral fitness have been developed, (as reviewed by Quinones-Mateu and Arts, 2001). *In vivo* assays are available that examine viral kinetics in a host environment such as in blood plasma. These have the advantage that an accurate estimate of viral fitness in the host environment is obtained, however this can produce inaccuracies for variants adapted to other compartments such as the CNS.

In addition, a number of *in vitro* assays have been developed which can be divided into growth kinetic assays or growth competition assays. Growth kinetic assays involve infection with single viruses carried out in parallel, with the amount of virus produced over time measured, for example by p24 ELISA. The amount of virus production can then be compared between viruses, although small differences in fitness may not be detected by this method. These assays can be carried out using replication competent viruses or with pseudovirions resulting in analysis of a single round of replication, as with the commercial Phenosense assay. Pseudovirions containing reporter genes such as luciferase enable more accurate measurment of the amount of virus production than in assays with replication competent viruses where p24 ELISA is used. However, single-cycle assays only include the stages of the viral life-cycle up to the transcription of viral genes, meaning that the efficiency of viral assembly is not represented.

Growth competition assays can only be used with replication competent viruses and involved direct competition between at least two viruses over multiple cycles of replication. Over time, the fittest of the viruses will outgrow the other(s) which enables a distinction between viruses with subtle differences in fitness. However, measurement of the proportion of each virus can be challenging with methods such as RT-PCR or use of different reporter genes being required (Quinones-Mateu and Arts, 2001).

1.7 Project overview

The determinants of PI resistance and susceptibility have been investigated for many years. However, most studies to date have utilised commercially available phenotypic susceptibility assays PhenoSense and Antivirogram which do not include co-evolved Gag from the virus of interest (Petropoulos et al., 2000). A number of recent studies have provided evidence that Gag can directly confer PI resistance in the absence of known resistance mutations (Dam et al., 2009; Nijhuis et al., 2007b; Parry et al., 2011). In addition, recent studies have shown that the inclusion of co-evolved Gag alongside protease in phenotypic assays can affect the susceptibility of the virus (Gupta et al., 2010; Parry et al., 2009).

To date, studies have examined the role of Gag from treatment-naïve subtype A and C viruses and PI-experienced subtype B viruses in phenotypic susceptibility assays. However, no data exist to indicate whether inclusion of co-evolved Gag would affect the phenotypic PI susceptibility of PI-naïve subtype B viruses. This report aimed to investigate the phenotypic PI susceptibility of full-length Gag-protease from a variety of PI-naïve subtype B viruses. Using a previously described single-cycle phenotypic PI susceptibility assay, this report demonstrates variation in PI susceptibility of both molecular clones and patient-derived subtype B viruses. This report describes the creation of a number of chimeric viruses containing fragments of Gag and protease to further investigate the regions of Gag and protease that contribute to variation in PI susceptibility of PI-naïve viruses. An investigation of particular amino acid positions conferring variation in PI susceptibility was carried out. In addition, this report describes the development of a protocol to investigate variation in the rates of maturation between pseudovirions using electron microscopy.

The second half of this project aimed to investigate the determinants of treatment failure on PI monotherapy, which almost always occurs in the absence of major PI resistance mutations. Plasma samples were available

from patients enrolled in the MONARK trial and randomised to the LPV/r monotherapy arm. A detailed investigation was carried out of five patients who failed PI monotherapy for which plasma samples were available at baseline and the time of treatment failure, which occurred in the absence of major resistance mutations. This report details the clonal analyses, phenotypic resistance and fitness testing and positive selection analyses carried out for each. In addition, similar data for three control patients who succeeded on PI monotherapy are included to enable an assessment of the clinical significance of any findings. This study is the first to assess PI susceptibility using full-length Gag-protease in paired samples at baseline and time of treatment failure, enabling an assessment of the potential role of reduced susceptibility in treatment failure.

2 Materials and methods

2.1 General microbiological techniques

2.1.1 Plasmid DNA preparation

2.1.1.1 Competent bacteria

The following strains of competent *E. Coli* cells were utilised: HB101 (Promega), JM109 (Promega) and XL10 Gold (Stratagene). HB101 cells were used for general cloning techniques and to produce plasmid DNA stocks, JM109s were utilised for cloning into pGEM as they enable colour screening of colonies and XL10 Gold cells were used for site-directed mutagenesis.

2.1.1.2 Transformation of bacteria

Typically, 2 µl of plasmid DNA was incubated with 45 µl of cells on ice for 30 minutes. XL10 Gold cells required an additional pre-incubation step with 2 µl β-methacapthanol for 2 minutes on ice before addition of DNA. Cells were heat shocked at 42°C for 45 seconds, followed by incubation for 2 minutes on ice. Cells were incubated, shaking, at 37°C for 1 hour with 450 µl Super Optimal Broth (SOC). Varying volumes of cells were plated out onto 10cm agar plates with 100 µg/ml ampicillin and incubated overnight at 37°C. Colour screening plates for JM109s were prepared with the addition of Bluo-Gal and IPTG (Invitrogen).

Following selection as necessary, colonies were seeded into 5 ml of LB with 100 µg/ml ampicillin and incubated, shaking at 37°C overnight. Cells were pelleted by centrifugation at 300 g for 3 minutes and the culture supernatant discarded. Cell pellets were either frozen, or processed immediately as detailed below. Where required, glycerol stocks of transformed bacteria were created by mixing 1 ml of bacterial culture with 200 µl glycerol solution (Sigma) in cryotubes and stored at -80°C.

2.1.1.3 Plasmid DNA preparation

Double stranded plasmid DNA was purified from bacteria using the Qiagen Miniprep Kit (Qiagen), as per manufacturer's protocol. Bacterial cells were lysed and clarified, then spun through a column containing silica membranes to which the plasmid DNA bound. The membranes were washed and the DNA subsequently eluted into 50 µl of elution buffer (10 mM Tris-Cl, pH 8.5). DNA concentration was measured using 2 µl of eluate on a Nanodrop 3100 Spectrophotometer.

2.1.2 Polymerase Chain Reaction (PCR)

PCR was performed using the Expand High Fidelity PCR system (Roche) when high fidelity DNA synthesis was required, or Taq polymerase (Roche).

2.1.2.1 Expand High Fidelity

Expand High Fidelity reactions contained 0.5 µM final concentration each of the sense and antisense primers. Typical cycling conditions were:

1	(x1)	94°C	2 minutes
2	(x35)	94°C	30 seconds
		53°C	30 seconds
		72°C	2 minutes
3	(x1)	72°C	7 minutes

The annealing temperature was reduced to 50°C when primers containing mismatched bases were used, such as GagNot+ and ProXhoR2. DNA primers used for PCR are listed in table 2.1.

2.1.2.2 Taq polymerase

Primers were used at a final concentration of 1 µM each and are detailed in table 2.1. Typical reaction conditions were as follows:

1	(x1)	94°C	2 minutes
2	(x20)	94°C	30 seconds
		53°C	30 seconds
		72°C	45 seconds
3	(x1)	72°C	2 minutes

2.1.2.3 Addition of polyA tails

Where cloning into the intermediate vector pGEM was performed (as detailed in 2.1.7.2), the addition of 3' A-overhangs to the PCR product from the Expand High Fidelity PCR system was required. PCR product from the Expand High Fidelity reaction was purified using the GFX PCR DNA and Gel Band purification kit (GE healthcare) as detailed in 2.1.6.1 and incubated with Taq polymerase and 0.1 µM dNTPs, to add 3' A overhangs. Cycling conditions were: 3 cycles of 53°C for 30 seconds and 72°C for 4 minutes, followed by 72° for 7 minutes.

2.1.3 Sequencing

DNA sequencing was carried out in-house by the Department of Bio-Analysis and Horizon Technologies (DBHT).

2.1.4 Primers

Throughout the project, various primers were utilised for PCRs and sequencing, listed in table 2.1. The areas of primers designed with mismatches to introduce restriction enzyme sites are underlined. Those used for site-directed mutagenesis are detailed later.

Table 2.1. Primers used for cDNA synthesis, nested PCR and sequence analysis

Primer	Sequence
GagNot+	GCGGCGGCCGCAAGGAGAGAGATGGGTGCG
Gag1F	CATTATCAGAAGGAGCCACC
Gag1.5R	TCTATCCCATTCTGCAGC
Gag2F	ATGATGACAGCATGTCAGGG
ProXhoR2	CTGGTACAGTCTCGAG**R**GGACT**R**AT**K**GG
KVL065[1]	TCCTAATTGAACYTCCCARAARTCYTGAGTTC
ProOutR	TTGGGCCATCCATTCCTGG
BKTO3	CGCAGGACTCGGCTTGC

[1]Van Laethem et al., 2006.

2.1.5 Restriction enzyme digestion

Restriction enzyme digestion was carried out using the appropriate endonucleases (NEB and Roche). Typical reaction conditions were: 1 µg of DNA, 5 U of *Not*I and 5 U of *Xho*I in a 20 µl reaction. Reaction was carried out at 37°C (or the temperature recommended by the manufacturer) for 2 hours. Heat inactivation of restriction enzymes was carried out at 65°C, depending on the onward use of DNA.

2.1.6 DNA purification

2.1.6.1 PCR/Gel band purification

DNA fragments were purified either directly from the PCR reaction, or from gel bands excised following agarose electrophoresis to exclude primer dimers. DNA was purified using the GFX PCR DNA and Gel Band Purification Kit (GE healthcare) as per manufacturer's protocol and eluted into 50 µl of elution buffer.

2.1.6.2 Ethanol precipitation

To further concentrate fragment DNA following purification, DNA was ethanol precipitated. 2-2.5 X volume 100% ethanol and 6 µl sodium acetate were added to eluted DNA, and incubated at -20°C overnight. The mix was centrifuged at 15000 g for 15 minutes to pellet the DNA, and the pellets were washed with 100 µl of 70% ethanol, then air dried. DNA pellets were re-suspended in water, typically 3-6 µl depending on the onwards ligation reaction.

2.1.7 Molecular cloning

Cloning of full-length Gag-protease from both molecular clones and clinical samples into the test vector p8.9NSX+ (described in more detail in 2.2.3) was carried out. For molecular clones, where only one clone was required, direct cloning into the p8.9NSX+ vector was sufficiently efficient. However, for cloning from clinical samples where the presence of quasispecies meant that many individual clones had to be investigated, a cloning method utilising an intermediate vector was required.

2.1.7.1 Direct cloning into test vector p8.9NSX+

Cloning of either full-length *gag*-protease or smaller fragments within this length into p8.9NSX+ was performed. The positions of the unique restriction sites present in p8.9NSX+ used for cloning in this project are shown in figure 2.1. Full-length *gag*-protease was cloned from the molecular clones into p8.9NSX+ using *Not*I and *Xho*I restriction sites located 13 nucleotides before the beginning

of the *gag* ORF and after the end of protease ORF, respectively. Other restriction sites utilised for cloning as part of this project were *Spe*I, located in the middle of *gag* within CA, *Apa*I, located at the C terminus of *gag*, and *Sma*I found at the end of *pol*.

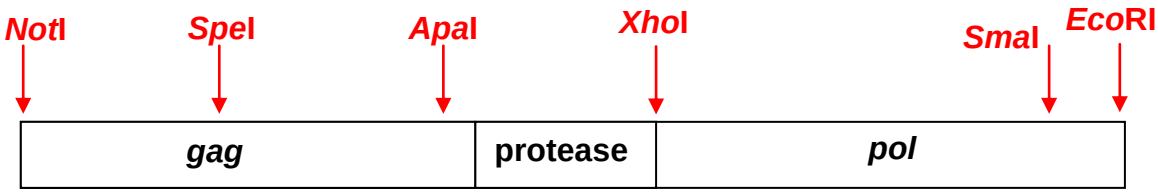


Figure 2.1. Unique restriction sites within p8.9NSX+ *gag-pol*

'Insert' DNA was derived either by PCR amplification to introduce restriction sites or by restriction digest from an existing p8.9NSX+ based plasmid. PCR amplification was carried out as in section 2.1.2.1, using the GagNot+ and ProXhoR2 primers to introduce *Not*I and *Xho*I restriction sites, respectively, into the full-length *gag*-protease fragment.

Insert was ligated into 'empty' p8.9NSX+ vector, prepared by restriction digest with appropriate enzymes and vector dephosphoryation, using the Rapid Dephos and Ligation kit (Roche). Typically ligation reaction was carried out with insert DNA resuspended in 6 µl water following ethanol precipitation, 2 µl empty, dephosphorylated vector and 1 µl DNA ligase. Reaction proceeded for 1 hour at room temperature before transformation into HB101 cells as detailed in section 2.1.1.2.

2.1.7.2 Cloning using pGEM

Cloning of *gag*-protease from clinical samples was carried out using an intermediate vector – pGEM-T Easy (Promega) – to enable clonal sequencing. Following RNA extraction, cDNA synthesis and nested PCR, the addition of 3' A-overhangs was carried out as detailed in section 2.1.2.3. PCR products were then TA cloned into the pGEM-T vector. Ligation proceeded at 4°C overnight in

10 µl total volume with 1 µl DNA ligase enzyme. DNA was transformed into XL10 Gold Competent cells as described above.

2.1.8 Site-directed mutagenesis

Site-directed mutagenesis was utilised to introduce restriction sites and amino acid substitutions. Site-directed mutagenesis was carried out using the Quik-Change Lightning Site-Directed Mutagenesis Kit (Stratagene) as per manufacturer's protocol or the Quik-Change Lightning Multi Site-Directed Mutagenesis Kit (Stratagene), where multiple changes were introduced simultaneously. The primers designed for mutagenesis are listed in table 2.2.

Table 2.2. Primers designed for site-directed mutagenesis.

Primers	**Sequence**
MutProSpe	GCTATAGGTACAGTACTAGTAGGACCTACACC
MutProSpeR	GGTGTAGGTCCTACTAGTACTGTACCTATAGC
Y_GagApaF	GCCAAAAATTGCAGGGCCCCTAGGAAAAAGGGC
Y_GagApaR	CGGTTTTTAACGTCCCGGGGATCCTTTTTCCCG
JRFL_R30K_F	GGCCAGGAGGAAAGAAAAAATATAAATTAAAACATA TAGTATGGGC
JRFL_R30K_R	CCGGTCCTCCTTTCTTTTTTATATTTAATTTTGTATAT CATACCCG
YU2_R30K_F	GGCCAGGGGGAAAGAAACAATATAAATTAAAACATA TAGTATGGGC
YU2_R30K_R	CCGGTCCCCCTTTCTTTGTTATATTTAATTTTGTATA TCATACCCG
YU2_JRFL_E102D_F	CCAAGGAAGCTTTAGACAAGATAGAGGAAGAGC
YU2_JRFL_E102D_R	GGTTCCTTCGAAATCTGTTCTATCTCCTTCTCG
p2T71A	CTGTGGACATAAAGCTGTAGGTACAGTGTTAATAGG ACC

p2T71A_V13I	GGCAACGACCCCTTGTCACAATAAAAGTAGGGGGGC
p2T71A_V13I_P63L	GACAGTATGAACAGATACTCATAGAAATCTGTGGAC
p4T71A	CTGTGGACATAAAGCTATAGGTACAGTATTAGTAGGACCTACACC
p4T71A_V13I	CCCCTCGTCACAATAAGGATAGGGGGACAGC
p4T71A_V13I_P63L	GACAGTATGATCAGGTACTCATAGAAATCTGTGGAC

Reactions included 25 ng of DNA template and 125 ng of each primer – sense and antisense. Typical cycling conditions were as follows:

1	(x1)	95ºC	2 minutes
2	(x18)	95°C	20 seconds
		60°C	10 seconds
		68°C	6 minutes 15 seconds (30 seconds per kb of template DNA)
3	(x1)	68°C	5 minutes

2.1.9 RNA extraction

Viral RNA was manually extracted from plasma samples using one of two methods, depending on the viral load of the samples. For clinical samples with a viral load >1000 copies/ml, viral RNA was extracted from 200 µl of plasma using the QIAmp UltraSens Virus Kit (Qiagen), following the manufacturer's protocol. For clinical samples with a viral load <1000 copies/ml, 1 ml of plasma was ultracentifuged at 55,000 g for 1 hour at 4°C. The resulting pellet was resuspended in DMEM and AVL buffer added (Qiagen) and washing with ethanol carried out. Following addition to columns from the QIAmp UltraSens Virus Kit (Qiagen) kit, samples were processed as per manufacturers instructions.

2.1.10 cDNA synthesis

10 µl RNA was denatured for 10 minutes at 80°C prior to the cDNA synthesis reaction. cDNA synthesis was carried out using the antisense primers ProOutR or KVL065 (Van Laethem et al., 2006), RNAseOut (Invitrogen) and SuperScriptIII enzyme (Invitrogen) in a 20 µl reaction at 53° for 60 minutes, followed by an RT inactivation step at 70°C for 15 minutes. Resultant cDNA was either used immediately for nested PCR as detailed above, or stored at -80°C.

2.1.11 Agarose gel electrophoresis

Varying volumes of DNA, between 5 µl and 18 µl, depending on the DNA concentration were visualised on a 1% agarose gel (Invitrogen) containing 1:16000 RedSafe DNA Stain (Chembio). Samples were loaded with 10X Blue Juice loading buffer and 1 kb DNA ladder to confirm band size (Invitrogen). Gels were run for 30-45 minutes at 110 to 150 V, depending on the size of the gel tank used. The gel was either photographed using a Gel Dock XR+ Imager (BioRad), or bands were visualised using a UV transilluminator (UVP) and the insert DNA band excised using a scalpel.

2.2 HIV-1 viruses

2.2.1 Molecular clones

The following six subtype B molecular clones were obtained from the NIH AIDS Research and Reagent Program: pYU2, pSF2, pJRFL, p89.6, pHXB2 and pNL4-3.

2.2.2 Clinical samples

2.2.2.1 Treatment-naïve, subtype B clinical samples

Plasma samples from HIV-1 infected patients were obtained from the HIV drug resistance genotypic testing service at PHE, Colindale. Anonymised samples were selected from newly diagnosed patients based on high viral load (above 10,000), subtype (B) and absence of major drug resistance mutations in protease and RT (to reduce the possibility of the presence of transmitted drug resistance mutations). The viral loads for each patient are shown in table 2.3.

Table 2.3. Viral loads for treatment-naïve patients infected with subtype B viruses.

Patient Number	Sample Viral Load (RNA copies/ml)
1	2,599,453
2	17,000
3	16,066
4	103,999

2.2.2.2 MONARK trial samples

Plasma samples from HIV-1 infected patients enrolled in the MONotherapy AntiRetroviral Kaletra (MONARK) clinical trial were obtained. MONARK was the first PI monotherapy trial to compare LPV/r monotherapy with LPV/r based HAART in previously treatment-naïve patients. For this study, five patients who experienced virological failure in the PI monotherapy arm were investigated. Therapy failure was defined as viral load >400 copies/ml at week 24 or viral load >50 copies/ml by week 48 of the trial. Patients were only included if genotypic resistance testing carried out as part of clinical care during the trial did not identify any major PI resistance mutations. For each patient, plasma samples from baseline before treatment initiation and from the time of treatment failure were obtained. In addition, 10 patients were selected from the PI monotherapy arm who succeeded on PI monotherapy, achieving a sustained virological response, as control patients. For these patients only baseline plasma samples were included in this study. The details for each of the patient samples are shown, including viral load and virus subtype, in table 2.4.

Table 2.4. Sample details for MONARK clinical trial patients

Treatment Outcome	Patient number	Time point	Subtype	Viral Load (copies/ml)
Failure	508 (SP)	Screening	B	37800
	508 (SP)	Failure	B	25300
	1403 (KON)	Screening	CRF02_AG	44600
	1403 (KON)	Failure	CRF02_AG	24000
	1404 (DIO)	Screening	CRF02_AG	166000
	1404 (DIO)	Failure	CRF02_AG	1080
	3204 (HG)	Screening	B	23800
	3204 (HG)	Failure	B	603
	4201 (SO)	Screening	G	79500
	4201 (SO)	Failure	G	342
Success (Controls)	110 (FRD)	Screening	CRF02_AG	25600
	4003 (MD)	Screening	CRF02_AG	29000
	2112 (WA)	Screening	B	78800
	4202 (LR)	Screening	G	36500
	909 (TH)	Screening	B	16200
	509 (MBF)	Screening	B	40800
	515 (MF)	Screening	CRF02_AG	67500
	1702 (BOY)	Screening	CRF02_AG	50500

2.2.3 Viral vectors

p8.9NSX+ is a ~12.2 kb HIV-based retroviral vector that contains Gag-Pol from HIV-1 (Parry *et al.*, 2009). Nucleotide 1592 to 2325 of p8.9NSX+ (correlating to the beginning of *gag* to the middle of the CA subunit) is derived from the molecular clone HXB2 and the rest of the HIV-1 *gag-pol* derived fragment is from the NL4-3 molecular clone. p8.9NSX+ contains the following unique restriction sites: *Not*I site upstream of the *gag* open reading frame (position 1592), *Spe*I and *Apa*I sites in *gag* (positions 2325 and 2825 respectively), *Xho*I site at the end of protease open reading frame (3380) and a *Sma*I site in integrase (5221) which allows for the cloning of *gag-pol* sequences, as shown in figure 2.1 (Parry et al., 2009). It contains an ampicillin resistance gene and an SV40 origin of replication which increases the copy number of plasmids once transfected into cells expressing the SV40 large T antigen (such as 293T cells). The expression of the HIV-1 derived genes within cells is driven by a CMV promoter. The vector was originally derived from pCMVΔR9 which encoded the full-length HIV-1 genome with the reading frames of Vpu and Env blocked (Naldini *et al.*, 1996).

As part of this project, to avoid the additional screening of clones after the cloning of clinical samples, the p8.9Δpro vector was generated by deleting a fragment, including part of *gag* and most of protease, from p8.9NSX+. This was achieved by introducing a *Spe*I site in protease at position 3297 by site directed mutagenesis using the Quik-change Site-Directed Mutagenesis Kit (Stratagene) with the sense primer MutProSpe and antisense primer MutProSpeR, table 2.2. *Spe*I restriction digest was then carried out to remove the fragment between the new and existing *Spe*I sites, followed by gel extraction and purification of the vector backbone. The vector backbone was then self-ligated using T4 ligase (Roche), which was confirmed by sequencing, to produce p8.9Δpro. See figure 4.1 and section 4.2.1 for further details.

The HIV-1 based retroviral packaging vector pCSFLW expresses the firefly luciferase reported gene under the spleen focus-forming virus (SFFV) promoter and has the HIV-1 packaging signal, both flanked by HIV-1 long terminal repeats (Wright et al., 2008). pCSGW has the same promoters and packaging signal as pCSFLW, but instead expresses the green fluorescence protein (GFP) (Demaison et al., 2002). pMDG expresses the vesicular stomatitis virus envelope glycoprotein (VSV-G). pGEM-T vector for the TA cloning of PCR products was purchased from Promega.

As part of this project, a large number of p8.9NSX+ based vectors were created by cloning techniques in which *gag* and protease were derived from various sources, including molecular clones, patient-derived viruses and the assay reference strain, in different combinations. In addition, a number of vectors containing mutations at specific amino acids of interest were generated. The molecular clone-based vectors are listed in table 2.5, those derived from patients infected with subtype B viruses in table 2.6.

Table 2.5. p8.9NSX+ based vectors derived from subtype B molecular clones, generated during this study

Vector Name	Description
p8.9NSX+	Resistance test vector, used as the assay reference strain control throughout this project
p8HXB2gagpro	p8.9NSX+ with Gag-protease derived from HXB2
p8NL4-3gagpro	p8.9NSX+ with Gag-protease derived from NL4-3
p8YU2gagpro	p8.9NSX+ with Gag-protease derived from YU2
p8JRFLgagpro	p8.9NSX+ with Gag-protease derived from JRFL
p8SF2gagpro	p8.9NSX+ with Gag-protease derived from SF2
p89.6gagpro	p8.9NSX+ with Gag-protease derived from 89.6
p8YU2gag	Chimeric virus with Gag derived from YU2, and protease and pol from p8.9NSX+
p8JRFLgag	Chimeric virus with Gag derived from JRFL, and protease and pol from p8.9NSX+
p8SF2gag	Chimeric virus with Gag derived from SF2, and protease and pol from p8.9NSX+
p8.89.6gag	Chimeric virus with Gag derived from 89.6, and protease and pol from p8.9NSX+
p8YU2pro	Chimeric virus with protease derived from YU2, and Gag and pol from p8.9NSX+
p8JRFLpro	Chimeric virus with protease derived from YU2, and Gag and pol from p8.9NSX+
p8SF2pro	Chimeric virus with protease derived from YU2, and Gag and pol from p8.9NSX+
p8.89.6pro	Chimeric virus with protease derived from YU2, and Gag and pol from p8.9NSX+
p8YU2gagN	Chimeric virus with the N terminus of Gag (MA and half of CA) derived from YU2, and the remainder of Gag-pol from p8.9NSX+
p8YU2gagC	Chimeric virus with the C terminus of Gag (half of CA to x) derived from YU2, and the remainder of Gag-pol from p8.9NSX+
p8JRFLgagN	Chimeric virus with the N terminus of Gag (MA and half of CA) derived from JRFL, and the remainder of

	Gag-pol from p8.9NSX+
p8JRFLgagC	Chimeric virus with the C terminus of Gag (half of CA to x) derived from JRFL, and the remainder of Gag-pol from p8.9NSX+
p8YU2_R30K	p8YU2, with position 30R of Gag reverted to consensus B sequence K
p8YU2_E102D	p8YU2, with position 102E of Gag reverted to consensus B sequence D
p8YU2_2M	p8YU2, with both R30K and E102D reversions to consensus B sequence
p8JRFL_R30K	p8JRFL, with position 30R of Gag reverted to consensus B sequence K
p8JRFL_E102D	p8JRFL, with position 102E of Gag reverted to consensus B sequence D
p8JRFL_2M	p8JRFL, with both R30K and E102D reversions to consensus B seqeunce
p8YU2gN_R30K	p8YU2gagN, with position 30R of Gag reverted to consensus B sequence K
p8YU2gN_E102D	p8YU2gagN, with position 102E of Gag reverted to consensus B sequence D
p8YU2gN_2M	p8YU2gagN, with both R30K and E102D reversions to consensus B sequence
p8JRFLgN_R30K	p8JRFLgagN, with position 30R of Gag reverted to consensus B sequence K
p8JRFLgN_E102D	p8JRFLgagN, with position 102E of Gag reverted to consensus B sequence D
p8JRFLgN_2M	p8JRFLgagN, with both R30K and E102D reversions to consensus B sequence

Table 2.6. Subtype B, treatment-naïve, patient-derived, p8.9NSX+ based vectors generated during this study.

Vector Name	**Description**
p8Pt1gagpro	p8.9NSX+ with Gag-protease derived from the majority variant from patient 1
p8Pt2.1gagpro	p8.9NSX+ with Gag-protease derived from the first of

	the majority variants from patient 2
p8Pt2.2gagpro	p8.9NSX+ with Gag-protease derived from the second of the majority variants from patient 2
p8Pt3.1gagpro	p8.9NSX+ with Gag-protease derived from the first of the majority variants from patient 3
p8Pt3.2gagpro	p8.9NSX+ with Gag-protease derived from the second of the majority variants from patient 3
p8Pt4gagpro	p8.9NSX+ with Gag-protease derived from the majority variant from patient 1
p8Pt2.1gag	Chimeric virus with Gag derived from p8Pt2.1gagpro, and protease and pol from p8.9NSX+
p8Pt4gag	Chimeric virus with Gag derived from p8Pt4gagpro, and protease and pol from p8.9NSX+
p8Pt2.1pro	Chimeric virus with protease derived from p8Pt2.1gagpro, and Gag and pol from p8.9NSX+
p8Pt4pro	Chimeric virus with protease derived from p8Pt4gagpro, and Gag and pol from p8.9NSX+
p8Pt2.1_1M	p8Pt2.1gagpro with T71A in protease, reversion of polymorphism to consensus B sequence
p8Pt2.1_2M	p8Pt2.1gagpro with T71A and V13I in protease, reversion of polymorphisms to consensus B sequence
p8Pt2.1_3M	p8Pt2.1gagpro with T71A, V13I and P63L in protease, reversion of polymorphisms to consensus B sequence
p8Pt4_1M	p8Pt4gagpro with T71A in protease, reversion of polymorphism to consensus B sequence
p8Pt4_2M	p8Pt4gagpro with T71A and V13I in protease, reversion of polymorphisms to consensus B sequence
p8Pt4_3M	p8Pt4gagpro with T71A, V13I and P63L in protease, reversion of polymorphisms to consensus B sequence

2.3 Tissue culture methods

2.3.1 General cell culture

293T cells were maintained at 37°C and 5% CO_2 in DMEM (Invitrogen), supplemented with 10% Foetal Calf Serum, 1% Penicillin Streptomycin and 1% Sodium Pyruvate (Invitrogen). Cells were maintained in 10 cm tissue culture dishes and passaged every two or three days. Cells were split between 1:4 and 1:8 depending on cell density. Media was removed, cells were washed with 10 ml PBS and 1.5 ml trypsin (Invitrogen) was used to dislodge cells.

2.3.2 Transfection

Transfection of 293T cells for PI susceptibility assay was typically carried out in 6 cm tissue culture dishes with 3 x 10^6 cells. 293T cells were plated and co-transfected after 5 hours with 1.1 µg total DNA, 300 ng p8.9NSX+ (or derivatives), 300 ng pCSFLW and 500 ng pMDG, using 6 µl Fugene 6 transfection reagent (Roche) and 70 µl Optimem (Invitrogen). Transfection conditions used for other plate sizes for both the PI susceptibility and single-round infectivity assays are detailed in table 2.7.

Table 2.7. Transfection conditions

Plate Size	Number of Cells	Volume of Fugene (µl)	Amount of DNA (µg)	Harvest Volume (ml)
24 well plate	2.5 x 10^5	1.2	0.22	N/A
6 well plate	1.25 x 10^6	6	1.1	8
6 cm plate	3 x 10^6	6	1.1	22
10cm plate	1.25 x10^7	18	3.5	N/A

Pseudovirion production for electron microscopy required a higher transfection efficiency and virion yield. Cells were plated in 6 well plates and cover slips were included for immunofluorescence. Transfection conditions were optimised as part of this study, as discussed in section 3.2.8.1.

2.3.3 Protease inhibitor susceptibility assay

PI susceptibility was determined using a previously described phenotypic drug susceptibility assay (Gupta et al., 2010; Parry et al., 2009). Briefly, 18 hours after transfection 293T cells were harvested into 22 ml of media for a 6 cm tissue culture dish, see table 2.8. Pseudovirion production was carried out in 96 well plates containing three-fold serial dilutions of PI, as shown in figure 2.2. The top concentration of drug used for each PI can be seen in table 2.8. The first row of the plate (1) was a no-cell control used to calculate the background levels of luciferase and row 12 a no-drug control, used to control for variation in transfection efficiency and replicative capacity between viruses.

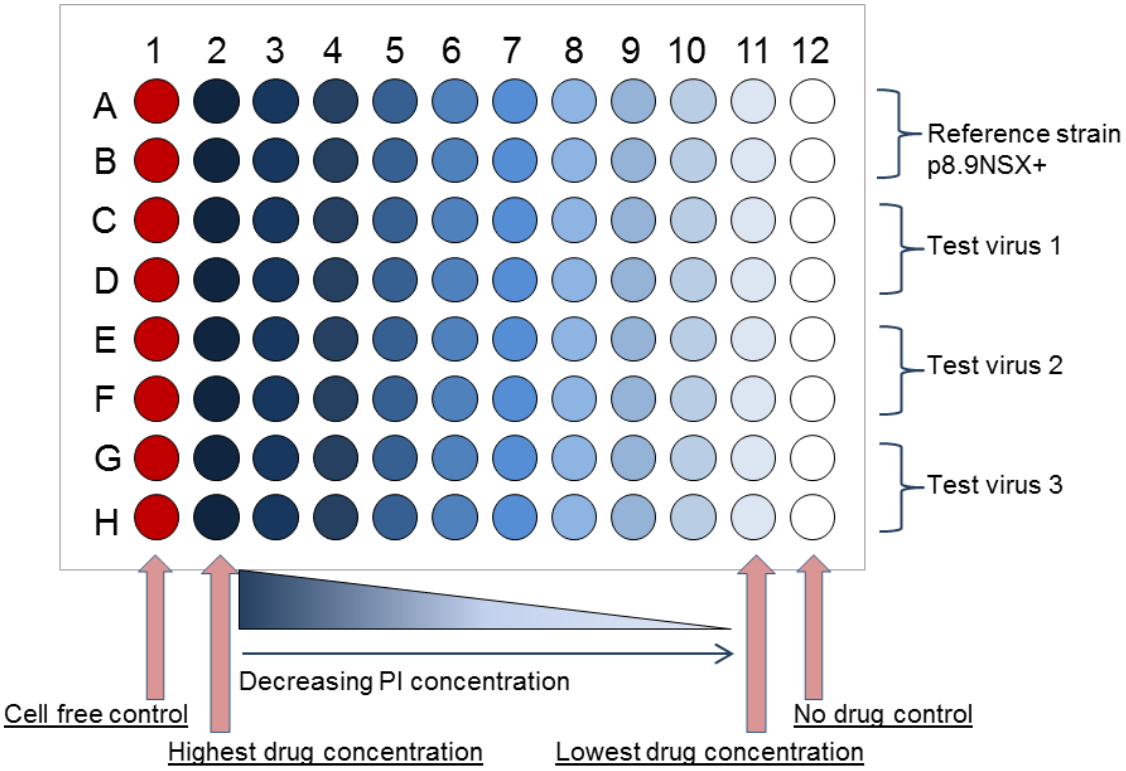


Figure 2.2. Layout of 96-well plate for PI susceptibility assays.

Table 2.8. Top drug concentrations used in phenotypic drug susceptibility assay

PI	Top concentration (nM)
APV	1000
ATV	100
DRV	200
LPV	500
SQV	200
TPV	1000

150 µl of cells were added to each well, containing 150 µl of media with the required concentration of PI. After 24 hours, media containing pseudovirions were harvested and used to infect fresh 293T cells in 96 well plates (layout as shown in figure 2.2). Infectivity was then measured 48 hours after virus transfer, using SteadyGlo luciferase substrate (Promega) and a Glomax luminometer to determine luciferase activity levels. From these data, EC_{50} values were determined by linear regression and data were expressed as a fold difference of the EC_{50} of the reference strain (p8.9NSX+). The experiment was repeated in duplicate within a single assay repeat (for example rows A and B) and mean values used for the calculation of fold-difference in EC_{50}. Assays were performed either in duplicate or in triplicate for each virus construct.

2.3.3.1 Antiretroviral drugs

The six PIs used in this study amprenavir (APV), atazanavir (ATV), darunavir (DRV), lopinavir (LPV), saquinavir (SQV) and tipranavir (TPV) were obtained from the NIH AIDS Research and Reference Reagent Program.

2.3.4 Single-round infectivity assay

2.3.4.1 Titrations

Single-round infectivity was determined as previously described by titration of serial pseudovirion dilutions incubated with 293T cells in the absence of drug (Gupta et al., 2010; Parry et al., 2009). Pseudovirus was harvested 48 hours post-transfection and passed through a 0.45 µm filter to remove cellular debris. 100 µl virus was used immediately for titration, and the remaining volume stored -80°C until required. Titration was carried out in 96 well, white plates with each virus in a single row as shown in figure 2.3. Virus was plated using two-fold dilutions across the plate, resulting in a range from 50 µl volume of virus in the first column to 0.025 µl of virus in the twelfth. A blank row was included in each run to measure background luminescence. 48 hours post-titration, infectivity was measured using SteadyGlo luciferase substrate (Promega) and a Glomax luminometer.

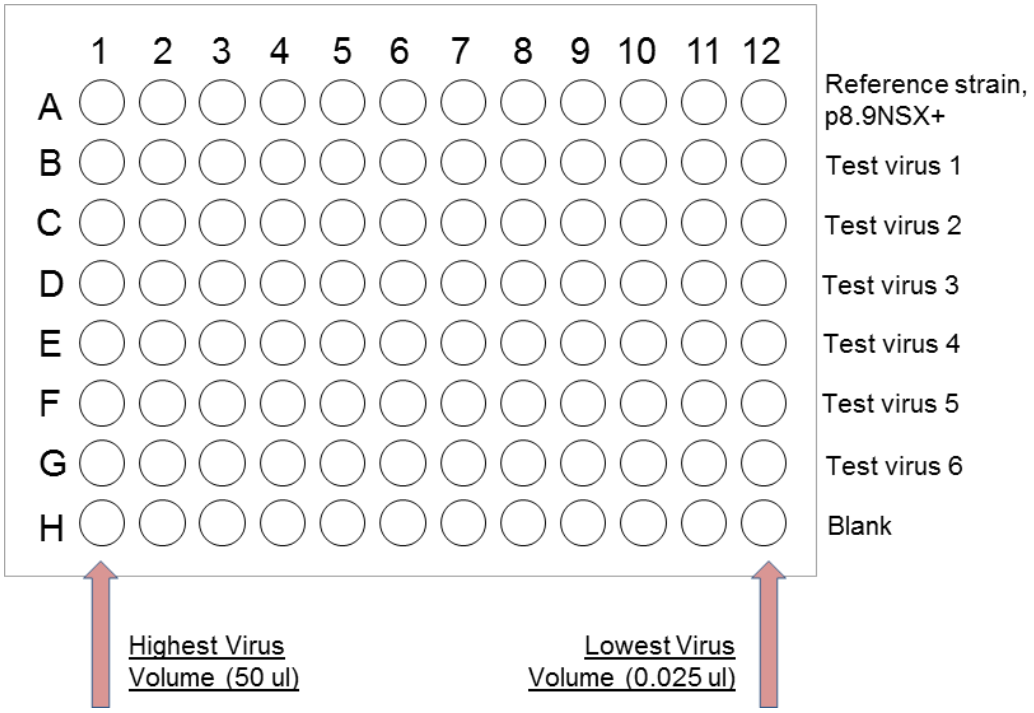


Figure 2.3. Layout of 96-well plate for single-round infectivity assay.

2.3.4.2 p24 ELISA

For the molecular clone-derived vectors, the titration readouts were adjusted based on the amount of p24 present in the harvested pseudovirus sample as previously described (Parry *et al.*, 2009). However, for the MONARK vectors adjusting for p24 was not carried out to capture the effect on PI susceptibility of all life cycle stages. In addition, this has been previously shown not to significantly affect the susceptibility data (Gupta *et al.*, 2010).

The amount of p24 was measured for each pseudovirus using the Genscreen HIV-1 Ag Assay (BIORAD), an enzyme immunoassay for the detection of HIV-1 p24 antigen. The assay was carried out as per manufacturers protocol with OD measured at 450 nM. For each test virus, the assay was carried out for two dilutions – 1:900 and 1:2700 – and negative controls of tissue culture media were included. A p24 standard curve was created using two-fold dilutions

between 100 pg/ml and 6.25 pg/ml to enable a calculation of the amount of p24 in each virus preparation. This value was expressed as a fold-difference of that of the p8.9NSX+ reference strain included on the same p24 ELISA plate and this value used to adjust the single-round infectivity value obtained from titration.

2.4 Imaging methods

2.4.1 Immunofluorescence (IF)

To visualise the presence of Gag following transfection, IF imaging techniques were utilised. Cells were fixed on coverslips in 3 % paraformaldehyde for 15 minutes at room temperature. Coverslips were quenched and permeabilised in PBS with 50 nM NH_4Cl and 0.2% saponin for 10 minutes before rinsing in IF buffer PGAS (PBS with 0.2% fish skin Gelatin, 0.02% sodium azide and 0.02% saponin). Coverslips were stained with a monoclonal antibody targeting Gag CA (p24) #4121 obtained from the AIDS Research and Reference Reagent Program (Simm et al., 1995). Coverslips were incubated with primary antibody at a concentration of 1:250 in PGAS for 30-40 minutes. Coverslips were washed three times in PGAS before incubation with secondary antibody Alexa 488 coupled anti-mouse (Molecular probes) for 30 minutes, diluted 1:300 in PGAS with 1 µg/ml bisbenzimide. Post-staining, coverslips were washed with PGAS, PBS and water before mounting on coverslips with Mowiol 4-88 (Calbiochem). Imaging was carried out using Leica SP8 Laser scanning confocal microscope.

Visualisation of GFP in live cells to monitor transfection efficiency was carried out using a Zeiss Axiovert 200 microscope.

2.4.2 Electron microscopy (EM)

Electron microscopy was carried out with the help of Electron Microscopist, Dr Matthew Hannah, based at PHE.

Following transfection in 6 well plates as described above and addition of PI 18-hours post-transfection where applicable, cell media containing pseudovirions was harvested 48 hours post-transfection. Media was centrifuged at 300 g for 10 minutes to remove cellular debris before ultracentrifugation at 100,000 g for 1 hour at 4°C to concentrate the pseudovirions into a pellet. The pseudovirus pellet was fixed with 2% paraformaldehyde and 1.5% glutaraldehyde in 100 mM sodium cacodylate pH 7.2 for 20 minutes. After washing, the fixed pellet was postfixed with 1% osmium tetroxide, further contrasted with 1% w/v tannic acid and the pellet (*in situ* in the tube following centrifugation) embedded in epoxy resin after dehydration through a graded series of ethanol solutions. The epon blocks for each pellet were carefully removed from the tube and re-embedded in the same orientation to avoid the introduction of bias during sectioning. Serial ultrathin sections, between 60 and 70 nm, were cut from each pellet using a Leica UC7 and Diaotome diamond Knife, mounted on pioloform coated slot grids and stained using Reynolds lead citrate. Pseudovirions were visualised using a JEOL JEM 1400 TEM and images acquired with an AMT 11 megapixel digital camera at 10,000-20,000 x magnification.

2.5 Bioinformatic analyses

2.5.1 Sequence analysis

Sequencing data were aligned and analysed using DNAdynamo software (BlueTractor Software). Sequences were trimmed and manually checked for any discrepancies between overlapping primers. Alignment of multiple sequences for further analyses, for example clonal sequences within a patient, was carried out using MEGA 4.1 or MEGA 5.0 software (Tamura *et al.*, 2011). Virus subtype was determined using the REGA HIV-1 Subtyping tool provided by Stanford University (http://dbpartners.stanford.edu/RegaSubtyping/) (de Oliveira et al., 2005). Analyses for PI resistance mutations within protease was carried out using the Stanford HIV Drug Resistance Database (http://hivdb.stanford.edu/) which reports major and minor resistance mutations and other polymorphisms (Liu and Shafer, 2006). Reference to the IAS-USA list of drug resistance mutations was also made (Johnson *et al.*, 2013). Gag changes were identified manually by reference to table 1.2 containing all previously described mutations in Gag associated with PI exposure or resistance.

2.5.2 Genetic distance calculation

Mean pairwise genetic distance (MPWGD) calculation was carried out in MEGA using standard parameters, based on amino acid alignment performed using the ClustalW algorithm. Data were expressed as mean amino acid substitutions per site.

2.5.3 Positive selection analysis

Selective pressure testing was performed to test for evidence of positive selection using a number of methods available on the datamonkey webpage (www.datamonkey.org). Analyses were performed for all clonal sequences from each patient, from both baseline and failure time points, in alignments. To

test for alignment wide evidence of selection the PARRIS method was used (A PARtitioning approach for Robust Inference of Selection) (Scheffler *et al.*, 2006). For identification of particular sites under diversifying/purifying selection the FUBAR (Fast Unbiased Bayesien AppRoximation) method was used (as discussed by Kosakovsky Pond and Frost, 2005).

2.5.4 Molecular Modeling

Molecular modeling of the position of amino acid changes within the crystal structures of proteins was carried out using RasMol software. Models were obtained from the Research Collaboratory for Structural Bioinformatics (RCSB) protein database (www.rcsb.org) (Berman et al., 2000). For protease, two molecular models were utilised: 7HVP (protease homodimer) and 3ELI (protease structure when bound to the PI ATV) (King et al., 2012; Swain et al., 1990).

2.5.5 Phylogenetics

Nucleotide sequences were aligned in MEGA 5.0 using the ClustalW algorithm, as above, and imported into the PHYLIP program for phylogeny construction using the Maximum Likelihood method under the GTR model of nucleotide substitution. Additional maximum likelihood trees were constructed using PhyML 3.0 software (Guindon et al., 2010) with the confidence of the tree tested using 500 bootstrap replications. Phylogenetic trees were viewed using FigTree v1.3.1 and MEGA 5.0 software.

2.5.6 Ancestral site reconstruction

Ancestral state reconstruction was performed using a maximum parsimony approach, with the software McClade version 4.07 by Stephane Hue, UCL (Maddison and Maddison, 2001). Amino acid substitutions were mapped along maximum likelihood phylogenies reconstructed for each patient with the software FastTree v2.1.7 (Prince et al., 2010).

2.5.7 Statistical analyses

Statistical analyses were performed using Minitab16 software and formulae in Microsoft Excel. In most cases, unpaired t-tests were used to compare the means of two populations, with $P < 0.05$ considered statistically significant. Mann Whitney U rank sum testing for the MONARK chapter was performed by John Gregson (London School of Hygiene and Tropical Medicine).

3 Investigation of the role of co-evolved Gag and protease in PI susceptibility of HIV-1 subtype B molecular clones

3.1 Introduction

At present, commercial phenotypic PI susceptibility assays include only protease and a short length of the C terminus of Gag from the patient (the portion of Gag included is p6 which overlaps with the protease in the Gag-Pol transframe region), with the rest of the genome coming from the test vector (Hertogs et al., 1998; Petropoulos et al., 2000). To date, studies investigating variation in phenotypic susceptibility of patient-derived viruses to PIs have used one of these assays – Antivirogram or Phenosense (Parkin et al., 2004; Vergne et al., 2006). However, recent *in vitro* studies using phenotypic drug susceptibility assays have demonstrated that inclusion of Gag from patients alongside co-evolved protease can affect susceptibility to PIs in comparison to protease with a wild-type Gag. This has been shown in patients infected with subtype CRF01_AE viruses (Jinnopat et al., 2009), in subtype A and C molecular clones and viruses derived from treatment-naïve patients, and in an extensively PI treated patient infected with subtype B virus (Gupta et al., 2010; Parry et al., 2009).

It is now well established that changes in Gag can compensate for the reduced fitness caused by major PI resistance mutations found in protease (Zhang et al., 1997). However, more recent data have indicated that Gag can affect protease inhibitor resistance, independent of this compensatory effect (Dam et al., 2009; Jinnopat et al., 2009; Nijhuis et al., 2007b; Parry et al., 2009). Changes in Gag cleavage sites, in particular within the C terminus of Gag, have been shown to affect PI susceptibility both at levels beyond compensation (Maguire et al., 2002; Zhang et al., 1997) and directly in the absence of co-evolved protease

(Nijhuis et al., 2007b). Data from our lab and others have shown the role of changes in the N terminus of Gag on PI susceptibility, at sites distant from the cleavage sites – positions 76, 79 and 81 within MA and 165 within CA (Jinnopat et al., 2009; Kameoka et al., 2010; Parry et al., 2009; Parry et al., 2011).

Given the mounting evidence for the role of Gag in protease inhibitor susceptibility, an assessment of the contribution and variability of full-length Gag to PI susceptibility in *in vitro* phenotypic PI susceptibility assays was required. This study was undertaken to further our understanding of the role of Gag in PI susceptibility by examining the variation in PI susceptibility of treatment-naïve subtype B HIV-1. Subtype B viruses were selected for several reasons. Firstly, subtype B HIV-1 has a high prevalence in the UK, Western Europe and the USA, where routine pre-treatment susceptibility testing is performed. Secondly, a small study looking at subtypes A and C had already been carried out so investigation of the other prevalent subtype was prudent (Gupta et al., 2010). Given the highly polymorphic nature of Gag, and that some of the changes reported to be associated with resistance are present in PI-naïve viruses, we wanted to investigate whether natural variation in Gag present in PI-naïve viruses may cause variation in susceptibility to PIs.

<u>Chapter aims:</u>

1) Determine PI susceptibility of full-length Gag-protease from six subtype B molecular clones utilising a single replication-cycle phenotypic PI susceptibility assay

2) Investigate any variation in PI susceptibility to determine the contribution of Gag and protease separately

3) Identify regions or particular amino acid changes contributing to the variation in susceptibility

4) Investigate the single-round infectivity of all vectors generated for the investigation of PI susceptibility

5) Investigate the mechanisms by which the change in susceptibility may occur

3.2 Results

3.2.1 Variation in PI susceptibility of full-length Gag-protease observed between six different subtype B molecular clones

Six subtype B molecular clones namely, pHXB2, pNL4-3, pYU2, pJRFL, p89.6 and pSF2 were obtained from the NIH AIDS Research and Reference Reagent Program. The protease sequences of the molecular clones indicated that none had any major or minor PI resistance mutations, as determined by the Stanford HIVdb Genotypic Resistance Interpretation Algorithm and comparison to the IAS list of Resistance Mutations (Johnson et al., 2013; Liu and Shafer, 2006). Full-length *gag*-protease from the six molecular clones was cloned into p8.9NSX+, the Gag-Pol expression vector, using the *Not*I restriction site located before the *gag* open reading frame and the *Xho*I restriction site located after the protease/RT cleavage site within p8.9NSX+, as shown in figure 2.1. *Not*I and *Xho*I restriction sites were introduced into the molecular clone *gag*-protease fragments using standard PCR techniques with primers GagNot+ and ProXhoR2. This led to the creation of six vectors, designated HXB2gagpro, NL4-3gagpro, YU2gagpro, SF2gagpro, 89.6gagpro and JRFLgagpro, as detailed in table 2.5. The susceptibility of pseudovirions derived from these six constructs was then determined using the single replication-cycle PI susceptibility assay, as described in chapter 2, for six PIs: APV, ATV, DRV, LPV, SQV and TPV.

Differences in susceptibility to each of the six PIs were observed for the molecular clone-derived vectors in comparison to that of the assay reference strain, p8.9NSX+, as shown in figure 3.1. Each of the molecular clone-derived vectors displayed reduced PI susceptibility in comparison to that of the reference strain, and the reduction was statistically significant to at least one PI for each molecular clone-derived vector ($P < 0.05$). The reductions in susceptibility were most significant for PIs APV, ATV and LPV, with up to 9-fold reductions in EC_{50} in comparison to assay reference strain.

In particular, full-length Gag-protease of three molecular clones (89.6, YU2 and SF2) showed statistically significant reductions in PI susceptibility in comparison to the assay reference strain to all six of the PIs tested and a fourth molecular clone (JRFL) showed statistically significant reductions to five of the PIs. Using the previously determined significance cut-off value in phenotypic assays of ≥ four-fold (Gong et al., 2000), these four molecular clones displayed significantly decreased susceptibilities to three of the PIs (APV, ATV and LPV) to varying degrees in the order: YU2>JRFL>89.6>SF2 ($P < 0.01$; two sample t test). The largest reductions in PI susceptibility were observed for YU2gagpro (for APV, ATV and LPV), JRFLgagpro (for ATV) and 89.6gagpro (for APV) of up to 9-fold, 9-fold and 8.5-fold, respectively. For molecular clones HXB2 and NL4-3, the EC_{50} values ranged from 1- to 5-fold reductions in comparison to that of assay reference strain for each of the PIs. The reduction in PI susceptibility was only statistically significant for two PIs for NL4-3 (APV and SQV) and three PIs for HXB2 (APV, ATV and TPV).

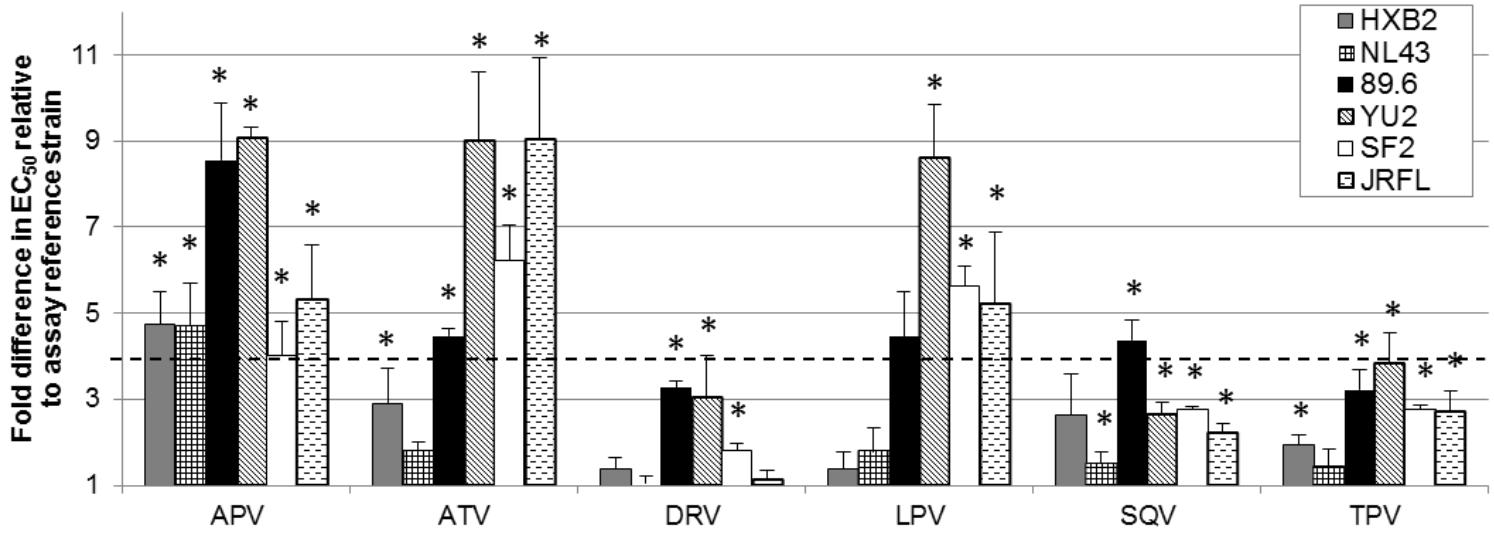


Figure 3.1. Variation in PI susceptibility of full-length Gag-protease derived from six subtype B molecular clones. VSV-g pseudotyped viruses encoding luciferase and containing full-length Gag-protease from six subtype B molecular clones were produced by co-transfection of 293T cells. PI susceptibility of pseudovirions was determined using a single replication-cycle drug susceptibility assay, with infectivity determined by measuring luciferase activity with SteadyGlo. Data displayed are fold difference in EC_{50} values in comparison with that of the reference strain control, p8.9NSX+, for each of six PIs: APV, ATV, DRV, LPV, SQV and TPV. Viruses for which the raw EC_{50} values were statistically different from that of assay reference strain as measured by an unpaired t-test ($P < 0.05$) are denoted with an asterisk(*). Error bars show the standard error of the mean of three independent experiments performed in duplicate. The dashed line shows the previously reported cut off for a significant reduction in susceptibility of greater than four-fold in comparison with the assay reference strain, derived using a different assay system (Gong et al., 2004).

3.2.2 Confirmation of the reproducibility of variation in PI susceptibility conferred by full-length Gag-protease from subtype B molecular clones

The reproducibility of the observed reductions in PI susceptibility was further investigated to exclude the possibility that the differences were caused by variation between individual DNA preparations. Thus, the PI susceptibility of two additional DNA preparations of YU2gagpro and 89.6gagpro was tested in parallel with the original DNA stock used for the production of the previous data set. YU2 and JRFL were selected as they displayed the greatest difference in PI susceptibility in comparison to the assay reference strain. The susceptibility to the two most variable PIs ATV and APV was determined using the phenotypic drug assay, as previously described.

These data showed no statistically significant differences in PI susceptibility between any of the DNA preparations for each molecular clone to either PI, as shown in figure 3.2 ($P < 0.05$). The pseudovirions derived from molecular clone 89.6 displayed some variation in susceptibility to APV (between 5- and 8-fold difference EC_{50}) and ATV (between 6- and 11-fold difference in EC_{50}), but this fell within the limits of inter-assay variability of approximately two-fold and fold-difference and EC_{50} values were not significantly different statistically ($P <0.05$; pairwise two-sample t-test). The DNA preparations of vectors derived from YU2 displayed little variation in susceptibility to APV (all DNA preparations showed around 9-fold difference in EC_{50}) and some variation in susceptibility to ATV (between 6- and 12-fold difference in EC_{50}). The differences in susceptibility between DNA preparations were not statistically significant for either PI ($P <0.05$; pairwise two-sample t-test).

These data demonstrate that the variation in susceptibility of full-length Gag-protease observed between the six subtype B molecular clones was not an

artefact of individual DNA preparations, confirming the validity of the previous observations.

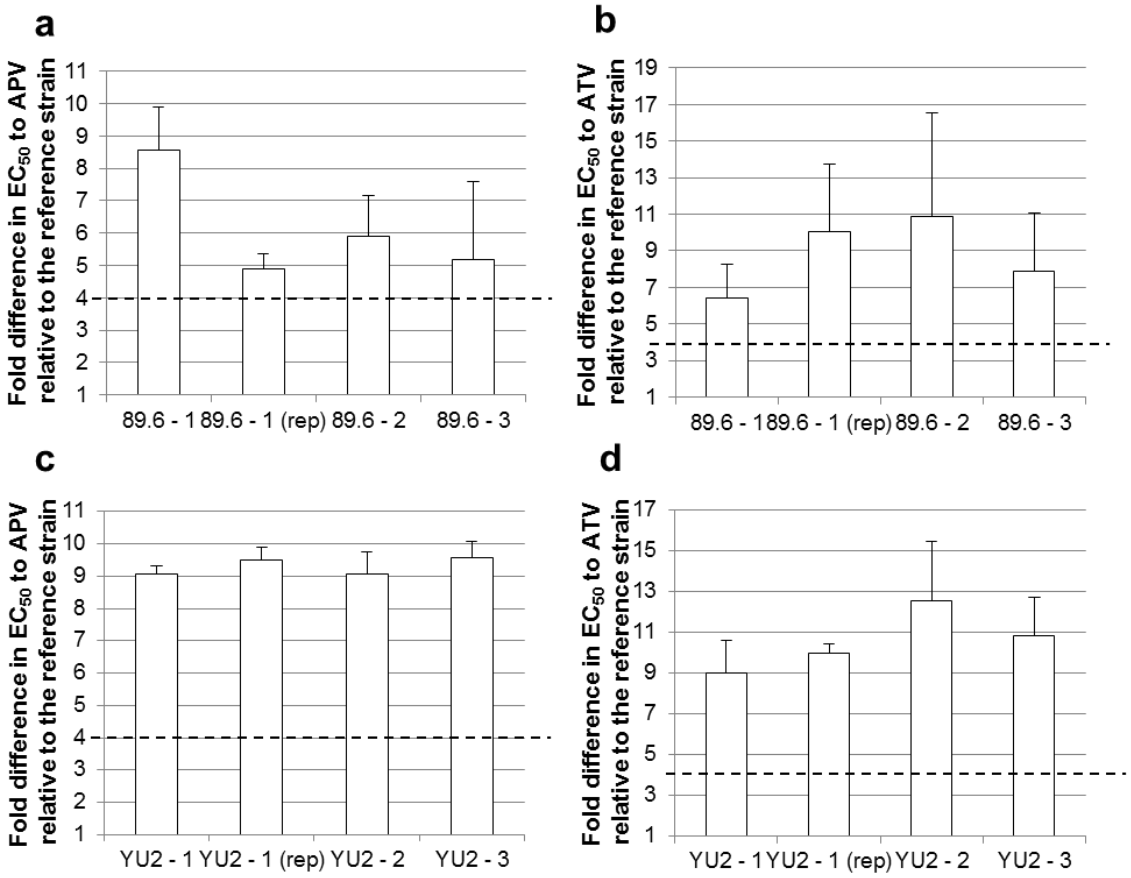


Figure 3.2. Reproducibility of PI susceptibility data shown by testing PI susceptibility of multiple DNA preparations. VSV-g pseudotyped viruses encoding luciferase were produced from multiple DNA preparations of p8MCgagpro. Their PI susceptibility was measured using a single replication-cycle phenotypic assay and expressed as a fold difference in EC_{50} values in comparison to that of the assay reference strain, p8.9NSX+, for the PIs APV and ATV. None of the differences in susceptibility to each PI of different preparations of the same vector were statistically significant, as measured by unpaired t-test ($P < 0.05$). Error bars represent the standard error of the mean of three independent experiments. The dashed line shows the previously reported cut off for a significant reduction in susceptibility of greater than four-fold in comparison with the assay reference strain (Gong et al., 2004).

3.2.3 Contribution of Gag to reduction in PI susceptibility observed for full-length Gag-protease of subtype B molecular clones

Having confirmed the validity of our observations, we investigated the role of Gag and protease separately in the reduced PI susceptibility observed for pseudovirions containing full-length Gag-protease derived from molecular clones. In this analysis, we included full-length Gag-protease vectors for which at least four-fold reduction in susceptibility was observed to at least three PIs. These criteria resulted in the inclusion of pseudovirions derived from YU2, JRFL, SF2 and 89.6 in the analysis, the molecular clones for which Gag-protease conferred the greatest reduction in susceptibility in comparison to the assay reference strain.

To determine the PI susceptibility of the *gag* and protease genes separately, chimeric vectors containing *gag* derived from the molecular clone and protease from p8.9NSX+, and vice versa, were generated using standard cloning techniques for each of the molecular clones of interest, as shown in figure 3.3 and detailed in table 2.5. The 'Gag only' vector contained the fragment between the *Not*I restriction site (located upstream of Gag) and the *Apa*I site flanking amino acids 406 and 407 of Gag (located within the NC subunit of Gag) from the molecular clone, and the remainder of Pol derived from the reference strain. The 'protease only' vector contained the fragment between the *Apa*I site and the *Xho*I site (located downstream of protease) of the molecular clones. Use of the *Apa*I site for the generation of these chimeric vectors results in the protease only vector containing the C terminus of NC, p1 and p6 from Gag in addition to protease from the molecular clone, with the remainder derived from the assay reference strain, as shown in figure 3.3a. Conversely, the Gag only chimeric virus did not contain full-length Gag derived from the molecular clone with the C terminal portion of NC, p1 and p6 derived from the reference strain. Since the *Apa*I restriction site was not present in molecular clone YU2 it was introduced by site-directed mutagenesis before cloning into p8.9NSX+. The PI

susceptibility of the chimeric pseudovirions to the three PIs APV, ATV and LPV was determined using the single replication-cycle PI susceptibility assay, as previously described.

For molecular clones JRFL and YU2, the Gag only chimeric pseudovirions displayed significant reductions in PI susceptibility in comparison to assay reference strain, whereas the protease only chimeric pseudovirions had PI susceptibilities similar to that exhibited by the assay reference stain (Figure 3.3b and 3c). The difference in the fold-difference EC_{50} values for the Gag only and protease only chimeric viruses were statistically significant to APV and ATV for JRFL, and to all three PIs for YU2 ($P < 0.05$). These data indicate that for molecular clones JRFL and YU2, Gag solely conferred the reduced PI susceptibility observed for the full-length Gag-protease fragment and provide evidence that the effect of Gag on PI susceptibility can be conferred in the absence of co-evolved protease.

In contrast, both the Gag only and the protease only chimeric pseudovirions derived from molecular clones 89.6 and SF2 showed reduced susceptibilities of up to 4-fold in EC_{50} in comparison to assay reference stain (Figure 3.3d and 3e). For both SF2 and 89.6, the Gag only chimeric virus displayed a statistically significant reduction in PI susceptibility in comparison to the protease only chimeric virus to the PI LPV ($P < 0.05$). Overall, both Gag and protease of molecular clones 89.6 and SF2 independently contribute to the reduced susceptibility observed for full-length Gag-protease. However, Gag independently conveys greater reductions in susceptibility than protease, in particular for molecular clone 89.6.

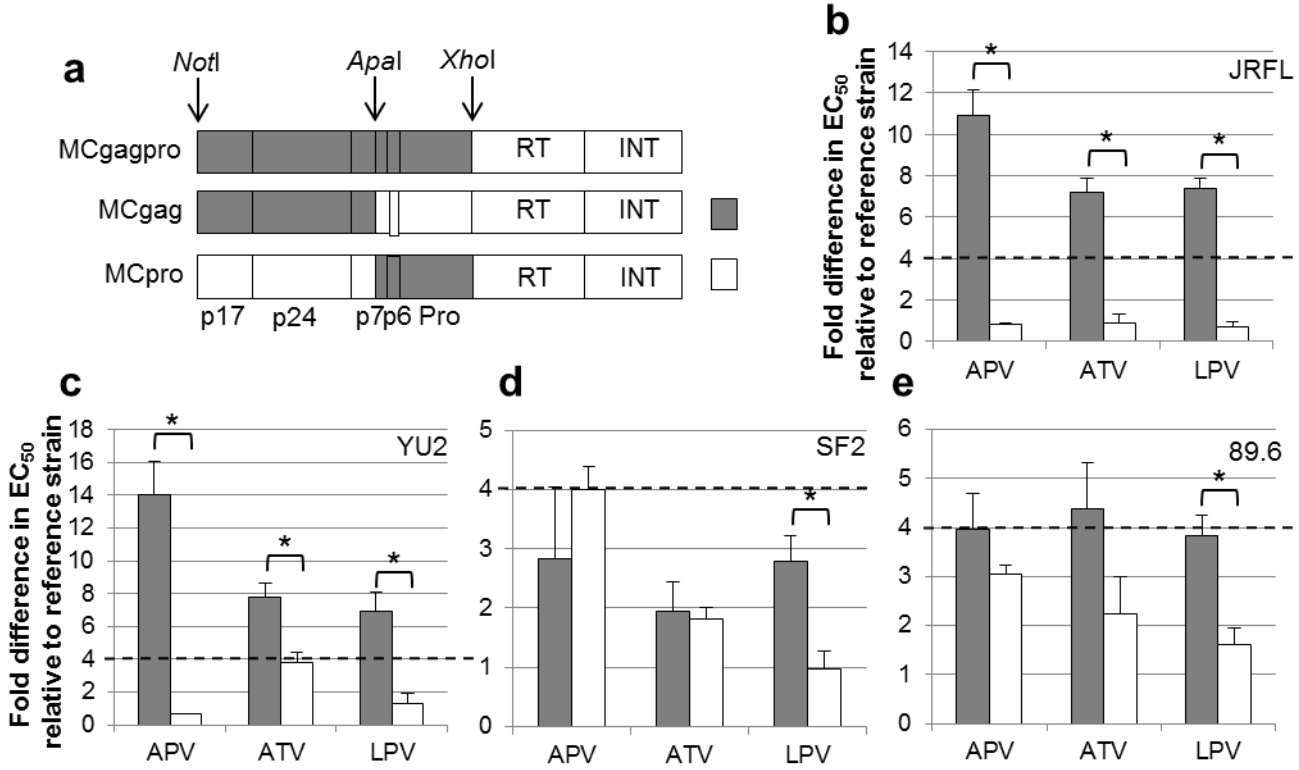


Figure 3.3. Reduced PI susceptibility of molecular clone Gag-protease is independently conferred by both the *gag* and *protease* genes. Chimeric vectors containing only Gag or only protease from the molecular clone of interest were generated using cloning techniques, as shown in (a). Segments of the chimeric viruses derived from the assay reference strain, p8.9NSX+, are depicted in white, and that from molecular clones shaded grey. The PI susceptibility of Gag only and protease only chimeric pseudoviruses was determined for molecular clones (b) JRFL, (c) YU2, (d) SF2 and (e) 89.6 using a single replication-cycle assay measuring luciferase activity with SteadGlo. Data are expressed as a fold difference in EC_{50} in comparison to that of the assay reference strain, p8.9NSX+. The PIs for which MCgag and MCpro displayed significantly different susceptibilities using an unpaired t-test ($P < 0.05$) are denoted with an asterisk(*). Error bars represent standard error of the mean of three independent experiments.

3.2.4 Role of the N terminus of Gag in reduced PI susceptibility exhibited by JRFL and YU2 Gag-protease

To further investigate the reduced PI susceptibility conferred by JRFL and YU2 Gag, we sought to identify a specific region of Gag conferring the reduced susceptibility observed. Only JRFL and YU2 were included in this analysis as these were the molecular clones in which Gag was shown to solely confer the reduced susceptibility observed for full-length Gag-protease. Chimeric *gag* vectors were constructed using a unique *Spe*I restriction site, flanking amino acids 240 and 241 of Gag within the CA subunit. This led to the generation of two chimeric vectors for each molecular clone: gagN, containing the N terminus of Gag derived from the molecular clone and the C terminus of Gag and protease from p8.9NSX+, and gagC, containing the C terminus of Gag derived from the molecular clone and the N terminus of Gag and protease from p8.9NSX+, as shown in figure 3.4a and table 2.5. In the resulting chimeric pseudovirions, the N terminal region included MA and partial CA, with the C terminal region encompassing the remainder of CA, p1, p2 and part of NC. The remainder of NC, p6 and protease were derived from the reference strain p8.9NSX+. The susceptibility of these vectors was then determined using the single replication-cycle drug susceptibility assay for PIs APV, ATV and LPV.

Reduced susceptibility to the three PIs APV, ATV and LPV was observed for the gagN-derived pseudovirions of up to 7- and 16-fold difference for JRFL and YU2, respectively, as shown in figure 3.4b and c. The gagC-derived pseudovirions of both JRFL and YU2 showed susceptibilities at levels similar to that of assay reference strain to all PIs with no significant fold-difference in EC_{50}. For each PI the difference in susceptibility between the gagN and gagC pseudovirions for each molecular clone was statistically significant ($P < 0.05$). These data demonstrate that for both YU2 and JRFL, the N terminal portion of Gag encompassing MA and the beginning of CA, is largely responsible for the reduced susceptibility observed for full-length Gag**.** There is no evidence that

the C terminal portion of Gag can confer reduced susceptibility independent of the N terminus.

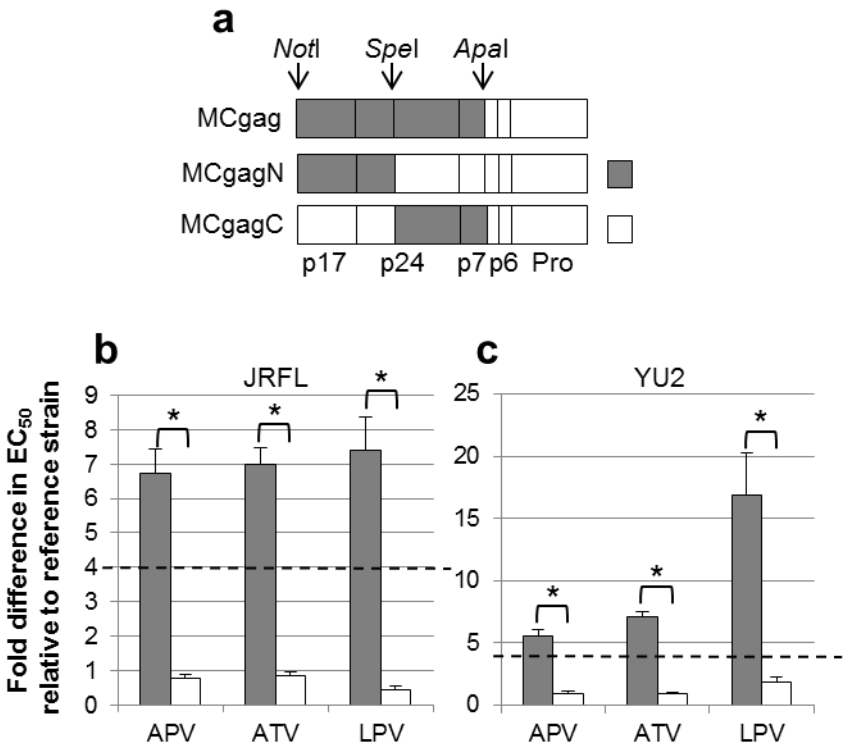
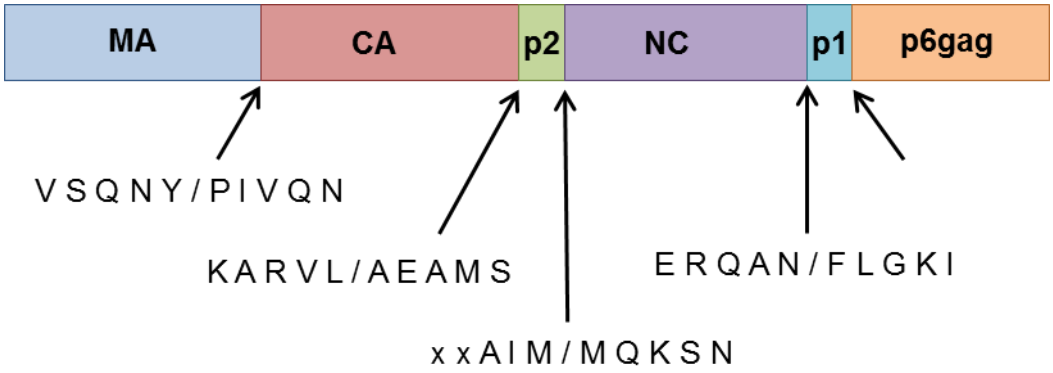


Figure 3.4. PI susceptibility of chimeric viruses derived from the N terminus and C terminus of molecular clone Gag demonstrates reduced susceptibility is conferred by the N terminus of Gag. (a) Schematic diagram of chimeric vectors created by standard cloning techniques containing either the N terminus (MCgagN) or the C terminus (MCgagC) from each molecular clone. Segments of the chimeric viruses derived from the reference strain, p8.9NSX+, are depicted in white, and those from the molecular clones shaded grey. The susceptibilities to three PIs (APV, ATV and LPV) of the pseudovirions generated from these chimeric vectors for two molecular clones are displayed (b) JRFL and (c) YU2. PI susceptibility was determined in a single-round assay by measuring luciferase activity. The PIs for which MCgagN and MCproC displayed significantly different susceptibilities using an unpaired t-test ($P < 0.05$) are denoted with an asterisk(*). Error bars represent the standard error of the means of three independent experiments. The dashed line shows the previously reported cut off for a significant reduction in susceptibility of greater than four-fold in comparison with the assay reference strain (Gong et al., 2004).

3.2.5 Genetic determinants in Gag responsible for the reduced susceptibility to PIs in the molecular clones YU2 and JRFL

Alignment and comparison of the molecular clone nucleotide sequences was performed to identify amino acid changes within Gag which could account for the reduced PI susceptibility observed. Sequence analysis of the Gag cleavage sites found no previously reported mutations that could account for this variation in susceptibility and only the p2/NC cleavage site, within the C terminus of Gag, showed sequence variation, as demonstrated in figure 3.5. The p2/NC is the most variable cleavage site, and the changes did not correlate with the molecular clones for which the reduced susceptibility of Gag was observed.



p8.9NSX+:	P	A	T	I	M	/	I	Q	R	G	N
YU2:	S	A	T	I	M	/	M	Q	R	G	N
89.6:	S	A	T	I	M	/	M	Q	R	G	N
JRFL:	P	A	T	I	M	/	M	Q	R	G	N
SF2:	P	A	N	I	M	/	M	Q	R	G	N
HXB2:	S	A	T	I	M	/	M	Q	R	G	N
NL4-3:	P	A	T	I	M	/	I	Q	R	G	N

Figure 3.5. The amino acid sequences of Gag cleavage sites for the six subtype B molecular clones included in this study demonstrate no correlation between genotype and phenotypic susceptibility. Consensus group M virus amino acid sequence at each Gag cleavage site is shown, and no variation was present at these sites in the molecular clones except at the p2/NC site. The amino acid sequence for each molecular clone is displayed for the p2/NC sequence.

In addition, nine amino acid changes in Gag that have been previously described as being associated with resistance or exposure to PIs were present in at least one of the molecular clones, as detailed in table 3.1. However, none of these correlated with the changes in phenotypic PI susceptibility observed as none were present in both molecular clones in which Gag solely conferred the reduced susceptibility, YU2 and JRFL, or in all four molecular clones in which significant reductions in susceptibility for full-length Gag-protease were observed, YU2, JRFL, 89.6 and SF2.

Table 3.1. Mutations in Gag of the molecular clones that have previously been associated with PI exposure or resistance do not correlate with phenotypic susceptibility.

Previously described Gag mutation	Present in (molecular clone)
R76K	89.6
H219Q	89.6
S373P	NL4-3, JRFL, SF2, p8.9NSX+
T375N	SF2
R380K	NL4-3, p8.9NSX+
I389T	NL4-3, YU2, SF2, 89.6
I401T	HXB2

Analysis of the amino acid sequence of the N terminus of Gag identified only one residue, at position 30, where an amino acid change occurs exclusively in YU2 and JRFL, the two molecular clones for which Gag solely contributed to reduced PI susceptibility observed for full-length Gag-protease, as demonstrated in figure 3.6. Another amino acid change at position 102 was present in both YU2 and JRFL and was also present in SF2, in which both Gag and protease contributed in part to the reduced PI susceptibility of full-length Gag-protease. Both of these positions are in the MA subunit within the N terminal segment of Gag shown to confer reduced susceptibility in these molecular clones. In addition, there were a number of unique changes from consensus B sequence in either YU2 or JRFL and within the N terminus of Gag these changes were T53S, Q69K, R76T, V94I, N109T, K112M and V159I for pJRFL and G10A, K28Q, N47D and R91K for pYU2.

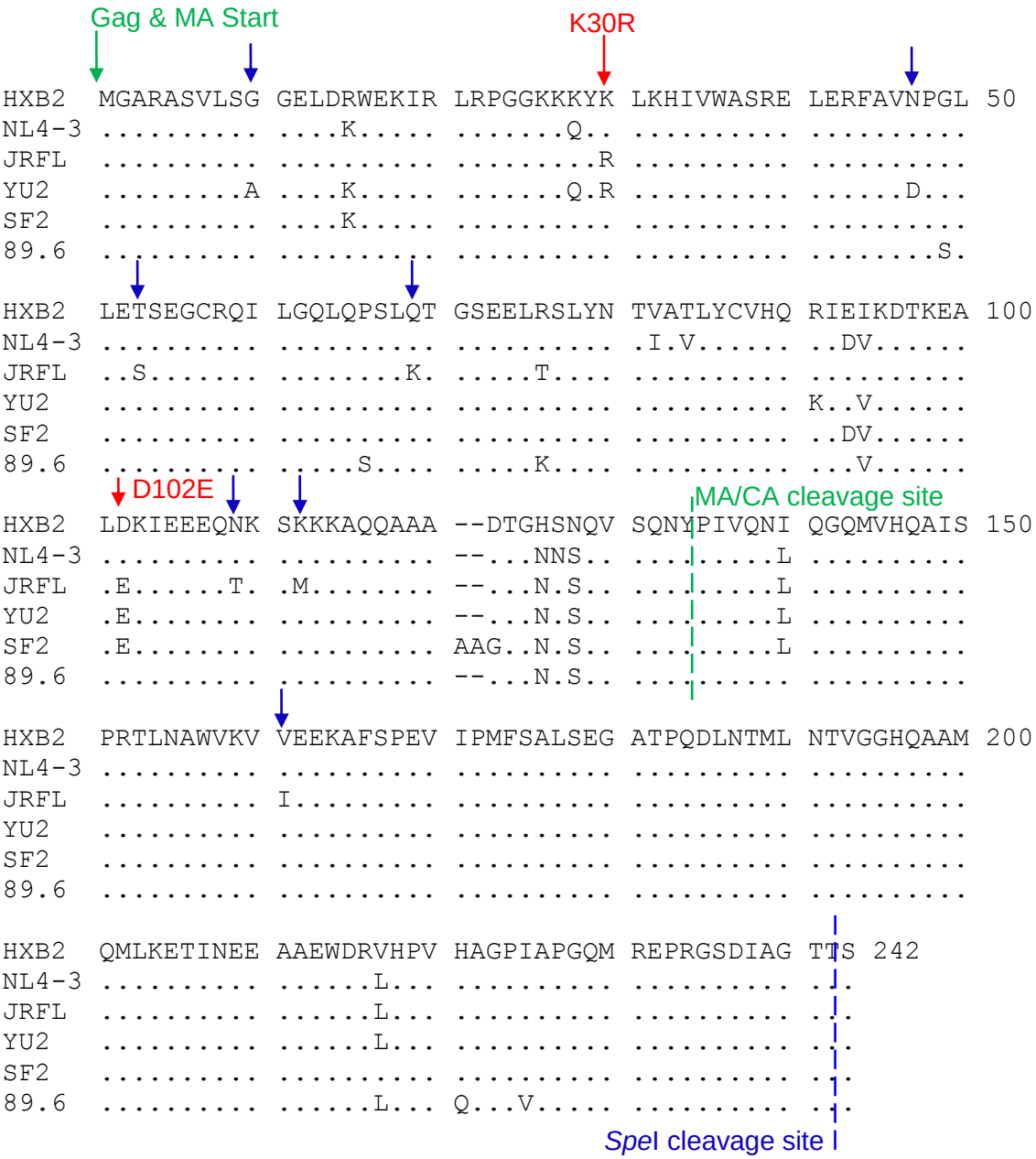


Figure 3.6. Amino acid sequences of the N terminus of Gag of six subtype B molecular clones. Schematic showing the amino acid sequence of the six subtype B molecular clones used for this study, up to the *Spe*I cleavage site which is the length contained within the N terminus chimeric virus. p8.9NSX+ is not shown separately as it is derived from HXB2 so is identical up to the *Spe*I cleavage site, The *Spe*I site is denoted by a blue dashed line and the MA/CA cleavage site by a green dashed line. The amino acid substitutions found in both JRFL and YU2, the molecular clones for which Gag solely confers the reduced susceptibility observed for full-length Gag-protease, are highlighted with a red arrow. The unique changes present in either YU2 or JRFL are highlighted with a blue arrow.

3.2.6 Role of amino acid changes in N terminal region of Gag on PI susceptibility

To investigate the role of the identified changes within the N terminus of Gag in the observed reduction in PI susceptibility, the two changes present in both JRFL and YU2 – K30R and D102E – were reverted to consensus B sequence using site-directed mutagenesis. Reversions were carried out in the vectors containing full-length Gag-protease derived from the two molecular clones leading to the creation of three vectors for each: pMC_R30K, pMC_E102D and pMC_2M, the latter which contained both the R30K and E102D reversions (figure 3.7a). In addition, the amino acid changes were introduced into the pMCgagN vectors to examine the effect on PI susceptibility in the absence of co-evolved C terminus of Gag and protease. These vectors were designated MCgagN_R30K, MCgagN_E102D and MCgagN_2M (figure 3.8a). The susceptibility of pseudovirions derived from these vectors was determined for three PIs, APV, ATV and LPV, using the phenotypic drug susceptibility assay as previously described.

In general, the reversion of at least one of the positions in both MCgagpro and MCgagN vectors increased the susceptibility of the virus to all three PIs towards that of the reference strain, as shown in figures 3.7 and 3.8. The differences in susceptibility between MCgagpro and the mutant pseudovirions were statistically significant for JRFL and YU2 in susceptibility to ATV, figure 3.7 ($P < 0.05$). Similarly, the differences in PI susceptibility between MCgagN and the mutant pseudovirions were significant for ATV, figure 3.8 ($P < 0.05$). For molecular clone JRFL, the reversion of K30R appeared to cause a greater increase in the PI susceptibility of the pseudovirions than the D102E reversion, although this observation was not statistically supported. Conversely, for YU2 the D102E appeared to cause the greatest increase in susceptibility of the two mutations, although again this was not statistically supported. For each molecular clone, the independent effects of the single mutants on PI

susceptibility were comparable to that of the double mutant suggesting an additive rather than synergistic effect of the two amino acid changes. Overall, these data indicate that the R30K and E102D changes contributed to the reduced susceptibility observed for full-length Gag-protease of JRFL and YU2. However, these reversions do not increase the PI susceptibility to levels similar to that of the assay reference strain Gag-protease, indicating the involvement of other as yet unidentified residues.

A p8.9NSX+ vector in which the R30K and E102D changes were introduced was generated, denoted p8.9SDMgag. However when transfected this vector did not produce any infectious pseudovirions in our system. Hence, it was not possible to determine the effect of the R30K and E102D amino acid changes on the PI susceptibility of the assay reference strain.

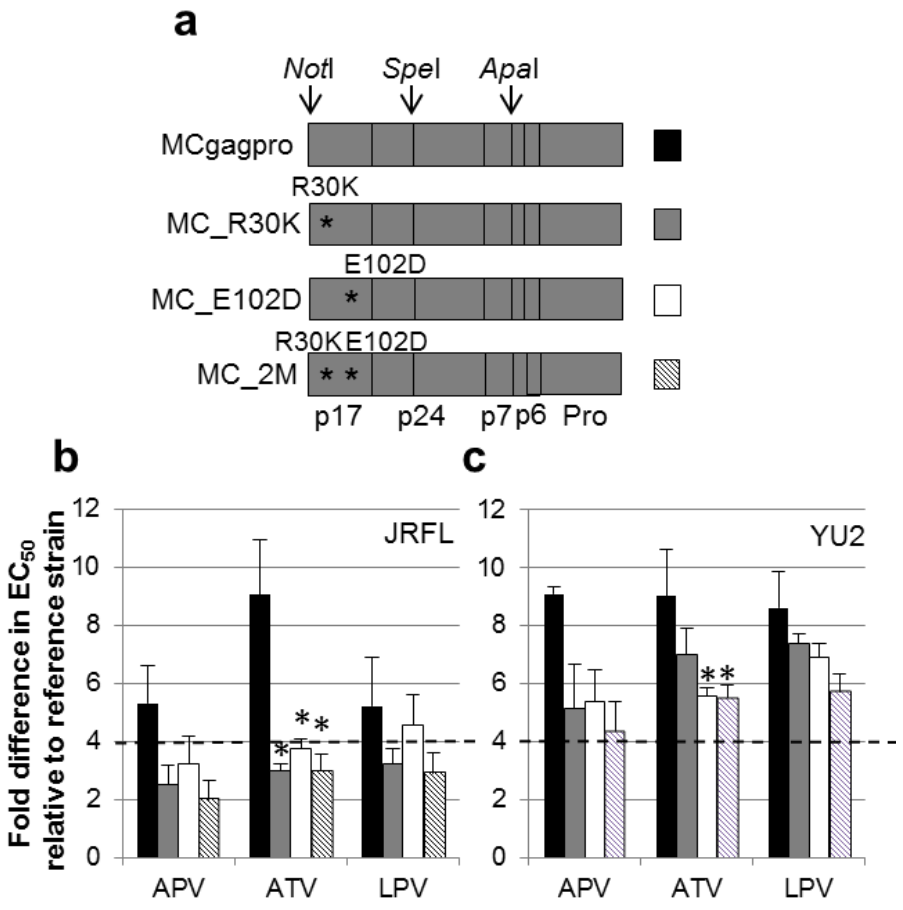


Figure 3.7. Role of amino acids 30 and 102 of Gag in variation in PI susceptibility of full-length Gag-protease. (a) Schematic diagram of MCgagpro dervied vectors containing reversions to consensus B sequence of mutations at positions 30 and 102 of Gag, generated using standard site-directed mutagenesis techniques. The position of the mutations introduced is marked with an asterisk (*) on the schematic. The susceptibilities of pseudovirions generated with the vectors to three PIs (APV, ATV and LPV) are displayed for (b) JRFL and (c) YU2. PI susceptibility of pseudovirions was determined in a single-cycle assay by measuring luciferase activity. Pseudovirions for which the fold-difference in EC_{50} value was statistically different from that of MCgagpro using an unpaired t-test ($P < 0.05$) are denoted with an asterisk(*). Error bars represent the standard error of the mean of three independent experiments. The dashed line shows the previously reported cut off for a significant reduction in susceptibility of greater than four-fold in comparison with the assay reference strain (Gong et al., 2004).

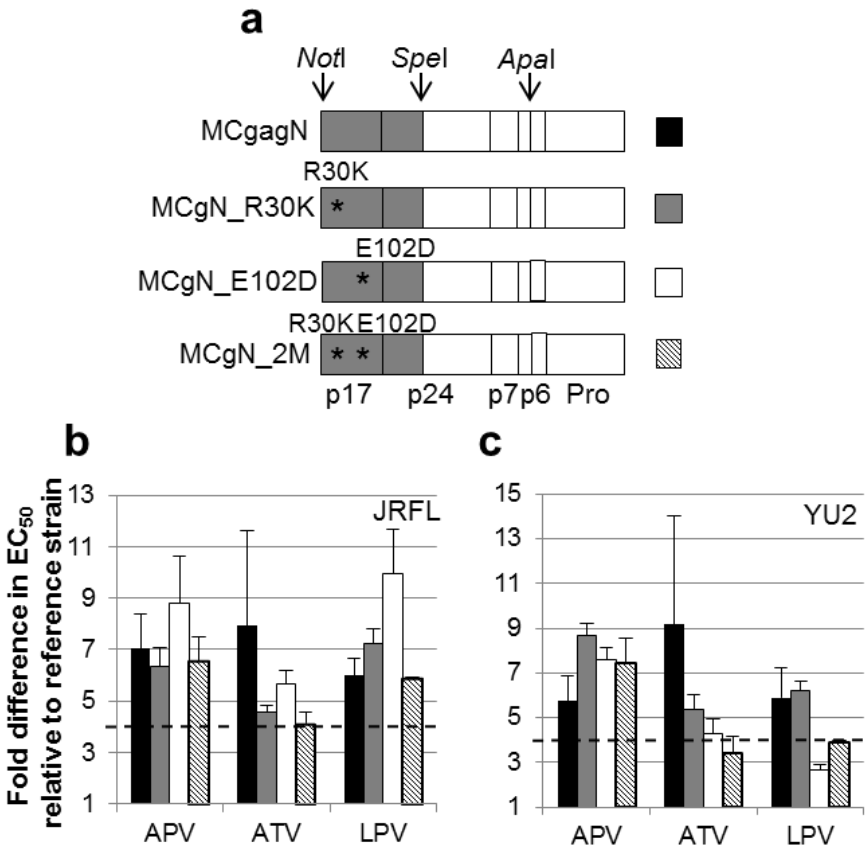


Figure 3.8. Role of amino acids 30 and 102 of Gag in variation in PI susceptibility of MCgagN derived pseudovirions. (a) Schematic diagram of MCgagN dervied vectors containing reversions to consensus B sequence of mutations at positions 30 and 102 of Gag, generated using standard site-directed mutagenesis techniques. The position of the mutations introduced is marked with an asterisk (*) on the schematic. The susceptibilities of pseudovirions generated with the vectors to three PIs (APV, ATV and LPV) are displayed for (b) JRFL and (c) YU2. PI susceptibility of pseudovirions was determined in a single-cycle assay by measuring luciferase activity. None of the differences in PI susceptibility from MCgagN were statistically significant using an unpaired t-test ($P < 0.05$). Error bars represent the standard error of the mean of three independent experiments. The dashed line shows the previously reported cut off for a significant reduction in susceptibility of greater than four-fold in comparison with the assay reference strain (Gong et al., 2004).

3.2.7 Single-round infectivity of the molecular clone-based vectors

Having explored the PI susceptibilities of vectors containing full-length Gag-protease from molecular clones, Gag and protease separately, the N terminus and C terminus of Gag separately and those with reversions at amino acids 30 and 102 of Gag within full-length Gag-protease, the single-round infectivity of each of the vectors was investigated. Pseudovirions were harvested 48 hours post-transfection from the supernatant of transfected cells and passed through a 0.45 μM filter to remove contaminants. Pseudovirions were then titrated in the absence of drug to measure infectivity and the results were normalised by the amount of p24 in the pseudovirion preparations, quantified by ELISA.

The single-cycle infectivity data for each of the vectors generated during this study are displayed in figure 3.9. The values for the JRFL-based vectors were all comparable to that of the assay reference strain (set as 100%), ranging from 76-112% as shown in figure 3.9a. The values for the YU2 based vectors were also similar to that of assay reference strain ranging between 87-100%, except for that of YU2gag which had a value 189% that of the reference strain (figure 3.9b). However, the single-round infectivity of YU2gag was not significantly different statistically from YU2gagpro, YU2gagN or YU2gagC ($P > 0.05$). The single-round infectivities of the SF2-based vectors were higher than that of the assay reference strain, ranging from 237% for SF2gagpro to 158% for SF2pro as shown in figure 3.9c. However the differences in value between SF2gagpro and SF2gag or SF2pro were not statistically significant, nor was the difference between SF2gag and SF2pro ($P > 0.05$). The single-round infectivities of the 89.6-based vectors were broadly similar to that of reference strain, ranging from 110-136% as shown in figure 3.9d. The differences between 89.6gagpro and 89.6gag or 89.6pro were not statistically significant, nor was the difference between 89.6gag and 89.6pro ($P > 0.05$).

For each molecular clone, the separation of co-evolved Gag and protease did not appear to affect the single-round infectivity of the p8.9NSX+ based vectors. In addition, there was no evidence that the reversion of changes at amino acid positions 30 and 102 of Gag, which were shown to affect PI susceptibility, affected the single-round infectivity significantly.

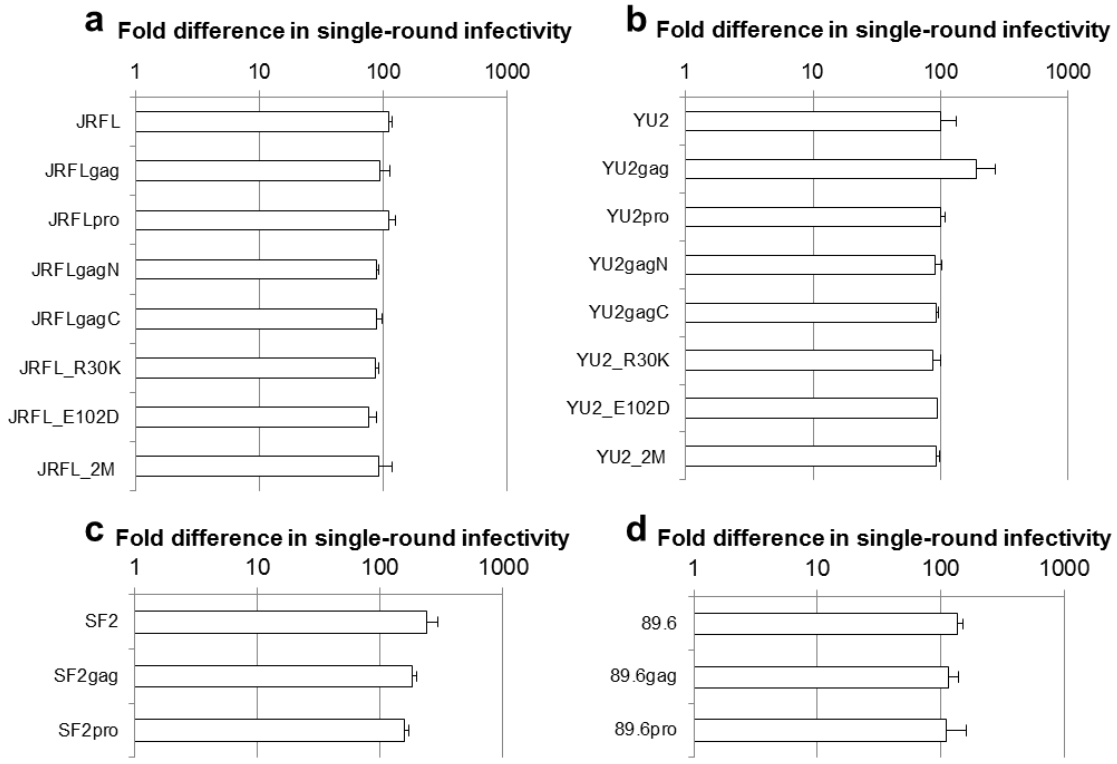


Figure 3.9. Limited variation between the single-round infectivity of molecular clone derived, p8.9NSX+ based vectors. The infectivity of pseudovirions produced from co-transfection of 293T cells was determined by titration of a single-cycle of infection, adjusted following p24 quantification. Luciferase activity was measured using SteadyGlo and expressed as a fold difference in comparison with that of the assay reference strain, p8.9NSX+. The single-round infectivity of all p8.9NSX+ based vectors derived from molecular clones (a) JRFL, (b) YU2, (c) SF2 and (d) 89.6 explored in this study are displayed. None of the differences in susceptibility between related vectors of the same molecular clone were statistically different using unpaired t-test ($P < 0.05$). Error bars are standard error of the means of two independent experiments.

3.2.8 Analysis of the mechanisms underlying the reduced susceptibility to PIs conferred by Gag-protease of molecular clones

HIV-1 Gag protein has been shown to control the proper assembly and morphology of HIV-1 nucleocapsid resulting in production of either immature (non-infectious) or mature (infectious) virus particles (as reviewed by Ono, 2010). To investigate whether an effect on virion morphology was contributing to the effect on PI susceptibility conferred by the Gag protein of molecular clones YU2 and JRFL, we used electron microscopy (EM) to visualise the morphology of nascent virions of reference strain- and molecular clone-derived virions, produced in the presence and absence of PI inhibitors.

3.2.8.1 Optimisation of transfection protocol for EM

Firstly, we undertook optimisation of our transfection protocol to ensure the best possible yield of pseudovirions for EM. In total six different transfection conditions were compared, each a variation on the transfection method used successfully for the generation of pseudovirions for the PI susceptibility and single-round infectivity phenotypic assays. The effect of three variables on transfection efficiency was tested, these being: volume of transfection reagent (Fugene-6), total amount of DNA and the ratios of the three assay plasmids (p8.9NSX+, pCSGW and pMDG). In addition, a second transfection reagent was tested, PEI. To enable visualisation of transfected and infected cells, pCSGW (which expressed GFP, green fluorescence protein) was utilised in the place of pCSFLW. The six transfection conditions tested are detailed in table 3.3. At high confluence, 293T cells do not adhere adequately to cover-slips as required for IF so we also tested a two-fold reduction in the number of cells seeded for transfection. Despite these optimisation steps the transfected cells still produced microvesicles and long membrane protrusions, as shown in figure 3.10. These structures were often difficult to distinguish from the real viral particles. It was hypothesised that this could be partly due to the toxicity of

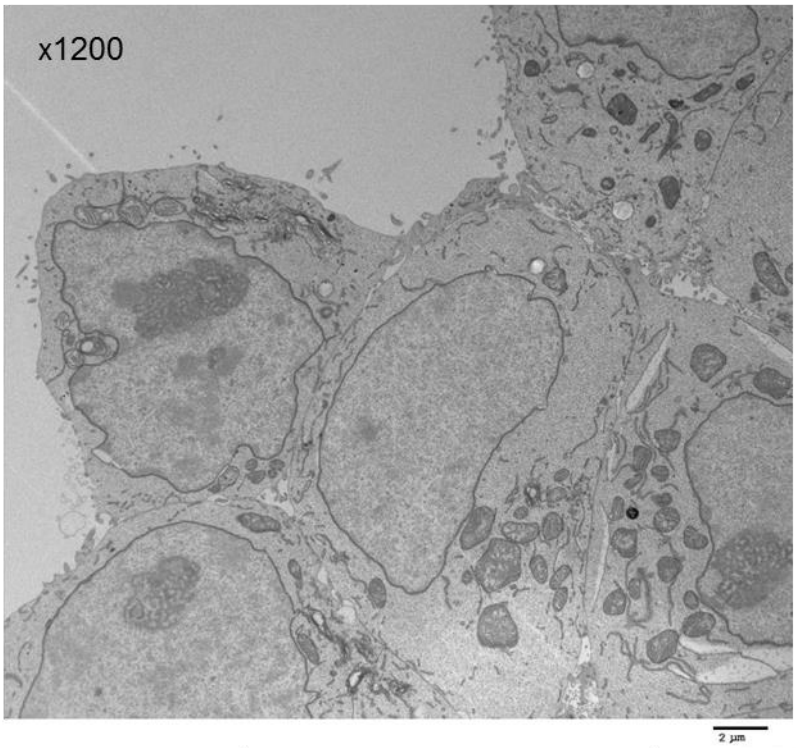
VSV-G envelope (Burns et al., 1993). Hence, we also carried out transfection with reduced amounts of VSV-G envelope vector, pMDG.

Table 2.3. Transfection optimisation conditions tested.

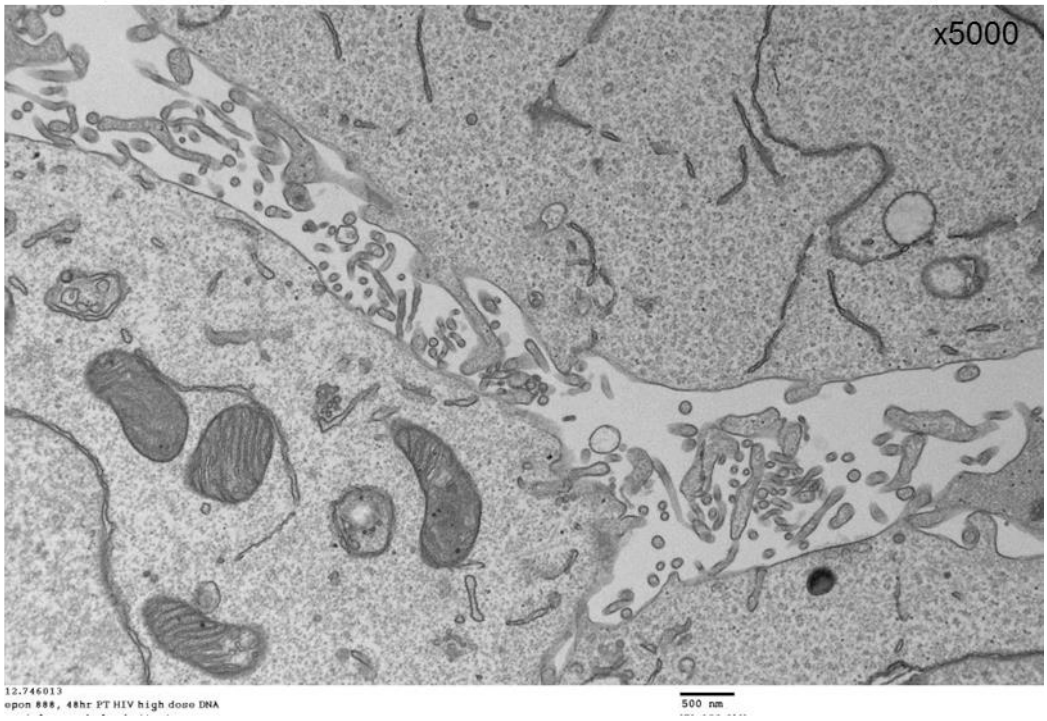
Condition #	**Volume Fugene (µl)**	**Total DNA (µg)**	**Plasmid amount (*ratio*)**		
			p8.9NSX+	**pMDG**	**pCSFLW**
1	6	1.1	300 ng	300 ng	500 ng
2	18	3	1 µg	1 µg	1 µg
3	6	1.1	480 ng (***2***)	120 ng (***1***)	360 ng (***3***)
4	6	1.1	600 ng (***2***)	150 ng (***1***)	300 ng (***2***)
5	18	3	1.44 µg (***2***)	360 ng (***1***)	1.08 µg (***3***)
6	none – PEI used	1.1	300 ng	300 ng	500 ng

Figure 3.10 (next page). Electron micrographs of 293T cells transfected using Fugene-6 reagent. 293T cells were plated on cover slips and co-transfected with p8.9NSX+, pMDG expressing VSV-g and pCSFLW, expressing luciferase reported gene. After 48 hours, cells were fixed, strained, sectioned and mounted viewing by electron microscopy. (a) An electron micrograph of transfected cells is shown, and pseudovirions budding from the surface cannot be seen. (b) The presence of microvesicles and membrane protrusions as a result of transfection is shown. Bar in (a) is 2 µm and (b) is 500 nm

a


x1200
2 µm

b


x5000
12.746013
epon 888, 48hr PT HIV high dose DNA
500 nm

Transfection efficiency was measured by a qualitative comparison of the number of cells expressing GFP and the levels of GFP expression using fluorescence microscopy. GFP expression was visualised 48-hours post-transfection using fluorescence microscopy and pseudovirions were harvested at this stage by filtration using a 0.45 μM filter to remove cellular debris. The pseudovirions were titrated and the levels of GFP expression for the top pseudovirus concentration (50 μl in 125 μl DMEM) and a two-fold dilution were visualised using fluorescence microscopy.

Representative microscopy pictures of transfected 293T cells taken 48-hours post-transfection for each set of transfection conditions are shown in figure 3.11. The effect of a higher volume of Fugene-6 and amount of total DNA are shown in transfection conditions 2 and 5, that displayed higher levels of GFP expression than the other transfection conditions. Transfection condition 6 utilised an alternative transfection reagent, PEI, which resulted in much lower GFP expression and thus transfection efficiency than those performed with Fugene-6. Plasmid ratios also affected GFP expression as comparison of conditions 1 (standard ratio: 300 ng p8.9NSX+, 500 ng pCSGW, 300 ng pMDG), 3 (3:2:1) and 4 (2:2:1) showed that the 2:2:1 plasmid ratio resulted in lower levels of GFP expression. Comparison of transfection conditions 2 (2:2:1) and 5 (3:2:1) demonstrated that the 3:2:1 plasmid ratio resulted in the highest levels of GFP expression. However, this may simply have occurred because less pCSGW was transfected and did not necessarily represent lower transfection efficiency and pseudovirion production. Finally, reduction of cell number as in transfection condition 7 resulted in markedly lower numbers of transfected cells and levels of GFP expression. Whilst visualizing GFP is a useful surrogate for transfection efficiency, it can overestimate the proportion of cells infected as pseudovirions can infect untransfected cells making them express GFP and hence appear transfected.

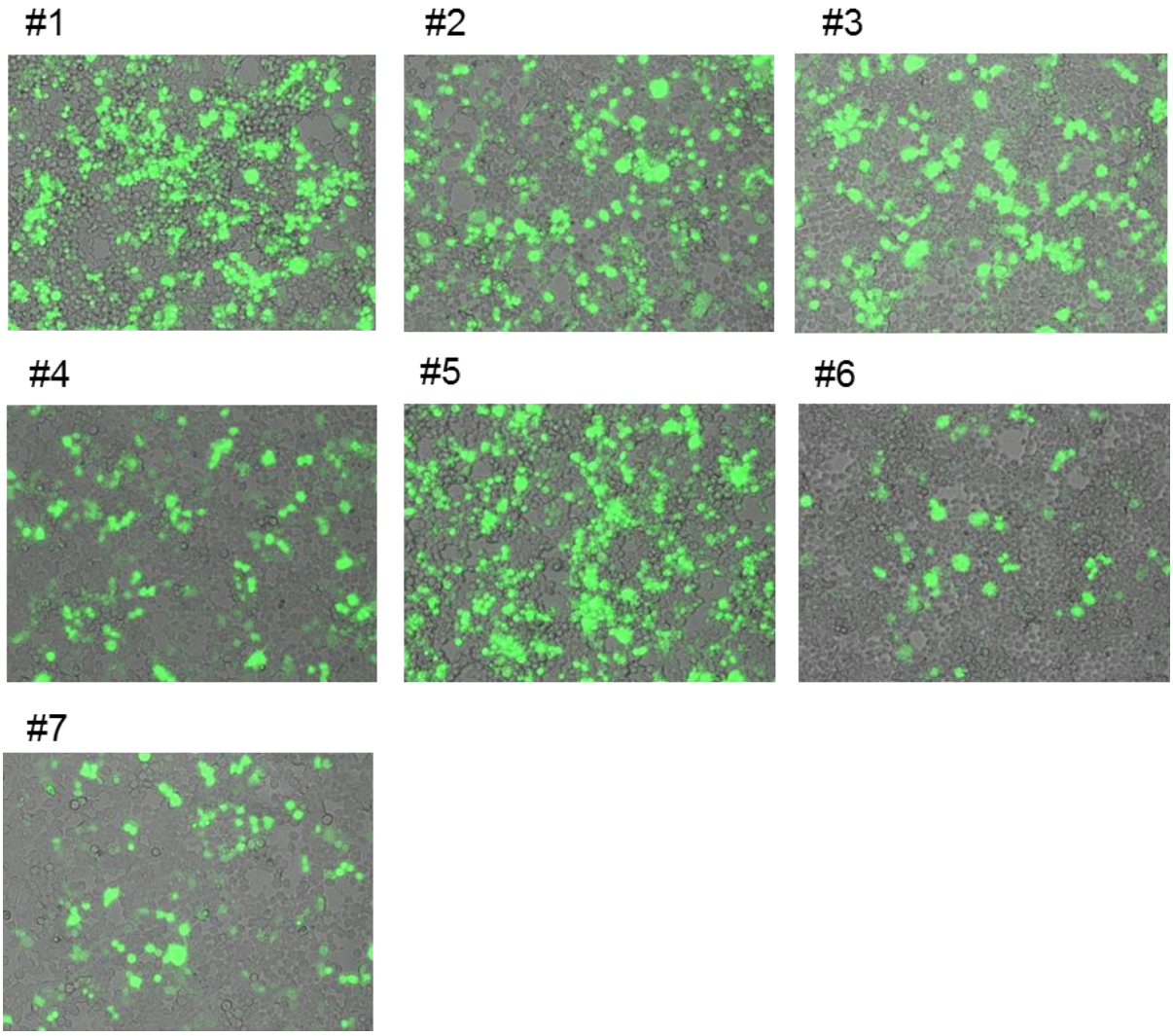
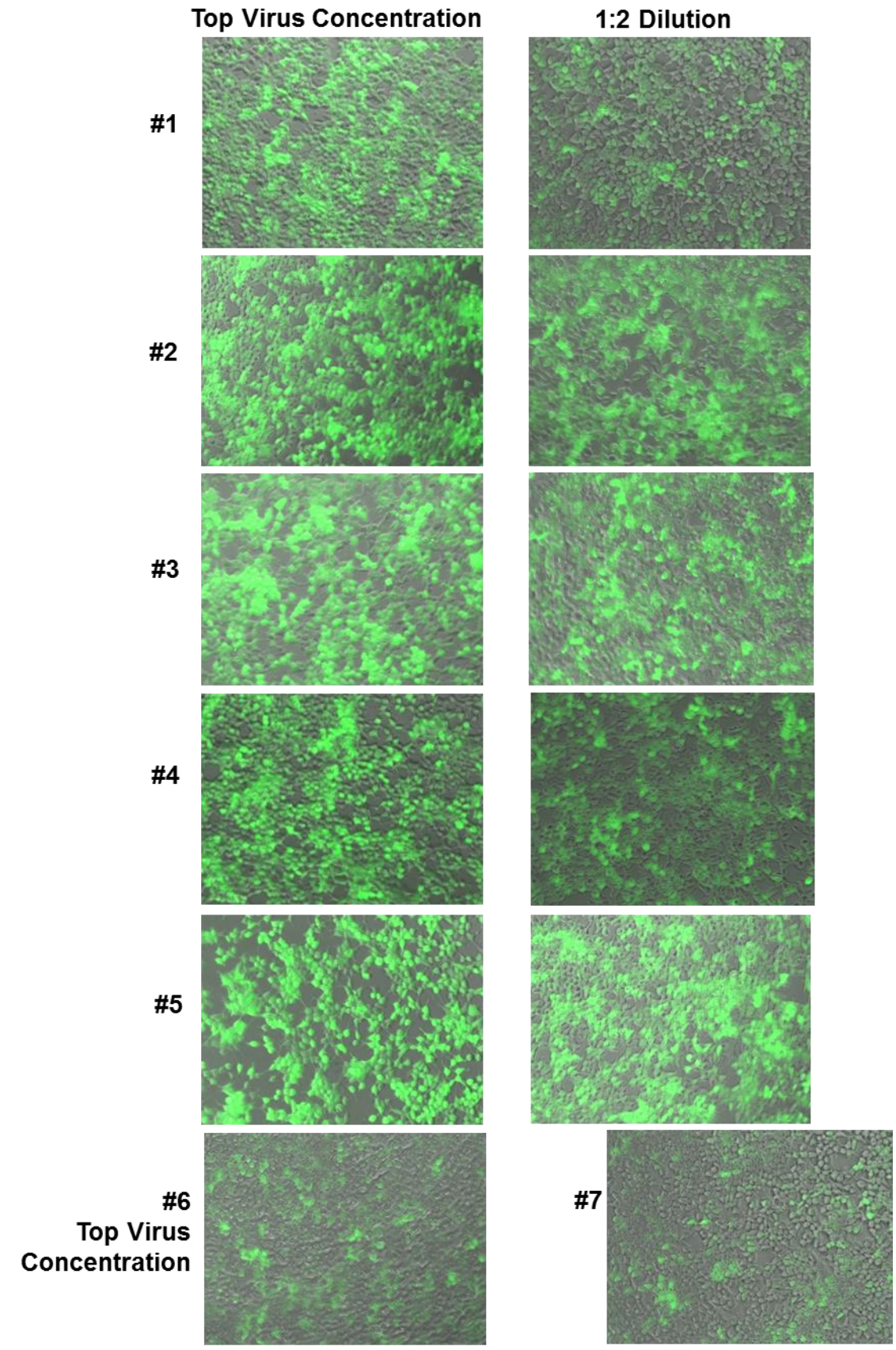


Figure 3.11. Transfection efficiency of different conditions, measured by GFP expression. 293T cells were co-transfected with p8.9NSX+, pCSGW and pMDG and GFP expression visualised 48-post transfection using fluorescence microscopy. Different volumes of the transfection reagent Fugene were tested (6 ul for 1, 3, 4 and 7 with 1.1 ug total DNA; 18 ul for 2 and 5 with 3 ug of total DNA), with PEI used for 7. In addition, different ratios of the three plasmids were tested as described fully in table 2.3. Pictures taken at x100 magnification.

Representative microscopy pictures taken following titration of pseudovirions are shown in figure 3.12, enabling an approximate quantification of the amount of pseudovirus production for each transfection condition. Levels of GFP expression and the number of cells expressing GFP at both the top pseudovirus concentration and the two-fold dilution are shown. At the top pseudovirus concentration, transfection conditions 1-5 appeared to result in similar levels of GFP expression, however conditions 6 and 7 resulted in far lower levels of GFP expression. At the two-fold dilution, GFP expression was lowest for conditions 1 and 3 and highest for 2 and 5. Overall transfection efficiency and virion production, as measured by GFP expression after titration, was optimal with 18 µl Fugene-6 and 3 µg total DNA. As a plasmid ratio of 3 p8.9NSX+: 2 pCSFLW : 1 pMDG resulted in the least amount of pMDG being used without compromising levels of pseudovirus production, these conditions were used for EM experiments.

Figure 3.12 (next page). Pseudovirus yield of different transfection conditions. Pseudovirions produced from different transfection conditions were harvested 48 hours post-transfection and infectivity determined by titration for each transfection condition, as detailed fully in table 2.3. GFP expression at the top concentration and 1:2 dilution are shown. Pictures were taken at x100 magnification. These data show that 18 ul of fugene with 3 ug was optimal (conditions 2 and 5), with a plasmid ratio of 3 pCSGW: 2 p8.9NSX+: 1 pMDG selected to minimise the amount of toxic pMDG used.


Top Virus Concentration
1:2 Dilution
#1
#2
#3
#4
#5
#6
Top Virus
Concentration
#7

3.2.8.2 Development of EM methods for visualisation of HIV-1 virions

EM methods for visualisation of HIV-1 pseudovirions were not established at PHE. We therefore developed these techniques in collaboration with the electron microscopist at PHE, Dr Matthew Hannah.

Previous reports have visualised HIV-1 virus and pseudovirus assembly and maturation by capturing the virions as they bud from the cells (Muller et al., 2009). Using sample preparation and EM techniques detailed in chapter 2, we visualised 293T cells 24- and 48-hours post-transfection. At both time points no pseudovirions were observed, although membrane protrusions and microvesicles were present, as shown in figure 3.12. IF confirmed the presence of Gag in approximately 30% of the cells, indicating that transfection efficiency was sufficient for virion production (figure 3.13). To accurately determine the proportion of mature pseudovirions, large numbers of pseudovirions would need to be counted and this preliminary experiment showed that this could not be achieved by sectioning of transfected cells.

However, pseudovirions should be present in large numbers in the supernatant of transfected cells following budding so we developed methods to perform EM on the cell media. Supernatants were collected 48-hours post-transfection, subjected to low-speed centrifugation to remove cellular debris and high-speed centrifugation to concentrate the pseudovirions into a pellet. Pellets were fixed and embedded in epoxy resin. Serial ultrathin section of the pellets were cut and mounted on slot grids before staining and visualisation on the electron microscope. The first experiment included only p8.9NSX+ derived pseudovirions produced in the presence and absence of the PI ATV and a mock transfection control. Electron micrographs in figure 3.14 showed the presence of many pseudovirions, both mature and immature. However, microvesicles of similar size to the pseudovirions were present in both the pseudovirion preparations and the mock transfected control, as shown in figure 3.14. Attempts to purify the pseudovirions away from the microvesicles using a

sucrose gradient were unsuccessful and further investigation revealed that the microvesicles co-purify with pseudovirions during ultracentrifugation (Greg Towers, personal communication). Thus, EM had to be carried out without purification of pseudovirions, but a stringent strategy was devised to prevent inclusion of the microvesicles in the final counts. The counting of immature virions was straightforward as they can be distinguished from mature virions and microvesicles by the uncleaved Gag polyrotein that is seen as a characteristic thick collar within the virion. However, mature virions can be similar morphologically to microvesicles, depending on the orientation at which the virion is sectioned. To avoid the accidental inclusion of microvesicles in the counts of mature pseudovirions, only virions with a conical capsid were included. This meant that mature pseudovirions sectioned in orientations that resulted in a circular or elliptical capsid were excluded, as shown in figure 3.15. Hence our mature pseudovirion counts were an underestimation, although still comparable between the molecular clones, as discussed in section 3.3 and shown in figure 3.20.

Using these counting criteria, approximately 60 virions were counted in the p8.9NSX+ preparation with and without ATV (figure 3.14) and a significant difference in the proportion of mature virions was observed between the pseudovirions produced in the absence of PI (25.4% immature and 74.6% mature) and those produced with ATV (51.6% immature and 48.4% mature). As a difference in the proportion of mature pseudovirions was detected between the pseudovirions produced in the presence and absence of PI, this method could be applied to the molecular clone based pseudovirions to enable a comparison of the effect of the Gag-protease derived from molecular clones on assembly and maturation of HIV-1 virions.

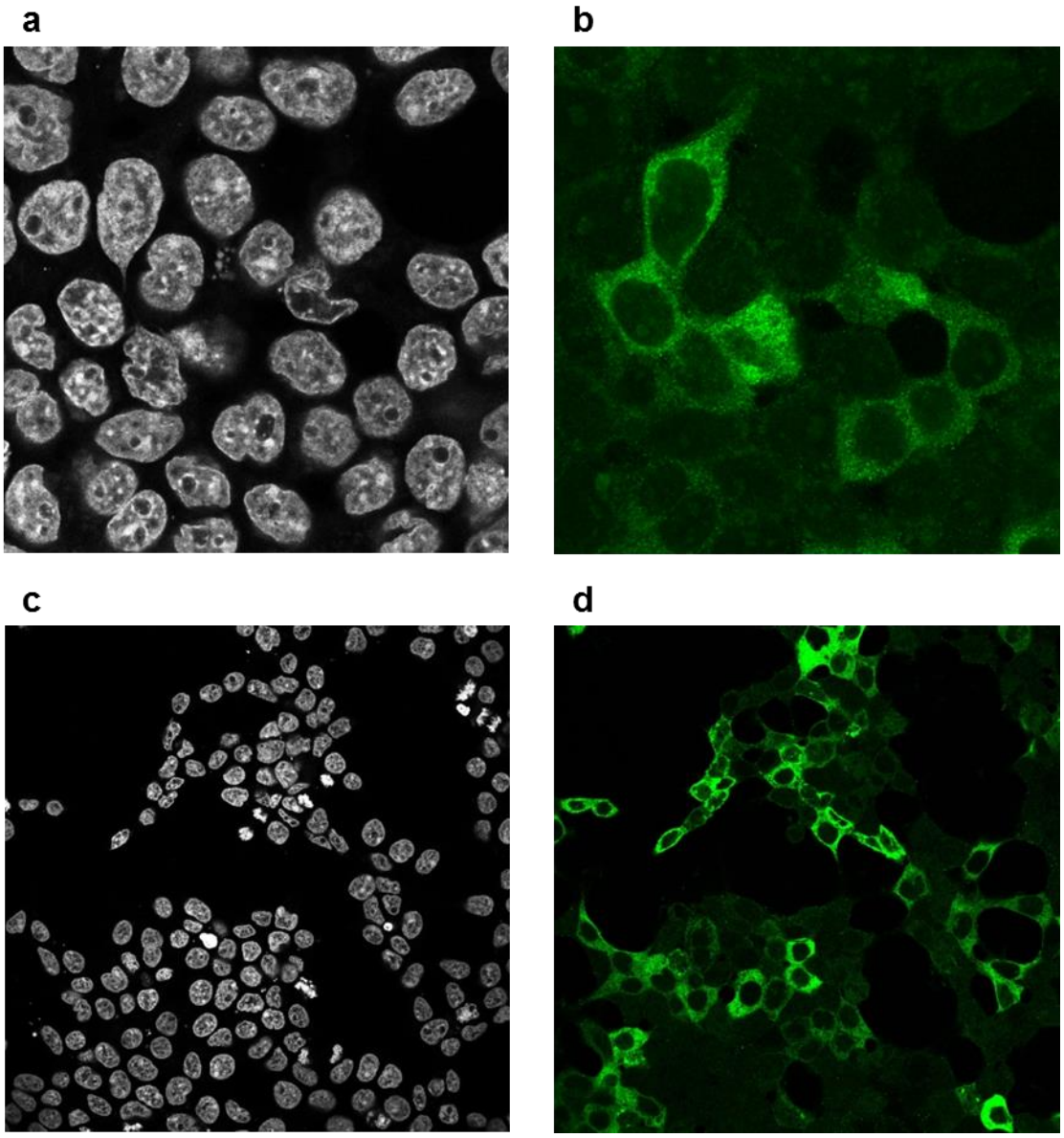


Figure 3.13. Immunofluorescence demonstrates successful transfection of 293T cells and Gag expression. 293T cells were plated on coverslips and co-transfected with p8.9NSX+, pMDG and pCSFLW. Cells were fixed 48-hours post transfection and IF staining for Gag protein was carried out, using a monoclonal antibody targeting Gag CA (p24) #4121 (section 2.4.1). Representative pictures of cells using light (a and c) and fluorescence (b and d) microscopy show expression of Gag in a significant proportion of cells**.**

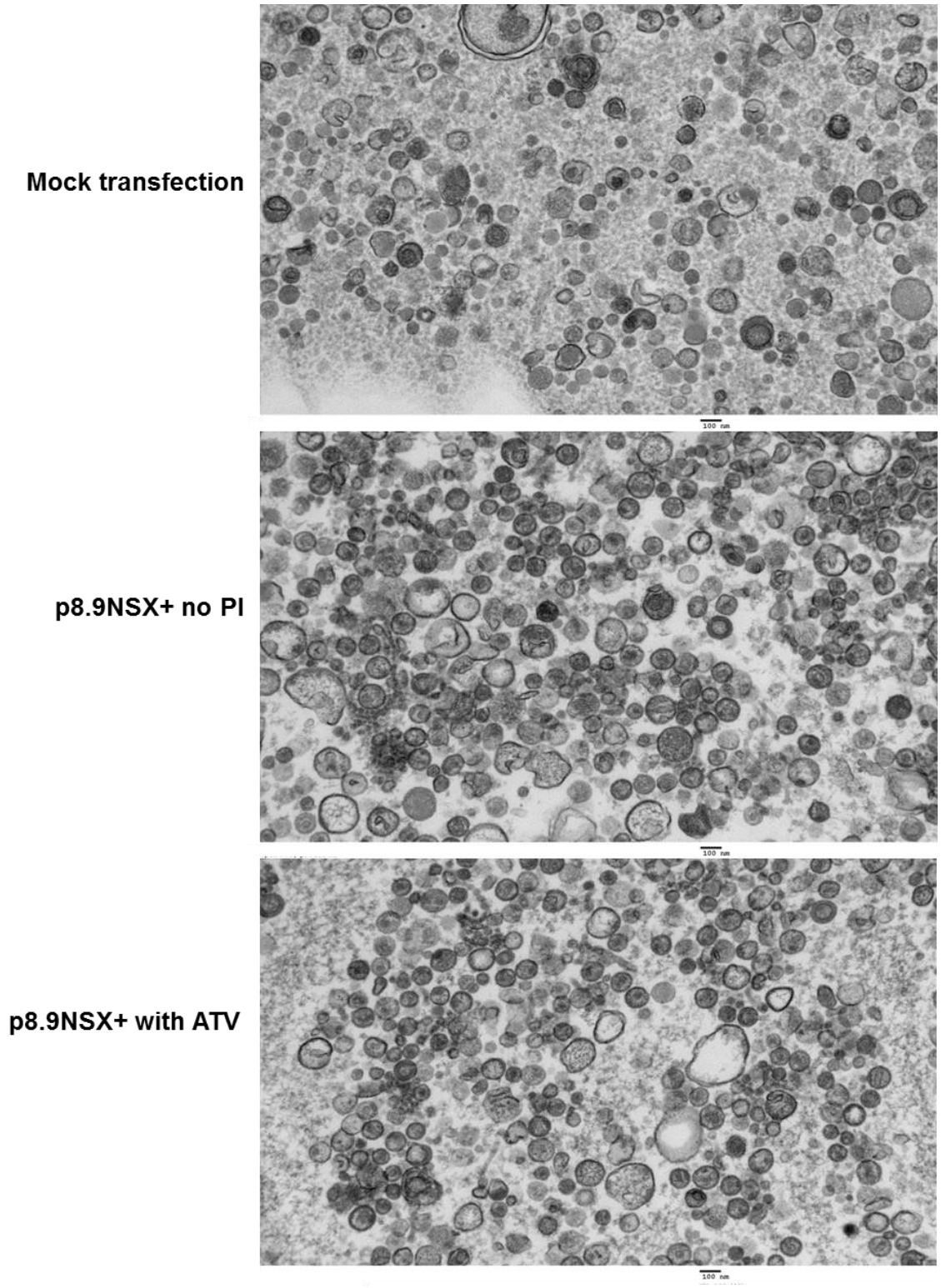


Figure 3.14: Electron micrographs of pseudovirions. 293T cells were co-transfected and pseudovirions harvested from supernatants and prepared for electron microscopy. Electron micrographs of supernatants from (a) mock transfection, (b) p8.9NSX+ pseudovirions produced without PI and (c) p8.9NSX+ with drug are shown. All at the same magnification x12,000, the bar represents 500 µm.

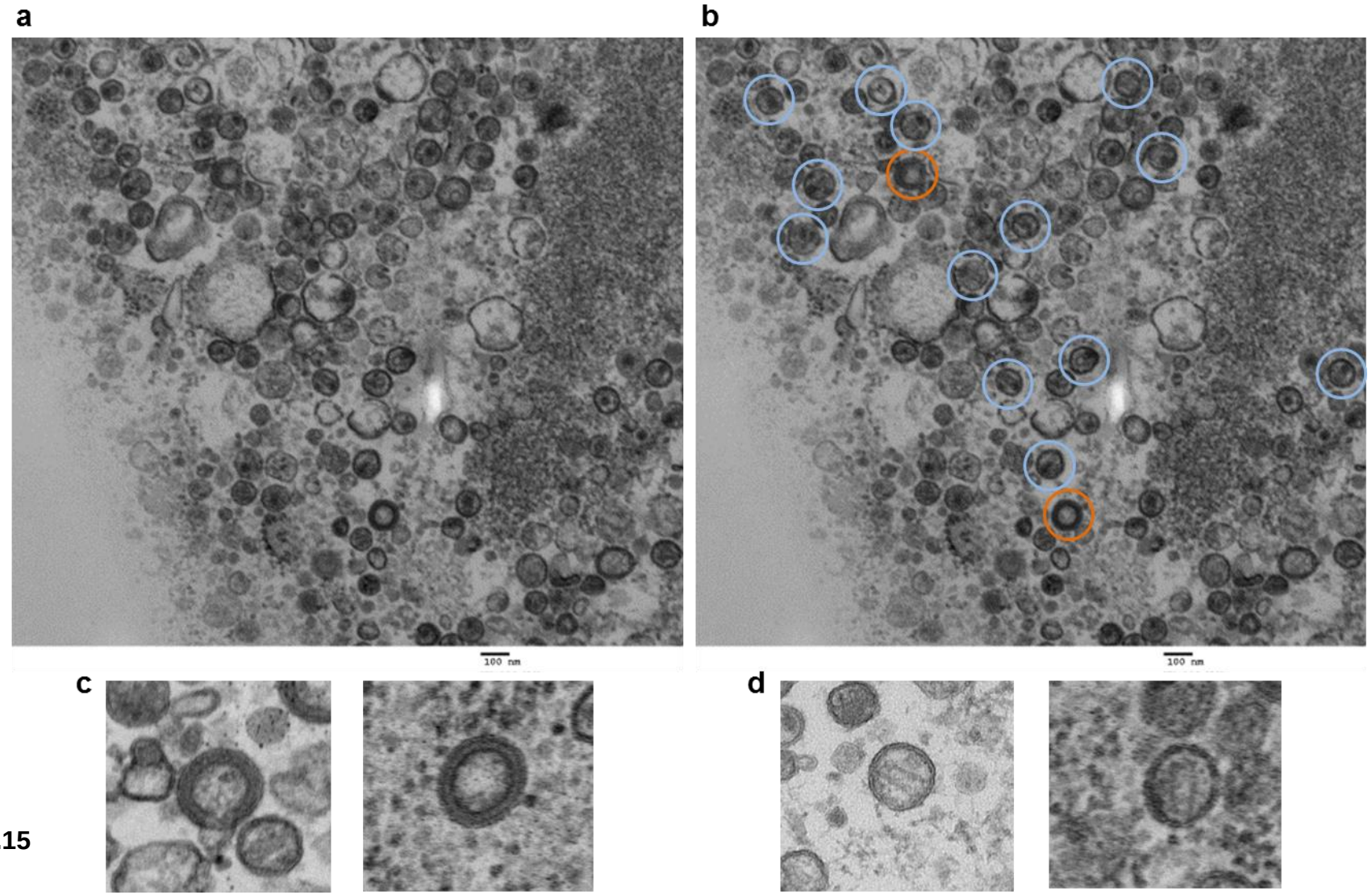


Figure 3.15

Figure 3.15. Counting strategy utilised for electron microscopy experiments. Following preliminary experiments demonstrating the presence of microvesicles of similar morphology to virions, a counting strategy was devised to avoid their accidental inclusion in the counts of the proportion of immature and mature pseuovirions. (a) A typical EM of pseudovirions at x15,000 magnification is shown and (b) the counting strategy employed demonstrated (mature pseudovirions counted highlighted in blue and immature in orange). Immature (c) and mature pseudovirions (d) are shown at higher magnification 25,000 x. Only mature pseudovirions with a clear conical-shaped capsid were included in the count.

3.2.8.3 Comparison of the proportion of mature pseudovirions in the presence of ATV for vectors p8.9NSX+, p8YU2gagpro, p8JRFLgagpro

Having shown reduced susceptibility of p8YU2gagpro and p8JRFLgagpro in comparison to p8.9NSX+, the mechanisms by which the reduced susceptibility was conveyed were investigated. Since PIs block Gag cleavage and virus maturation, we sought to use the EM techniques discussed above to investigate whether the presence of the EC_{90} of ATV for p8.9YU2gagpro (25 nM) resulted in different proportions of mature pseudovirions derived from vectors p8.9NSX+, p8YU2gagpro and p9JRFLgagpro. Transfections were performed in duplicate for each vector and 18 hours post-transfection, ATV was added to the media at 25 nM, the EC_{90} as determined by the PI susceptibility assays. 48 hours post-transfection, pseudovirions were processed for EM as previously described. Pseudovirions were counted on sections taken from two different parts of each pseudovirus pellet to reduce the possibility of an effect of the position within a pellet on the proportions of mature pseudovirions present. In addition, the experiment was performed in duplicate and the virus pellet created by centrifugation embedded in a different orientation and thus sectioned differently for each of the duplicate transfections, to further reduce the possibility of bias. 400 virions in total were counted for each virus with and without drug, 200 from different sections of the pellet.

Figure 3.16 shows the proportion of mature virions in each sample and representative electron micrographs for the assay reference strain and each molecular clone are shown in figures 3.17, 3.18 and 3.19. In the absence of ATV the proportion of mature pseudovirions was similar for each of the three molecular clones based vectors p8.9NSX+, p8YU2gagpro and p9JRFLgagpro, of between 83.9 and 87.7%. For p8.9NSX+ and p8JRFLgagpro, addition of ATV significantly reduced the proportion of mature pseudovirions to 47% and 52%, respectively ($P < 0.05$). However, addition of ATV did not significantly affect the proportion of mature pseudovirions for p8YU2gagpro, 80%.

These data demonstrate that at an ATV concentration of 25 nM, the maturation of pseudovirions derived from p8.9NSX+ and p8YU2gagpro is significantly inhibited, but this has no effect on the maturation of p8JRFLgagpro-derived pseudovirions.

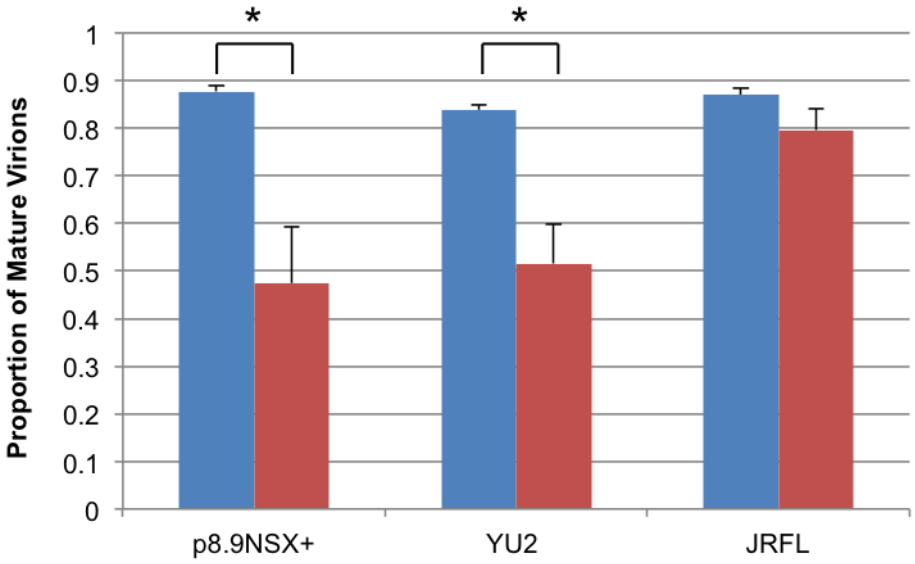


Figure 3.16. Presence of ATV affects the proportion of mature pseudovirions present for YU2, but not JRFL. Pseudovirions were produced in the presence and absence of ATV by co-transfection of 293T cells. Pseudovirions were harvested from the cell media, then fixed, stained and embedded for EM. Data are results of two independent experiments, with each performed in duplicate, to count 200 virions as either mature or immature (in total these data represent the proportion of 800 virions that were mature, counted from four separate EM sections, derived from two transfections). The proportion of mature pseudovirions in the absence (blue bars) and presence (red bars) of ATV is shown. Error bars are derived from the standard error of the mean of the two independent repeats. Molecular clones for which the proportion of mature virions was statistically different in the presence of absence of drug, as measured by an unpaired t-test ($p < 0.05$), are marked by an asterisk(*).

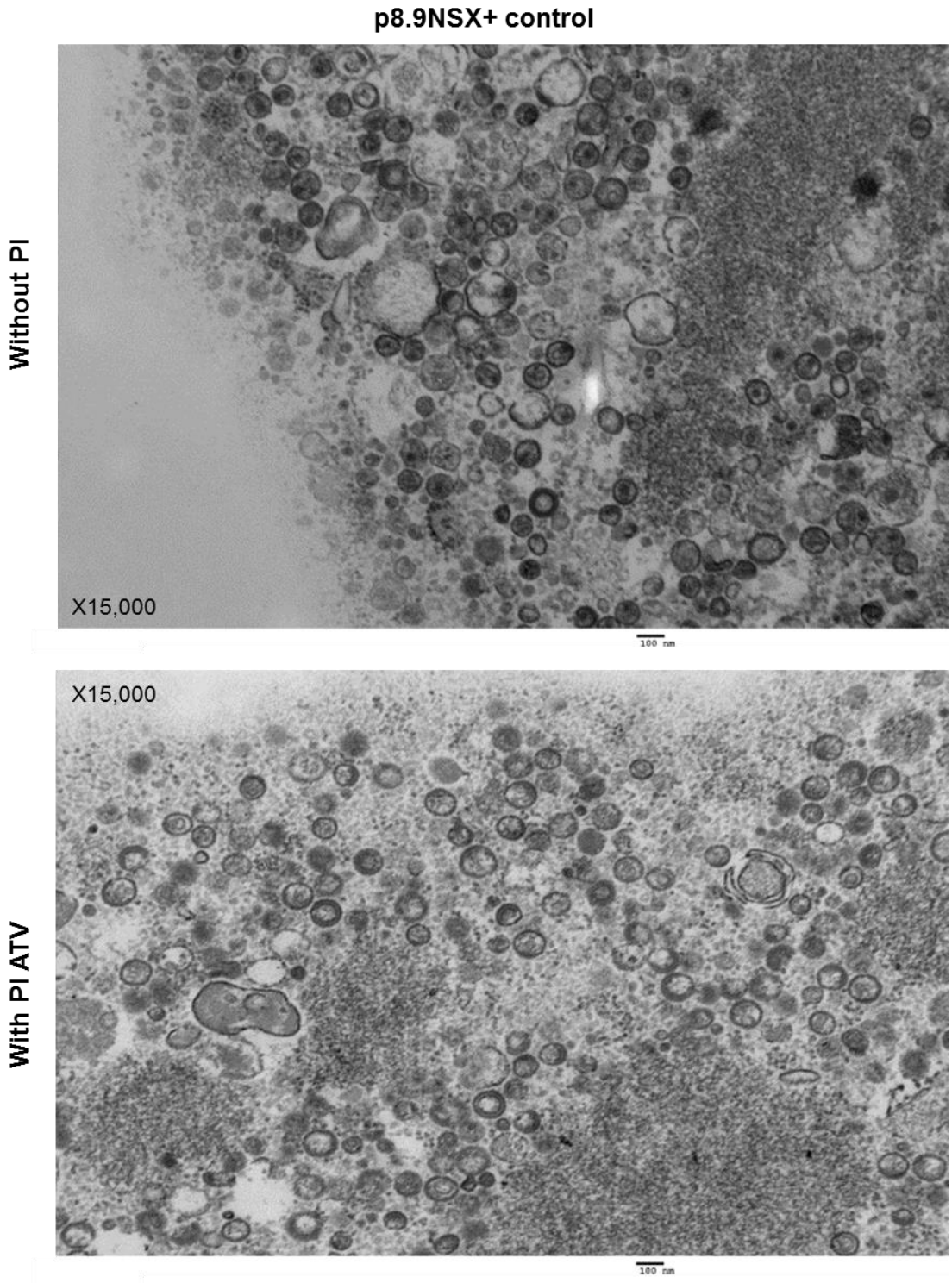


Figure 3.17. Electron micrographs of p8.9NSX+ pseudovirions assembled in the presence and absence of the PI ATV. Pseudovirions were produced in 293T cells in the presence and absence of the PI ATV. Pseudovirions were harvested from the media and fixed, stained and sectioned for electron microscopy. Both electron micrographs are at 15,000x magnificication and bars represent 100 nm.

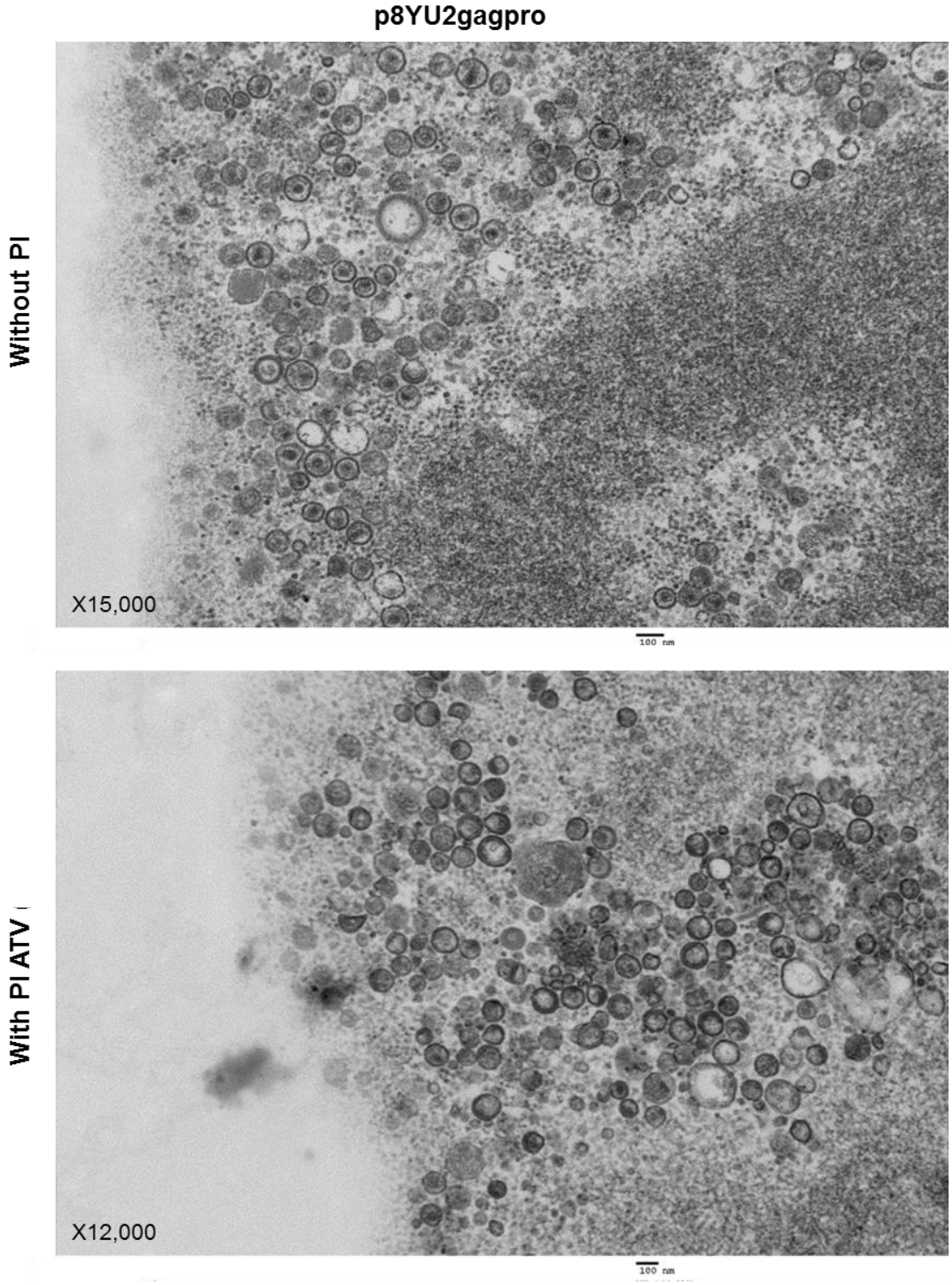


Figure 3.18. Electron micrographs of p8YU2gagpro pseudovirions assembled in the presence and absence of the PI ATV. Pseudovirions were produced by co-transfection of 293T cells in the presence and absence of the PI ATV. Pseudovirions were harvested from the media and fixed, stained and sectioned for electron microscopy. Both electron micrographs are at 15,000x magnificication and bars represent 100 nm.

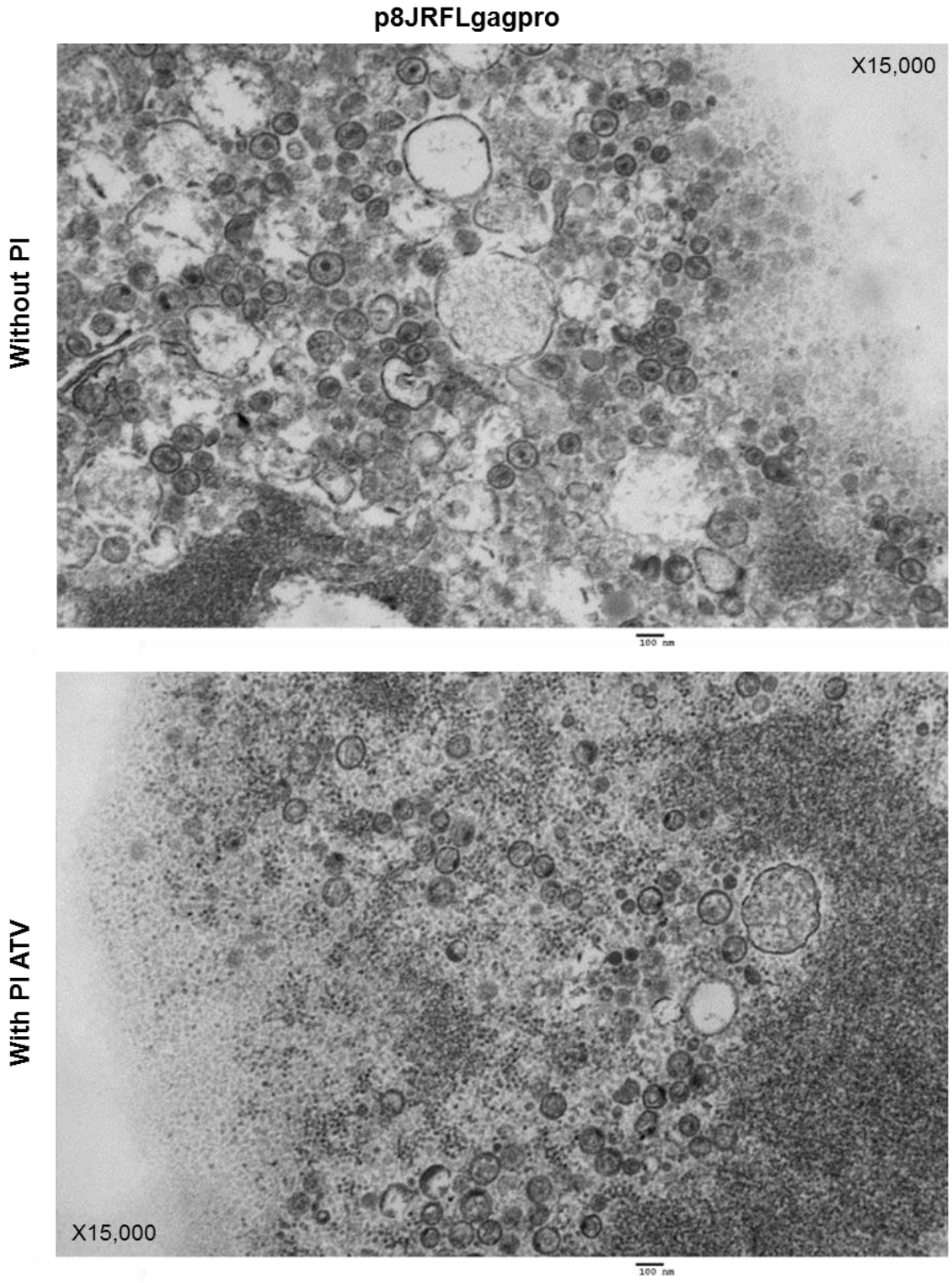


Figure 3.19. Electron micrographs of p8JRFLgagpro pseudovirions in the assembled in the presence and absence of the PI ATV. Pseudovirions were produced by co-transfection of 293T cells in the presence and absence of the PI ATV. Pseudovirions were harvested from the media and prepared for electron microscopy. Both electron micrographs are at 15,000x magnificication and bars represent 100 nm.

3.3 Discussion

Here we demonstrate variation in PI susceptibility in PI-naïve subtype B viruses when full-length co-evolved Gag-protease is included in a phenotypic PI susceptibility assay. We show reduced PI susceptibility of four widely used molecular clones YU2, SF2, JFRL and 89.6, to a number of PIs of up to 9-fold in comparison to assay reference strain. To our knowledge this is the first demonstration of variation in PI susceptibility of full-length Gag-protease from wild-type subtype B HIV-1 viruses. However, this is in keeping with another study that has demonstrated variation in PI susceptibility of full-length Gag-protease from PI-naïve subtype A and C viruses (Gupta et al., 2010).

Generation of chimeric pseudovirions provided evidence that for YU2, SF2, JFRL and 89.6, Gag independently contributes to the reduced PI susceptibility observed for full-length Gag-protease, at least in part. For two molecular clones, JRFL and YU2, the reduced susceptibility can be attributed solely to the *gag* gene (the region of Gag up to the *Apa*I cleavage site, encompassing MA, CA, p2 and part of NC). This is in agreement with other studies which have shown the role of Gag in PI susceptibility for subtype A and C molecular clones (Gupta *et al.*, 2010), subtype A and CRF01_AE viruses from treatment- naïve HIV-1 infected patients (Gupta et al., 2010; Jinnopat et al., 2009), subtype B from treatment-experienced HIV-1 infected patients (Dam et al., 2009; Parry et al., 2009) and from PI resistant viruses created by *in vitro* passage in the presence of PI (Nijhuis et al., 2007b). Current commercial resistance testing methods, both phenotypic and genotypic, include only the protease in their analysis alongside the 3' end of Gag that overlaps with the protease ORF. Therefore, these methods would not have captured this observed reduction in PI susceptibility.

Two of the six molecular clones selected for this study, HXB2 and NL4-3, did not display significant reductions in susceptibility to a number of PIs in comparison to the assay reference strain. The assay reference strain,

p8.9NSX+, is derived from HXB2 upstream of the *Spe*I site in Gag and from NL4-3 downstream of the *Spe*I site onwards. Hence, the little variation in susceptibility observed for these molecular clones was expected. However, both HXB2 and NL4-3 displayed reduced susceptibility in comparison to the reference strain to APV. This further emphasises the importance of considering full-length Gag alongside its cognate protease when investigating PI susceptibility as in the absence of the co-evolved Gag the susceptibility of both HXB2 and NL4-3 was reduced to APV. This indicates an effect of separating co-evolved Gag and protease even in closely related viruses.

The reason for the different roles of Gag and protease in PI susceptibility between each of the molecular clones remains unclear. JRFL was subjected to limited cell culture during its generation whereas YU2 was not subjected to any tissue culture propagation; however, the remaining molecular clones were all subjected to more extensive cell culture during their derivation (Koyanagi et al., 1987; Li et al., 1991). It is therefore possible that tissue culture adaptation resulted in changes in Gag in the other molecular clones which are not present in JRFL and YU2, explaining the difference in susceptibility between these two molecular clones and the other four. However, molecular clones YU2 and JRFL, in which reduced susceptibility are both conferred solely by Gag, were both derived directly from brain tissue which is not the case for the other molecular clones. Thus, an alternative explanation could be that adaptation to replication in the central nervous system resulted in changes in Gag that affect PI susceptibility.

The exclusion of the NC/p1/p6 cleavage sites from the Gag only chimeric virus further emphasises the role of other regions within Gag in PI susceptibility, as mutations in the NC/p1/p6 that affect PI susceptibility have already been described (Dam et al., 2009; Nijhuis et al., 2007b). The creation of further Gag chimeric viruses demonstrated that the N terminal end of Gag (the MCgagN vector, encompassing MA and the N terminus of CA), and not the C terminus,

conferred the reduced susceptibility observed for full-length Gag-protease in molecular clones JRFL and YU2. To date, two other studies have implicated the N terminus of Gag in contributing to reduced susceptibility in an extensively treated patient infected with HIV-1 subtype B and a treatment-naïve patient infected with subtype CRF01_AE HIV-1 (Jinnopat et al., 2009; Parry et al., 2009).

Amino acid alignment identified only two amino acids changes that were present in both JRFL and YU2, the molecular clones in which Gag solely conferred the reduced PI susceptibility observed for full-length Gag-protease. Use of site-directed mutagenesis provided evidence for the involvement of amino acid positions 30 and 102 in PI susceptibility, both in the context of co-evolved Gag-protease and with a reference strain Gag C-terminal and protease. Amino acid changes at positions 30 and 102 of Gag have not been previously reported to affect susceptibility to PIs. However, these positions do not solely confer the observed reduction in susceptibility indicating that additional unique changes identified in the N terminus of Gag of YU2 or JRFL must also independently or in combination contribute to the observed reduction in PI susceptibility.

The exact mechanism by which changes at position 30 and 102 of Gag may affect PI susceptibility remains unknown. However, positions surrounding amino acid 30 have been implicated in correct targeting of Gag to the plasma membrane, so it is possible that an increase in efficiency of Gag targeting may lead to an increase in the number of virions produced by JRFL and YU2 in comparison to other molecular clones (Ono and Freed, 2004). Of interest, position 30 of MA has been described as a key position for the switch in species specificity during evolution of SIV to HIV. In SIV viruses from chimpanzees and gorillas, position 30 is M, but in HIV viruses R or K is encoded and this is fairly well conserved (Wain et al., 2007). Amino acid changes R30K and E102D are

not uncommon in PI-naïve viruses, with Lysine (K) at position 30 in 27.9% of group M viruses and Glutamic acid (E) at 102 in 17% (http://www.hiv.lanl.gov/).

This study also included the development and use of EM methods to investigate the mechanisms by which the change in PI susceptibility between p8YU2gagpro, p8JRFLgagpro and p8.9NSX+ was conferred. Preliminary experiments indicate that the maturation of JRFLgagpro-derived pseudovirions was unaffected by relatively high ATV concentrations (EC_{90}). Conversely, p8.9NSX+ and YU2gagpro-derived pseudovirions displayed reduced levels of maturation and hence an increase in the proportion of immature pseduovirions generated in the presence of ATV. Although the infectivity of p8.9NSX+ and p8YU2gagpro are reduced by approximately 90% at this particular concentration of ATV, 50% of pseudovirions counted were mature and infectious, at least morphologically. Our data are perhaps not surprising, given that another study has shown that the morphology and cleavage of Gag do not account for differences in the infectivity of viruses produced in the presence of clinically suboptimal LPV concentrations (Muller et al., 2009).

For JRFL Gag-protease, these data provide evidence that changes present in the N terminus of Gag enable efficient cleavage of Gag and pseudovirion maturation in the presence of ATV. However, the mechanism by which these changes enable efficient Gag cleavage and maturation remains to be elucidated. Conversely, for YU2 Gag-protease no difference in maturation efficiency in comparison to that of p8.9NSX+ reference strain was observed. Previous studies have shown that PIs act at multiple life cycle stages including cell entry, uncoating and nuclear entry so it is possible that changes in the N terminus of Gag mitigate the effect of PI at one or more of these stages (Muller et al., 2009; Rabi et al., 2013). This study provides further evidence that the presence of PI affects other stages of the life cycle, in addition to the well-established effect on Gag cleavage and virion maturation.

The EM method employed here had a number of limitations, including the counting strategy utilised to avoid the counting of microvesicles as pseudovirions. The counting strategy will have excluded mature virions that were sectioned in orientations other than that which resulted in bullet shaped capsids, as illustrated in figure 3.20. It is also possible that defective and aberrant pseudovirions would not have been included in the count. This could have resulted in an underestimation of the effect of PI on maturation, as any block that did not result in an immature virion would not be captured. However, it was necessary to exclude these from the counting to avoid inclusion of microvesicles in the pseudovirion counts. In the future, CD45 depletion could be used to purify the pseudovirions and this counting strategy would not be required (Ott, 2009). Additional controls would be required if this work was to be continued, in particular a count of the proportion of pseudovirions produced in the absence of protease and with an extensively PI resistant protease. In the future, virus maturation could be measured more effectively using other methods such as western blotting, as discussed in section 7.3.1.

In conclusion, this study provides evidence of variation in PI susceptibility of subtype B molecular clones when full-length co-evolved Gag-protease is included in phenotypic assays. We show that Gag can directly affect PI susceptibility via mechanisms independent of co-evolved protease and that the N terminus of Gag, including amino acid positions in MA, are involved. Our data provide evidence that efficient Gag cleavage in the presence of PI contributes in part to the variation in PI susceptibility, but that other stage(s) of the virus life cycle may also be involved in the observed variation in PI susceptibility.

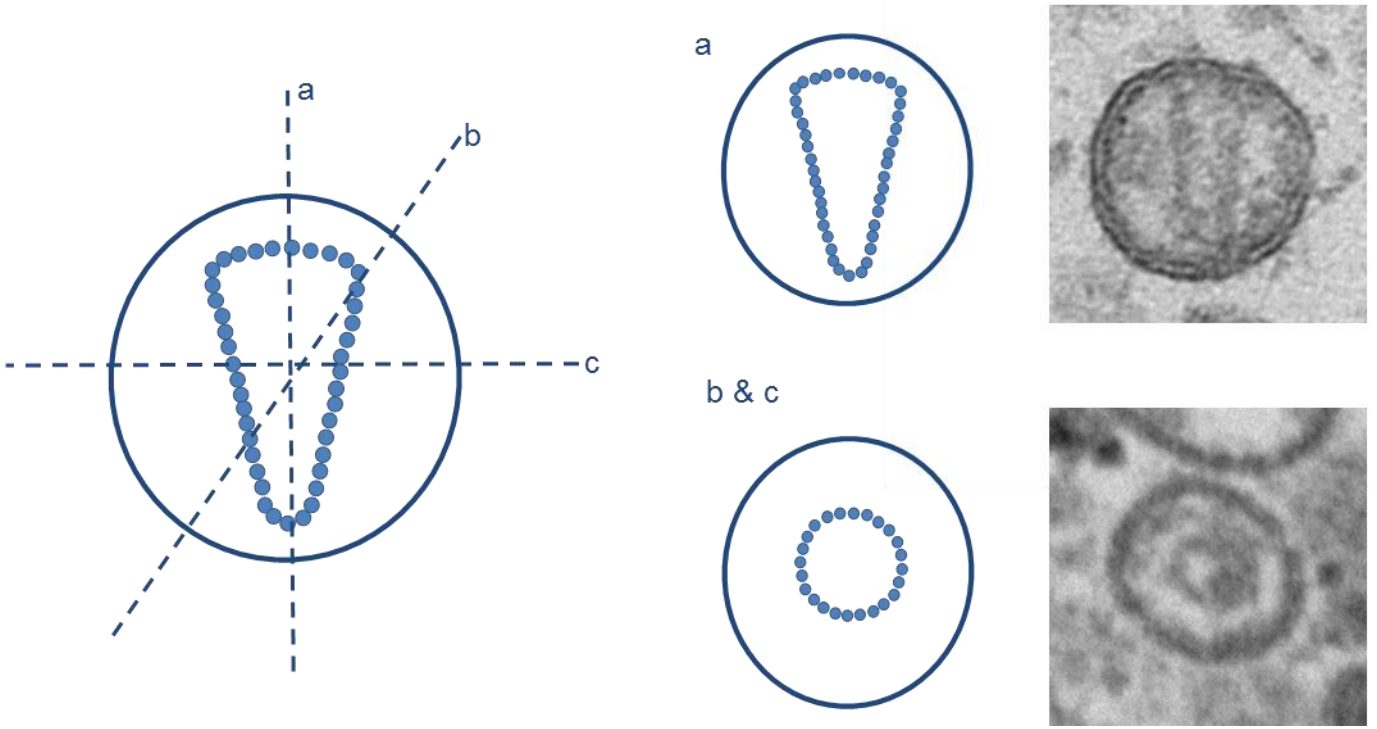


Figure 3.20. Limitations of the counting strategy utilised for electron microscopy experiments. Following preliminary experiments, a counting strategy was devised to avoid accidental inclusion of microvesicles. The schematics and electron micrographs highlight that only pseudovirions sectioned in the orientation ‘a’ would be captured by this counting strategy, and those sectioned in orientations ‘b’ or ‘c’ would be missed.

4 Investigation of the role of co-evolved Gag and protease in susceptibility to PIs of different HIV-1 subtype B viruses derived from treatment-naïve patients

4.1 Introduction

In chapter 3, results were presented that showed significant variation in PI susceptibility of full-length Gag-protease from subtype B molecular clones. This chapter describes an extension of this work examining viruses derived directly from patients. The generation of molecular clones can include extensive *in vitro* passage, and it has been shown that tissue culture adaptation of HIV-1 viruses can occur in short periods of time (Collman et al., 1992; Hahn et al., 1984; van Opijnen et al., 2007). This could affect Gag in particular, as it contains a large number of CTL epitopes and these have been shown to vary *in vivo* in the absence of the relevant HLA type, a condition that is replicated in tissue culture conditions where no immune system is present (Smith, 2004). For these reasons an assessment of the variation in PI susceptibility of full-length Gag-protease directly derived from patients was required to confirm direct clinical relevance of our previous findings.

To our knowledge, only two studies to date have examined the PI susceptibility of full-length Gag-protease from subtype B HIV-1 and in both studies the viruses were from extensively treated patients (Dam et al., 2009; Parry et al., 2009). The findings reported by Parry et al. (2009) were in keeping with those here for molecular clones, with Gag shown to confer up to 10-fold reduction in susceptibility to APV and ATV. As with our data, this reduction in susceptibility mapped to the N terminus of Gag. A second study has examined the PI susceptibility of full-length Gag-protease from PI-naïve viruses, although these were subtype A and C (Gupta *et al.*, 2010). This study included both molecular

clones of subtype A and C, and subtype A viruses derived directly from infected patients. No difference in the role of Gag in PI susceptibility between molecular clones and viruses derived directly from patients was reported. For each subtype A virus, inclusion of full-length Gag significantly affected the susceptibility of protease with up to 24-fold difference in susceptibility observed. In addition, for each virus Gag independently conferred reduced susceptibility of up to 8-fold to LPV in comparison to assay reference strain in the absence of its cognate protease. Given that subtype B is the most prevalent subtype in the UK, accounting for ~40% of infections and that genotypic resistance testing is routinely carried out at diagnosis in UK before initiation of treatment, an assessment of the variability in PI susceptibility conferred by full-length Gag-protease derived from treatment-naïve patients was of considerable interest (Dolling et al. 2013).

Chapter aims:

1) Adaption of the current single replication-cycle assay vector, p8.9NSX+, to enable efficient cloning of patient-derived full-length Gag-protease
2) Investigate intra-patient variability in Gag-protease to ensure adequate representation of the patient quasispecies in the single replication-cycle assay
3) Assess the variability in PI susceptibility of full-length Gag-protease in subtype B viruses derived directly from treatment-naïve, HIV infected patients
4) Identify and investigate the regions and/or specific amino acid positions conferring variation in PI susceptibility between viruses

4.2 Results

4.2.1 Modification of p8.9NSX+ vector for use with viruses from clinical samples

In order to investigate the variation in PI susceptibility of full-length Gag-protease derived directly from viruses from HIV infected individuals, modifications to the existing test vector p8.9NSX+ were required. The method used to clone into p8.9NSX+ from extracted viral RNA utilises a pGEM intermediate vector and requires screening of clones by DNA sequencing at two stages – once after the cloning into pGEM and again after cloning into p8.9NSX+. Colony PCR to check for the presence of a *gag*-protease fragment after sub-cloning into p8.9NSX+ is not sufficient as re-ligated p8.9NSX+ vector also gives a positive PCR result. In order to shorten the cloning protocol, the requirement for a second sequencing step was removed by the generation of a vector that did not yield a positive colony PCR when religated, using primers GagNot+ and Gag1.5R. A vector designated p8.9Δpro was created in which a fragment of *gag*-protease had been removed which included the Gag1.5R primer binding site, as shown in figure 4.1.

The creation of the p8.9Δpro vector was achieved by the introduction of an additional *Spe*I site flanking amino acids 75 and 76 of protease by site-directed mutagenesis with primers MutProSpe and MutProSpeR (see table 2.2). *Spe*I restriction digest was carried out to remove the fragment between the new (protease amino acids 75-76) and existing (Gag amino acids 240-241) *Spe*I sites, and the vector backbone was then re-ligated, as detailed in chapter 2. This resulted in the removal of the fragment between amino acids 241 of Gag and 75 of protease, meaning both the C terminal portion of *gag* and the majority of the protease gene were deleted (~1kb fragment). Plasmid maps of p8.9NSX+ and p8.9Δpro are shown in figure 4.1 and the restriction sites of interest present in p8.9NSX+ are detailed in table 4.1.

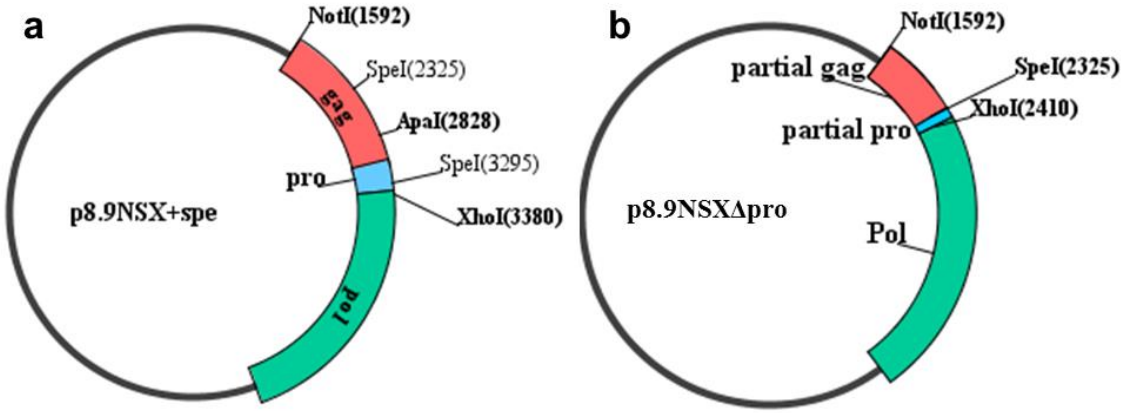


Figure 4.1. Modifications to p8.9NSX+ vector and creation of p8.9NSXΔpro
In order to increase the efficiency of our cloning protocol, a p8.9-based vector that did not produce a positive result in our colony PCR reaction when religated without the insert of interest was designed. (a) Existing restriction sites used for cloning in this study are shown, with the NotI and XhoI used for cloning in of full-length Gag-protease being of particular interest. The *Spe*I site introduced during this study at position 3295 of the vector is shown. The vector created following removal of the fragment encompassing part of Gag-protease by *Spe*I digest and re-ligation, named p8.9NSXΔpro, is shown (b).

Restriction Site	**Nucleotide position in vector**	**Amino acid position**
*Not*I (engineered)	1592	13 nt before Gag ORF
*Spe*I (naturally present)	2325	240/241 of Gag
*Apa*I (naturally present)	2828	406/407 of Gag
*Spe*I (engineered)	3295	75-76 of protease
*Xho*I (engineered)	3380	12 nt after end of protease ORF

Table 4.1. Unique restriction sites present in p8.9NSX+ utilised for cloning.
Restriction sites used for cloning of Gag and protease fragments derived from patients, for example *Not*I and *Xho*I were used for full-length Gag-protease fragments. The second *Spe*I site was added to enable the creation of the p8.9Δpro.

4.2.2 Clonal analysis of Gag-protease from four treatment-naïve patients, infected with subtype B HIV-1

Before an assessment of the variability of PI susceptibility of full-length Gag-protease subtype B viruses derived from treatment-naïve patients could be carried out, it was necessary to determine a method by which the quasispecies within a patient could be represented in the phenotypic assay. Whilst it would have been possible to test the phenotypic susceptibility of a large number of variants in a pool, this method was discounted as it would complicate further investigation of specific regions or changes contributing to any observed changes in susceptibility. Thus, it was decided that just a small selection of Gag-protease variants should be used to represent the viruses derived from each patient in the phenotypic susceptibility assay.

Viral RNA was manually extracted from patient plasma samples and cDNA synthesis using reverse primer ProOutR was carried out. Nested PCR to amplify full-length *gag*-protease was performed and PCR product cloned into pGEM, as described in chapter 2. Clonal analysis was carried out; colonies positive by colony PCR were cultured, the plasmid DNA purified and the *gag*-protease fragment sequenced. Sequences obtained that did not cover the full-length *gag* -protease fragment were discarded, but those containing frameshift PCR errors were included in the analysis after manual correction of the PCR error. Alignment of these clonal sequences enabled the construction of a consensus sequence for full-length *gag*-protease for each patient. Gag amino acid sequences were compared to a table of all mutations in Gag previously reported to be associated with PI resistance or exposure (table 1.2) and protease sequences were submitted to the Stanford Resistance Database to determine whether any previously reported PI resistance mutations were present.

An assessment of the variability of Gag-protease within the quasispecies of a patient was carried out for patient 1. We aimed to sequence 33 viral variants as this is the number that will result in the capture of any variants present at a

frequency greater than 10% with 95% confidence, according to Poisson distribution (Palmer et al., 2005). However, the efficiency of the cloning process, in particular fragment ligation and transformation into *E. coli*, limited the total number of variants to 27 that could be sequenced for patient 1. Alignment of these clonal sequences enabled the construction of a consensus sequence for full-length Gag-protease, as shown in figure 4.2 for Gag and figure 4.3 for protease. Of note, the consensus sequence from patient 1 contained a number of changes in Gag previously associated with PI resistance or exposure: E12K, Y79F, K113Q, V128A, V362I, T375 deletion and T375N (table 1.2). In addition within protease the L10I minor resistance mutation and polymorphisms G17E, L19I, E35D, M36I, L63T, E65D and I72T were present. Of the 27 clones for patient 1, 18 were identical and 9 contained one amino acid change from the patient consensus.

Given the relatively homogeneous nature of the sequences from the first patient analysed, with 18 of 27 variants being identical, it was decided that approximately ten clonal sequences from each patient would be sufficient to enable the selection of representative viral variant(s) for phenotyping. The number of clonal sequences obtained for the remaining patients was as follows: patient 2 (n=8), patient 3 (n=10) and patient 4 (n=9). Alignment of the clonal sequences resulted in consensus sequences shown in figures 4.2 for Gag and 4.3 for protease.

Viral clonal sequences derived from patient 2 could be divided into two groups of variants with similar amino acid sequences (variants 1, 4 and 5 in one group, with variants 2 and 3 forming a second group). Variants 6, 7 and 8 represented apparent recombinants between the first two groups.. Within Gag, there were five variable amino acids at positions 58, 127, 220, 341 and 480. Within protease, two variable amino acid positions were present: variant 8 displays a G57R change and position 72 is mixed I/V, with variants dividing into the groups described above for Gag. Of interest, the following changes in Gag previously associated with PI resistance or exposure were present: R76K, Y79F, H219Q (mixed H/Q within patient), V362I, S373P, I376V, K436R, L449P and T456S

(Table 1.2). In addition, protease polymorphisms I13V, I15V, P39E, D60E, L63P, A71T, I72V, V77I and I93L in comparison to consensus subtype B were present in viruses derived from patient 2.

The estimation of the average intrapatient viral evolutionary divergence, by computing the genetic distance between clonal amino acid sequence pairs, demonstrated that the Gag sequences from patient 3 were more variable than those from patients 1 and 2 with a value of 0.015 amino acid substitutions per site compared to 0.001 and 0.005, respectively. Eighteen variable amino acid positions were present within Gag and two within protease. Of interest, the following changes in Gag previously associated with PI resistance or exposure were present: R76K, S373P (mix of S/P), I376V, R380K and L449P – and a deletion of amino acid 375 within the p2/NC cleavage site (Table 1.2). In addition within protease the L33I minor PI resistance mutation and the following protease polymorphisms were present: T12S, N37C, I62V, L63P, H69Q and I93L.

The intrapatient viral divergence in patient 4 was comparable to that of patient 3, with a mean genetic distance of 0.016 amino acid substitutions per site. Analysis revealed 23 variable amino acid positions in Gag, two 5 amino acid indels present in Gag in some variants and 5 variable positions within protease. Of note, a duplication of the PTAP domain of Gag was present in variants 2, 8 and 9. The following changes in Gag previously associated with PI resistance or exposure were observed in patient 4: E12K (mix of E/K), R76K, Y79F, Y132F (Y/F mix), H219Q, R380K (R/K mix) and S451N. Within protease, there were 5 variable amino acids at positions 14, 18, 37, 77 and 87 and in total 8 polymorphisms relative to subtype B consensus were present in protease: I13V, K14R, E35D, N37T, R41K, I62V, L63P and A71T.

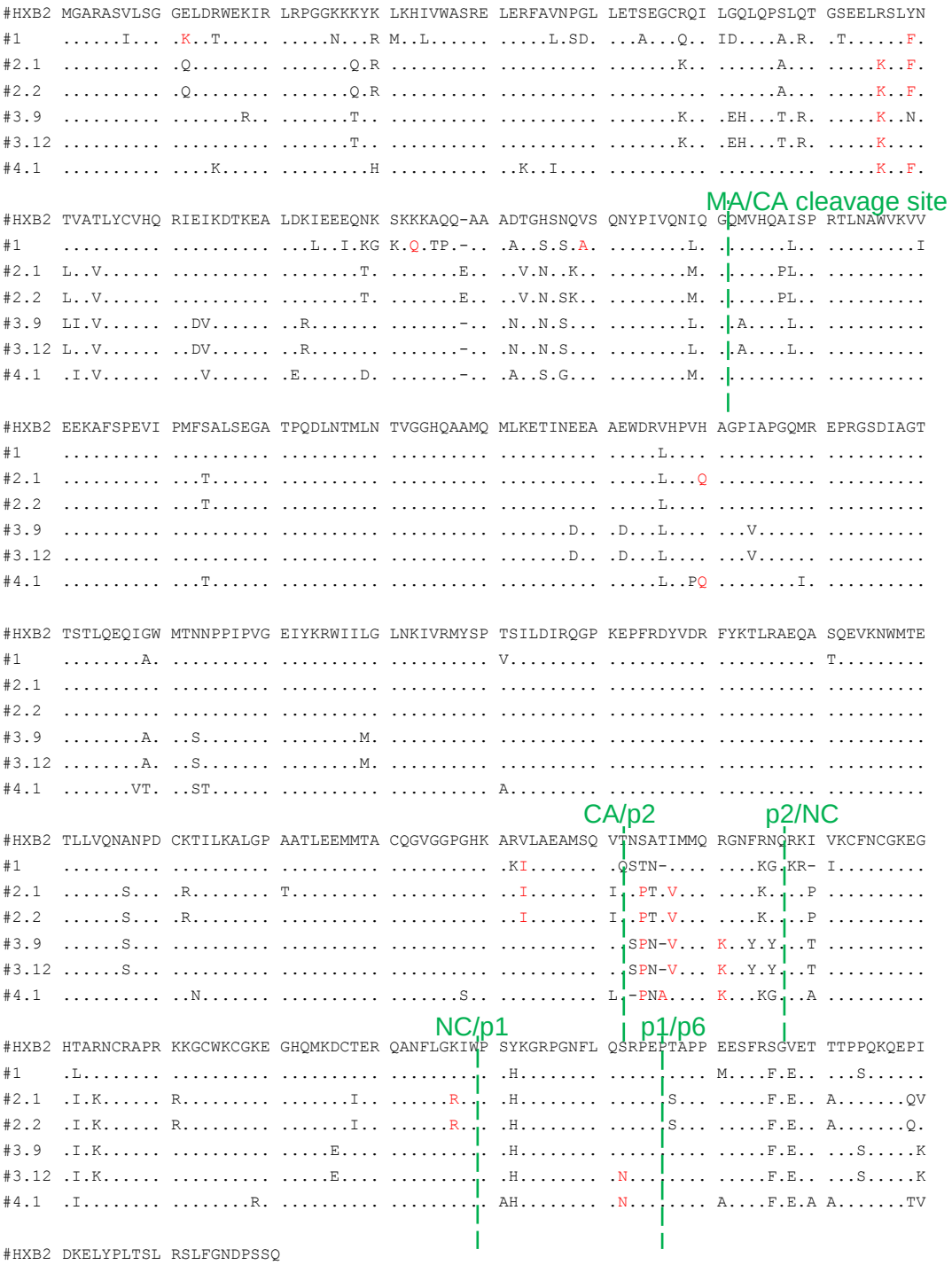


Figure 4.2. Amino acid alignment of full-length Gag from patients 1-4. An alignment of all viral variants subjected to phenotyping was produced using MEGA 4.0 and the ClustalW algorithm. Gag cleavage sites marked in green

```
#HXB2 PQVTLWQRPL VTIKIGGQLK EALLDTGADD TVLEEMSLPG RWKPKMIGGI GGFIKVRQYD

#1    ..I......I ......E.I. .......... ....DIN... .......... ..........

#2.1  ..I.......  ..V.V..... .......... ......N.E. .......... .........E

#2.2  ..I.......  ..V.V..... .......... ......N.E. .......... .........E

#3.9  ..I.......  .S........ .......... ..I...C... .......... ..........

#3.12 ..I.......  .......... .......... ..I...Y... .......... ..........

#4.1  ..I.......  ..VR...... .......... ....D.T... K......... ..........


#HXB2 QILIEICGHK AIGTVLVGPT PVNIIGRNLL TQIGCTLNF

#1    ..T.D.....  .T........ .......... .........

#2.1  ..P.......  TV....I... .......... ..L......

#2.2  ..P.......  T.....I... .......... ..L......

#3.9  .VP.....Q.  .......... .......... ..L......

#3.12 .VP.....Q.  .......... .......... ..L......

#4.1  .VP.......  T......... .......... .........
```

Figure 4.3. Protease amino acid alignment for viruses derived from patients 1-4. Protease amino acid sequences of all viral variants subjected to phenotyping were aligned in MEGA 4.0 software using the ClustalW algorithm. Amino acid changes correlating with reduced susceptibility are highlighted in red.

4.2.3 Variation in PI susceptibility of full-length Gag-protease from treatment-naïve HIV-1 infected patients

Following the sequencing of patient variants, we selected the variant(s) that would adequately represent the patient's viruses for use in the phenotypic PI susceptibility assay. When a variant identical to the consensus amino acid was present, this was selected for phenotyping. For patient 1, a single variant that was identical to the consensus was selected, however for patients 2, 3 and 4 a variant identical to the consensus was not among the clones sequenced. Therefore for patients 2, 3 and 4, the variant(s) with the least number of non-conserved amino acid changes from the consensus sequence determined after population-based sequencing was selected. This resulted in the selection of two variants for each of patient 2 and 3, as two variants were equally similar to the consensus, and only a single variant for patient 4. Once selected, variants were cloned from pGEM into p8.9NSX+ using the *Not*I and *Xho*I restriction sites as described in chapter 2. The PI susceptibility of pseudovirions containing full-length Gag-protease derived from the four patients viruses was determined, using the single replication-cycle PI susceptibility assay as described previously. Susceptibility to the six PIs in widespread clinical use was determined, these being: APV, ATV, DRV, LPV, SQV and TPV.

Variation in PI susceptibility of full-length Gag-protease derived from treatment-naïve patient viruses was observed, both between patients and between different PIs as shown in figure 4.4. Pseudovirions containing Gag-protease derived from all four patients exhibited reduced PI susceptibility in comparison to the assay reference strain, p8.9NSX+, to at least one PI. In the majority of cases this reduction in susceptibility was statistically significant ($P < 0.05$; two sample *t* test). Pseudovirions derived from patients 2 and 4 exhibited significant reductions in susceptibility to a number of PIs. This reduction in susceptibility was greatest for PIs APV, ATV, LPV and TPV ($P < 0.01$; two sample *t* test), but was particularly pronounced for ATV where a reduction of up to 17- and 16-fold

in susceptibility was observed for patients 2 and 4, respectively. For APV, LPV and TPV reduced PI susceptibilities of up to 6-, 8- and 6-fold difference in EC_{50}, respectively, were observed for patients 2 and 4. Conversely, full-length Gag-protease from patients 1 and 3 conferred PI susceptibilities that were much closer to that of the reference strain, although in some cases statistically different, for DRV, LPV, SQV and TPV. For the PIs APV and ATV small reductions in susceptibility of between 2- and 4-fold difference in EC_{50} were observed. Of, interest for patients where two variants were tested, both showed broadly similar PI susceptibility levels.

As the PI susceptibility of viruses derived from patients 1 and 3 was so close to that of the reference strain, they were not investigated further. However, viruses derived from patients 2 and 4 were further analysed, as these showed the greatest reduction in susceptibility in comparison to the assay reference strain to a number of PIs.

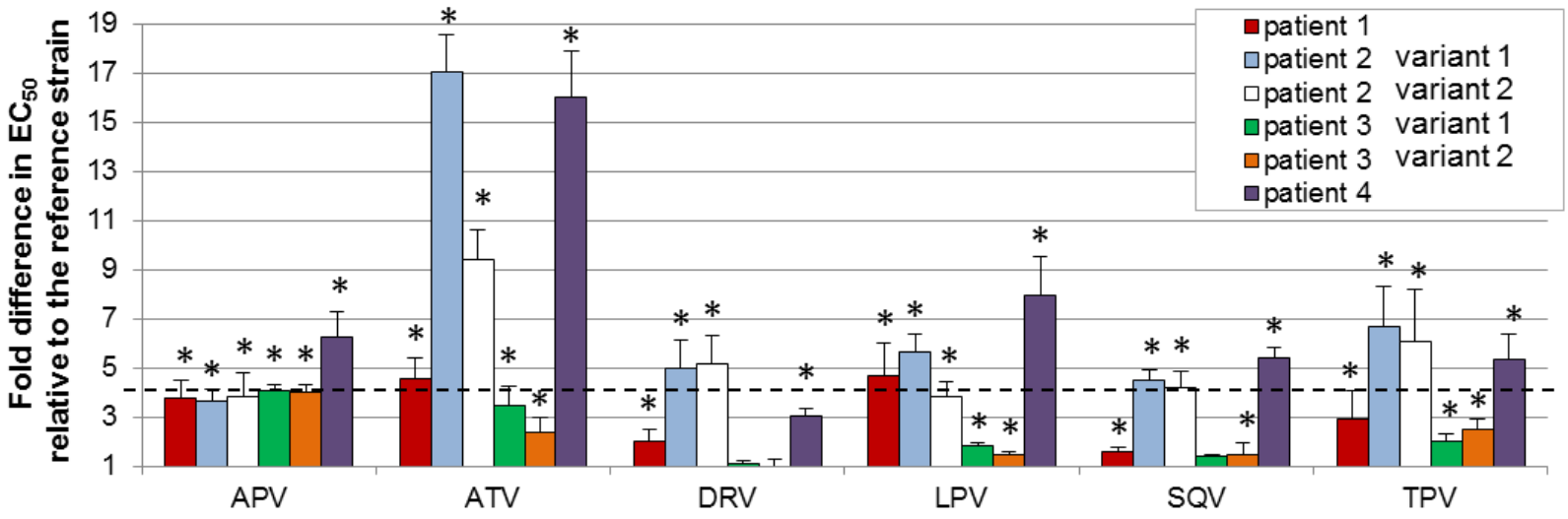


Figure 4.4. Variation in PI susceptibility of full-length Gag-protease from HIV-1 infected, treatment-naïve patients. Full-length Gag-protease was amplified from plasma samples and cloned into p8.9NSX+. VSV-g pseudotyped viruses encoding luciferase were produced by co-transfection in 293T cells. PI susceptibility of pseudovirions derived from each patient was determined using a single replication-cycle drug susceptibility assay as measured by luciferase activity. Data displayed are fold difference in EC_{50} values in comparison to that of the assay reference strain, p8.9NSX+, for each of six PIs: APV, ATV, DRV, LPV, SQV and TPV. Viruses for which the raw EC_{50} values were statistically different from assay reference strain using unpaired t-tests ($P < 0.05$) are denoted with an asterisk(*). Error bars represent the standard error of the mean of three independent experiments performed in duplicate. The dashed line shows the previously reported cut off for a significant reduction in susceptibility of greater than four-fold in comparison with the assay reference strain (Gong et al., 2004).

4.2.4 Contribution of protease to reduction in PI susceptibility observed for full-length Gag-protease in subtype B clinical viruses

To elucidate the role of Gag and protease individually in the reduced PI susceptibility observed for full-length Gag-protease from patients 2 and 4, the PI susceptibility of chimeric viruses containing Gag only or protease only from the patient was determined. For each patient two chimeric viruses were created: p8PtGag containing Gag derived from the patient virus and protease from p8.9NSX+ and p8PtPro, which contained Gag from p8.9NSX+ and protease from the patient, as depicted in figure 4.5a. Restriction sites *Not*I, *Apa*I and *Xho*I were used. The location of the *Apa*I site resulted in the inclusion of MA, CA, p2 and the N terminus of NC from the patient virus in the p8PtGag chimeric virus. The p8PtPro contained the C terminus of NC, p1 and p6 of Gag in addition to protease from the patient virus, as displayed in figure 4.5a. For patient 2, variant 1 was selected for further analysis as it showed the highest fold-difference in EC_{50} of the two variants. The PI susceptibility of these chimeric vectors to the four PIs APV, ATV, LPV and TPV, was determined using the single replication-cycle drug susceptibility assay.

The results in figure 4.5b and c show that the protease only chimeric viruses from patients 2 and 4 displayed reduced susceptibility in comparison to assay reference strain of up to 9- and 13-fold-difference in EC_{50} for patients 2 and 4, respectively (figure 4.5). Conversely, the Gag only chimeric viruses displayed PI susceptibility levels similar to that of the assay reference strain for all four PIs for which the susceptibilities were tested. For each patient the difference in susceptibility between the Gag and protease only chimeric viruses was statistically significant for both APV and ATV ($P < 0.05$).

There was no evidence that Gag from these patient-derived viruses confers reduced susceptibility to PIs when separated from its cognate protease, suggesting that in patients 2 and 4 the reduced susceptibility conferred by

protease alone largely accounts for the reduced susceptibility observed for full-length Gag-protease. This was contrary to the effect observed with the molecular clones in chapter 3, where the *gag* gene was shown to significantly contribute to the reduced PI susceptibility of full-length Gag-protease.

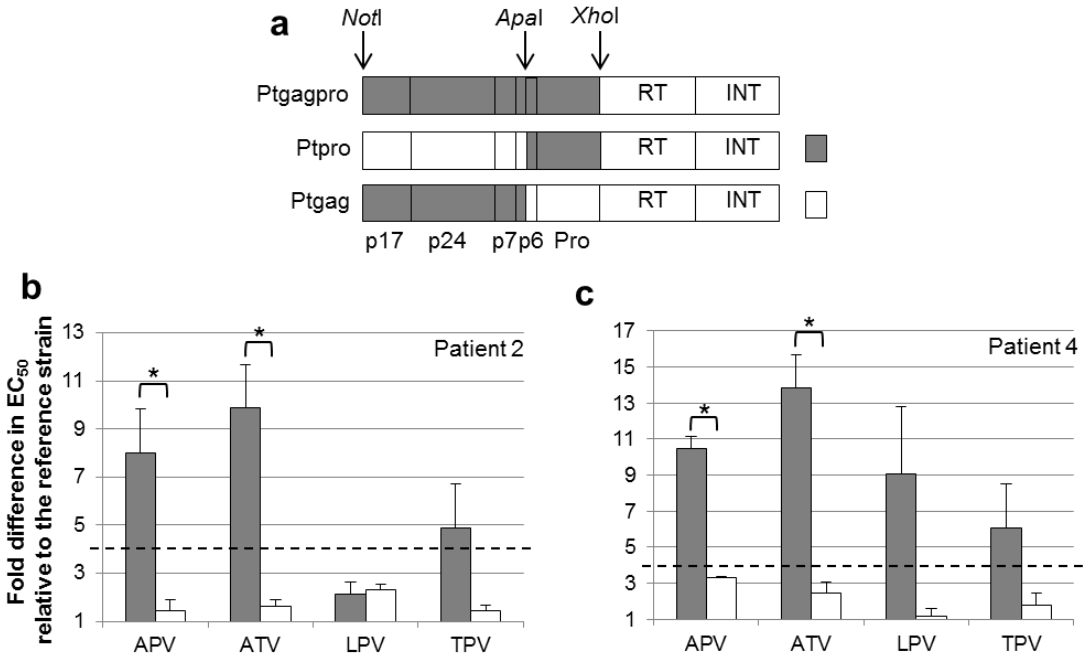


Figure 4.5. Variation in PI susceptibility of full-length Gag-protease conferred independently by Gag in a single replication-cycle assay. Chimeric vectors containing only Gag or only protease from the patient were generated using standard cloning techniques and a schematic is shown in (a). Schematic diagram of chimeric viruses containing either Gag or protease from viruses derived from each patient. Segments of the chimeric viruses derived from the assay reference strain, p8.9NSX+, are depicted in white, and that from patient-derived viruses shaded grey. Pseudovirions encoding luciferase were generated by co-transfection in 293Ts of the chimeric vectors for two patients (b) patient 2 and (c) patient 4 for viruses for which the fold difference in EC_{50} values for Gag only and protease only chimeric viruses were statistically different using an unpaired t-test ($P < 0.05$) are denoted with an asterisk(*). Error bars are standard error of the mean, derived from three independent experiments performed in duplicate. The dashed line shows the previously reported cut off for a significant reduction in susceptibility of greater than four-fold in comparison with the assay reference strain (Gong et al., 2004).

4.2.5 Effect of minor resistance mutations and polymorphisms in protease on PI susceptibility of treatment-naïve, patient-derived viruses

Having shown that the patient-derived protease largely conferred the reduced PI susceptibility observed for full-length Gag-protease, the determinants of the reduced susceptibility were investigated. Analysis of protease sequences of the viral variants from patients 2 and 4 using the Stanford HIVdb Genotypic Resistance Interpretation Algorithm (http://hivdb.stanford.edu/) and the IAS list of drug resistance mutations showed no major PI resistance mutations. However, it indicated the presence of two minor resistance mutations, A71T and L63P, and a polymorphism I13V in both patients as shown in Figure 4.6 (Johnson *et al.*, 2013). The Stanford Resistance Algorithm did not predict an effect of these minor mutations on phenotypic susceptibility, hence the inclusion of these patients in this study. In addition, there were a number of other amino acid changes from consensus subtype B protease in each of the patients. In contrast, patients 1 and 3 did not share all three of protease amino acid changes, A71T, I13V and L63P, with only the L63P mutation present in patient 3. However, patients 1 and 3 shared minor resistance mutations L10I and L33I and each had a number of other additional changes compared to consensus B protease sequence as detailed previously.

As the reduced PI susceptibility mapped to protease, we next investigated the contribution of the amino acid changes I13V, L63P and A71T to the reduction in PI susceptibility. The amino acids were sequentially reverted to the consensus subtype B amino acid using site-directed mutagenesis, leading to the creation of the following three vectors for each patient: 1M (T71A), 2M (T71A and V13I) and 3M (T71A, V13I and P63L), as shown in Figure 4.7a. A71T was chosen to be studied alone as it was a minor resistance mutation reported to affect susceptibility to ATV, IDV, LPV, NFV, SQV (Johnson *et al.*, 2011). The polymorphism I13V was reverted next as it was present less frequently in PI

naïve viruses than L63P, which is a common polymorphism considered a minor resistance mutation to LPV only. The susceptibility of the three vectors for each patient was then tested using the single replication-cycle drug susceptibility assay.

Overall, reversion of protease positions 13, 63 and 71 to consensus subtype B increased the susceptibility of the virus towards that of the reference strain (Figure 4.7b and c). This effect was statistically supported for patient 2 derived pseudovirions to ATV. However, for some PIs reversion of A71T conferred a further reduction in PI susceptibility (for Pt 2 to APV and for Pt 4 to ATV and LPV). In addition, for the PI APV reversion of protease positions to wild-type does not appear to increase the susceptibility of the viruses towards that of the reference strain, in fact reversion results in a virus significantly less susceptible for patient 2 derived viruses. These data provided evidence that the minor resistance mutations and polymorphism present in the protease of viruses derived from patients 2 and 4 contribute in part to the reduced susceptibility observed in the absence of major PI resistance mutations.

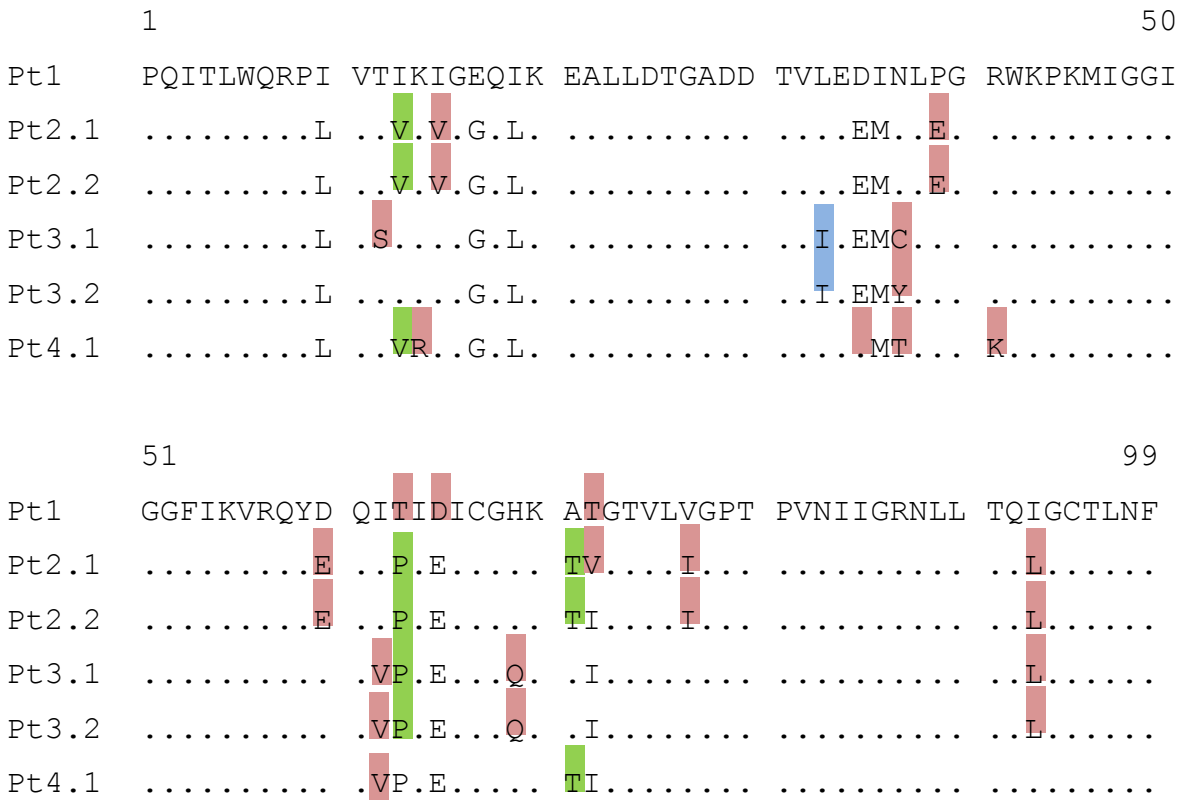


Figure 4.6. Protease variation of all patient derived viruses subjected to phenotypic PI susceptibility testing. An alignment of all protease amino acid sequences subjected to phenotypic testing was produced using the ClutalW algorithm – six viruses in total derived from four patients. Minor resistance mutations or protease polymorphisms identified by the Stanford Resistance Algorithm are highlighted in blue and pink, respectively. Amino acids 13V, 63P and 71T that correlated with reduced phenotypic susceptibility are highlighted in green.

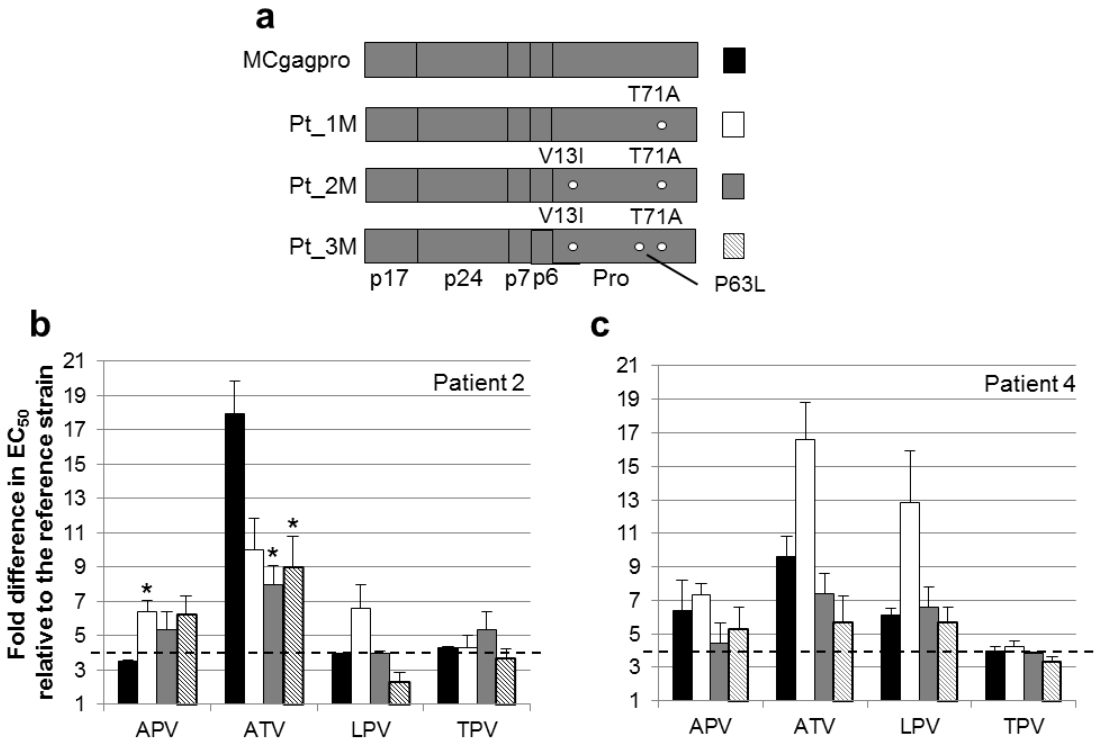


Figure 4.7. Protease changes I13V, L63P and A71T affect PI susceptibility.
Protease changes 13V, 63P and 71T were reverted to consensus amino acid using standard site-directed mutagenesis techniques. For each patient, three vectors were created containing one reversion, two reversions and all three changes reverted (a). For each vector, pseudovirions were produced by co-transfection in 293T cells and their PI susceptibility tested using a single replication-cycle assay. The susceptibilities to four PIs (APV, ATV, LPV and TPV) of the three reversion vectors are shown for (b) patient 2 and (c) patient 4. Data are expressed as a fold difference in comparison to that of the assay reference strain. Viruses for which the fold difference EC_{50} values were statistically different from that of Ptgagpro using an unpaired t-test ($P < 0.05$) are denoted with an asterisk(*). Error bars represent the standard error of the mean of three independent experiments performed in duplicate. The dashed line shows the previously reported cut off for a significant reduction in susceptibility of greater than four-fold in comparison with the assay reference strain (Gong et al., 2004).

4.2.6 Effect of amino acid changes in protease on PI susceptibility of assay reference strain

Having demonstrated the contribution of amino acid changes I13V, L63P and A71T to the reduced susceptibility observed for full-length Gag-protease derived from patients 2 and 4, an investigation of the effect of these changes on the susceptibility of the reference strain was required. I13V, L63P and A71T were introduced into the p8.9NSX+ reference strain using site-directed mutagenesis techniques, giving a vector denoted p8.9proSDM. The susceptibility of this vector to four PIs, APV, ATV, LPV and TPV, was determined using a single replication-cycle PI susceptibility assay.

Figure 4.8 demonstrates that the p8.9proSDM vector displayed small reductions in PI susceptibility to the PIs ATV, LPV and TPV in comparison to assay reference strain, p8.9NSX+. p8.9proSDM displayed a 3-fold reduction in EC_{50} in comparison to assay reference strain to the PI ATV, and 2-fold to PIs LPV and TPV. However, to PI APV no difference in susceptibility in comparison to reference strain was observed.

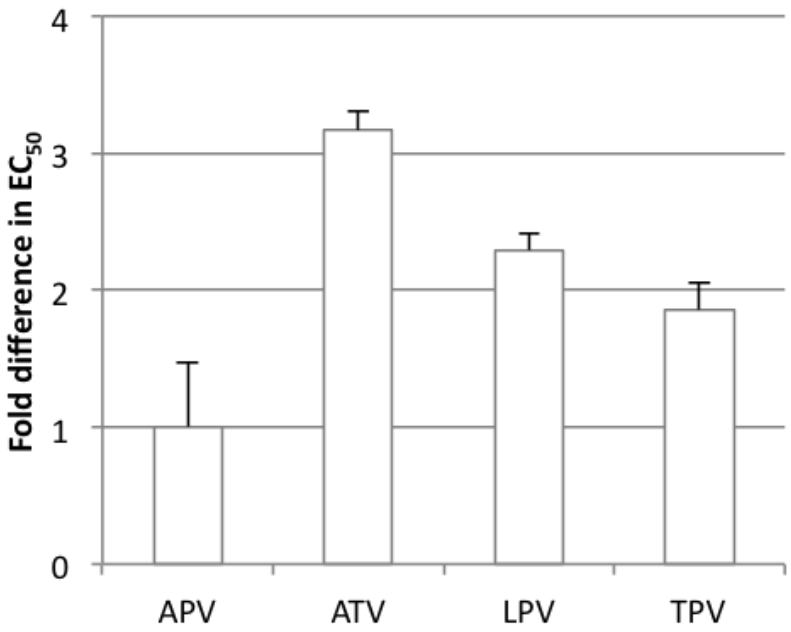


Figure 4.8. Amino acid changes I13V, L63P and A71T reduce PI susceptibility of reference strain, p8.9NSX+. Amino acid changes I13V, L63P and A71T were introduced into the assay reference strain p8.9NSX+ using site directed mutagenesis techniques. Pseudovirions were produced by co-transfection in 293Ts and PI susceptibility tested in a single replication-cycle assay to APV, ATV, LPV and TPV. Error bars represent the standard error of the mean of two independent experiments, each performed in duplicate.

4.2.7 Computer modelling of positions of amino acid changes within the protease homodimer

Having demonstrated an effect of changes at amino acid positions 13, 63 and 71 on PI susceptibility of patient-derived virus, the position of these changes within the structure of the protease homodimer was determined. Molecular modelling was carried out using RasMol software and protease structures 7HVP (the protease homodimer) and 3EL1 (protease homodimer structure in complex with ATV), obtained from the RCSB protein database (www.rcsb.org) (Berman et al., 2000; King et al., 2012; Swain et al., 1990).

Figure 4.9a shows the position of amino acids 13, 63 and 71 within the protease enzyme structure. The three sites are located relatively close to one another on the outside of protease and are not found near the enzyme active site. The position of the amino acids in protease is not significantly altered when ATV is bound to the protease active site, as shown in figure 4.9b. The positions of the other unique polymorphisms present within patients 2 and 4 are shown in figure 4.9c and 4.9d, respectively. The molecular modelling shows that none of these changes are present in the active site region of protease, instead they are found on the outside of the protease dimer. Only position 77 is found near the active site.

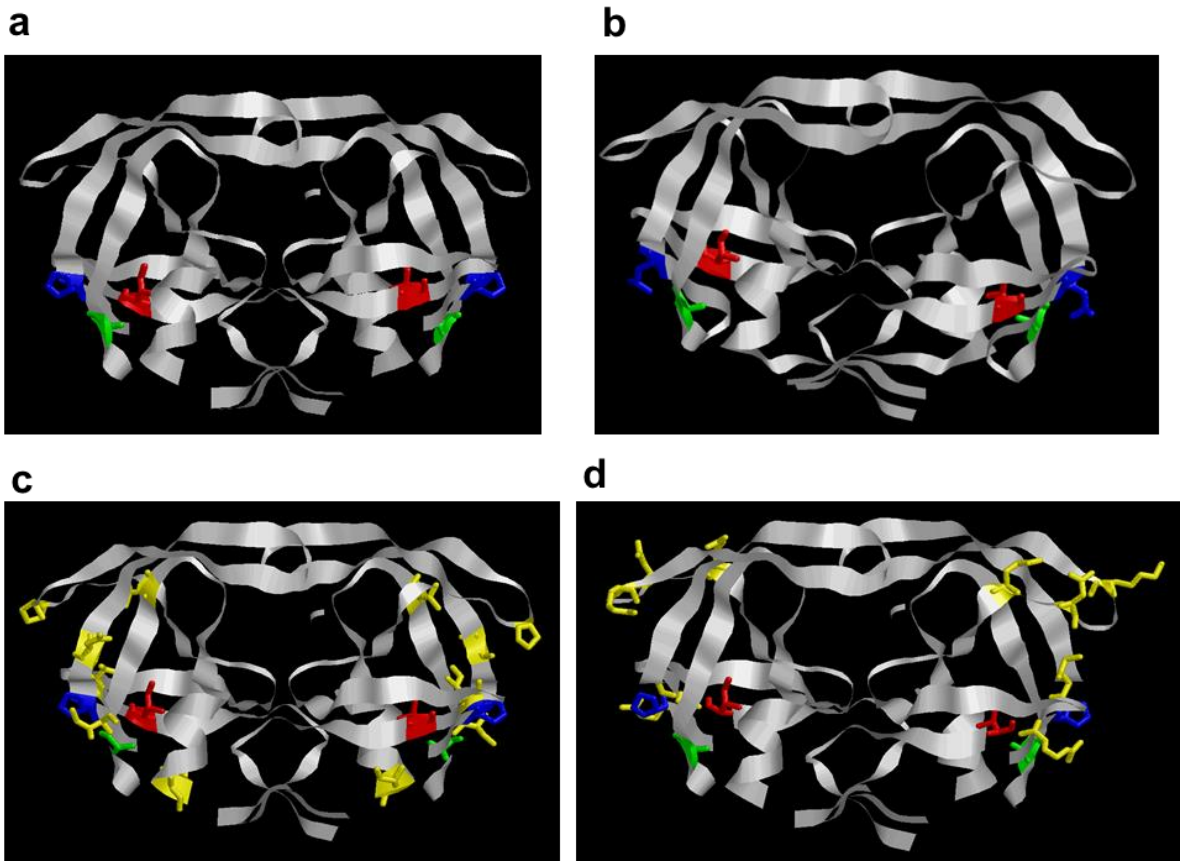


Figure 4.9. Computer modelling of the location of amino acids positions in protease shown to affect PI susceptibility. Molecular modelling was performed using RasMol software to determine the location of amino acid positions of interest within the 3D protein structure. The position of amino acids 13 (red), 63 (blue) and 71 (green) are shown within the protease homodimer, 7HVP, (a) and the structure of protease bound to the PI ATV, 3ELI, (b) from the RSCB PDB. Modelling demonstrates that these changes do not appear within the active site of the enzyme. In addition, positions 13, 63 and 71 are shown with the other amino acids in which changes were observed (yellow) for patient 2 (c) and patient 4 (d).

4.3 Discussion

Here, significant variation in PI susceptibility of full-length Gag-protease derived from four treatment-naïve patients infected with subtype B HIV-1 was observed. Whilst viruses derived from patients 1 and 3 displayed PI susceptibilities most similar to that of the reference strain, viruses from patients 2 and 4 displayed significant reductions in susceptibility to a number of PIs. This finding is in keeping with the data for molecular clones in chapter 3 where up to 9-fold reductions in PI susceptibility were observed, and that of previous studies investigating PI susceptibility of patient-derived viruses of subtypes A, C and CRF01_AE which have also reported significant levels of variation when full-length Gag-protease is incorporated (Gupta et al., 2010; Jinnopat et al., 2009). These data show that the reduction in susceptibility was particularly pronounced for ATV, with patient 2 showing a 17-fold reduction in EC_{50} in comparison to that of the reference strain. This finding is of particular importance because a recent change in PI usage in London guidelines means ATV is now the recommended first line PI (London HIV Consortium, 2012). The BHIVA guidelines also recommend ATV/r as an option for the third agent in first-line HAART (Williams et al., 2012). In addition, the use of ATV is increasing in the developing world where pre-treatment resistance testing is not widespread so this finding of potential variation in susceptibility in treatment-naïve patients merits further investigation.

Use of chimeric viruses demonstrated that protease solely conferred the reduced susceptibility observed for full-length Gag-protease in both patients 2 and 4. This is in contrast to the findings for molecular clones presented in chapter 3. Here, there is no evidence that Gag can independently confer variation in susceptibility in these patient-derived viruses, despite the presence of changes in Gag that have previously been reported to be involved in PI resistance or associated with PI exposure. In previous studies, Gag mutations were described and studied in combination and these particular combinations of

changes were not present in the patients in this study, which may explain why they did not confer reduced PI susceptibility (Aoki et al., 2009; Callebaut et al., 2007; Gatanaga et al., 2002; Mammano et al., 1998). For example, a previous study reported the role of three mutations at positions 76, 79 and 81 in PI susceptibility, but also demonstrated that they had little or no effect on susceptibility when present singly (Parry *et al.*, 2011). A number of these previously reported Gag mutations have been described in non-B subtypes and their role in subtype B viruses on PI susceptibility is unknown (Jinnopat et al., 2009; Kameoka et al., 2010).

Whilst not observed in this study, an independent effect of Gag derived from PI-naïve patients infected with non-B subtype viruses was shown to result in the reduction of susceptibility to LPV, with a greater effect for Gag alone than for full-length Gag-protease in some cases (Gupta *et al.*, 2010). Although our study did not provide evidence of an independent role of Gag in susceptibility, for some PIs protease alone did not confer the same reduction in susceptibility observed for full-length Gag-protease, for example for patient 2 to ATV, LPV and TPV. Under certain conditions a 'wild-type', p8.9NSX+ derived Gag does not appear to be sufficient to enable patient-derived protease to confer the same level of reduced susceptibility observed when present with co-evolved Gag. This suggests that the matching of co-evolved Gag and protease is important. Just as it has been shown that cleavage-site mutations in Gag can occur to compensate for reductions in fitness caused by major PI resistance mutations, it is possible that co-evolution at other sites may occur between Gag and protease that has not yet been described. A similar effect was seen in an earlier study in which the fold-difference EC_{50} conferred by full-length Gag-protease was greater than the total of the fold-difference EC_{50} values of Gag and protease independently (Parry *et al.*, 2009). This finding is potentially significant as commercially available assays which have been used for much of the phenotypic drug susceptibility research and clinical service to date do not

include co-evolved Gag, potentially resulting in an inaccurate determination of PI susceptibility.

Analysis using the Stanford HIVdb Genotypic Resistance Interpretation Algorithm and the IAS list of drug resistance mutations identified three protease changes that were common in both patients 2 and 4: A71T, I13V and L63P (Johnson et al., 2013; Liu and Shafer, 2006). Here, we have shown that the protease mutations A71T, I13V and L63P contributed to the reduced susceptibility observed for viruses derived from patients 2 and 4. In addition, we demonstrate that introduction of these changes to the assay reference strain confers small reductions in susceptibility, of up to 3-fold. Our data indicate that the presence of a combination of minor resistance mutations and polymorphisms in protease is able to convey reduced susceptibility to PIs, in the absence of major PI resistance mutations. To our knowledge, only one study examining the role of these three protease mutations in the absence of other minor resistance mutations in protease on PI susceptibility *in vitro* has been carried out (Kempf *et al.*, 2001). The study reported a small FC EC_{50} to LPV when the three mutations were found together (0.8 fold), but did not examine the effect on susceptibility to the three other PIs studied here (APV, ATV and TPV). In addition, two studies have been carried out that demonstrate a role of these three protease mutations in an increased rate of PI resistance development during passage *in vitro* in the presence of drug (Lisovsky et al., 2010; Vergne et al., 2006). *In vivo*, a number of studies have reported an association between these protease mutations and reduced rates of virological response in patients when present in baseline sequences (Baxter et al., 2006; Johnson et al., 2005; Marcelin et al., 2007; Pellegrin et al., 2006; Pellegrin et al., 2008).

These amino acid changes are not uncommon within PI-naïve subtype B viruses, with I13V present in 14%, L63P present in 57% and A71T present in 6% (Rhee et al., 2003). In addition, each of these is present at an increased

frequency in PI-experienced viruses in comparison to PI-naïve viruses: I13V in 25%, L63P in 78% and A71T in 10%. L63P was also identified as one of the amino acid changes present at higher frequency in viruses from patients failing ATV therapy in comparison to those from PI-naïve patients in a UK cohort as described below, providing evidence for a role in PI resistance (Dolling et al. 2013). However our modeling data show that amino acids 13, 63 and 71 are not located near the active site of the protease enzyme, so the exact mechanism by which they affect PI susceptibility is unknown. It is possible that these amino acid changes induce a conformational change in protease that results in reduced affinity of protease for the drug and/or a higher affinity for the natural Gag substrate. Taken together, these data indicate that the precise mechanisms of development of resistance to PIs and the role of protease polymorphisms remain to be fully determined.

The most significant reduction in PI susceptibility and the greatest effect on PI susceptibility of reversion of the minor mutations and protease polymorphisms was observed for ATV. This led us to further investigate the role of protease polymorphisms in ATV susceptibility and resistance. To date, studies have mainly concentrated on the development of major mutations as a marker for changes in susceptibility, but it has been shown that major resistance mutations are rarely found in patients failing ATV therapy, in both PI-naïve patients and those with prior PI exposure (Scherrer *et al.*, 2012). It has been assumed that treatment failure in the absence of resistance mutations is a consequence of poor adherence, but we hypothesised that reduced susceptibility in the absence of major resistance mutations could also contribute to treatment failure, for example conferred by combinations of minor mutations and polymorphisms in protease. To further investigate, a collaboration was established with the UK HIV Drug Resistance Database to establish whether there was any variation in the frequency of amino acids within protease between ATV-naïve patients and those failing on ATV-based therapy (Dolling et al., 2013b). The study utilised

the UK National HIV Resistance Database, which contains all genotypic resistance tests carried out by public labs in the UK coupled with clinical information from the UK Collaborative HIV Cohort (CHIC), to compare two groups of patients: those who had experienced virological failure on ATV based therapy and a PI-naïve control group. The study reported that only six of the 322 resistance tests included contained major PI resistance mutations. In addition, the study reported multiple novel mutations associated with treatment failure including L19T, K43T, L63P/V, K70Q, V77I and L89I/T/V, providing further evidence that the exact genetic determinants of ATV susceptibility remain to be fully defined.

In conclusion, we demonstrate significant variation in PI susceptibility using a single replication-cycle phenotypic assay encompassing full-length Gag-protease from subtype B HIV-1 viruses derived from treatment-naïve patients. We show that protease can confer significant reductions in susceptibility in the absence of major protease mutations, an effect that is particularly pronounced for ATV. Furthermore, this reduction in susceptibility would not be captured by traditional genotyping methods and partnership with its co-evolved Gag in phenotypic assays is necessary to measure the full effect on PI susceptibility. Further research is required to fully elucidate the role of minor resistance mutations and polymorphisms in protease on PI susceptibility and to confirm the clinical significance of these findings.

5 Investigation of the genotypic and phenotypic correlates of virological failure in treatment-naïve patients receiving LPV/r monotherapy

5.1 Introduction

Although efficacious, long-term HAART treatment entails both a high financial cost and a risk of significant side-effects, making exploration of alternative treatment regimens prudent. Simplification strategies have been investigated in an effort to reduce the number of antiretroviral drugs required as part of treatment regimens without compromising treatment efficacy. To date two main mechanisms of treatment simplification have been explored: firstly the removal of one or more ART drugs from patients with a suppressed viral load on HAART (Arribas et al., 2010; Gilks et al., 2012) and secondly by initiating treatment-naïve patients on just one ART drug (Delfraissy *et al.*, 2008). Early clinical trials have shown that reverse transcriptase inhibitors cannot be used successfully as monotherapy, although they are used in certain situations including the prevention of mother-to-child transmission in the developing world (Guay et al. 1999). However, PIs are a potent class of inhibitors and to date a number of trials have been carried out to examine their efficacy when used as monotherapy

5.1.1 Previous PI simplification clinical trials

The first PI simplification trials conducted in the 1990s investigated the use of the first generation of PIs NFV, SQV and IDV (Havlir et al., 1998; Reijers et al., 1998). The ADAM study recruited ART-naïve patients and started with an induction therapy comprising d4T, NFV, 3TC and SQV, before a randomised switch to maintenance therapy after 26 weeks of either d4T and NFV, SQV and NFV, or continuation on the induction therapy. This study was discontinued after interim analyses due to inferiority of the d4T/NFV and SQV/NFV arms

(Reijers et al., 1998). A second study recruited patients with suppressed viral load <200 copies/ml on IDV, 3TC and AZT before randomisation to either a continuation of triple therapy, IDV monotherapy or AZT and 3TC dual therapy, with 106, 103 and 107 patients randomised to each arm respectively. This trial reported 23% patients experiencing virological failure in both the IDV monotherapy and AZT/3TC dual therapy arms, compared to only 4% in the triple therapy arm (Havlir et al., 1998). These early trials showed significant inferiority of PI simplification strategies with the early PIs in comparison to triple therapy.

5.1.1.1 ATV monotherapy trials

Several clinical trials to examine the efficacy of ATV/r monotherapy have been carried out, but to date they have included small numbers of patients, between 15 and 36. The first study recruited 36 patients with suppressed viral loads on a PI-containing regimen. Patients were switched to an ATV/r-containing regimen and after 6 weeks the NRTIs were removed from the regimen, leaving patients with ATV/r monotherapy. After 24 weeks, 91% patients had maintained virological suppression. Three patients failed on PI monotherapy, but none had major resistance mutations and 2/3 had low or undetectable ATV/r levels in the blood at the time of treatment failure (Swindells et al., 2006). Another small, single-centre trial aimed to recruit 30 patients but was discontinued at only 15, when 5 patients experienced virological failure. However, none of the patients failing monotherapy had any major PI resistance mutations (Karlstrom et al., 2007). A third recruited 34 patients with suppressed viral loads, whose therapy was simplified to monotherapy. Thirty of the 34 patients maintained virological suppression and none of the 4 patients failing therapy had a major PI resistance mutation (Wilkin et al., 2009). The ATARITMO trial recruited 30 patients with suppressed viral loads on HAART and patients were swapped to receive only ATV/r monotherapy. The trial showed good efficacy, however there was evidence of ongoing viral replication in the CSF and semen despite suppression

in the plasma (Vernazza et al., 2007). A larger ATV/r monotherapy trial – MODAt - has enrolled 342 patients and is currently ongoing (Castagna et al., 2013).

5.1.1.2 DRV/r monotherapy trials

Several large clinical trials showing the efficacy of DRV/r monotherapy have been reported to date. The MONET trial took place in Spain and included 256 patients with viral loads <50 copies/ml. Patients were randomised to DRV/r monotherapy (n=127) or the DRV/r based triple therapy arm with two NRTIs (n=129). The trial reported non-inferiority of the monotherapy arm, with 86.2% and 87.8% maintaining viral suppression at week 48 for the monotherapy and triple therapy arms respectively (Arribas et al., 2010). Of those failing therapy, only one patient in each treatment arm developed major PI resistance mutations. The MONOI-ANRS 136 trial recruited 225 patients with viral loads <400 copies/ml, and randomised them to DRV/r monotherapy arm or DRV/r based triple therapy. Virological suppression of 99% and 94% was seen in the triple therapy and monotherapy arms, respectively. Of the three patients who failed, none had developed any major PI resistance mutations (Katlama et al., 2010b). Another DRV/r monotherapy trial is currently underway in the UK, the PIVOT trial. It includes several sub-studies to investigate the penetration of DRV/r into the CNS.

5.1.1.3 LPV/r monotherapy trials

Currently, LPV/r is the only PI available in most of the developing world particularly sub-Saharan Africa, so investigation of its efficacy as monotherapy is particularly important. The Only Kaletra (OK) trial involved 21 patients who had been virologically suppressed for at least 6 months before trial enrollment. At trial end, 14/21 patients remained suppressed and 5 experienced virological rebound. However, no major PI resistance mutations were present in these patients and suppression was achieved with the addition of two NRTIs (Pulido et al., 2008). The SARA study was another LPV/r monotherapy study that was

conducted as a sub-study of the DART trial. The DART trial took place in Uganda and Zimbabwe and aimed to establish whether ARV could be successfully introduced in the absence of viral load and CD4 count monitoring. Patients on second-line therapy in DART were recruited to the SARA study and randomised to continue therapy (CT) or simplify therapy to PI monotherapy. At entry to the SARA study, patients had already received ARV for a mean 4.4 years and only 77% had a suppressed viral load, <50 copies/ml. After 24 weeks, in those patients who had suppressed viral loads at entry, CT was superior to monotherapy with 85% and 66% of patients with viral load <50 copies/ml, respectively. Genotypic resistance testing carried out in 12 patients failing therapy identified major PI resistance mutations in two patients, I54V in one patient and both M46IM and V82AV in the second (Gilks et al., 2012).

5.1.2 The MONARK trial

The MONARK trial investigated the efficacy of boosted lopinavir (LPV/r) monotherapy in comparison to LPV/r-containing HAART in treatment-naïve patients. In total, 83 patients were randomised to the LPV/r monotherapy arm, with 53 patients in the HAART control arm who received LPV/r with AZT and 3TC. Patients in the monotherapy arm were followed for up to 96 weeks, with CD4+ T cell counts and viral load monitoring carried out. Primary endpoints were defined as proportion of individuals achieving a virological response – a viral load of <400 copies/ml after 24 weeks and <50 copies/ml after 48 weeks. The MONARK trial was undertaken in France, Spain and Poland with the proportion of subtypes as follows: 68% subtype B, 16% CRF02_AG, and 16% other subtypes and recombinants (A, CRF01_AE, C and G).

The trial reported inferiority of LPV/r monotherapy in comparison to the LPV/r-containing HAART regimen after 48 weeks, with virological response achieved in 64% and 75% of patients respectively (P = 0.19) (Delfraissy *et al.*, 2008). A significant proportion of patients randomised to the LPV/r monotherapy arm

experienced virological failure (23/83), defined as viral load >400 copies/ml from week 24 or viral load >50 copies/ml by week 48 (Delfraissy *et al.*, 2008). By week 96, 23/83 patients had experienced virological failure (a viral load >50 copies/ml) (Ghosn *et al.*, 2010). Of patients experiencing failure in the LPV/r monotherapy arm, genotypic resistance testing identified major PI resistance mutations in only 5, with none occurring in the triple therapy arm. Phenotypic PI susceptibility testing was performed for patients failing with major PI resistance mutations utilising the Monogram Phenosense Assay. This showed a mean increase of 2.2 fold-difference in LPV/r EC_{50} in viruses at the time of treatment failure in comparison to baseline viruses from the same patient (Delaugerre *et al.*, 2009). However, as previously discussed the Phenosense assay does not include full-length Gag from the patient virus and our data have shown the importance of its inclusion when determining phenotypic susceptibility accurately (see chapters 3 and 4, Gupta et al., 2010; Parry et al., 2009).

Flandre et *al.* (2009) reported a number of prognostic factors associated with virological response in the MONARK trial including a lower baseline viral load and viral subtype. 87% of patients infected with subtype B virus achieved virological response in comparison to 65% of those infected with non-B viruses, although the significance of this is complicated by the finding that patients infected with non-B subtype viruses reported missing more doses that those infected with subtype B viruses. Analysis of the amino acid sequences at the Gag cleavage sites of viruses from all patients in the PI monotherapy arm reported CSMs in 81/82 patients at baseline including Gag changes K463R, I437V, L449F and P453L. Furthermore, Gag changes V484G/I/P/S and S451G/N/R were associated with virological failure and success, respectively (Ghosn *et al.*, 2011). Comparison of viral sequences at baseline and failure showed the development of additional CSMs at failure in 11/23 patients and conversely reversions to wild-type in 9/23 patients. Thus, in the patients failing

therapy without major PI resistance mutations the exact determinants of treatment failure remain unknown.

Recent studies have provided evidence for the role of Gag in the determination of PI susceptibility, with amino acid changes located outside the Gag cleavage sites reported to affect PI susceptibility and correlate with treatment failure and PI exposure (Aoki et al., 2009; Gatanaga et al., 2002; Parry et al., 2011). In addition, the inclusion of full-length patient-derived Gag alongside its co-evolved protease in *in vitro* phenotypic assays has been shown to affect PI susceptibility in both treatment-naïve and treatment-experienced patients (Gupta et al., 2010; Jinnopat et al., 2009; Parry et al., 2009). In chapters 3 and 4 of this study, we have shown variation in susceptibility to PIs in the absence of previously reported resistance mutations when full-length Gag-protease from the test virus was included in phenotypic assays. We have provided evidence that both Gag and protease can contribute to this variation in susceptibility. Thus, this study sought to investigate the determinants of treatment failure in the MONARK trial by studying co-evolved, full-length Gag-protease both genotypically and phenotypically from patient viruses at baseline and time of treatment failure.

<u>Chapter aims:</u>

1) Amplify full-length Gag-protease from paired plasma samples taken at baseline and at the time of treatment failure from patients in the LPV/r monotherapy arm of the MONARK trial

2) Perform clonal analysis of full-length Gag-protease of viruses derived from each time point for each patient to identify amino acid changes previously reported to be involved in or associated with PI resistance

3) Use bioinformatic methods to examine sequences for evidence of positive selection and evolution.

4) Examine the PI susceptibility and single-round infectivity of representative viral variants from each patient at baseline and failure time points.

5.2 Results

5.2.1 Clinical and virological information of patients experiencing virological failure on LPV/r monotherapy, selected for this study.

We had access to paired baseline and failure samples from 23 patients experiencing virological failure on LPV/r monotherapy up to week 96 of the MONARK trial. Full-length Gag-protease from both baseline and failure plasma samples was successfully amplified from five of these patients by performing RNA extraction, DNA synthesis and nested PCR as previously described (see materials and methods section). These patients formed the basis of this study (figure 5.1) and their clinical and virological information including viral load, subtype and time of treatment failure is shown in table 5.1. These data show that the patients selected for this study were representative of all those enrolled in the trial. Patients randomised to the LPV/r monotherapy arm were infected with the following subtypes: 68% B, 16% CRF02_AG, 2% A, 4% G and 10% other subtypes (Ghosn *et al.*, 2011). The five patients detailed here were infected with three subtypes, two with B, two with CRF02_AG and one with G, as determined by the REGA subtyping tool (de Oliveira et al., 2005). The patients' viral loads at time of failure ranged between 342 and 25,300 RNA copies/ml. In addition, the five patients experienced virological failure at different times spanning the length of the trial – ranging from week 24 (failure to achieve the primary endpoint of viral load <400 copies/ml by week 24) to week 96 (the total length of trial follow up), as shown in table 5.1 and figure 5.2.

Therapy adherence data from compliance assessments carried out at nine time points and LPV trough concentrations measured in the blood plasma at three

time points, are presented in table 5.2 and figure 5.2, respectively. Analysis of adherence data and LPV trough concentrations at the time of treatment failure, when available, enabled an assessment of the potential contribution of suboptimal plasma drug concentrations to treatment failure. For this study, LPV/r trough concentrations below 3,000 ng/ml were considered suboptimal (Solas et al. 2004).

Table 5.1. Clinical and virological information for the five patients included in this study.

Patient Number	Subtype	Screen VL (copies/ml)	Failure time (weeks)	Fail VL (copies/ml)
1403 (KON)	CRF02_AG	44,600	24	24,000
3204 (HG)	B	23,800	40	603
1404 (DIO)	CRF02_AG	166,000	42	1080
4201 (SO)	G	79,500	48	342
508 (SP)	B	37,800	96	25,300

Table 5.2. Self-reported adherence to trial medication by week.

Patient #	Adherence report from compliance assessment (by week)								
	4	8	12	16	20	24	32	40	48
1403 (KON)	0	0	0	0	-	1	-	-	-
3204 (HG)	0	0	0	0	0	0	0	0	1
1404 (DIO)	0	0	0	0	1	0	0	1	1
4201 (SO)	0	0	0	0	0	1	0	0	1
508 (SP)	0	0	0	0	1	0	0	0	0

Table shows the self-reported adherence of each patient, as assessed by compliance evaluation at each study visit. Input was either reported as no doses of therapy missed (0) or at least one dose missed (1). For some patients, data are missing (-)

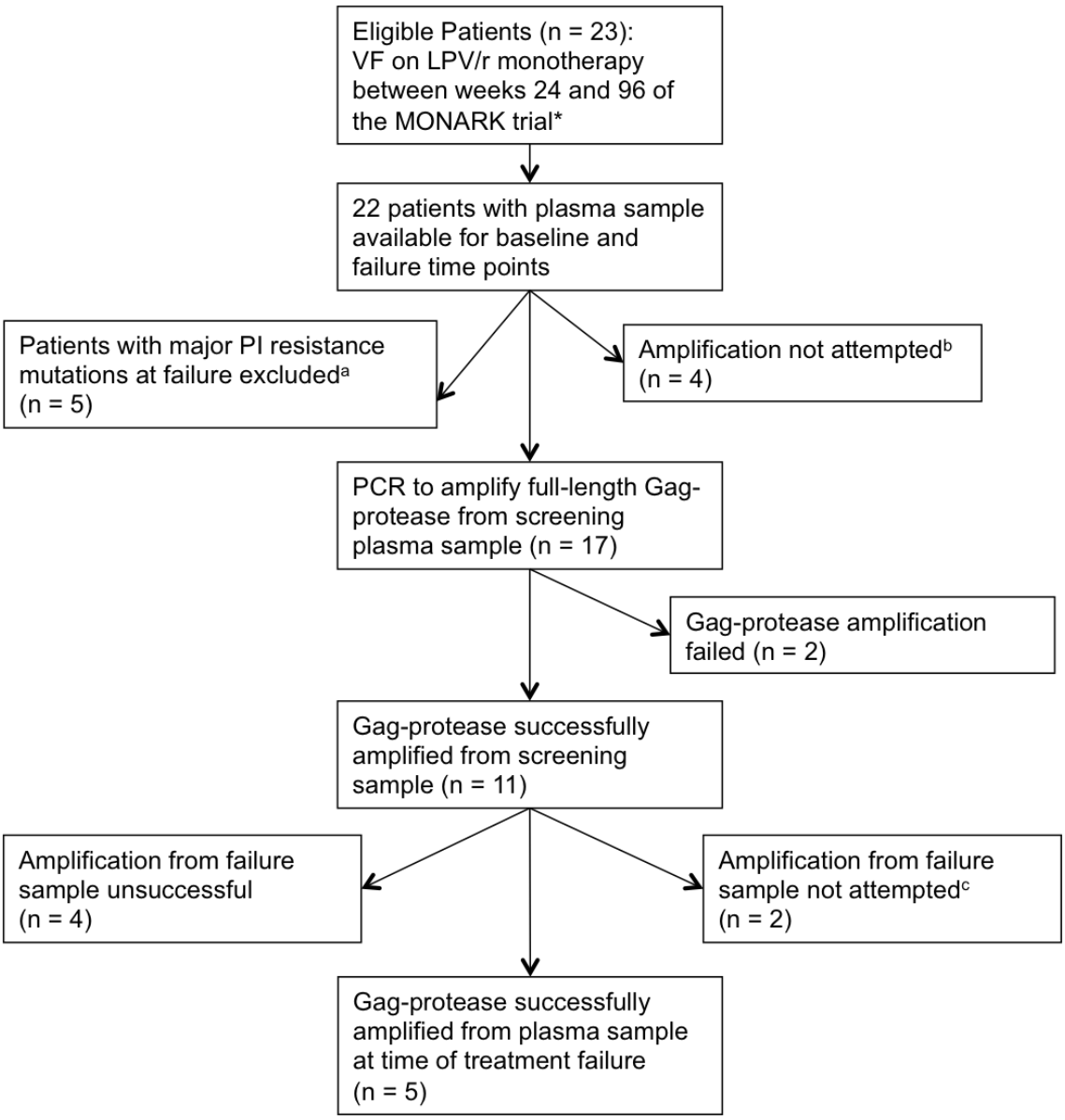


Figure 5.1. Selection criteria for patients included in this study. The patients included in the study are shown, with the reasons stated for those that experienced virologic failure on LPV/r monotherapy, but were excluded. *As defined by trial primary endpoints. [a] As detected by clinical testing during the trial and investigated by Delfraissy *et al.*, 2008. [b] Patients with a subtype present at low frequency in the MONARK trial (A1 and CRF01_AE) or VL < 200 copies/ml at failure were excluded. [C] Viral load < 200 copies/ml, below limit of detection for our PCR.

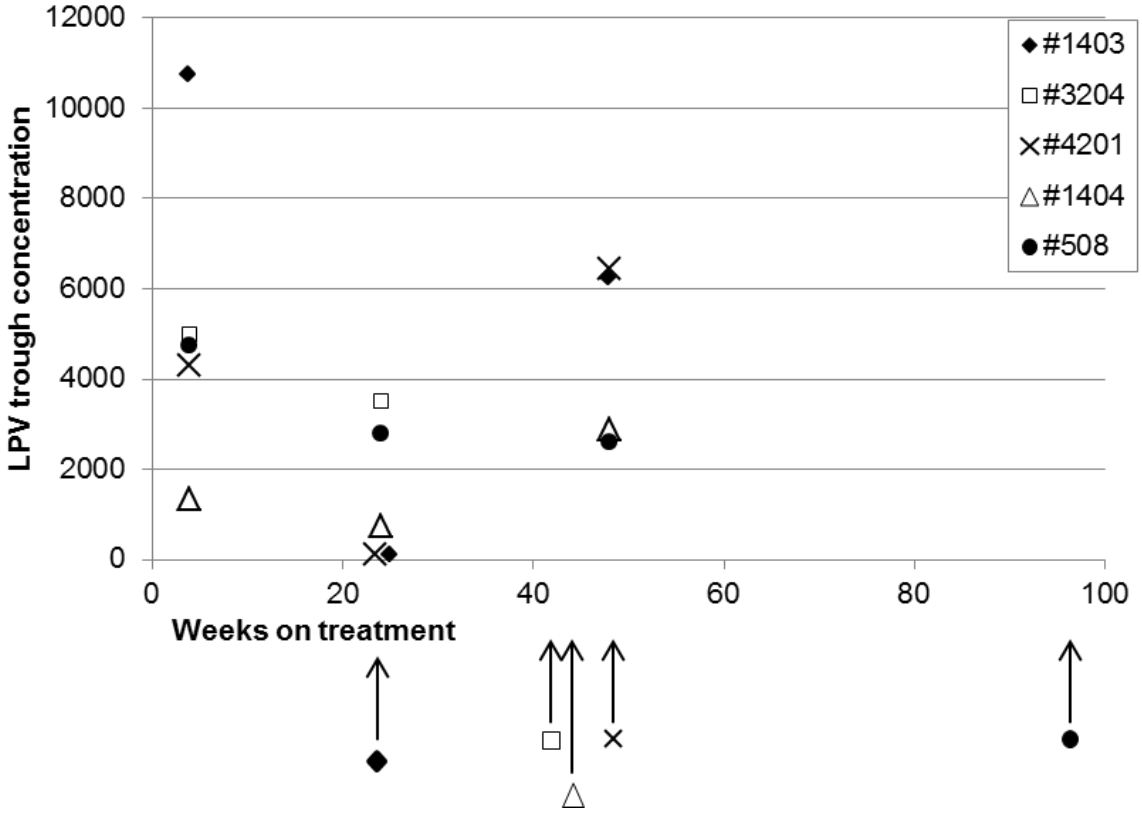


Figure 5.2. LPV trough concentrations and time to treatment failure for each patient included in this study. The plasma LPV trough concentrations (ng/ml) at intervals since the start of therapy (measured in weeks) are shown, with assessments performed at weeks 4, 24 and 48 after treatment initiation. LPV concentrations are shown for each of the five patients experiencing virological failure included in this study and the time of virological failure is represented by the symbols and arrows below the X axis. The minimum therapeutic dose is considered to be 3000 ng/ml, hence at least 3 patients exhibited sub-optimal LPV drug concentrations in their plasma, particularly at week 24.

5.2.2 Clonal analysis of Gag-protease from patients experiencing treatment failure

Clonal analysis of viral variants from each patient at each time point was performed to identify amino acid mutations in Gag or protease that have been previously described to affect PI susceptibility. In addition, comparison of sequences at baseline and failure could enable the identification of mutations associated with PI exposure. Following amplification of full-length Gag-protease by cDNA synthesis and nested PCR, the PCR product was cloned into the pGEM vector and transformed into JM109 competent cells. Colony PCR to identify clones containing full-length Gag-protease was performed, colonies were cultured and plasmid DNA purified and sequenced. To represent the quasispecies within a patient at each time point, sequence analysis of at least 10 clonal variants for each sample was performed. Viral variants were numbered 1-10 for each sample, and the sequences labelled S (screening sample at baseline) or F (failure time point sample). Protease sequences were submitted to the Stanford HIVdb Resistance Algorithm for analysis and compared to the IAS list of Resistance Mutations (Johnson et al., 2013; Liu and Shafer, 2006). The results of the protease analysis are shown in table 5.3. Gag sequences were compared to table 1.2 containing the mutations in Gag previously reported to be associated with PI resistance and/or PI exposure, both *in vitro* and *in vivo* (see appendix). The transframe-protease changes D5N, L9P, K12E and N55S/D/V/E were also considered, but they were present at such high frequency in these patients that no correlation with phenotype was possible and hence they were not considered further (McKinnon et al., 2011).

Patients #1403 (KON) and #1404 (DIO) were both infected with CRF02_AG viruses and patient #4201 with a subtype G virus. Viruses from these patients at both time points shared a number of Gag mutations and protease polymorphisms, although differences between the patients were present. In particular Gag mutations E12K and R76K were present in viral variants from all

three patients at both time points, along with Y79F which was present in only a few viral variants from each patient. Of note, Y79F appeared at time of treatment failure correlating with PI exposure, in failure variants 6 and 9 from patient #1403 (KON). Significant amino acid variation was present in the p2/NC Gag cleavage site – in particular amino acid positions 370-376. V370A was present in viral variants from both time points in patients #1404 (DIO) and #4201 (SO) and T371Q in variants from patient #1404 (DIO). A deletion of position 373 was present in viruses from patients #1403 (KON) and #1404 (DIO), and in variants from patient #4201 (SO) the S373A mutation was present. In patients #1403 (KON) and #4201 (SO), baseline variants contained amino acids A and T at position 374, but only T was present at treatment failure. There was also a change at this position in patient #1404 (DIO) to N. Viral variants derived from patient #1403 (KON) contained two additional previously reported mutations – T375N and I376V. Of interest I376V was present in the viral variants from patient #1403 (KON) at baseline, but had reverted to consensus amino acid by the time of treatment failure in variants 1, 2 and 6. Previously described mutations in the p1/p6 cleavage site were also present with L449P in variants from both time points for patients #1403 (KON) and #1404 (DIO), and S451N in patient #4201 (SO).

Of particular importance, the V82A major PI resistance mutation was detected in patient #1403 (KON) in failure variant 6 which had not been detected by previous genotypic resistance testing performed as part of clinical care during the trial. Viral variants derived from both screening and failure samples for patient #1404 (DIO) contained the A71T minor resistance mutation that was analysed phenotypically in chapter 4. No PI resistance mutations were present in viral variants derived from patients #1403 (KON), #1404 (DIO) and #4201 (SO) at either time point, however the subtype CRF02_AG and G viruses contained a number of protease polymorphisms. Changes K20I, M36I and H69K were present in viral variants from all three patients, of which K20I and

M36I have previously been associated with PI exposure and treatment failure (Baxter et al., 2006; Bertoli et al., 2006; Descamps et al., 2009). Changes at position 13 and 89 of protease were present in viral variants from all three patients. In addition, the R41K polymorphism was present in variants from patients #1404 (DIO) and #4201 (SO). Each of the protease polymorphisms present represent consensus amino acids in recombinant CRF02_AG and subtype G viruses (Kuiken et al., 2012).

Whilst a number of changes present in variants from patient #1403 (KON) correlated with PI exposure (Gag Y79F and protease V82A), none of the changes present in Gag and protease in viral variants from patients #4201 (SO) and #1404 (DIO) correlated fully with PI exposure.

Patients #3204 (HG) and #508 (SP) were infected with subtype B viruses. The R380K mutation was present in viral variants from both patients at both time points, and the S451N p1/p6 CSM was present in a minority of viral variants from both patients. For patient #508 (SP), the presence of these two changes only in variants from the failure sample correlates with their development following PI exposure. As observed for the recombinant CRF02_AG and subtype G viruses, variation in the p2/NC cleavage site of the subtype B viruses was present with changes at position 370 to A in patient #508 (SP) variants and L in those from patient #3204 (HG). In addition, viral variants from patient #3204 (HG) contained the previously reported T375N and a minority of screening variants contained S373P and I376V. Of note, these previously described mutations at positions 373 and 376 were not present at treatment failure and hence do not correlate with PI exposure. This is also true for the PTAP duplication present as a mixture at baseline but not present in the failure variants of patient #3204 (HG). Conversely, the T81A change does appear at time of treatment failure in all variants from patient #3204 (HG).

The protease polymorphism L63P, which was phenotypically examined in chapter 4 of this thesis, was present in viral variants derived from both patients #3204 (HG) and #508 (SP) at both time points. In addition, viral variants from both patients contained V771I and I93L changes in protease. Of note, patient #3204 (HG) had the M36I polymorphism present in all failure variants, and the mutation has previously been associated with PI resistance when present with other changes in subtype B viruses (Baxter *et al.*, 2006).

Table 5.3. Drug resistance mutations and polymorphisms present in protease of patient-derived viruses.

Patient	Protease Inhibitor Resistance mutations[+]		Protease polymorphisms[+]
	Major	Minor	
#1403 (KON)	V82A (F6)		I13V, K14R, G16E, K20I, E35D, M36I, H69K, K70R, L89M
#508 (SP)			T31I (F1), N37S, P39Q, S47N (S2), Q49L (S5), I62V, L63P, V77I, I93L
#1404 (DIO)		A71T	I13A, K20I, M36I, R41K, I64V (S1,2,3,4,6,7,9), H69K, I72M (F1-10), L89I, I93M
#3204 (HG)			E35D, M36I (F1-10), N37S (S1,2,3,4,5,7,8,10), L63P, I72V, V77I, I93L
#4201 (SO)			I13V, K14R, K20I, E35Q, M36I, R41K, G52R (S8), R57K, Q61N, Q61D (S7), I64M (S1,2,3,4,5,6,8,9,10), C67E, H69K, I72V (S2,4,6), V82I, L89M

Protease sequences were assessed for resistance mutations and polymorphisms using the Stanford Resistance algorithm. Major and minor resistance mutations are shown, along with any polymorphisms in comparison with a group M consensus sequence. Changes reported are present in all viral variants from a patient, except where specific variants are shown in brackets after the change.

5.2.3 Positive selection analyses

The viral variants from each patient were subjected to selective pressure testing to identify evidence of positive selection as a result of PI exposure using various methods available on www.datamonkey.org (Delport *et al.*, 2010). All viral variants from a single patient, derived from both screening and failure samples, were aligned in MEGA software using the ClustalW algorithm and nucleotides 1450-1 removed to ensure Gag and protease were in a single reading frame. Aligned nucleotide sequences were submitted to datamonkey in FASTA format. To test for genome wide evidence of selection the PARRIS method (Scheffler *et al.*, 2006) was used and for identification of particular sites under diversifying/purifying selection the FUBAR method (as discussed by Kosakovsky Pond and Frost, 2005) was selected.

The PARRIS method detected evidence of genome-wide positive selection at $P < 0.1$ in patients #3204 (HG), #508 (SP) and #1404 (DIO). No evidence of positive selection was reported in alignments for patients #1403 (KON) and #4201 (SO). The sites identified by the FUBAR analysis as undergoing positive selection within each patient are shown in table 5.4, along with the amino acid present at that position in each of the viral variants. In patient #1404 (DIO), no sites were identified as undergoing positive selection, but in the remaining patients at least one amino acid position was identified. For two patients, #1404 (KON) and #508 (SP), none of the amino acid changes identified by the analysis correlated with PI exposure. For example for patient #1404 (KON) at position 109 the N/K mix was present in variants from both baseline and failure time points, indicating that selection at this amino acid position as a result of PI exposure was unlikely.

Most of the amino acid positions identified by the analysis in patient #4201 (SO) contained a mixture of amino acids in the screening variants but at failure a single amino acid was present in all variants. For example position 75 was an

I/L mix at baseline, but at failure all viral variants contained L. At position 418 viral variants were R/Q mix at baseline, but at failure all variants contained K.

In viral variants derived from patient #3204 (HG) two amino acid positions of particular interest were identified by the positive selection analysis – position 12 of Gag and 36 of protease. The E12K mutation in Gag has been previously reported following *in vitro* PI exposure, and in this patient a mixture of K and Q was present in baseline variants but K was present in 9/10 failure variants (Aoki et al., 2009). Protease mutation M36I appeared in all variants from this patient at time of treatment failure, but was not present at baseline. M36I has been previously associated with PI resistance in subtype B viruses, and in particular to the PI TPV/r (Baxter *et al.*, 2006).

Of particular note, a significant number of the amino acid positions identified by positive selection analyses were located within the MA subunit of Gag – position 75 and 91 in patient #4201 (SO), 15 and 55 in patient #508 (SP) and positions 12, 63 and 70 for patient #3204 (HG).

Table 5.4. Positively selected sites in Gag and protease (next page). Full-length Gag-protease sequences from all viral variants isolated at baseline and failure for each patient were used. Positive selection analysis was performed using the FUBAR method available on the Datamonkey website tool. Positively selected sites are shown, along with the amino acid residue present in each viral variant at that site. Positively selected sites are numbered according to the patient sequence, with the HXB2 numbering shown in brackets where different.

Table 5.4. Positively selected sites in Gag and protease (legend on previous page)

Patient #	Patient Initials	Positively selected site+	Screening Variants										Failure Variants									
			1	2	3	4	5	6	7	8	9	10	1	2	3	4	5	6	7	8	9	10
1403	KON	28	Q	Q	Q	Q	Q	Q	Q	Q	Q	Q	R	R	Q	Q	R	R	Q	Q	R	R
		34	L	L	L	L	L	L	L	L	L	L	I	I	I	L	I	L	L	L	L	I
		109	N	N	N	N	K	K	N	N	K	N	N	N	N	K	K	N	N	K	K	K
1404	DIO	None																				
4201	SO	75	L	I	I	L	L	I	L	I	I	L	L	L	L	L	L	L	L	L	L	L
		91	R	K	K	R	R	K	R	R	K	R	R	R	R	R	R	R	R	R	R	R
		107	I	I	I	R	R	R	R	I	I	I	I	I	I	I	I	I	I	I	I	I
		215	I	L	I	I	I	I	L	I	I	I	L	L	L	L	L	L	L	L	L	L
		418	R	R	Q	R	R	R	Q	Q	Q	R	K	K	K	K	K	K	K	K	K	K
508	SP	15	Q	Q	Q	Q	R	Q	Q	Q	Q	Q	K	K	T	Q	T	Q	Q	S	T	Q
		55	I	I	G	G	E	G	A	A	G	A	D	D	A	A	A	I	A	A	A	D
		121	D	D	D	D	G	G	G	D	D	G	G	G	G	G	G	D	G	G	G	D
		312	D	D	D	D	D	D	E	E	E	D	E	E	D	D	D	D	D	D	D	D
		403	K	R	K	K	K	K	K	R	K	K	K	K	K	K	K	K	R	R	K	K
		480	D	D	N	N	D	D	D	D	N	D	D	D	D	N	D	D	D	D	D	N
3204	HG	12	K	Q	K	K	Q	Q	K	K	K	K	K	Q	K	K	K	K	K	K	K	K
		63	Q	H	H	Q	H	H	Q	Q	H	Q	Q	Q	Q	Q	Q	Q	Q	Q	Q	Q
		70	T	A	A	T	A	A	T	T	A	T	T	T	T	T	T	T	T	T	T	T
		81	A	T	T	T	T	T	T	T	T	T	A	A	A	A	A	A	A	A	A	A
		313 (310)	Q	Q	Q	Q	Q	Q	Q	Q	P	Q	S	Q	S	S	S	S	S	S	S	S
		378 (376)	I	I	V	I	I	I	V	I	V	V	V	I	V	V	V	V	V	V	V	V
		pro 36	M	M	M	M	M	M	M	M	M	M	I	I	I	I	I	I	I	I	I	I

5.2.4 Phylogenetic and genetic distance analysis

In order to further study the evolution of the viral variants within each patient, phylogenetic analysis was performed. Nucleotide sequences from both baseline and time of treatment failure were included for each patient. Alignment was performed in MEGA using the ClustalW algorithm and imported into the PHYLIP program for phylogeny construction using the Maximum Likelihood method under the GTR model of nucleotide substitution. In addition, ancestral reconstruction was performed using HyPhy software and inferred nucleotide sequences for each node used to identify amino acid changes correlating with treatment failure (Figure 5.3). Mean pairwise genetic distances (MPWGD) between the amino acid sequences of the viral variants were calculated in MEGA to compare genetic diversity at screening and time of treatment failure.

Figure 5.3 shows a maximum-likelihood tree constructed using all Gag-protease variants from both baseline and failure time points from each of the five patients studied here, with the confidence of the tree tested using 500 bootstrap replications. The amino acid changes correlating with treatment failure and the node from which they were identified is shown. Figure 5.4 and 5.5 show maximum-likelihood trees constructed using only the *gag* sequence and the protease sequence of each clonal variant, respectively. These trees enabled an examination of the relationship between the baseline and failure viral variants from each of the patients. In each tree, the sequences from each patient formed a monophyletic cluster confirming that no contamination between patient samples had occurred. As expected, the viruses grouped by subtype, with subtype B variants from patients #508 (SP) and #3204 (HG) forming a distinct cluster, subtype CRF02_AG viruses from patients #1404 and #1403 (DIO and KON) forming another with the subtype G viruses from patient #4201 (SO) being more closely related to the CRF02_AGs than the Bs. The branches separating the interpatient sequences were strongly supported with 100% bootstrap support values.

For phylogeny derived from Gag-protease, the screening and failure variants grouped separately on the tree for four patients; #1404, #1403, #4201 and #3204. The emergence of genetically distinct variants at failure indicated ongoing evolution of the viruses during PI monotherapy. For these patients, amino acid changes were identified using ancestral reconstruction that correlate with PI exposure and treatment failure. Mutations in Gag were identified in all four patients and changes in protease identified in #4201 (position 64) and #5-8 (M36I). The majority of positions identified by the ancestral site reconstruction are located in Gag, indicating that PI pressure results in the development of changes within Gag. However, for one patient #508 the baseline and failure viral variants were not phylogenetically distinct, indicating the re-emergence of genetically similar viruses at failure. The bootstrap values were low for the nodes between the branches within a single patient, but this is likely due to the close relatedness of the variants derived from a single patient. However in three patients, the branch lengths for the failure variants were much shorter than those for screening variants – patients #1404, #4201 and #3204. This indicated that viral diversity was reduced at time of treatment failure in comparison to baseline.

The emergence of genetically distinct viruses at failure is also supported by phylogeny based on Gag only, figure 5.4, and protease only, figure 5.5. In both trees separation of viruses at baseline and failure was present for patients #4201, #1404 and #3204.

The MPWGD values, showed that overall genetic diversity was significantly reduced at time of treatment failure in comparison to baseline (P = 0.0345, unpaired t test). In patients #1404 (DIO), #4201 (SO) and #3204 (HG) significant reductions in MPWGD were present at failure in comparison to baseline, from 0.014 to 0.001, 0.009 to 0.002 and 0.019 to 0.001 amino acid substitutions per site for baseline and failure time points, respectively. However, for patients #1403 (KON) and #508 (SP) no significant change in

diversity at baseline and failure was present. MPWGD values at baseline and failure were 0.007 and 0.009 for #1403 (KON) and 0.011 and 0.010 for #508 (SP).

We also investigated the relationship between genetic diversity and viral load. Correlation analyses of both factors including both failure and baseline samples showed no correlation between viral load and MPWGD, indicating that the difference in MPWGD is not an artefact of reduced viral load at failure ($R = 0.132$).

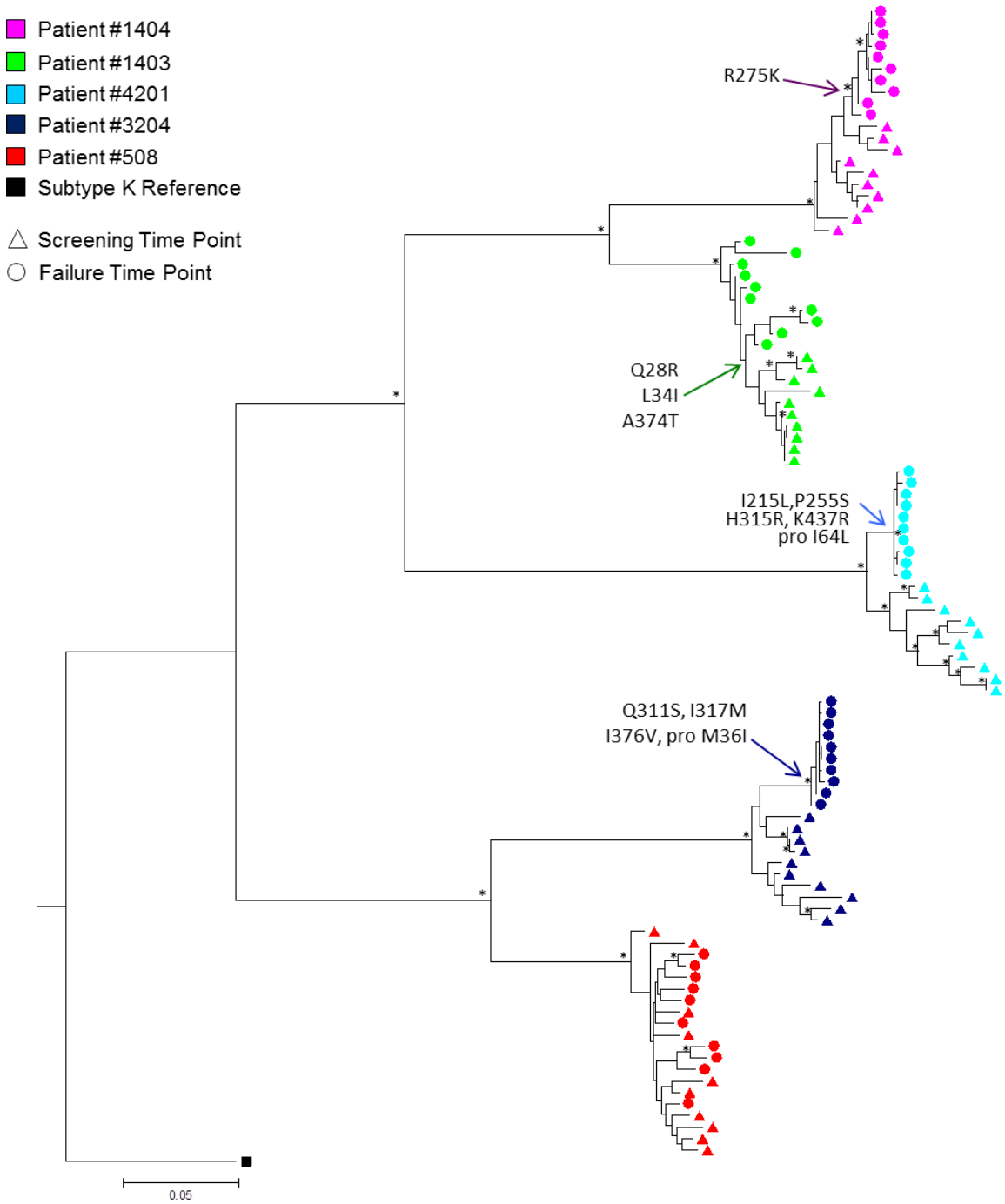


Figure 5.3. (Title and figure on next page)

Figure 5.3. Maximum-likelihood phylogeny of Gag-protease viral variants from baseline and failure time points. A maximum-likelihood tree constructed under the GTR model of nucleotide substitution in PhyML using an alignment of all variants from baseline and failure time points from five patients. Variants from each patient are represented colour, the screening variants by triangles and those from failure with circles. Nodes supported by >75% bootstrapping (at least 350/500) are marked by an asterisk(*). The tree shows the emergence of genetically distinct variants at failure, indicating ongoing replication during PI exposure in four of five patients. In a fifth patient, the re-emergence similar variants at the time of failure is in-keeping with poor adherence. Mutations correlating with treatment failure are marked at the node at which they were identified from the most recent common ancestor, inferred using ancestral reconstruction using HyPhy software.

Figure 5.4. Maximum Likelihood phylogeny of full-length Gag from five patients at baseline and time of treatment failure, derived by clonal analysis. A maximum-likelihood tree constructed under the GTR model of nucleotide substitution in PhyML using an alignment of all variants from baseline and failure time points from five patients. Variants from each patient are represented colour, the screening variants by triangles and those from failure with circles. As for the tree based on Gag-protease, there is evidence for the emergence of genetically distinct viruses at the time of treatment failure for a number of patients (4201, 1404 and 3204).

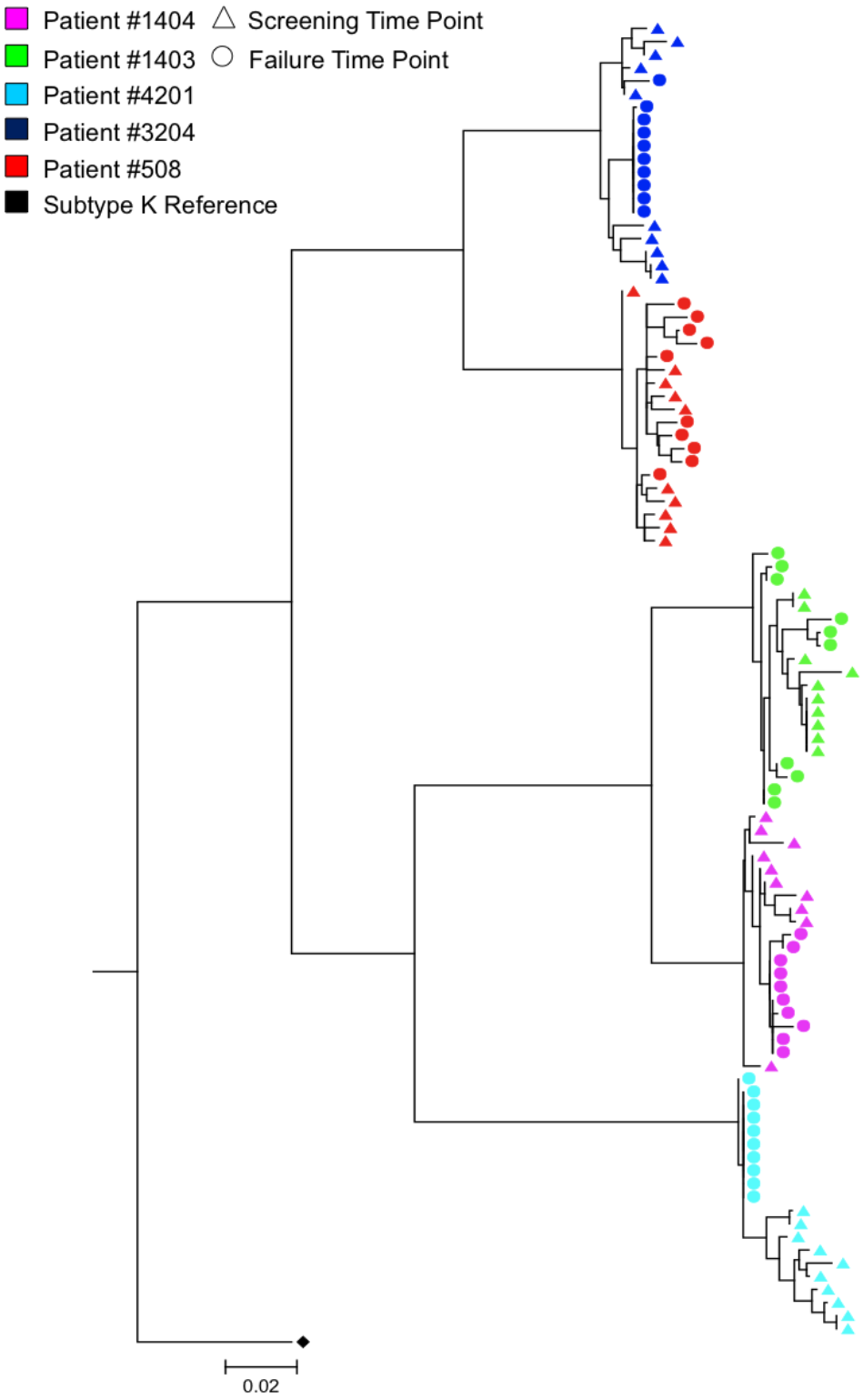


Figure 5.4. Maximum Likelihood phylogeny of full-length Gag from five patients at baseline and time of treatment failure, derived by clonal analysis. (Legend on previous page).

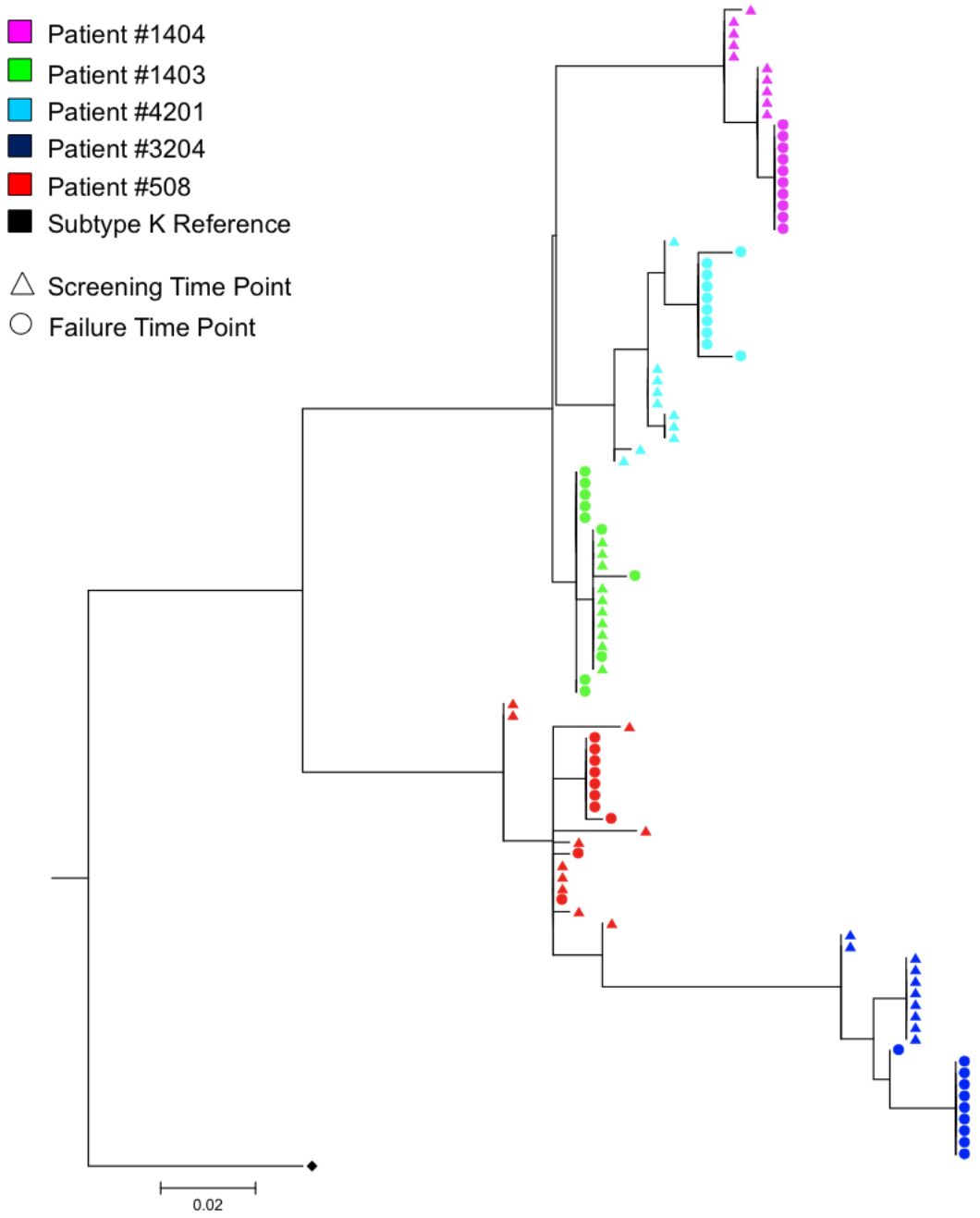


Figure 5.5. Maximum Likelihood phylogeny of the protease gene from five patients at baseline and time of treatment failure, derived by clonal analysis. Trees were derived using the same method as in figure 5.3 and 5.4. As for tree 5.4 based on Gag, there is evidence for the emergence of genetically distinct viruses at the time of treatment failure for three patients (4201, 1404 and 3204)

5.2.5 Phenotypic PI susceptibility and single-round infectivity of Gag-protease derived from patients receiving LPV/r monotherapy

To elucidate the mechanisms contributing to treatment failure in the absence of known PI resistance mutations, we determined the phenotypic PI susceptibility and single-round infectivity in comparison to reference strain p8.9NSX+ for viral variants representative of each patient sample at baseline and failure. Approximately six viral variants were chosen for each patient from both baseline and failure time points. At each time point the viral variant most similar to the consensus amino acid sequence was selected, defined as the variant with the least number of non-conserved amino acid changes from the consensus sequence when an identical variant was not present. Additional viral variants were selected based on the presence or absence of amino acids of interest identified in the genotypic analyses. Full-length *gag*-protease from the selected variants was cloned from the pGEM intermediate vector into the reference strain p8.9NSX+, using the *Not*I and *Xho*I restriction sites. The phenotypic susceptibility to the PIs ATV, DRV, LPV and SQV, and the single-round infectivity were determined for each variant using single-cycle assays, as previously described (see materials and methods).

For patient #1403 (KON) seven viral variants in total were subjected to phenotypic testing, two derived from the baseline sample (S1, S5) and five from the time of treatment failure (F2, F5, F6, F7, F9), as shown in figure 5.6. Of note the F5 variant containing the major PI resistance mutation V82A displayed significantly reduced susceptibility to LPV of 17-fold. For the remaining variants, reduced susceptibility in comparison to the reference strain was observed to both ATV and LPV of up to 10 fold-difference in EC_{50} for S1, F2 and F5. However, F6 and F9 displayed susceptibilities similar to the assay reference strain to ATV and LPV, with only up to four fold-difference in EC_{50}. Susceptibility levels similar to that of the reference strain were observed for all variants from patient #1403 (KON) to PIs DRV and SQV. All viral variants displayed reduced

single-round infectivity in comparison to the reference strain of between 60% and 90% (figure 5.7). Two viral variants in particular displayed further reductions to only 30% that of the reference strain, these being F6 and F9.

A total of seven viral variants were subjected to phenotyping for patient #3204 (HG), five from baseline (S1, S2, S6, S8, S9) and two from the time of treatment failure (F1, F2). Figure 5.6 shows that for this patient, the failure variants displayed greater reductions in susceptibility to PIs ATV, DRV and LPV than the screening variants. This was most pronounced for PIs ATV and LPV where up to 16-fold and 13-fold reduction in susceptibility in comparison to assay reference strain was observed, approximately two-fold higher than the reduction observed for screening variants of up to 8- and 7-fold, respectively. Reduced susceptibility to a lesser extent was observed for DRV, of up to 5-fold difference for the failure viral variants. For the PI SQV reduced susceptibility was observed for both screening and failure variants of up to 9-fold for S6. Each of the variants displayed reduced infectivity of between 40-80% that of the reference strain, as shown in figure 5.7.

Six viral variants in total were tested for patient #1404 (DIO), four from baseline (S4, S5, S6, S8) and two from the failure time point (F1, F4). For this patient, reduced PI susceptibility to ATV and LPV were observed with the screening variants and F1, which displayed similar fold-differences in EC_{50} values of up to 11- and 8-fold, respectively, as shown in figure 5.6. Screening variants remained susceptible to the PI DRV, although variant F4 displayed a 5-fold reduction in DRV susceptibility. Reduced susceptibility to SQV of up to 5-fold was observed for screening variants and up to 9-fold for failure variants. Each of the viral variants displayed a reduced single-round infectivity in comparison to the assay reference strain. Figure 5.7 shows that the single-round infectivity for the screening variants ranged from 30-70%; however, variant F4 displayed a pronounced reduction to only 5% that of reference strain.

For patient #4201 (SO) five variants in total were analysed phenotypically: four variants derived from baseline (S1, S3, S4, S5) and one from failure (F2), as shown in figure 5.6. The baseline viral variants S1 and S5 displayed significant reductions in susceptibility to a number of PIs, most notably to LPV where at least 16-fold reduction in susceptibility in comparison to reference strain was observed (figure 5.6). Reduced susceptibility to ATV (up to 11-fold), DRV (up to 5-fold) and SQV (up to 4-fold) were also observed. This reduction in susceptibility was also observed for the viral variant derived from the failure time point, F2. The other variants derived from the baseline time point, S3 and S4, exhibited less pronounced reductions in susceptibility of 8- and 11-fold to LPV, respectively. All variants derived from patient #4201 (SO) exhibited pronounced reductions in single-round infectivity in comparison to the reference strain of 14-22% (figure 5.7).

For patient #508 (SP), phenotypic susceptibility was measured for six variants in total – three each from baseline (S1, S2, S3) and the time of treatment failure (F2, F3, F9). Figure 5.6 shows that modest reductions in susceptibility to the PIs ATV and SQV were seen of up to 7- and 5.5-fold, respectively. For LPV, most viral variants remained susceptible with up to 4-fold reduction in EC_{50} in comparison to assay reference strain. In addition, viral variants were fully susceptible to the PI DRV. Viral variants isolated from patient #508 (SP) displayed similar RCs to the reference strain, ranging from 80-120% (figure 5.7).

For each patient the fold-difference in EC_{50} values of the viral variants most similar to consensus at baseline and failure were compared using t tests. In each case there were no statistically supported differences in susceptibility at baseline and treatment failure within a patient. However, the trend was towards reduced susceptibility at failure in comparison to baseline – in particular for patient #3204 (HG). Comparison of LPV susceptibility of HG variants S9 and F1

by t-test returned a P value of 0.068. Analysis for each patient using EC_{90} data showed the same pattern as observed for the EC_{50} data, as shown in figure 5.9.

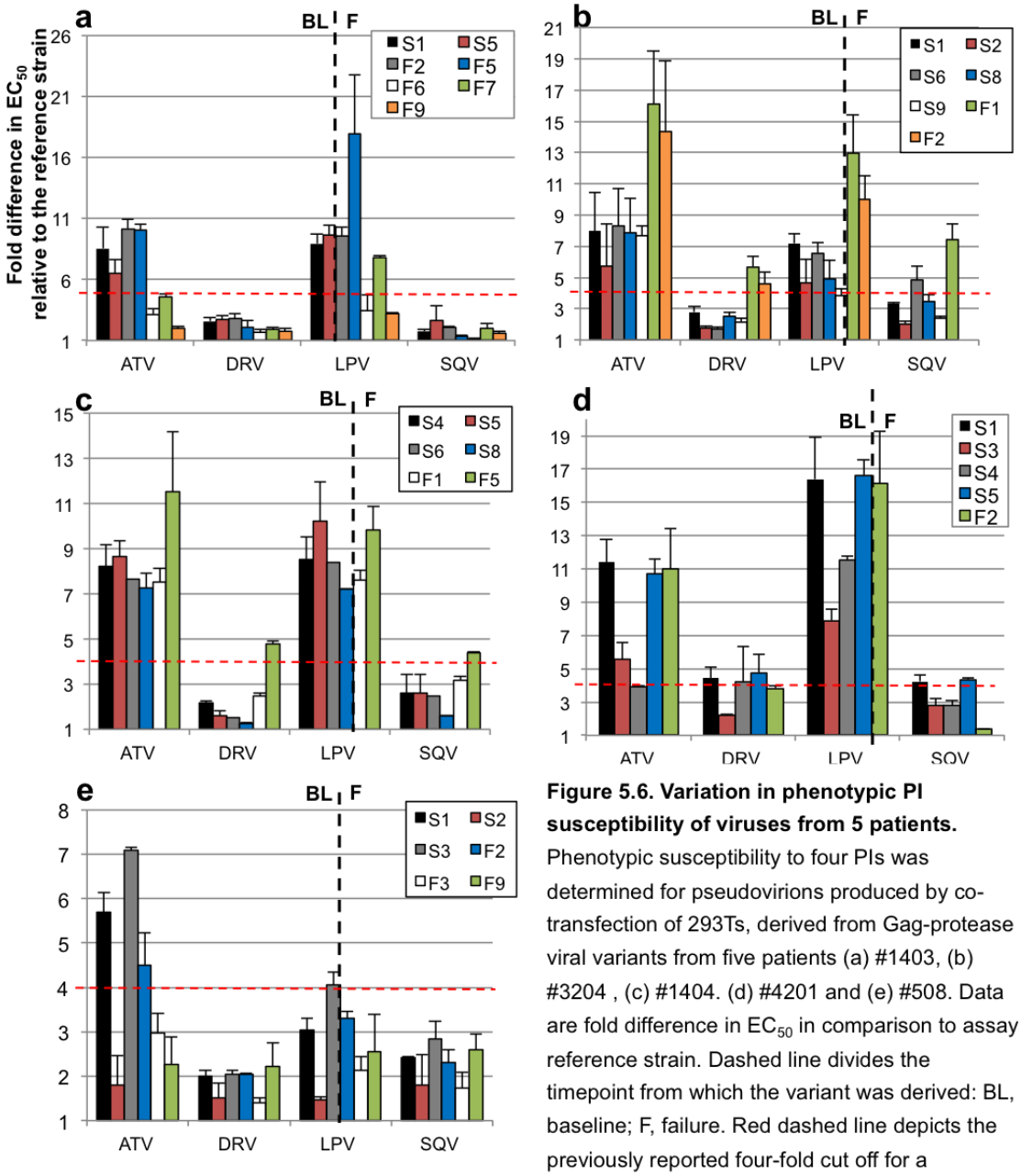


Figure 5.6. Variation in phenotypic PI susceptibility of viruses from 5 patients. Phenotypic susceptibility to four PIs was determined for pseudovirions produced by co-transfection of 293Ts, derived from Gag-protease viral variants from five patients (a) #1403, (b) #3204 , (c) #1404. (d) #4201 and (e) #508. Data are fold difference in EC_{50} in comparison to assay reference strain. Dashed line divides the timepoint from which the variant was derived: BL, baseline; F, failure. Red dashed line depicts the previously reported four-fold cut off for a significant reduction in susceptibility (Gong et al., 2004). Error bars represent the standard error of the mean of two independent experiments performed in duplicate.

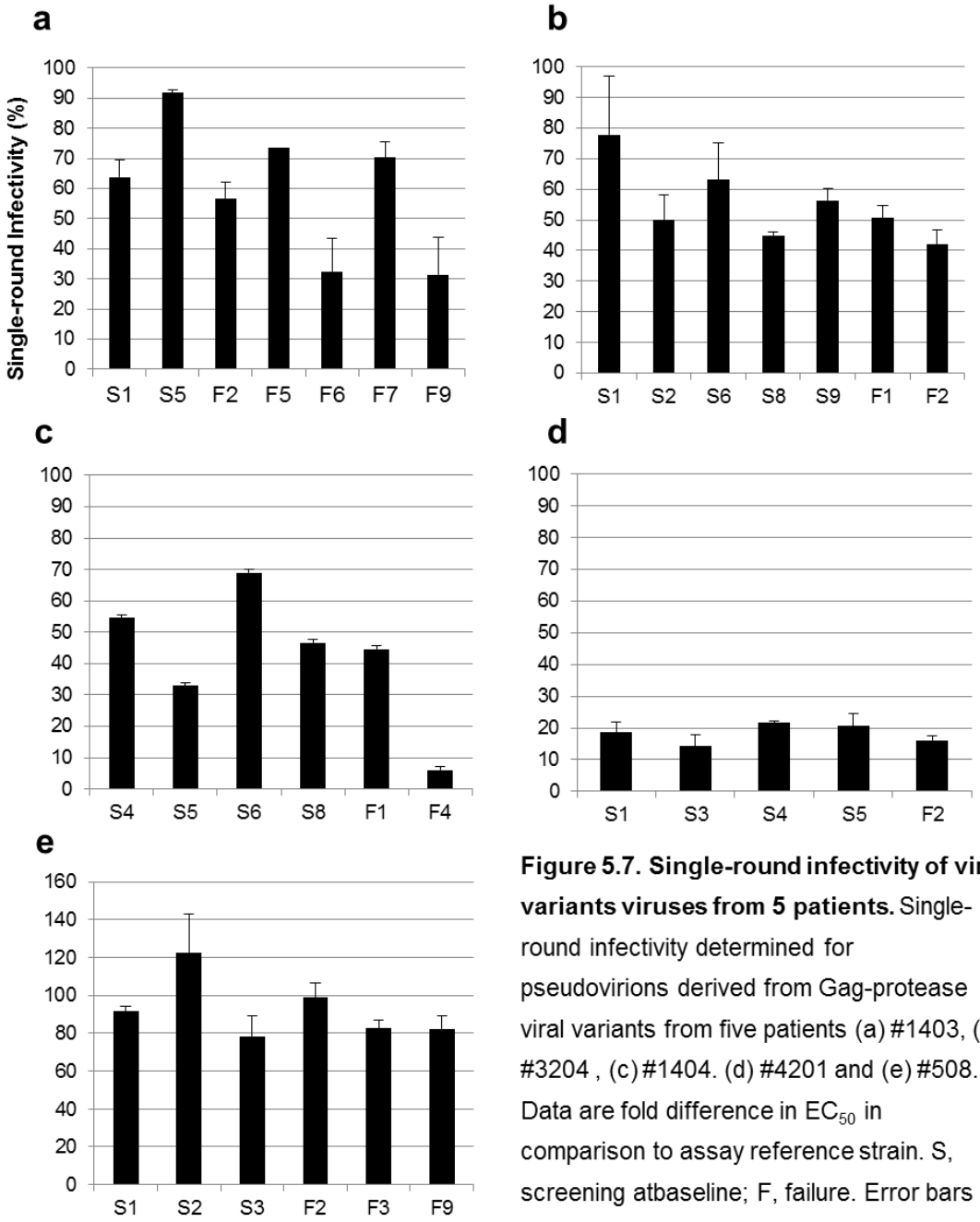


Figure 5.7. Single-round infectivity of viral variants viruses from 5 patients. Single-round infectivity determined for pseudovirions derived from Gag-protease viral variants from five patients (a) #1403, (b) #3204 , (c) #1404. (d) #4201 and (e) #508. Data are fold difference in EC_{50} in comparison to assay reference strain. S, screening atbaseline; F, failure. Error bars are standard error of the mean of two independent repeats performed in duplicate

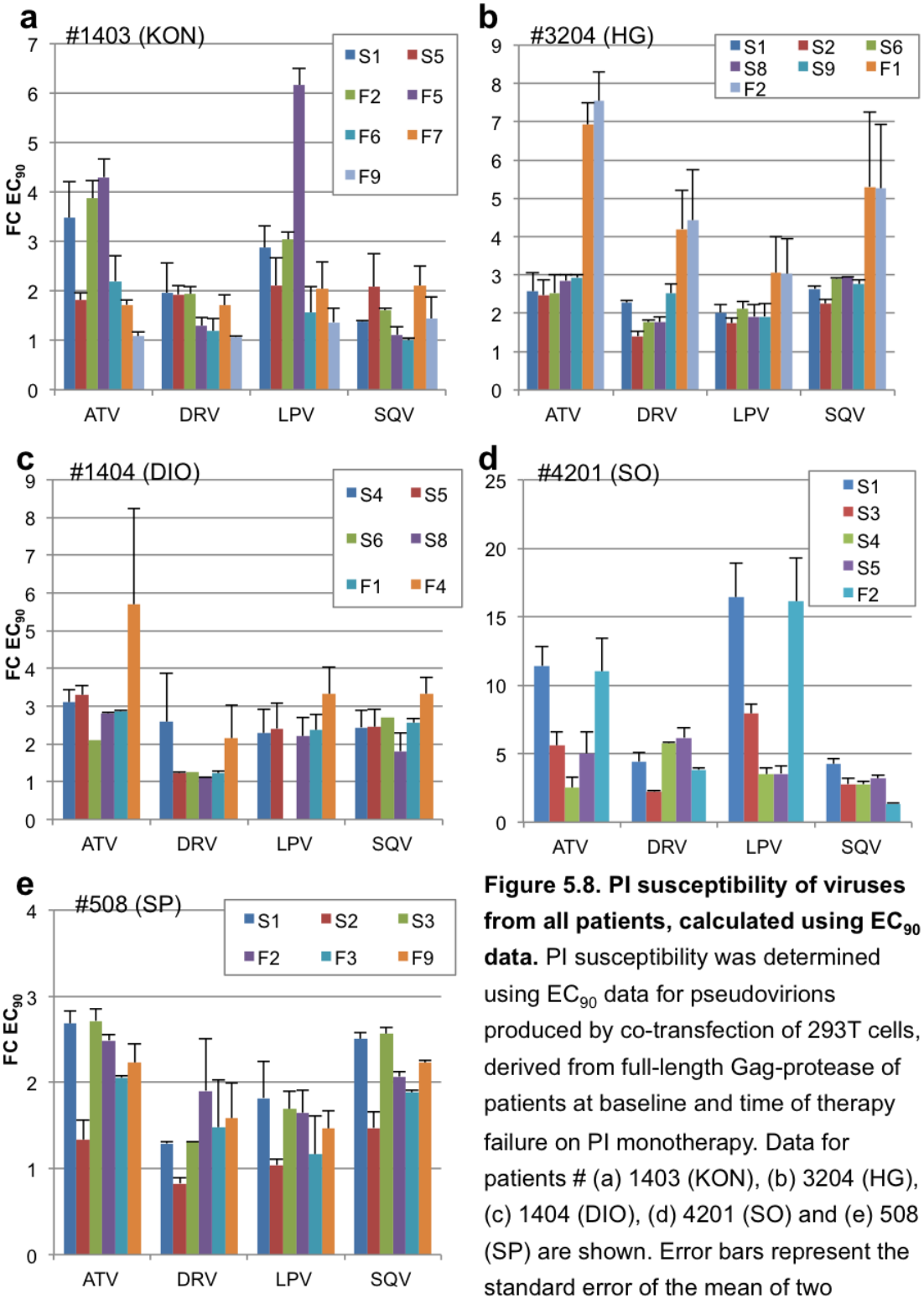


Figure 5.8. PI susceptibility of viruses from all patients, calculated using EC_{90} data. PI susceptibility was determined using EC_{90} data for pseudovirions produced by co-transfection of 293T cells, derived from full-length Gag-protease of patients at baseline and time of therapy failure on PI monotherapy. Data for patients # (a) 1403 (KON), (b) 3204 (HG), (c) 1404 (DIO), (d) 4201 (SO) and (e) 508 (SP) are shown. Error bars represent the standard error of the mean of two independent experiments performed in duplicate.

Table 5.5. Summary of clinical and virological information and data for each patient (legend on next page)

Patient	Subtype	VL Failure	MPWGD at failure	Mean LPV susceptibility		Mutations correlating with treatment failure
				Baseline	Failure	
#1403 (KON)	CRF02_AG	24,000	0.009	9.28	8.39	Gag Y79F, V186I, A375T, E452K, protease V82A
#3204 (HG)	B	603	0.001	5.43	11.48	Gag E12K, T81A, Q311S, M317I, I376V, D432N, D470N, protease M36I
#1404 (DIO)	CRF02_AG	1080	0.002	8.53	8.74	Gag H144Y
#4201 (SO)	G	342	0.001	13.11	16.1	Gag P255S, A339P, R461C, protease M64L
#508 (SP)	B	25,300	0.010	2.86	2.66	Gag R380K and S451N (each in one variant)

Table 5.5. Summary of clinical and virological information and data for each patient. A summary of the data generated in this chapter is shown in this table. MPWGD was calculated in MEGA 4.0 software using alignments of all viral variants from a each patient. LPV susceptibility was measured using VSV-g pseudotyped virions derived from patient Gag-protease in a single replication-cycle phenotypic susceptibility assay. The mean LPV susceptibility was calculated from the susceptibility data of all viral variants tested from the patient at either the baseline or failure time points. Mutations identified by ancestral site reconstruction, performed using a maximum parsimony approach, positive selection analysis and analysis of alignments by eye as correlating with treatment failure are shown.

5.3 Discussion

Here we report the first detailed analysis of the determinants of treatment failure in a PI monotherapy clinical trial, as summarised in table 5.5. The MONARK clinical trial offered an opportunity to examine the effects of PI monotherapy in previously treatment and PI-naïve patients, as opposed to other PI monotherapy trials which concentrated on the simplification of regimens for patients with a suppressed viral load on HAART. To date, phenotypic analysis using the commercial Phenosense assay had only been performed on viruses derived from patients experiencing virological failure in the LPV/r monotherapy arm of the MONARK trial exhibiting major PI resistance mutations (Delaugerre *et al.*, 2009). This analysis showed only a mean 2.2-fold increase in EC_{50} at treatment failure in comparison to baseline. A second study has examined the Gag cleavage sites from all patients, but neither of these studies conclusively identified the determinants of treatment failure (Flandre *et al.*, 2009). Given the evidence for the importance of the inclusion of full-length Gag alongside its co-evolved protease in phenotypic assays, we set out to examine whether testing phenotypic PI susceptibility of full-length Gag-protease from patients failing therapy in the absence of PI resistance mutations could shed light on the causes of treatment failure. The paired nature of these samples, one taken before treatment and one at time of treatment failure, enabled a comparison of

PI susceptibility, single-round infectivity and amino acid sequence of viruses from a single patient before and after PI monotherapy.

As with the other subtype B viruses we have studied in chapters 3 and 4, here we have observed significant variation in PI susceptibility to PIs ATV and LPV of subtype B, CRF02_AG and G PI-naïve viruses derived from baseline time points when full-length Gag-protease is included in a single-cycle phenotypic assay. In total, full-length Gag-protease of numerous viral variants from 5 patients were tested, two infected with subtype B, two with CRF02_AG and one with subtype G viruses. This showed reduced susceptibility of up to 17-fold to the PIs ATV and LPV. To date, no other study has examined the PI susceptibility of subtype CRF02_AG and G viruses using an assay which encompasses full-length Gag-protease. However, these data are in keeping with other studies and our own data, in chapters 3 and 4, demonstrating significant PI variation in PI susceptibility of full-length Gag-protease of other viral subtypes (Gupta et al., 2010; Jinnopat et al., 2009; Parry et al., 2009).

Our data provide evidence for a number of factors that separately or in combination contribute to treatment failure amongst patients in the MONARK LPV/r monotherapy trial. The first of these was suboptimal LPV/r trough concentrations in the plasma and the data from the trial indicate that 3/5 patients had LPV/r trough concentrations below the 3,000 ng/ml considered the minimum effective therapeutic LPV concentration for at least one of the three time points where drug levels were measured. Previous analysis has shown that the mean LPV/r trough concentrations for patients who succeeded on monotherapy were 6466 ng/ml at week 4, 4518 ng/ml at week 24 and 5149 ng/ml at week 48 (Flandre *et al.*, 2009). At weeks 4 and 24 of the trial in particular, the five patients failing monotherapy included in this study had lower LPV/r levels than these values. In addition, each of the patients reported missing at least one dose during the course of the trial but as compliance assessment did not quantify the number of doses missed it was not possible to

gauge the percentage adherence for each patient. For example ‘1’ could indicate a single missed dose or no drug taken at all in the previous 4-8 weeks. Although studies on the minimum adherence required to suppress viral replication successfully during LPV/r monotherapy have not been performed, studies examining LPV/r based HAART have shown that reduction to 90-95% adherence does not significantly increase the risk of treatment failure (King et al., 2005; Shuter et al., 2007). However, as only one active agent is present during PI monotherapy in comparison to three for HAART it is likely that this may be more sensitive to poor adherence than HAART regimens.

Data from two patients in particular, #1403 (KON) and #508 (SP), indicated the significant contribution of poor adherence and/or suboptimal LPV trough concentrations to subsequent LPV/r monotherapy failure. Patient #1403 (KON) experienced therapy failure at the earliest possible point in the trial – failure to achieve the primary endpoint of a viral load <400 copies/ml by week 24 – with a high viral load of 24,000 copies/ml. Very low levels of LPV were present at time of treatment failure, with only 75 ng/ml LPV in plasma when 3,000 ng/ml is considered the minimum effective therapeutic level in plasma. Thus, we cannot exclude the possibility that the V82A mutation was present at low frequency at baseline but was not detected by clonal analysis. Given that at failure V82A did not dominate the viral variants, being present in only 1/10 clones, it seems unlikely that outgrowth of this pre-existing resistant variant caused treatment failure. However, a recent study reported PI resistance mutations at low frequency using ultra-deep sequencing in a minority of patients failing PI-based HAART that were not identified by Sanger sequencing (9/36 patients) (Lataillade et al., 2012). Next generation sequencing would aid this, as discussed in section 7.3.3, although the clinical significance of low level viral variants remains unknown (Lataillade et al., 2010).

Patient #508 (SP) also appears to have experienced treatment failure primarily due to poor adherence. The viral variants from this patient were the most

susceptible of all five failure patients at both baseline and time of treatment failure, with only up to 4-fold reduction in PI susceptibility to LPV. Treatment failure occurred with a high viral load of 25,300 RNA copies/ml after full viral suppression for 96 weeks. Up to 48 weeks, when drug adherence monitoring was performed, LPV trough concentrations were above or around the 3,000 ng/ml minimum and reported adherence was relatively good. In addition, there was no evidence of positive selection using the PARRIS method and amino acids identified by FUBAR as undergoing positive selection did not correlate with PI exposure. Phylogenetic analysis demonstrated that the variants present at failure were closely related to those at screening indicating the re-emergence of the baseline viruses without viral evolution during the 96 weeks on therapy. Taken together, there is no evidence for any other reason for failure than a slip in adherence towards the end of trial and other studies have shown significant reductions in adherence over time (Parienti *et al.*, 2013).

This study also provides evidence for a second factor in LPV/r monotherapy failure – the development of reduced PI susceptibility over the course of treatment in the absence of major PI resistance mutations in protease. Data from one patient in particular, #3204 (HG), provided evidence that reduced PI susceptibility developed over the course of LPV/r monotherapy in the absence of known PI resistance mutations and that this may have contributed to the subsequent treatment failure. Patient #3204 (HG) experienced therapy failure at week 40 despite demonstrating adequate LPV trough concentrations and reporting good adherence. Screening variants displayed reduced PI susceptibilities of up to 9-fold to ATV and LPV, but reduced PI susceptibilities of up to 16- and 13-fold were observed for ATV and LPV, respectively, for variants derived from the time of treatment failure.

Positive selection analysis using the PARRIS model provided evidence of evolution of the virus under PI pressure and a number of amino acid positions undergoing positive selection in both Gag and protease were identified,

including E12K and M36I. In addition, ancestral reconstruction identified four amino acid chnages correlating with PI exposure and treatment failure: Gag 311, 317, 376 and protease 36. Several of these mutations that have been previously described in relation to PI resistance – 376 in Gag and M36I in protease – hence they may have contributed to the reduced PI susceptibility observed at treatment failure. Recent studies have shown an association between M36I and ATV exposure and virological failure on LPV/r containing regimens (Dolling et al., 2013b; Santos et al., 2012). In addition, our data in chapter 4 have demonstrated that the presence of polymorphisms can confer reduced PI susceptibility in the absence of major resistance mutations. However, molecular modelling demonstrates that residue 36 is not located near the protease active site, so the mechanism by which it may affect PI susceptibility remains unknown, figure 5.9. E12K has been previously described as being associated with exposure to PI *in vitro* and shown to affect PI susceptibility and single-round infectivity when present in combination with other changes (Aoki et al., 2009).

In addition to E12K, a number of other changes at positions previously described to be involved in PI resistance are present at treatment failure within the MA subunit of Gag – E62D and T81A. Positive selection analyses additionally identified mutations at two other positions within MA, 63 and 70. Our data from chapter 3 showed that the N terminus of Gag, a region encompassing MA and the N terminus of CA, can confer reduced PI susceptibility and we hypothesise that changes in this region within patient #3204 (HG) are contributing to the reduced PI susceptibility observed and hence LPV/r monotherapy failure. Future work should investigate which of these changes directly affect susceptibility, and whether they are required in combination or can have an effect when present singly. In addition, the mechanism by which these mutations affect PI susceptibility should be explored, as discussed in chapter 7.

Whilst drug levels and the development of reduced susceptibility played a role in treatment failure in these patients, our data do provide evidence for a third factor in failure on LPV/r monotherapy; the contribution of reduced PI susceptibility conferred by full-length Gag-protease of viral variants at baseline. Data for one patient in particular, #4201 (SO), provided evidence that a significant reduction in PI susceptibility at baseline may contribute to treatment failure. The baseline variants for this patient displayed up to 16-fold reduction in susceptibility to LPV that was also present at treatment failure. This reduction in PI susceptibility was the same as was observed for variant F6 of patient #1403 (KON) which contained the V82A major PI resistance mutation. Patient #4201 (SO) failed at week 48 of the clinical trial, at which time a good LPV trough concentration was present in the plasma and self-reported adherence was good at most appointments. Positive selection analysis did not provide evidence of evolution during PI monotherapy, although phylogenetic analysis did show the emergence of a genetically distinct virus at treatment failure. Ancestral reconstruction identified a number of amino acid positions correlating with treatment failure: Gag positions 215, 255, 315, 437 and protease 64. None of these amino acid positions have been previously described and as the PI susceptibility is unchanged between baseline and failure variants, they do not appear to have contributed to treatment failure. Molecular modelling demonstrates that residue 64 is not located near the active site of the enzyme, so the mechanism by which it may affect PI susceptibility remains unknown, figure 5.9. Given that virological failure occurred with acceptable LPV/r plasma concentrations and with a low viral load, this indicates that the reduced PI susceptibility present at baseline may have enabled ongoing viral replication in the presence of PI monotherapy. The role of reduced PI susceptibility in treatment failure is further explored in chapter 6 of this thesis.

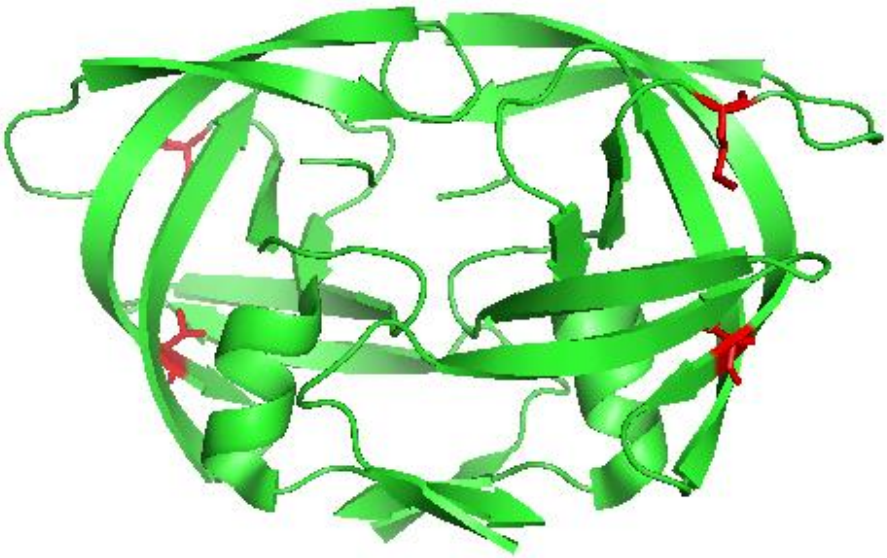


Figure 5.9. Location of residues 36 and 64 in the structure of the protease protein. Residues 36 and 64 were identified by ancestral site reconstruction as evolving under PI exposure, each in a single patient experiencing therapy failure on the MONARK trial. The location of these residues was mapped using PyMol software and the 3ELI protease structure from the RBSC PDB. Neither residue is located near the active site of the enzyme.

The use of phylogenetic analysis to study the evolution of the viruses within each patient also provides useful insight into the mechanisms of treatment failure in the PI monotherapy trial. In four patients #1403 (KON), #3204(HG), #1404 (DIO) and #4201 (SO), the failure variants are phylogenetically distinct from the screening variants indicating the presence of a genetic bottleneck resulting in the selection of one or a few variants at treatment failure. This provides evidence for the selection of a minority population at failure as a consequence of drug selective pressure and indicates the involvement of genetic determinants. This suggest the existence of genetic determinants enabling the emergence of variants under PI pressure, as identified by positive

selection analysis and ancestral reconstruction. Conversely, the failure variants from patient #508 (SP) are not phylogenetically distinct from the screening variants. This supports the hypothesis that poor adherence likely caused this patient to experience virological failure as there is no evidence of a genetic bottleneck before the emergence of virus at treatment failure.

Whilst the use of positive selection analysis is very informative in the study of the evolution of Gag-protease in these patient-derived viruses, the interpretation of the data are limited. As the viral loads at time of treatment failure were low for three of the five patients in comparison to screening, it is possible that some of the amino acids identified may be a result of reduced diversity and not truly undergoing positive selection. It is difficult to determine if this is the case, but a number of the amino acid positions identified here have been previously described in the literature as associated with PI resistance or exposure. Future work to examine whether these amino acid changes affect PI susceptibility directly will help to inform whether they are truly undergoing positive selection under PI-pressure.

We conclude that a number of factors, both virological and therapeutic, contributed to these patients experiencing virological failure on LPV/r monotherapy. Whilst suboptimal drug levels were obviously a contributing factor for therapy failure in some of these patients, either due to poor adherence or other biological factors, this study provides evidence that variation in PI susceptibility also contributed to treatment failure. This finding is novel, and of particular importance as it may explain PI therapy failure in the absence of major PI resistance mutations. Our data here, in particular for patient #3204 (HG), provide evidence that reduced PI susceptibility can develop in response to PI therapy in the absence of major PI resistance mutations, contributing to subsequent treatment failure. We also provide evidence that reduced susceptibility at baseline, particularly in patient #4201 (SO), may have contributed to subsequent therapy failure.

6 Genotypic and phenotypic comparison between patients achieving a sustained virological response and those experiencing virological failure on LPV/r monotherapy

6.1 Introduction

In chapter 5 we examined the determinants of therapy failure in five patients experiencing virological failure on LPV/r monotherapy as part of the MONARK clinical trial. MONARK was a randomized control trial to compare the efficacy of LPV/r monotherapy in comparison to LPV/r based HAART. After 48 weeks, inferiority of the LPV/r monotherapy regimen was reported with 23/83 (24%) patients experiencing therapy failure in the monotherapy treatment arm (Delfraissy et al., 2008). Major PI resistance mutations were present in only 5/23 patients and examination of Gag cleavage site mutations did not fully explain the determinants of treatment failure in the remaining patients (Delaugerre et al., 2009; Ghosn et al., 2011). In chapter 5 we investigated full-length Gag-protease from paired samples in patients who experienced virological failure on LPV/r monotherapy, one sample from baseline and one from the time of therapy failure. We examined full-length Gag-protease genotypically and phenotypically in the context of clinical information and our data provided evidence for a number of factors that contributed to virological failure. We presented evidence for a role of suboptimal LPV trough concentrations, most likely due to poor adherence. In addition, we showed the development of reduced PI susceptibility during PI exposure in the absence of major resistance mutations in one of our study patients that may have contributed to therapy failure (patient #3204 (HG)).

Reduced PI susceptibility of baseline viral variants from a number of the failure patients was also observed (see chapter 5). Full-length Gag-protease from one

patient in particular, patient #4201 (SO), demonstrated significant levels of reduced baseline PI susceptibility, indicating that reduced PI susceptibility at baseline may have contributed to virological failure on LPV/r monotherapy. In viral variants from patient #4201 (SO), reduced susceptibility of up to 17-fold to LPV was observed, which was the same reduction conferred by the V82A major mutation in patient #1403 (KON). This was a potentially important observation, hence further investigation of the potential correlation between reduced PI susceptibility at baseline and subsequent therapy failure was required.

In order to determine the clinical significance of the reduced PI susceptibility observed at baseline in some patients who went on to experience therapy failure, baseline samples from patients who achieved a sustained virological response were obtained. The PI susceptibility and single-round infectivity of baseline viral variants from the two groups of patients were measured and compared. In this analysis, we use the term ‘non-controller’ for patients experiencing virological failure in the LPV/r monotherapy arm and ‘controllers’ for the patients who had a sustained virological response.

Chapter aims**:**

1) Amplify full-length Gag-protease from baseline samples of controller patients, matched to the non-controller patients described in chapter 5
2) Genotypic analysis for the presence of amino acid positions previously described to be associated with PI exposure or resistance in viruses derived from controller patients
3) Determine PI susceptibility and single-round infectivity of representative viral variants from each controller patient
4) Compare the PI susceptibility and single-round infectivity of baseline variants from non-controller and controller patients

5) Use statistical modeling to determine factors associated with treatment outcome in the MONARK patients included in this study

6.2 Results

6.2.1 Clinical and virological information for controller patients selected for analysis

Controller patients were selected by matching to the non-controller patients included in chapter 5 using the following baseline criteria: viral subtype > viral load > CD4 count. Virological and clinical information for the five non-controller and eight controller patients is presented in table 6.1. Statistical analysis demonstrated no significant differences in viral load or CD4+ T cell count between the non-controller and controller patients.

Table 6.1. Clinical and virological information for patients studied.

Treatment Outcome	Patient number	Subtype	Viral Load (copies/ml)	CD4+ T cell count (cells/ml)
Non-Controller (Failure)	508 (SP)	B	37800	250
	3204 (HG)	B	23800	318
	1403 (KON)	CRF02_AG	44600	262
	1404 (DIO)	CRF02_AG	166000	326
	4201 (SO)	G	79500	168
Controller (Success)	2112 (WA)	B	78800	331
	909 (TH)	B	16200	249
	509 (MBF)	B	40800	422
	110 (FRD)	CRF02_AG	25600	237
	4003 (MD)	CRF02_AG	29000	235
	515 (MF)	CRF02_AG	67500	144
	1702 (BOY)	CRF02_AG	50500	211
	4202 (LR)	G	36500	341

6.2.2 Clonal analysis of Gag-protease from controller patients

Full-length *gag*-protease was amplified from the baseline plasma samples using the same method as used for the non-controller patients. Viral RNA was manually extracted from the plasma and cDNA synthesis and nested PCR performed. PCR products were cloned into pGEM and subjected to clonal sequencing analysis as for the non-controller patients, with approximately 10 clones sequenced for each patient. Clonal nucleotide sequences were aligned in MEGA using the ClustalW algorithm. Protease sequences were analysed for PI resistance mutations using the Stanford HIVdb (Liu and Shafer, 2006) and comparison to the IAS List of Drug Resistance Mutations (Johnson *et al.*, 2013). The results of these analyses are presented in table 6.2. The clonal Gag amino acid sequences were compared to the table of Gag mutations previously reported to be associated with PI resistance or exposure (table 1.2). The changes present in Gag for each patient are presented in the appendix and discussed below.

Within Gag, the E12K mutation was present in all recombinant CRF02_AG and subtype G viruses. Position 62, at which G is consensus amino acid, was particularly variable and G was only present in viruses from one patient #909 (TH), with viral variants from the other patients containing E, S, L or V. Gag mutations previously reported by studies in our laboratory were also present in several patients, R76K in patients #4003 (MD), #909 (TH), #4202 (LR), #509 (MBF) and #515 (MF) and Y79F in patients #4003 (MD), #909 (TH), #4202 (LR) and #1702 (BOY) (Parry et al., 2011). The p2/NC cleavage site was variable in the subtype CRF02_AG and G viruses with changes at positions 373, 374 375 and 376. In particular, T375N was present in patients #110 (FRD), #4003 (MD) and #1702 (BOY) and T375A in patient #515 (MF). The other cleavage site at which variation was present was the p1/p6 site, with the L449P change present

in patients #110 (FRD), #4003 (MD), #909 (TH) and #4202 (LR), and with patient #515 (MF) containing an L449V change. In addition, the S451N mutation was present in viruses from patients #4003 (MD), #4202 (LR) and #515 (MF).

Neither major nor minor PI resistance mutations were identified in the viral variants derived from any of the patients, although a number of polymorphisms were identified by the Stanford HIVdb present in all subtype CRF02_AG and G viruses, these being I13V, K20I and L89M. K20I, predicted to confer low level resistance to NFV only, was identified in all subtype CRF02_AG and G viruses, those from patients #110 (FRD), #4003 (MD), #515 (MF), #1702 (BOY) and #4202 (LR) (Vermeiren et al., 2007). In addition, R41K was present in viruses from patients #110 (FRD), #4003 (MD), #515 (MF), #4202 (LR) and one subtype B infected patient #509 (MBF). However, each of these amino acid changes are considered the consensus in subtype CRF02_AG and G viruses (Kuiken et al. 2012).

Amino acid position 63 of protease was polymorphic in the majority of patients regardless of virus subtype, with amino acid P present in viruses from patients #110 (FRD), #2112 (WA), #509 (MBF) (as a mixture with S and T) and #1702 (BOY). Viruses from patient #909 (TH) contained an L63S change and from patient #515 (MF) L63V/M. Position 69 was also highly polymorphic with amino acid K present in patients #110 (FRD), #4003 (MD), #1702 (BOY) and #515 (MF), and Q in patient #2112 (WA).

Table 6.2. Protease polymorphisms in viruses derived from non-controller patients.

Patient	Protease Polymorphism
#110 (FRD)	I13V, K14R, K20I, M36I, P39L, R41K, L63P, H69K, L89M
#2112 (WA)	T12A (v4)†, L33V, N37Y, I62V, L63P, I64L, H69Q
#4003 (MD)	I13V, K20I, M36I, R41K, H69K, K70R, L89M
#4202 (LR)	I13V, K14R, K20I, D25N (v8 & v10), E35Q, M36I, N37D, R41K, G48R (v8 & v10), G49E (v5), C67E, K70R, V82I, L89M, Q92H
#909 (TH)	L10I, E35D, M36I, R57K (except v5), D60E (except v5), I62V, L63S, I64V, K70T, V77I (except v5), I72T (v5), I93L
#509 (MBF)	P39S (v2,6,9,10), R41K (v1,3,4,5,7,8), L63P (v1,2,6,7,9,10), L63S (v3,4), L63T (v5,8), I64V (v.2), V77I
#1702 (BOY)	I13V, K14R (v2,4,5,7,9), I15V, K20I, M36I, I62V (v4), L63P, G68E, E68D (v10), H69K, L70R, L89M
#515 (MF)	L10V (v.1,2,3,8), T12K (v5), I13V, K14R (v1,2,3,7,8,10), K14Q (v4,5,6,9), I15V, Q18K, Q18R (v2), L19I (v4,5,6,8), K20I, M36I, N37D, R41K, L63V (v1,3,4,5,6,8), L63M (v2,7,9,10), I64M, C67S, H69K, K70R (v10), L89M

Protease sequences were assessed for resistance mutations and polymorphisms using the Stanford Resistance algorithm. Major and minor resistance mutations are shown, along with any polymorphisms in comparison with a group M consensus sequence. Changes reported are present in all viral variants from a patient, except where specific variants are shown in brackets after the change.

6.2.3 PI susceptibility and single-round infectivity of viral variants derived from controller patients

The PI susceptibility and single-round infectivity of viral variants derived from controller patients was measured, with between one and three viral variants from each patient selected for phenotyping. From each patient one viral variant identical, or most similar to, the consensus amino acid sequence was selected, along with one or two others containing amino acid changes of interest, where applicable. Full-length *gag*-protease from the selected variants was cloned from

the pGEM intermediate vector into the reference strain p8.9NSX+, using the *Not*I and *Xho*I restriction sites. The PI susceptibility to four PIs, ATV, DRV, LPV and SQV, and the infectivity were determined using single-cycle phenotypic assays.

The PI susceptibilities of each of the controller patient viral variants subjected to phenotyping are shown in figure 6.1. Significant variation in the PI susceptibilities of viral variants derived from different patients was observed, with reduced PI susceptibility of viral variants derived from some patients to the PIs ATV and LPV. Variants from all patients appeared fully susceptible to both DRV and SQV, with fold-difference EC_{50} values of less than three-fold. To the PI ATV, significant reductions in susceptibility of up to 9-fold were observed for viral variants from patients #110 (FRD) and #4003 (MD). Intermediate reductions in ATV susceptibility of between 5- and 7-fold were shown for viral variants from patients #1702 (BOY), #909 (TH) and #4202 (LR). Variants from three patients did not display significant reductions in PI susceptibility to ATV, with only up to four-fold difference in EC_{50}, #2112 (WA), #515 (MF) and #509 (MBF). For LPV, variants derived from three patients displayed significant reductions in PI susceptibility of up to 9-fold, #110 (FRD), #4003 (MD) and #4202 (LR). Conversely, the remaining patients did not display significant reductions in PI susceptibility of higher than four-fold.

The infectivity of each of the viral variants derived from the controller patients was determined using a single-cycle infectivity assay, as shown in figure 6.2. The infectivity of variants from each patient were not significantly different from that of the assay reference strain ranging from 65-115%.

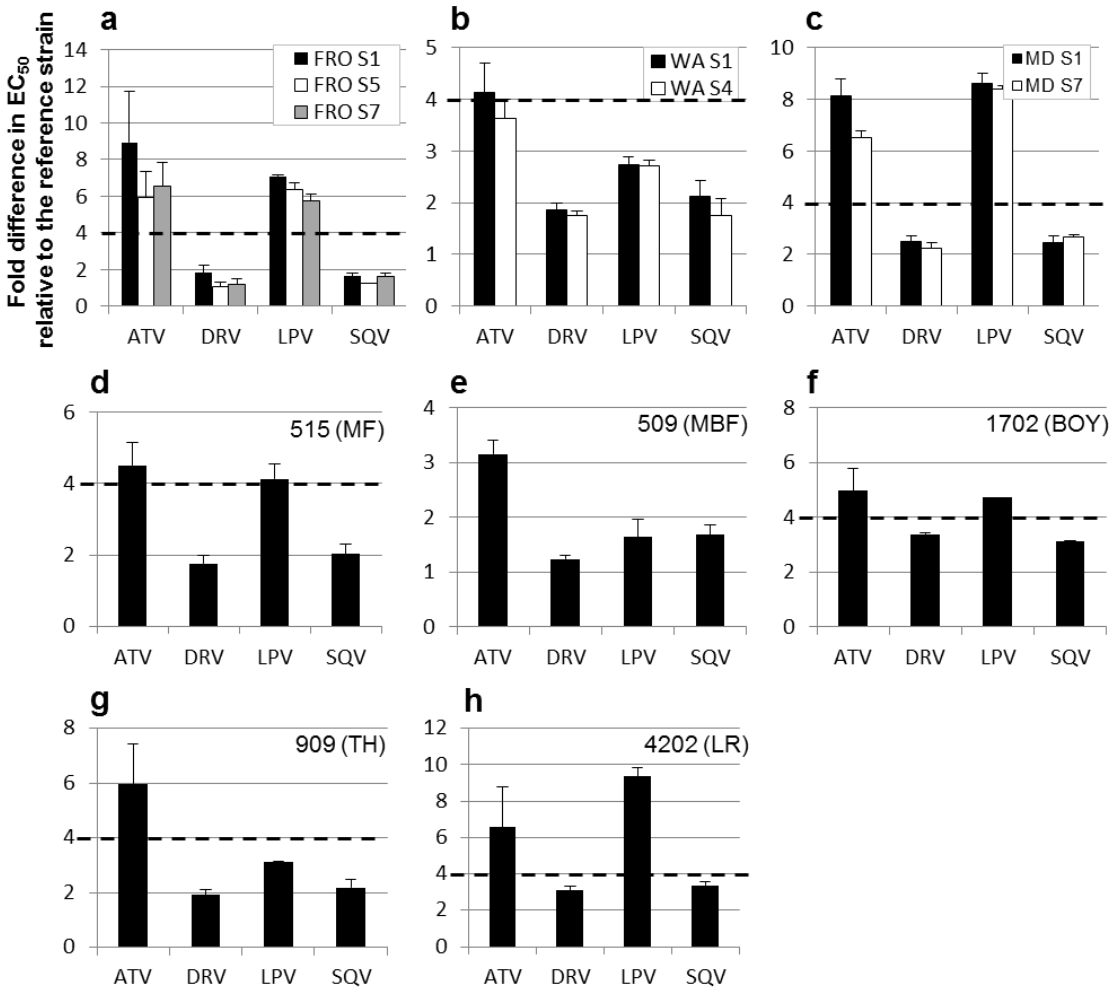


Figure 6.1. Limited variation in PI susceptibility of baseline viral variants from controller patients. Full-length Gag-protease was amplified from baseline plasma samples for 8 controller patients (a-h) and cloned into p8.9NSX+. VSV-g pseudotyped viruses encoding luciferase were produced by co-transfection in 293T cells. PI susceptibility of pseudovirions derived from each patient was determined using a single replication-cycle drug susceptibility assay as measured by luciferase activity. Data displayed are fold difference in EC_{50} values in comparison to that of the assay reference strain, p8.9NSX+, for each of six PIs: APV, ATV, DRV, LPV, SQV and TPV. Viruses for which the raw EC_{50} values were statistically different from assay reference strain using unpaired t-tests ($P < 0.05$) are denoted with an asterisk(*). Error bars represent the standard error of the mean of two independent experiments performed in duplicate. The dashed line shows the previously reported cut off for a significant reduction in susceptibility of greater than four-fold in comparison with the assay reference strain (Gong et al., 2004).

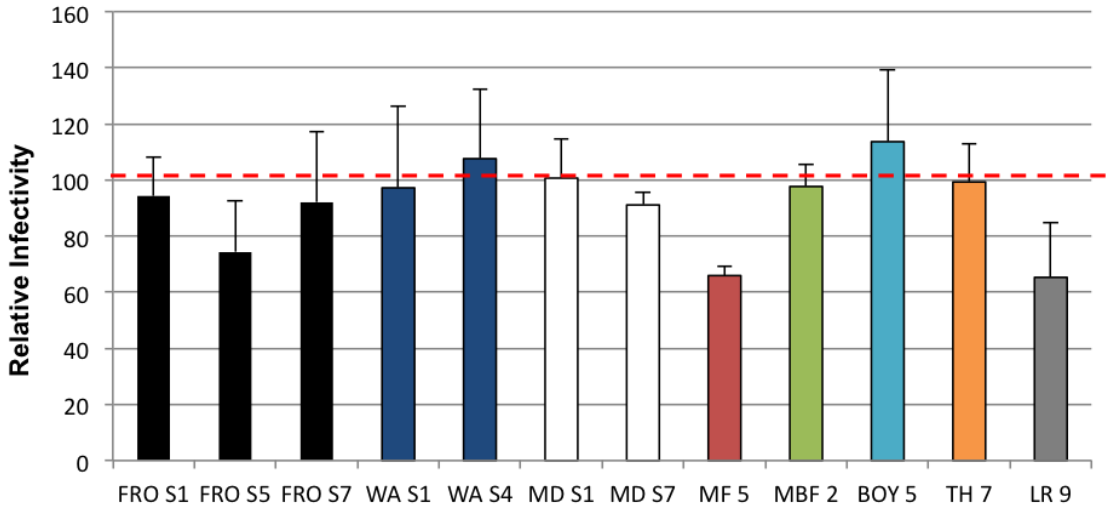


Figure 6.2. Single-round infectivity of baseline viral variants from controller patients. The single-round infectivity of pseudovirions produced from co-transfection of 293T cells was determined by titration of viruses using a single-cycle of infection. Luciferase activity was measured using SteadyGlo and expressed as a fold difference in comparison with that of the assay reference strain, p8.9NSX+. The single-round infectivity of consensus baseline viral variants from 8 controller patients is shown, with each patient represented by a different colour. Error bars are standard error of the means of two independent experiments, each performed in duplicate.

6.2.4 Association between PI susceptibility at baseline and therapy failure

Having determined the PI susceptibility of viral variants derived from non-controller patients (chapter 5) and controller patients, comparison of the two patient groups was performed. This would enable the identification of any correlation between baseline PI susceptibility and subsequent treatment outcome. The susceptibility of baseline viral variants derived from patients in both treatment outcome groups for all four PIs is displayed in figure 6.3.

The viral variants derived from non-controller patients showed reductions in susceptibility to PIs ATV, LPV and SQV in comparison to those who achieved a sustained virological response. No viral variants from controller patients

displayed fold-difference EC_{50} values above 10 to LPV, but 4/18 non-controller patient variants displayed reduced LPV susceptibility above 10-fold difference in EC_{50} in comparison to assay reference strain. There was no evidence of reduced SQV susceptibility in any viral variants derived from controller patients. 8/12 controller viral variants displayed below 2-fold reductions in EC_{50} and all variants conferred less than 4-fold reductions in SQV susceptibility. Conversely, 11/18 viral variants from non-controllers demonstrated reduced SQV susceptibility of above 4-fold in EC_{50} and one variant displayed a 10-fold reduction.

For the PI ATV, 7/12 controller viral variants and only 5/18 variants from non-controller patents displayed fold-difference in EC_{50} values < 6. Controller patient variants displayed up to 9-fold reductions in EC_{50} whereas variants derived from non-controllers displayed up to an 11-fold reduction in ATV susceptibility. 9/18 non-controller viral variants displayed approximately 8-fold reductions in ATV susceptibility. For the PI DRV, none of the viral variants displayed significant reductions in PI susceptibility.

Despite these observations, reduced PI susceptibility was not significantly associated with treatment failure using Mann-Whitney U rank sum tests at the $P = 0.05$ significance level (ATV $P = 0.08$, DRV $P = 0.46$, LPV $P = 0.31$, SQV $P = 0.31$).

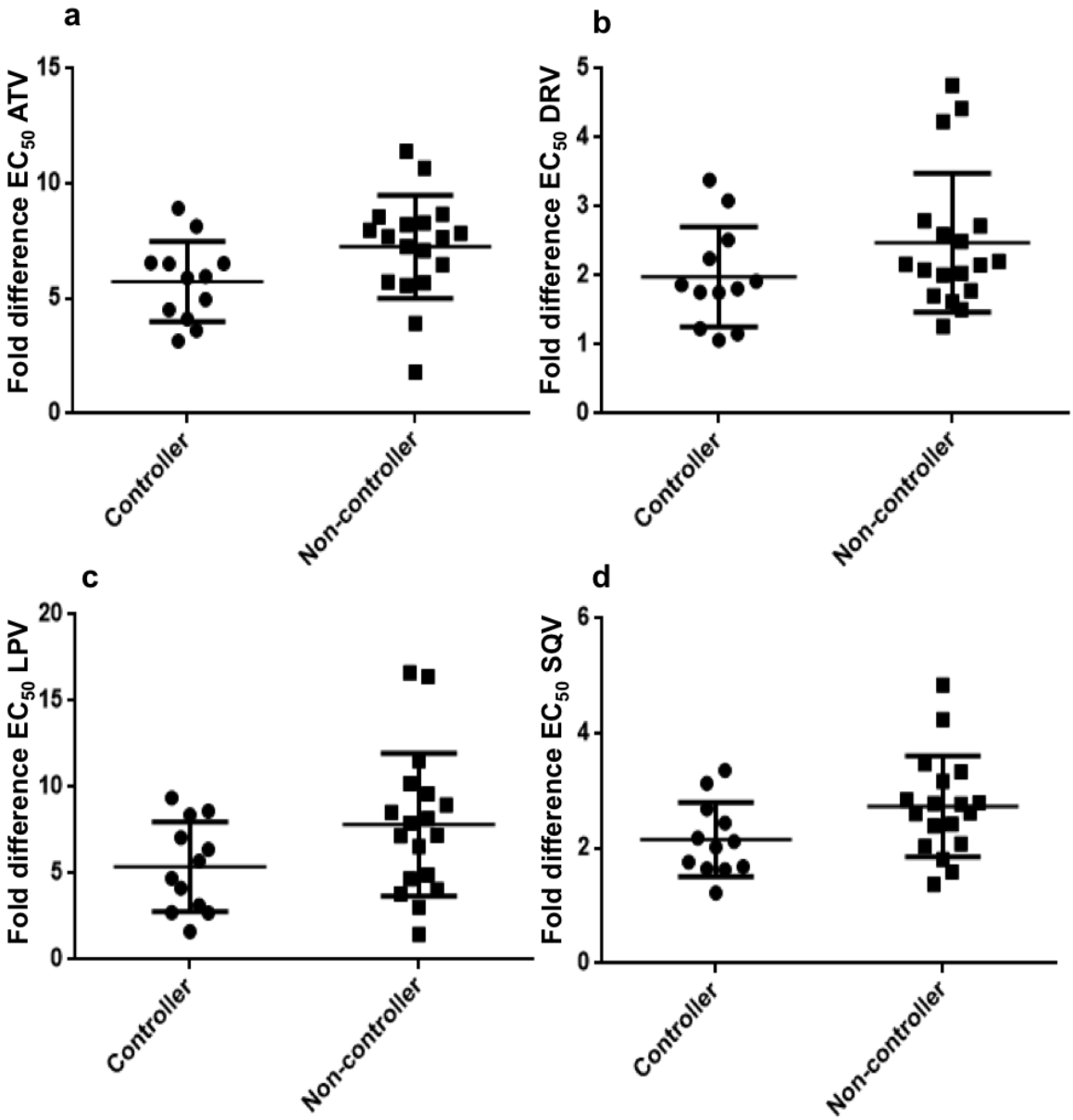


Figure 6.3. Comparison of PI susceptibility of viral variants from non-controller and controller patients shows a trend towards greater reductions in susceptibility for non-controllers. The phenotypic susceptibility of VSV-g pseudotype viruses derived at baseline from 8 controller and 5 non-controller patients was determined measuring luciferase activity. Susceptibility to four PIs was measured: (a) ATV, (b) DRV, (c) LPV and (d) SQV, data are expressed as a fold-difference in comparison to p8.9NSX+. The mean and standard deviation for each treatment outcome group are shown. The difference in susceptibility between controller and non-controllers was not statistically significant for any PI ($p = 0.05$, unpaired T test).

6.2.5 Reduced single-round infectivity was associated with sustained virological response

The single-round infectivity of the consensus viral variants for each controller and non-controller patient compared to the assay reference strain are shown in figure 6.4. Significant variation in single-round infectivity between non-controller and controller derived variants was present, with non-controller variants displaying significantly lower values compared to controllers. Of the five variants from non-controllers, three demonstrated infectivity values around 60% and a fourth a value of just 20% that of assay reference strain. The fifth displayed 90% that of the assay reference strain. Only 2/8 variants from controller patients demonstrated values below 90% that of reference strain, with the other patients displaying values between 90 and 100%. These differences were statistically supported when the mean single-round infectivity of variants from non-controllers at 60.4%± 25.3% was compared to that of the controllers 90.6%±16.8% (P = 0.0078, unpaired t test). In addition, lower single-round infectivity was significantly associated with treatment failure in a Mann-Whitney U rank sum test using mean of within individual log-fold differences (P = 0.04).

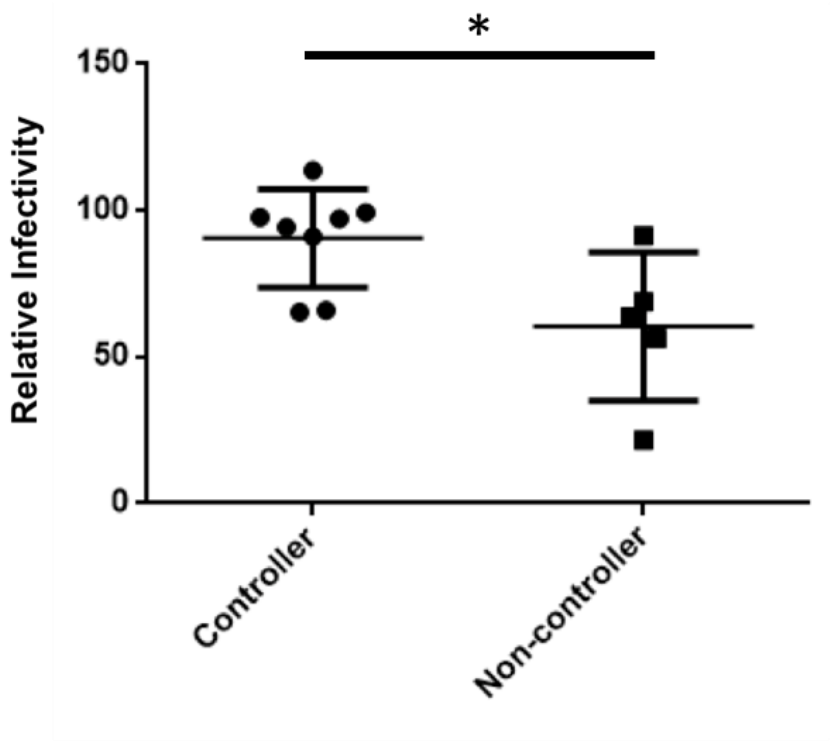


Figure 6.4. Reduced single-round infectivity at baseline correlates with treatment failure. The single-round infectivity of pseudovirions produced from co-transfection of 293T cells was determined by titration of viruses using a single-cycle of infection. Luciferase activity was measured using SteadyGlo and expressed as a fold difference in comparison with that of the assay reference strain, p8.9NSX+. The single-round infectivity of consensus baseline viral variants from 8 controller patients and 5 non-controller patients is shown. Mean and standard error for each treatment outcome group is shown. The relative infectivity of controllers was significantly different at the p>0.05 significance level (p=0.0078, unpaired t test). In addition, lower single-round infectivity was significantly associated with treatment failure in a Mann-Whitney U rank sum test using mean of within individual log-fold differences (*P* = 0.04).

6.2.6 Exploration of genetic correlates of reduced PI susceptibility

Having shown a correlation between reduced PI susceptibility at baseline and treatment failure, investigation of the genetic determinants of reduced PI susceptibility and hence treatment failure was performed. Subtype specific alignments containing the viral variants subject to phenotypic analysis were generated using MEGA. Each alignment was searched for genetic changes correlating with treatment outcome, defined as being present in all non-controller variants and not in any controller variants.

Alignment of the subtype B controller and non-controller consensus viral variants revealed a total of 78 amino acids in Gag that were variable between patients out of 498 amino acids in total, and 3 indels of at least three amino acids (see appendix figure 2). Of these, variation at three amino acid positions correlated with virological failure: 34, 93 and 490 (HXB2 position #473). At position 34, L was present in non-controllers and I in controllers. For positions 93, E was present in non-controllers and D in controllers. At position 473 amino acid P correlated with treatment failure, and S with virological suppression. Within protease 16 variable positions were present for subtype B viruses, but none of these correlated with treatment outcome.

Alignment of consensus viral variants from non-controller and controller patients infected with the recombinant CRF02_AG viruses revealed the presence of 97 variable amino acids in Gag. Of these, only position 287 (HXB2 position #286) correlated with treatment outcome – with R present in patients #1403 (KON) and #1404 (DIO) and K in controller patients. Within protease 19 amino acids were variable, but none correlated with treatment outcome. Analyses were also

performed with only one non-controller patient #1404 (DIO), as patient #1403 (KON) was thought to fail primarily due to poor adherence (see chapter 5). In this analysis, 8 amino acid positions in Gag correlated with treatment outcome – 75, 164, 166, 192, 225, 235, 287 and 327. In addition, three polymorphisms were present only in protease of patient #1404 (DIO): 71T, 89I and 93M.

Alignment of the subtype G consensus viral variants from the non-controller (#4201, SO) and controller (#4202, LR) patients revealed 42 variable amino acids in Gag and two indels of at least two amino acids. In addition, five variable positions in protease were present – 37, 57, 61, 64 and 70. As only one non-controller and one controller subtype G patient were included, filtering based on treatment outcome was not possible. However, one of these Gag positions, 93, was identified as correlating with treatment outcome in both subtype G and subtype B viruses. At this position, E correlated with treatment failure and D with virological suppression.

6.3 Discussion

Here we report a correlation between PI susceptibility and infectivity of VSV-G pseudotyped viruses in a single-round of infection, with reduced PI susceptibility and single-round infectivity at baseline associated with subsequent virological failure on LPV/r monotherapy. The variation in PI susceptibility observed in PI-naïve viruses is in keeping with other studies and our own data in chapters 3, 4 and 5 using single replication-cycle assays encompassing full-length co-evolved Gag-protease (Gupta et al., 2010; Jinnopat et al., 2009). Studies using commercial phenotypic assays have not reported the same degree of variation in PI susceptibilities of PI-naïve viruses as described here, but these did not include full-length co-evolved Gag from the patient viruses (Parkin et al., 2004; Vergne et al., 2006). Given the data showing that the inclusion of co-evolved Gag affects the PI susceptibility of protease, the inclusion of full-length patient-derived Gag-protease was prudent and justified (Gupta et al., 2010; Parry et al., 2009; chapter 3).

Our data provide the first evidence that reduced PI susceptibility and reduced single-round infectivity of patient-derived Gag-protease in the absence of major resistance mutations correlate with treatment failure on LPV/r monotherapy. This is a potentially clinically significant finding, as it indicates that screening at baseline may enable a prediction of patients who are suitable for these simplification strategies. However, we have been unable to elucidate the genetic correlates of reduced susceptibility, single-round infectivity and hence treatment failure.

Subtype-specific genotypic analysis identified a number of amino acid changes in Gag and protease that were only present in non-controller patient viruses, correlating with treatment outcome, these being 34, 93, 287 and 473 of Gag and 16 in protease. However, the small number of viruses in this study limited the confidence with which these can be associated with therapy outcome. Given the significant levels of variability in Gag between patients infected with the same subtype, it is possible that the amino acid positions are associated with treatment failure solely by chance. However, position 93 in Gag was identified in the separate analyses for both subtype B and G analyses with the amino acid E correlating with therapy failure. This amino acid position in Gag has not previously been associated with PI exposure or resistance, so further investigation of its role in PI susceptibility would be prudent (table 1.2). Additionally, it is located within the MA subunit of Gag, a region shown to confer reduced PI susceptibility in subtype B molecular clones (chapter 3) and in an extensively PI-experienced subtype B virus (Parry et al. 2009).

Whilst our data indicate that reduced baseline susceptibility may be associated with treatment failure, four of the five non-controller patients did initially achieve virological suppression to a viral load below 400 copies/ml on LPV/r monotherapy by week 24, before experiencing virological failure after 40 weeks of the trial. As virological suppression was initially achieved, the reduced PI susceptibility present at baseline in these patients was not sufficient to prevent

the PI from initially suppressing viral replication. However, in each of the non-controller patients examined, suboptimal LPV trough concentrations were present and compliance assessments reported missed doses for at least one time point. We hypothesise that the reduced PI susceptibility present in the baseline viral variants from non-controller patients reduces the tolerance to suboptimal adherence, thus rendering these patients more likely to experience virological failure when adherence is not optimal. This is supported by the comparison of the EC_{90} for each viral variants with the LPV trough concentrations. For example, for screening variants from patient SO the EC_{90} was 65 ng/mL and the LPV concentration 121 ng/mL, indicating that small levels of viral replication would be possible.

Although studies on the minimum adherence required to suppress viral replication successfully during LPV/r monotherapy have not been performed, those examining LPV/r based HAART have shown that reduction to 90-95% adherence does not significantly increase the risk of treatment failure (King et al., 2005; Shuter et al., 2007). However, as only one active agent is present during PI monotherapy it is likely that this regimen may be more sensitive to poor adherence compared to HAART regimens. This is supported by data from the OK LPV/r monotherapy trial showing that the addition of two NRTIs enables virological suppression in patients experiencing virological failure on PI monotherapy in the absence of major resistance mutations (Pulido et al., 2008). It is possible that patients with viruses demonstrating reduced PI susceptibility at baseline may be better suited to HAART, which is likely to be more forgiving of reduced adherence, as three active agents are present. Alternatively, identification of these patients before treatment initiation would enable interventions to increase adherence and reduce this risk of failure, such as additional adherence counseling.

In conclusion, we report an association between reduced PI susceptibility and single-round infectivity at baseline and subsequent virological failure on LPV/r

monotherapy in patients enrolled in the MONARK trial. This is a potentially important finding as it indicates that it may be possible to predict treatment outcome on LPV/r monotherapy from baseline PI phenotyping. We hypothesise that reduced baseline PI susceptibility renders patients more vulnerable to virological rebound when their adherence is not optimal. Extension of this study to include more patients is warranted.

7 Final discussion and future work

7.1 Final discussion

HIV causes a significant global burden, with an estimated 35.3 million HIV positive individuals and 1.6 million AIDS related deaths worldwide in 2012 (WHO, 2013). However, the introduction of HAART to treat HIV has greatly improved the prognosis of HIV-infected individuals. An optimal treatment regimen will result in a decrease in viral load to 'undetectable' levels (< 50 copies/ml) together with an increase in patient CD4+ T cell count. In recent years, efforts and funds to ensure worldwide access to antiretrovirals have been stepped up and by the end of 2011, approximately 8 million patients were receiving ART.

Although HAART has greatly improved the prognosis of HIV positive individuals, virological failure and the development of resistance remains a significant problem. Philips et al. (2007) estimated that after 10 years of HAART treatment, 9.2% of patients will have developed resistance to all three drugs included, and that the risk of death within 5 years of the development of resistance is 10.6% (Phillips et al., 2007). PIs are a potent class of inhibitors and it has been reported that use of a PI as the third drug in HAART leads to significantly fewer patients experiencing virological failure than HAART using an NNRTI as the third agent (Gupta *et al.*, 2008). Nonetheless, therapy failure during treatment with PIs remains a significant issue. Virological failure in patients receiving PI-based therapy often occurs in the absence of major PI resistance mutations in protease and hence the exact determinants of treatment failure on PIs remain largely unknown. This is particularly important as the number of patients receiving PIs in the developing world increases, where traditional laboratory monitoring including regular viral load testing and CD4+ T cell counts cannot always be performed (Mugyenyi et al., 2010).

Resistance to PIs primarily occurs via the accumulation of resistance mutations in protease. Major mutations reduce the affinity of the protease to the PI, but also reduce the affinity of protease to the Gag substrate, hence they cause reduced viral fitness (Condra et al., 1995; Zennou et al., 1998). Minor mutations occur within protease and Gag to compensate for the reduced fitness conferred by the major resistance mutations (Croteau et al., 1997). However, recent studies have shown that mutations in Gag can directly contribute to PI resistance in mechanisms independent of the compensation of reduced RC conferred by the resistance protease (Dam et al., 2009; Jinnopat et al., 2009; Nijhuis et al., 2007b; Parry et al., 2011). It has also been demonstrated that inclusion of co-evolved Gag alongside its cognate protease in phenotypic assays can directly affect PI susceptibility (Gupta et al., 2010; Parry et al., 2009). Given the increasing use of PIs, mainly ATV and LPV, in second line therapies in the developing world an understanding of the exact mechanism and determinants of treatment failure and virological resistance is vital.

Given the mounting evidence for a role of Gag in PI susceptibility, we hypothesised that PI-naïve Gag may contain polymorphisms that can affect PI susceptibility. In addition, given the extensive evidence for the co-evolution of Gag and protease, a detailed investigation of phenotypic susceptibility of co-evolved Gag-protease was required (as reviewed by Fun et al., 2012). This thesis has investigated the determinants of PI susceptibility utilising a single-cycle phenotypic assay, with the aim of answering two main questions. The first was whether natural amino acid sequence variation present in full-length Gag-protease of PI-naïve subtype B HIV viruses can confer variation in PI susceptibility. To date, no assessment of variation in phenotypic PI susceptibility between different PI-naïve subtype B viruses has been performed using assays encompassing full-length Gag alongside its co-evolved protease. The second objective was to investigate whether variation in PI susceptibility of full-length Gag-protease may have contributed to treatment failure in the PI monotherapy

arm of the MONARK clinical trial. Our study was the first to perform such a detailed virological analysis of patients failing PI monotherapy, and in particular, no previous studies have utilised phenotypic assays encompassing full-length co-evolved Gag.

We have shown significant variation in PI susceptibility of full-length Gag-protease from PI-naïve viruses of subtypes B, CRF02_AG and G, for both molecular clones and viruses derived directly from patients. Reduced susceptibility of up to 17-fold in EC_{50} in comparison to assay reference strain to the PIs ATV and LPV was observed. Reduced susceptibility was consistently observed to four of the six PIs tested: APV, ATV, LPV and TPV. This is a novel finding and the first demonstration of variation in susceptibility of PI-naïve, subtype B, CRF02_AG and G viruses using an assay encompassing full-length, co-evolved Gag-protease. These data are in keeping with other reports based on full-length Gag-protease of viruses of other subtypes and those from PI-experienced patients (Gupta et al., 2010; Jinnopat et al., 2009; Parry et al., 2009). Other studies based on large numbers of viruses assessing variation in PI susceptibility have not reported such a range, but these utilised commercial assays in which only protease from the test virus was included alongside reference strain Gag (Parkin et al., 2004; Vergne et al., 2006). We provide further evidence of the importance of the inclusion of co-evolved Gag alongside protease in phenotypic PI susceptibility assays. However, the limited number of samples included in our study means that further and more extensive research to confirm our findings in a wider panel of viruses is required.

Our second main finding is that variation in PI susceptibility can be independently conferred by the N terminus of Gag, specifically MA and N terminus of CA. Our data show the involvement of changes distant from the Gag cleavage sites in reduced susceptibility – positions 30 and 102 of Gag. This is one of a handful of reports that have provided evidence for the importance of changes outside the cleavage sites within the N terminus of Gag

in affecting PI susceptibility, but ours is the first to demonstrate the role of changes in MA in PI-naïve viruses. Of these reports, Parry et al (2011) demonstrated that changes within MA could directly affect PI susceptibility in a PI-experienced subtype B virus, specifically changes at positions 76, 79 and 81. The second report was a PI-naïve subtype CRF01_AE virus, in which position 165 of Gag located in CA was shown to affect PI susceptibility (Jinnopat et al., 2009; Kameoka et al., 2010). Positive selection analysis and correlation between genotypic and phenotypic data amongst non-controller patients in the MONARK trial identified a number of amino acid positions within MA that were associated with PI exposure and reduced susceptibility, providing further evidence for the role of MA in PI susceptibility. Overall, our data show the direct involvement of the MA subunit of Gag in PI susceptibility and we conclude that further investigation of the role of MA in PI resistance and therapy failure is required.

We have shown that significant reductions in PI susceptibility can be conferred by protease in the absence of major PI resistance mutations, with up to 13-fold reduction in EC_{50} observed for the PI ATV. Furthermore, this study provides evidence that minor resistance mutations and polymorphisms present in combination can confer reduced PI susceptibility. Data for three amino acid positions in protease are presented here, namely 13, 63 and 71, but we hypothesise the involvement of other changes that when present in combination confer reduced PI susceptibility. These data are in keeping with that of other studies which have shown that PI polymorphisms affect PI susceptibility and fitness, including at amino acid positions 10, 20, 33 and 71 (Nijhuis et al., 1999; Vermeiren et al., 2007). A number of reports have also shown as association between the presence of protease polymorphisms and poorer treatment outcomes (Kempf et al., 2001; Pellegrin et al., 2006; Vora et al., 2006). In addition, the study performed using the UK HIV Drug Resistance Database identified a number of other protease changes that are present at higher

frequencies in patients failing ATV therapy than in PI-naïve patients (Dolling et al., 2013b). Further research to fully elucidate the role of protease polymorphisms in PI resistance is required, particularly as the roll-out of HAART to areas of the world where non-B subtypes are most prevalent occurs.

Of particular interest was the difference in the independent role of Gag and protease in PI susceptibility in molecular clones (chapter 3) and patient-derived viruses (chapter 4). Our data show that Gag solely conferred the reduced susceptibility observed in molecular clones YU2 and JRFL, but that protease solely conferred the reduced susceptibility observed in the patient-derived viruses. The exact reason for this difference is unknown, although another study that included both molecular clones and patient viruses did not report a difference in the role of Gag and protease (Gupta et al., 2010). It is possible that the relatively small number of viruses included in this study could account for this difference. The fact that the two molecular clones in which Gag conferred reduced susceptibility were not subjected to extensive *in vitro* passage, YU2 and JRFL, appears to discount the hypothesis that adaptation to tissue culture during the generation of the molecular clones may explain the difference in the role of Gag in comparison with patient-derived viruses (Collman et al., 1992; Hahn et al., 1984). Another possible explanation is that the effect of Gag on PI susceptibility is only present in viruses derived from the CNS, as YU2 and JRFL were both isolated from the brain. Adaptations in Gag could occur to favour viral replication in the brain that by chance also affect the susceptibilities of these viruses to PIs. Further investigation is required to elucidate the clinical relevance of these findings, in particular the study of other molecular clones derived from brain tissue. Given the concern of ongoing viral replication in the brain during PI therapy and the question of the penetration of PIs into the CNS, this could be of potential importance (Valero et al., 2014).

Here, we have performed the first detailed virological analysis of full-length Gag-protease in patients experiencing virological failure on PI monotherapy.

We have shown that reduced PI susceptibility of full-length Gag-protease in the absence of major resistance mutations contributed to treatment failure in at least some of the patients studied in the MONARK trial by two main mechanisms. Firstly, reduced PI susceptibility at baseline may render the patient more likely to experience subsequent virological failure and secondly reduced PI susceptibility developed during therapy in the absence of major resistance mutations. We hypothesise that reduced PI susceptibility alone did not cause treatment failure, but coupled with suboptimal adherence led to ongoing viral replication in these patients. This is supported by the findings of another report, in which escalation of these patients onto HAART lead to successful viral suppression (Pulido et al., 2008). If confirmed by studies of larger groups of patients, this would offer an explanation for the causes of treatment failure in other patients on PI monotherapy who fail without major PI resistance mutations (Gilks et al., 2012; Pulido et al., 2008).

To date, PI monotherapy trials have largely been performed in the developed world in patients closely monitored with regular laboratory testing including CD4+ T cell counts and viral loads (Arribas et al., 2010; Castagna et al., 2013; Delfraissy et al., 2008; Pulido et al., 2008). All of these trials have reported very low rates of failure with PI resistance mutations which would compromise future treatment options (Arribas et al., 2010; Delaugerre et al., 2009; Pulido et al., 2008). Given that the vast majority of those requiring treatment are in the developing world, in particular sub-Saharan Africa where laboratory testing is not routinely performed, it is important to understand the utility of PI monotherapy in this setting. Recent data from the ERNEST trial have shown that in the absence of viral load testing, 18% of patients on LPV/r monotherapy developed major resistance mutations after 96 weeks of therapy in comparison to just 2% on LPV/r based HAART (Paton et al. 2013). In many parts of Africa, only two lines of treatment are available and LPV-based HAART forms the second, and final, option for patients experiencing failure on first line regimens.

Our data, along with other studies, indicate that PI monotherapy may have a place in treatment of patients with fully-susceptible viruses at baseline, good adherence and close virological monitoring. In these patients the reduced financial cost, pill-burden and side-effects merit the consideration of PI monotherapy as a treatment option, especially given that it is unlikely to compromise future treatment by the development of major PI resistance mutations. However, PI monotherapy appears to have limited utility in sub-Saharan Africa where there is little to no viral load or resistance testing and LPV/r is the final treatment option for many patients.

7.2 Study limitations

This study had a number of limitations, as imposed by the methods used to perform the analysis. Although the single-cycle phenotypic assay used to measure PI susceptibility has a number of advantages justifying its selection, it also imposes limitations on the interpretation of our data. The phenotypic assay was carried out at containment level 2, as opposed to multiple-cycle assays performed with replication-competent viruses which require containment level 3 facilities. In addition, it has been reported that single-cycle assays have increased sensitivity and reproducibility over multiple-cycle ones, although they are further removed from the conditions in vivo. Another important advantage is that single-cycle assays offer no opportunity for evolution and selection under tissue culture conditions unlike multiple-cycle assays (Petropoulos et al., 2000). The assay used in this study and its modified versions have been used in a series of published studies, so the assay has already undergone peer-review (Gupta et al., 2010; Koning et al., 2013; Martin et al., 2012; Parry et al., 2009). The assay has also been subjected to extensive measures of variation, as well as the additional confirmations performed here (Gupta et al., 2010). This assay also controlled for variation in infectivity and transfection efficiency between viruses as each EC_{50} value was normalised using a no-drug control luciferase reading for each virus.

Whilst the assay offers important advantages over similar multiple-cycle assays, it also has limitations. A previous study has shown that the results of single-cycle and multiple-cycle assays are not always concordant (Dykes et al., 2010). The most important limitation is the use of pseudoviruses with a VSV-G envelope glycoprotein, as opposed to an HIV envelope. A recent study has demonstrated the PIs block viral entry and that this block is only observed when an HIV-1 envelope is used and is not found when VSV-g and MLV pseudotyped viruses are used (Rabi et al. 2013). Our data indicate the role of MA in PI susceptibility, and given that MA interacts with HIV Env for the incorporation of Env in the assembling virion and the correct positioning of Env in the mature plasma membrane, it is important to confirm our data in the future using an assay system with HIV Env (as discussed in 7.3.1). In addition, VSV-g enters cells using a different pathway than native HIV virions, which could affect downstream pathways such as the transport of the capsid across the cytoplasm, uncoating and nuclear entry, which have been shown to be affected by the presence of PI (Muller et al. 2009).

The assay reference strain, p8.9NSX+, is derived from molecular clone HXB2 up to the *Spe*I site in Gag and from NL4-3 from the *Spe*I site onwards to the end of Pol. However, both HXB2 and NL4-3 displayed reduced susceptibility in comparison to the reference strain to APV, indicating that p8.9NSX+ may exhibit hypersusceptibility to the PI APV. This does not significantly limit the interpretation of our data as all fold-difference values were derived using the same reference strain so susceptibility and single-round infectivity values for each virus were directly comparable. In addition, as this reference strain has been used in other published studies with this assay system, direct comparisons to these viruses can also be performed (Gupta et al. 2010; Parry et al. 2009). Whilst the use of p8.9NSX+ does not enable direct comparison with other assays using different reference strain, it is similar to the reference strains used in commercial phenotypic assays Antivirogram and Phenosense,

that use HXB2 and NL4-3, respectively. In particular, clinical cut-offs derived for other assays cannot be directly applied to our assay. The lack of clinical cut-offs for this assay is an obvious limitation, but their derivation would require phenotyping of thousands of patient derived viruses (Winters et al., 2009).

Another disadvantage of this assay is the use of luciferase activity as a marker for infectivity. The luciferase activity measurement is affected by factors other than the percentage of cells infected, in particular the reading is altered by cells infected multiple times. Conversely, the use of a GFP reporter measured by FACS would have enabled the determination of the percentage of cells infected.

The single-round infectivity assay used here offers similar advantages as the phenotypic susceptibility assay in that it is single-cycle and hence carried out in containment level 2 laboratories. However, as p24 quantification occurs after virion assembly this assay will not capture any differences in viral fitness conferred during virion production – including transcription/translation of viral genes, virion assembly and budding. Gag has been shown to drive the process of virion assembly and in particular amino acid changes surrounding position 30 of Gag have been shown to affect targeting to the plasma membrane (Ono and Freed, 2004). The single-round infectivity assays performed without p24 adjustment for the MONARK study enabled capturing of all life cycle stages, but do not control for variation in transfection efficiency. In the future, different methods to measure fitness over multiple-cycles of replication could be used such as competition assays to compare the fitness of each of the molecular clones.

Use of clonal analysis limited the representation of the quasispecies within patients for both the PI-naïve subtype B viruses and the patients from the MONARK trial. This method has a number of limitations as it is laborious and time consuming, and is vulnerable to recombination events during the PCR step. As recombination between viral genomes cannot be ruled out, linkage of

mutations cannot be inferred from sequences generated by clonal analysis. Sequencing of just ten variants from each time point limited the depth to which variation was examined and may not have detected low level variants which could have had differing PI susceptibilities. Different experimental methods to overcome these issues are discussed below (section 7.3.3).

7.3 Future work

7.3.1 Variation in PI susceptibility of subtype B molecular clones

We have shown that the N terminus of Gag conferred variation in PI susceptibility observed for full-length Gag-protease of some subtype B molecular clones in our pseudotyped virus assay. In the future, we need to confirm that this effect is also present with full-length virus given the interactions that take place between HIV Env and the MA subunit of Gag. Our data demonstrated the involvement of two novel changes in Gag in PI susceptibility – K30R and D102E. In addition, examining the role of positions 30 and 102 in the susceptibility of full-length viruses with HIV Env is important, given that PIs have been shown to affect entry and that this is dependent on the interaction between MA and Env. One finding of interest was the introduction of these two changes into the reference strain prevented the production of infection pseudovirus. Further investigation of this finding, in particular whether the changes would also cause a loss of infectivity when present singly, and the life cycle stage at which this block occurred would be of interest. It is possible that these experiments could help elucidate how PIs block infectivity as well as the role of these two changes in PI resistance.

Whilst the identification of changes K30R and D102E is of interest, they did not fully account for the reduced PI susceptibility observed and the remaining genetic determinants were unknown. Further work to identify these would concentrate on the amino acid changes present within the N terminus of Gag, but these are different in the two molecular clones studied, YU2 and JRFL. The

unique changes from consensus B sequence in the N terminus of Gag were T53S, Q69K, R76T, V94I, N109T, K112M and V159I for JRFL and G10A, K28Q, N47D and R91K for YU2. Further site-directed mutagenesis and PI susceptibility assays could explore the role of these changes in PI susceptibility. Given the small effect on PI susceptibility of the changes at positions 30 and 102, it is likely that a number of these changes in combination account for the reduced susceptibility observed for gagN in each molecular clone.

The other main way in which this study could be extended is further investigation of the mechanisms by which the reduced PI susceptibility of Gag-protease in conferred. We have already performed preliminary experiments here using electron microscopy techniques which showed that for molecular clone JRFL, pseudovirion maturation is not significantly affected by the presence of the EC_{90} of ATV. Conversely the maturation of p8.9NSX+ and YU2 is inhibited by ~50%. Examination of the chimeric pseudovirions (MCgag, MCpro, MCgagN and MCgagC) using EM would enable confirmation that the difference in maturation was in fact conferred by the N terminus of molecular clone Gag as was assumed here, given that the N terminus of Gag solely conferred the change in PI susceptibility. Further experiments using the same EM methods, but with more concentrations of PI would be beneficial. In addition the inclusion of more molecular clones in the analysis, in particular one of 89.6 or SF2 in which both Gag and protease contributed to the reduced PI susceptibility observed, would shed further light on the role of the N terminus of Gag.

To understand the determinants of the variation in PI susceptibility between different viruses, western blotting to monitor Gag cleavage could be used. Western blots would show whether the proportion of Gag cleaved varies between the different viruses and if these differences are small this may not be obvious in the proportion of mature virions. Whilst western blotting was attempted during the course of this study, a method sensitive enough for use

with pseudovirus in cell media proved technically challenging and could not be developed in time. As an alternative, the cleavage of radio-labelled Gag-Pol protein by wild-type protease could be examined, as in previous studies (Pettit et al., 2005). A quantitative examination of the Gag cleavage products present at various time points (12, 24 and 48 hours post-transfection), and at varying concentrations of ATV would be required. This could provide additional evidence as to whether a difference in the rates of Gag cleavage between the different molecular clones and reference strain is present and whether the proportion of pseudovirions achieving maturation is affected. In addition, comparison of the efficiency of Gag cleavage at each site between each molecular clone could be performed using an enzymatic cleavage assay. Van Maarseveen et al. (2012) described a method involving the synthesis of peptides mimicking each of the Gag cleavage sites, which were cleaved by wild-type protease enzyme under controlled conditions. Cleavage products were identified and quantified using HPLC (high-performance liquid chromatography) which would enable a comparison of cleavage efficiencies between molecular clones and the reference strain (van Maarseveen et al., 2012).

At the concentration of ATV used here for EM, viral infectivity was reduced by 90% but no difference in maturation was present for JRFL, indicating that this reduced infectivity must be conferred at another stage of the life cycle. A recent study has examined the effect of PIs on the efficiency of each life cycle step and in particular has shown that PIs significantly block cellular entry (Rabi et al. 2013). Another study has examined the effect and mechanisms of varying concentrations of the PI LPV on viral infectivity. This reported that incompletely cleaved Gag, and in particular a lack of full release of the CA subunit, exerts a strongly trans-dominant, negative effect on viral infectivity. The authors hypothesised that within a virion which appears mature morphologically, the presence of relatively small amounts of incompletely cleaved CA in the viral

core lead to incorrect uncoating post-entry, preventing the formation of the reverse transcriptase complex (Muller et al., 2009). Hence, it is possible that a difference in the efficiency of viral uncoating in the presence of PI may contribute to the different PI susceptibilities of the JRFL, YU2 and p8.9NSX+. In addition, mutations at positions 89 and 90 of Gag have been shown to affect infectivity by the inhibition of interactions with host factors cyclophilin A and Nup358, necessary for entry and uncoating steps (Schaller et al., 2011).

This efficiency of various viral life cycle stages including reverse transcription, nuclear entry and integration could be tested using qPCR based assays and compared between the molecular clones and p8.9NSX+ (Butler et al., 2001; Mbisa et al., 2009). This assay can distinguish early, intermediate and late RT products as well as products present following nuclear import, LTR circles and integrated cDNA, by the design of primers targeting the regions present exclusively at each stage. For example primers-probe sets covering the U5-R-U3 regions to target early RT products as this region is only present during minus DNA strand synthesis following the first DNA strand jump. For the purposes of this study, comparison of the amount of early and late RT products, and integrated cDNA (performed using primers targeting the R-U5 of HIV and the Alu repeats present throughout the human genome) would enable the identification of reduced efficiency of reverse transcription and integration. If required, qPCR to distinguish a surrogate marker of nuclear import (LTR circles) and integrated cDNA could also be performed.

7.3.2 Variation in PI susceptibility of subtype B, patient-derived PI-naïve viruses

This study examined the variability in PI susceptibility of PI-naïve subtype B viruses derived directly from patient viruses. A major limitation was that only four patients were included, but the MONARK study detailed in chapter 5 also included an additional five patients infected with subtype B viruses. These data

from the MONARK patients support the findings of the PI-naïve patients from PHE, with both demonstrating variation in PI susceptibility in particular to the PIs ATV and LPV. Unfortunately, clinical information was not available for the PHE patients as this would have been useful in determining the treatment outcomes for any who subsequently received PI containing regimens.

In this study, we have shown variation in PI susceptibility conferred by protease in the absence of major resistance mutations. Whilst we were able to identify three polymorphisms that contributed to this reduced susceptibility – I13V, L63P and A71T – they did not fully account for the reduction in susceptibility. Experiments to determine the other genetic determinants are important, as a change affecting PI susceptibility may also confer resistance. As proposed for the identification of further Gag amino acid changes in section 7.3.1, site-directed mutagenesis and PI susceptibility experiments could shed light on the role of the other protease polymorphisms present in each of the two patients: I15V, P39E, D60E, I72V, V77I and I93L in patient 2 and K14R, E35D, N37T, R41K and I62V in patient 4. The same methods could be used to determine whether any of the changes identified at higher frequency in patients failing ATV using the UK HIV Drug Resistance Database have a direct effect on PI susceptibility, either separately or in combination (Dolling et al., 2013b).

In the future, work to identify the mechanisms by which protease polymorphisms affect PI susceptibility should be carried out. Computer modelling performed here has shown that the three polymorphisms analysed phenotypically were not located near the substrate binding cleft so are unlikely to directly affect the binding of the PI to the protease enzyme, but further experiments could be performed. The efficiency of the protease enzyme from each patient virus and from the reference strain could be measured using a previously described assay involving the cleavage of chromogenic substrates by protease proteins produced in *E. coli* (Velazquez-Campoy et al., 2001). This assay could also be used to further investigate the effect of the polymorphisms

identified in collaboration with the UK HIV Drug Resistance Database (Dolling et al., 2013b).

There is evidence that some of the protease polymorphisms studied phenotypically may affect the pathways to resistance to PIs (Lisovsky et al., 2010; Vergne et al., 2006). In the future, *in vitro* passage experiments with replication-competent viruses to examine the evolution of resistance to ATV should be performed. To date, only one study has described *in vitro* passage in the presence of ATV and the authors only described the data from one experimental repeat (Gong et al., 2000). *In vitro* passage with ATV could be performed using methods that have been previously described (Patick et al., 1995). Viruses containing the polymorphisms analysed phenotypically in combination, I13V, L63P and A71T, could be included which would show whether they affect the development of other resistance mutations. In addition the novel protease mutations present at higher frequency in patients failing ATV therapy could also be examined in this way (Dolling et al., 2013b). It would be particularly interesting to passage a number of different subtype B viruses from different patients with ATV and see whether any of these novel protease mutations described here appear over time.

7.3.3 Determinants of treatment failure in LPV/r monotherapy arm of the MONARK trial

This study examined 13 patients in total; 5 non-controllers who subsequently experienced treatment failure and 8 controllers who achieved a sustained virological response. This study could have been improved by the inclusion of more patients, but the sensitivity of the PCR methods employed limited the number of failure samples with low viral loads from which Gag-protease could be amplified. Another study has described a PCR that resulted in amplification from a diverse range of subtypes and from viral loads down to 500 copies/ml (Van Laethem et al., 2006). However, in 13/23 non-controller patients the viral

load at treatment failure was below 500 copies/ml. The reverse primer used for cDNA synthesis in this study was taken from this publication, KVL065, but in our hands the method described by Van Laethem et al. (2006) was less sensitive than our existing one. Plans are in place to extend the study so that the is a 1:1 ratio of non-controllers to controllers.

The small number of patients was further complicated by the fact that patients were infected with viruses of different subtypes with three subtypes represented here: subtype B, CRF02_AG and subtype G. The genotypic and phenotypic differences between the viruses of different subtypes further complicated the analysis of our data. For example, M36I in protease and E12K in Gag have been described to affect PI susceptibility in subtype B viruses, but are consensus amino acids in subtypes CRF02_AG and G (Aoki et al., 2009; Baxter et al., 2006). In the future, further statistical analyses including Mann Whitney U Rank Sum tests and conditional logistic regression will be performed. This will test whether any factors are significantly associated with treatment outcome, for example virus subtype or LPV susceptibility.

As well as an expansion of the number of patients, a more detailed study of the patient viruses already characterised is important. In particular, whilst the ancestral site reconstruction has identified changes that correlate with PI exposure and therapy failure, we have not directly shown the effect of particular mutations on PI susceptibility, as in chapters 3 and 4. Sites correlating with treatment failure were identified in patients #1403 (KON – 3 sites), #4210 (SO – 4 sites) and #3204 (HG – 7 changes). Site-directed mutagenesis techniques could be used to revert these mutations to consensus B sequence in the patient vectors and introduce these changes into the reference strain. Further PI susceptibility assays could then be used to determine their exact role in PI susceptibility, and hence the potential role in treatment failure.

Another of the limitations of our study was the use of clonal analysis to examine the quasispecies within a patient. Other PCR and sequencing techniques could be used in the future which would enable sequencing to a greater depth and the linkage of mutations on a single viral genome. Single Genome Analysis (SGA), a terminal dilution PCR assay that results in the amplification of a PCR product from a single RNA genome, could be used (Palmer et al., 2005). A SGA assay covering full-length *pol* is already in use in our lab and a SGA assay covering Gag has been previously described and could be implemented (McKinnon et al., 2011). This would offer an important advantage over the existing method as linkage between mutations on a single viral genome could be shown and the frequency of recombination would be greatly reduced.

Next generation sequencing of certain patient samples would provide insight, in particular those from patient #1403 (KON) which would shed light on the development of the V82A mutation at failure, as discussed in section 5.3. In particular the frequency of resistance mutations or other amino acid changes of interest within the viral quasispecies could be determined with greater accuracy. Certain next generation sequencing assays are now able to detect resistance mutations present in as low as 1% of viral variants, although the exact effect of low level viral variants on treatment outcome is yet to be determined (Lataillade et al., 2010).

It would be interesting to examine other PI monotherapy studies to further explore the utility of PI monotherapy, in particular those including patients with other subtypes and those using different PIs. Of particular interest would be the more recent studies performed with DRV/r monotherapy, a more potent PI than LPV with a high genetic barrier to resistance, and trials performed in the absence of close virological monitoring. In the future, the use of ancestral site reconstruction as will help to identify novel mutations that occur under PI pressure that could may contribute to treatment failure.

7.4 Final comments

Here, we have investigated the determinants of PI susceptibility utilising a phenotypic PI susceptibility assay encompassing full-length, co-evolved Gag-protease from the virus of interest. We have demonstrated significant variation in PI susceptibility in PI-naïve viruses of subtypes B, CRF02_AG and G, showing the direct role of Gag and protease independently in this variation. We have shown evidence for the contribution of variation in susceptibility of Gag-protease in the absence of major protease resistance mutations to treatment failure on PI monotherapy, an important and novel finding. Our data further support the importance of considering Gag and protease together when investigating PI susceptibility.

7.4.1.1 Reference List

Abdool,K.Q., Abdool Karim,S.S., Frohlich,J.A., Grobler,A.C., Baxter,C., Mansoor,L.E., Kharsany,A.B., Sibeko,S., Mlisana,K.P., Omar,Z., Gengiah,T.N., Maarschalk,S., Arulappan,N., Mlotshwa,M., Morris,L., and Taylor,D. (2010). Effectiveness and safety of tenofovir gel, an antiretroviral microbicide, for the prevention of HIV infection in women. Science *329*, 1168-1174.

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

Adamson,C.S., Waki,K., Ablan,S.D., Salzwedel,K., and Freed,E.O. (2009). Impact of human immunodeficiency virus type 1 resistance to protease inhibitors on evolution of resistance to the maturation inhibitor bevirimat (PA-457). J. Virol. *83*, 4884-4894.

Ait-Khaled,M., Rakik,A., Griffin,P., Stone,C., Richards,N., Thomas,D., Falloon,J., and Tisdale,M. (2003). HIV-1 reverse transcriptase and protease resistance mutations selected during 16-72 weeks of therapy in isolates from antiretroviral therapy-experienced patients receiving abacavir/efavirenz/amprenavir in the CNA2007 study. Antivir. Ther. *8*, 111-120.

Amico,P., Aran,C., and Avila,C. (2010). HIV Spending as a Share of Total Health Expenditure: An Analysis of Regional Variation in a Multi-Country Study. PLoS One *5*, e12997.

Aoki,M., Venzon,D.J., Koh,Y., Aoki-Ogata,H., Miyakawa,T., Yoshimura,K., Maeda,K., and Mitsuya,H. (2009). Non-cleavage site gag mutations in amprenavir-resistant human immunodeficiency virus type 1 (HIV-1) predispose HIV-1 to rapid acquisition of amprenavir resistance but delay development of resistance to other protease inhibitors. J. Virol. *83*, 3059-3068.

Arien,K.K., Vanham,G., and Arts,E.J. (2007). Is HIV-1 evolving to a less virulent form in humans? Nat. Rev. Microbiol. *5*, 141-151.

Arribas,J.R., Horban,A., Gerstoft,J., Fatkenheuer,G., Nelson,M., Clumeck,N., Pulido,F., Hill,A., van,D.Y., Stark,T., and Moecklinghoff,C. (2010). The MONET trial: darunavir/ritonavir with or without nucleoside analogues, for patients with HIV RNA below 50 copies/ml. AIDS *24*, 223-230.

Arts,E.J. and Hazuda,D.J. (2012). HIV-1 antiretroviral drug therapy. Cold Spring Harb. Perspect. Med *2*, a007161.

Asaoka,K., Ikeda,K., Hishinuma,T., Horie-Inoue,K., Takeda,S., and Inoue,S. (2005). A retrovirus restriction factor TRIM5alpha is transcriptionally regulated by interferons. Biochem. Biophys. Res. Commun. *338*, 1950-1956.

Bacheler,L., Jeffrey,S., Hanna,G., D'Aquila,R., Wallace,L., Logue,K., Cordova,B., Hertogs,K., Larder,B., Buckery,R., Baker,D., Gallagher,K., Scarnati,H., Tritch,R., and Rizzo,C. (2001). Genotypic correlates of phenotypic resistance to efavirenz in virus isolates from patients failing nonnucleoside reverse transcriptase inhibitor therapy. J. Virol. *75*, 4999-5008.

Bailey,H., Townsend,C., Cortina-Borja,M., and Thorne,C. (2011). Insufficient antiretroviral therapy in pregnancy: missed opportunities for prevention of mother-to-child transmission of HIV in Europe. Antivir. Ther. *16*, 895-903.

Baldauf,H.M., Pan,X., Erikson,E., Schmidt,S., Daddacha,W., Burggraf,M., Schenkova,K., Ambiel,I., Wabnitz,G., Gramberg,T., Panitz,S., Flory,E., Landau,N.R., Sertel,S., Rutsch,F., Lasitschka,F., Kim,B., Konig,R., Fackler,O.T., and Keppler,O.T. (2012). SAMHD1 restricts HIV-1 infection in resting CD4(+) T cells. Nat. Med *18*, 1682-1687.

Bally,F., Martinez,R., Peters,S., Sudre,P., and Telenti,A. (2000). Polymorphism of HIV type 1 gag p7/p1 and p1/p6 cleavage sites: clinical significance and implications for resistance to protease inhibitors. AIDS Res. Hum. Retroviruses *16*, 1209-1213.

Barber,T.J., Harrison,L., Asboe,D., Williams,I., Kirk,S., Gilson,R., Bansi,L., Pillay,D., and Dunn,D. (2012). Frequency and patterns of protease gene resistance mutations in HIV-infected patients treated with lopinavir/ritonavir as their first protease inhibitor. J. Antimicrob. Chemother. *67*, 995-1000.

Barbour,J.D., Hecht,F.M., Wrin,T., Liegler,T.J., Ramstead,C.A., Busch,M.P., Segal,M.R., Petropoulos,C.J., and Grant,R.M. (2004). Persistence of primary drug resistance among recently HIV-1 infected adults. AIDS *18*, 1683-1689.

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

Baxter,J.D., Schapiro,J.M., Boucher,C.A., Kohlbrenner,V.M., Hall,D.B., Scherer,J.R., and Mayers,D.L. (2006). Genotypic changes in human immunodeficiency virus type 1 protease associated with reduced susceptibility and virologic response to the protease inhibitor tipranavir. J. Virol. *80*, 10794-10801.

Berkhout,B. and Jeang,K.T. (1992). Functional roles for the TATA promoter and enhancers in basal and Tat-induced expression of the human immunodeficiency virus type 1 long terminal repeat. J. Virol. *66*, 139-149.

Berman,H.M., Westbrook,J., Feng,Z., Gilliland,G., Bhat,T.N., Weissig,H., Shindyalov,I.N., and Bourne,P.E. (2000). The Protein Data Bank. Nucleic Acids Res. *28*, 235-242.

Bertoli, A., Santoro, M, Lorenzini, P, Ceccherini-Silberstein, F, Lazzarin, A., Di Perri, G, Esposito, D, Caramello, P, Cargme, A, Cargnel, A, Narciso, P, Rizzrdini, G, Filice, G, Minoli, L, Carosi, G, Antinori, A, and Perno, C. F. Different patterns of mutations involved in tghe genotypic resistance score for atazanavir versus atazanavir unboosted in multiplying fialing patients. HIVDRW2006 . 2006.
Ref Type: Abstract

Bishop,K.N., Verma,M., Kim,E.Y., Wolinsky,S.M., and Malim,M.H. (2008). APOBEC3G inhibits elongation of HIV-1 reverse transcripts. PLoS. Pathog. *4*, e1000231.

Bour,S., Geleziunas,R., and Wainberg,M.A. (1995). The human immunodeficiency virus type 1 (HIV-1) CD4 receptor and its central role in promotion of HIV-1 infection. Microbiol. Rev. *59*, 63-93.

Boyd,M.A., Siangphoe,U., Ruxrungtham,K., Duncombe,C.J., Stek,M., Lange,J.M., Cooper,D.A., and Phanuphak,P. (2005). Indinavir/ritonavir 800/100 mg bid and efavirenz 600 mg qd in patients failing treatment with combination nucleoside reverse transcriptase inhibitors: 96-week outcomes of HIV-NAT 009. HIV. Med *6*, 410-420.

Brennan,C.A., Bodelle,P., Coffey,R., Devare,S.G., Golden,A., Hackett,J., Jr., Harris,B., Holzmayer,V., Luk,K.C., Schochetman,G., Swanson,P., Yamaguchi,J., Vallari,A., Ndembi,N., Ngansop,C., Makamche,F., Mbanya,D., Gurtler,L.G., Zekeng,L., and Kaptue,L. (2008). The prevalence of diverse HIV-1 strains was stable in Cameroonian blood donors from 1996 to 2004. J. Acquir. Immune. Defic. Syndr. *49*, 432-439.

Brenner,B.G., Routy,J.P., Petrella,M., Moisi,D., Oliveira,M., Detorio,M., Spira,B., Essabag,V., Conway,B., Lalonde,R., Sekaly,R.P., and Wainberg,M.A. (2002). Persistence and fitness of multidrug-resistant human immunodeficiency virus type 1 acquired in primary infection. J. Virol. *76*, 1753-1761.

Briggs,J.A. and Krausslich,H.G. (2011). The molecular architecture of HIV. J. Mol. Biol *410*, 491-500.

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

Burns,J.C., Friedmann,T., Driever,W., Burrascano,M., and Yee,J.K. (1993) Vesicular stomatitis virus G glycoprotein pseudotyped retroviral vectors: Concentration to very high titer and efficient gene transfer into mammalian and nonmammalian cells. Proc. Natl. Aca. Sci. *90,*8033-8037.

Butler,S.L., Hansen,M.S., and Bushman,F.D. (2001). A quantitative assay for HIV DNA integration in vivo. Nat. Med. *7*, 631-634.

Buzon,M.J., Massanella,M., Llibre,J.M., Esteve,A., Dahl,V., Puetas,M.C., Gatell,J.M., Domingo,P., Paredes,R., Sharkey,M., Palmer,S., Stevenson,M., Clotet,B., Blanco,J., and Martinez-Picado,J. (2010). Nat Med. *16*, 460-465.

Callebaut,C., Stray,K., Tsai,L., Xu,L., Lee,W., and Cihlar,T. (2007). In vitro HIV-1 resistance selection to GS8374, a novel phosphonate protease inhibitor: comparison with lopinavir, atazanavir and darunavir. Antivir. Ther. *12*, S18.

Carrillo,A., Stewart,K.D., Sham,H.L., Norbeck,D.W., Kohlbrenner,W.E., Leonard,J.M., Kempf,D.J., and Molla,A. (1998). In vitro selection and characterization of human immunodeficiency virus type 1 variants with increased resistance to ABT-378, a novel protease inhibitor. J. Virol. *72*, 7532-7541.

Castagna,A., Spagnuolo,V., Galli,L., Vinci,C., Nozza,S., and Carini,E. (2013). 48 weeks outcomes of atazanavir/ritonavir monotherapy as maintenance strategy in HIV-1 treated subjects with viral suppression: interim analysis results of the MODAt Study. EACS 2013.

Chatterji,U., Bobardt,M.D., Gaskill,P., Sheeter,D., Fox,H., and Gallay,P.A. (2006). Trim5alpha accelerates degradation of cytosolic capsid associated with productive HIV-1 entry. J. Biol Chem. *281*, 37025-37033.

Chen,P., Hubner,W., Spinelli,M.A., and Chen,B.K. (2007). Predominant mode of human immunodeficiency virus transfer between T cells is mediated by sustained Env-dependent neutralization-resistant virological synapses. J. Virol. *81*, 12582-12595.

Chin-Hong,P.V., Deeks,S.G., Liegler,T., Hagos,E., Krone,M.R., Grant,R.M., and Martin,J.N. (2005). High-risk sexual behavior in adults with genotypically proven antiretroviral-resistant HIV infection. J. Acquir. Immune. Defic. Syndr. *40*, 463-471.

Chojnacki,J., Staudt,T., Glass,B., Bingen,P., Engelhardt,J., Anders,M., Schneider,J., Muller,B., Hell,S.W., and Krausslich,H.G. (2012). Maturation-dependent HIV-1 surface protein redistribution revealed by fluorescence nanoscopy. *338*, 524-8.

Christ,F., and Debyser,Z. (2013). The LEDGF/p75 integrase interaction, a novel target for anti-HIV therapy. Virology. *435*, 102-9.

Cihlar,T. and Ray,A.S. (2010). Nucleoside and nucleotide HIV reverse transcriptase inhibitors: 25 years after zidovudine. Antiviral Res. *85*, 39-58.

Clavel,F. and Mammano,F. (2010). Role of Gag in HIV Resistance to Protease Inhibitors. Viruses. *2*, 1411-1426.

Coffin,J.M., Hughes,S.H., and Varmus,H.E. (1997). Retroviruses. Cold Spring Harbor Laboratory Press).

Cohen,M.S., Chen,Y.Q., McCauley,M., Gamble,T., Hosseinipour,M.C., Kumarasamy,N., Hakim,J.G., Kumwenda,J., Grinsztejn,B., Pilotto,J.H., Godbole,S.V., Mehendale,S., Chariyalertsak,S., Santos,B.R., Mayer,K.H., Hoffman,I.F., Eshleman,S.H., Piwowar-Manning,E., Wang,L., Makhema,J., Mills,L.A., de,B.G., Sanne,I., Eron,J., Gallant,J., Havlir,D., Swindells,S., Ribaudo,H., Elharrar,V., Burns,D., Taha,T.E., Nielsen-Saines,K., Celentano,D., Essex,M., and Fleming,T.R. (2011). Prevention of HIV-1 infection with early antiretroviral therapy. N. Engl. J. Med *365*, 493-505.

Collins,J.R., Burt,S.K., and Erickson,J.W. (1995). Activated dynamics of flap opening in HIV-1 protease. Adv. Exp. Med Biol *362*, 455-460.

Collman,R., Balliet,J.W., Gregory,S.A., Friedman,H., Kolson,D.L., Nathanson,N., and Srinivasan,A. (1992). An infectious molecular clone of an unusual macrophage-tropic and highly cytopathic strain of human immunodeficiency virus type 1. J. Virol. *66*, 7517-7521.

Colonno,R., Rose,R., McLaren,C., Thiry,A., Parkin,N., and Friborg,J. (2004). Identification of I50L as the signature atazanavir (ATV)-resistance mutation in treatment-naive HIV-1-infected patients receiving ATV-containing regimens. J. Infect. Dis. *189*, 1802-1810.

Colonno,R.J., Thiry,A., Limoli,K., and Parkin,N. (2003). Activities of atazanavir (BMS-232632) against a large panel of human immunodeficiency virus type 1 clinical isolates resistant to one or more approved protease inhibitors. Antimicrob. Agents Chemother. *47*, 1324-1333.

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

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

Cote,H.C., Brumme,Z.L., and Harrigan,P.R. (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.

Coutsoudis,A., Pillay,K., Spooner,E., Kuhn,L., and Coovadia,H.M. (1999). Influence of infant-feeding patterns on early mother-to-child transmission of HIV-1 in Durban, South Africa: a prospective cohort study. South African Vitamin A Study Group. Lancet *354*, 471-476.

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

Craigie,R., and Bushman,F.D. (2012). HIV DNA Integration. Cold Spring Harb Perspect Med. *2*, a006890.

Croteau,G., Doyon,L., Thibeault,D., McKercher,G., Pilote,L., and Lamarre,D. (1997). Impaired fitness of human immunodeficiency virus type 1 variants with high-level resistance to protease inhibitors. J. Virol. *71*, 1089-1096.

Dam,E., Quercia,R., Glass,B., Descamps,D., Launay,O., Duval,X., Krausslich,H.G., Hance,A.J., and Clavel,F. (2009). Gag mutations strongly contribute to HIV-1 resistance to protease inhibitors in highly drug-experienced patients besides compensating for fitness loss. PLoS. Pathog. *5*, e1000345.

Daniels,S.I., Davis,D.A., Soule,E.E., Stahl,S.J., Tebbs,I.R., Wingfield,P., and Yarchoan,R. (2010). The initial step in human immunodeficiency virus type 1 GagProPol processing can be regulated by reversible oxidation. PLoS. One. *5*, e13595.

Davis,M.R., Jiang,J., Zhou,J., Freed,E.O., and Aiken,C. (2006). A mutation in the human immunodeficiency virus type 1 Gag protein destabilizes the interaction of the envelope protein subunits gp120 and gp41. J Virol. *80*, 2405-17.

Davis,D.A., Soule,E.E., Davidoff,K.S., Daniels,S.I., Naiman,N.E., and Yarchoan,R. (2012). Activity of human immunodeficiency virus type 1 protease inhibitors against the initial autocleavage in Gag-Pol polyprotein processing. Antimicrob. Agents Chemother. *56*, 3620-3628.

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

de Oliveira,T., Deforche,K., Cassol,S., Salminen,M., Paraskevis,D., Seebregts,C., Snoeck,J., van Rensburg,E.J., Wensing,A.M., van de Vijver,D.A., Boucher,C.A., Camacho,R., and Vandamme,A.M. (2005). An automated genotyping system for analysis of HIV-1 and other microbial sequences. Bioinformatics. *21*, 3797-3800.

De,M.S., Vangeneugden,T., Van,B.B., De,P.E., van,M.H., Picchio,G., Lefebvre,E., and de Bethune,M.P. (2008). Resistance profile of darunavir: combined 24-week results from the POWER trials. AIDS Res. Hum. Retroviruses *24*, 379-388.

Delaugerre,C., Flandre,P., Chaix,M.L., Ghosn,J., Raffi,F., Dellamonica,P., Jaeger,H., Shurmann,D., Cohen-Codar,I., Van,P.N., Norton,M., Taburet,A.M., Delfraissy,J.F., and Rouzioux,C. (2009). Protease inhibitor resistance analysis in the MONARK trial comparing first-line lopinavir-ritonavir monotherapy to lopinavir-ritonavir plus zidovudine and lamivudine triple therapy. Antimicrob. Agents Chemother. *53*, 2934-2939.

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

Delfraissy,J.F., Flandre,P., Delaugerre,C., Ghosn,J., Horban,A., Girard,P.M., Norton,M., Rouzioux,C., Taburet,A.M., Cohen-Codar,I., Van,P.N., and Chauvin,J.P. (2008). Lopinavir/ritonavir monotherapy or plus zidovudine and lamivudine in antiretroviral-naive HIV-infected patients. AIDS *22*, 385-393.

Delport,W., Poon,A.F., Frost,S.D., and Kosakovsky Pond,S.L. (2010). Datamonkey 2010: a suite of phylogenetic analysis tools for evolutionary biology. Bioinformatics. *26*, 2455-2457.

Demaison,C., Parsley,K., Brouns,G., Scherr,M., Battmer,K., Kinnon,C., Grez,M., and Thrasher,A.J. (2002). High-level transduction and gene expression in hematopoietic repopulating cells using a human immunodeficiency [correction of imunodeficiency] virus type 1-based lentiviral vector containing an internal spleen focus forming virus promoter. Hum. Gene Ther. *13*, 803-813.

Derdowski,A., Ding,L., and Spearman,P. (2004). A novel fluorescence resonance energy transfer assay demonstrates that the human immunodeficiency virus type 1 Pr55Gag I domain mediates Gag-Gag interactions. J. Virol. *78*, 1230-1242.

Descamps,D., Lambert-Niclot,S., Marcelin,A.G., Peytavin,G., Roquebert,B., Katlama,C., Yeni,P., Felices,M., Calvez,V., and Brun-Vezinet,F. (2009). Mutations associated with virological response to darunavir/ritonavir in HIV-1-infected protease inhibitor-experienced patients. J. Antimicrob. Chemother. *63*, 585-592.

Devereux,H.L., Emery,V.C., Johnson,M.A., and Loveday,C. (2001). Replicative fitness in vivo of HIV-1 variants with multiple drug resistance-associated mutations. J. Med Virol. *65*, 218-224.

Diaz-Griffero,F., Li,X., Javanbakht,H., Song,B., Welikala,S., Stremlau,M., and Sodroski,J. (2006). Rapid turnover and polyubiquitylation of the retroviral restriction factor TRIM5. Virology *349*, 300-315.

Didigu,C.A. and Doms,R.W. (2012). Novel approaches to inhibit HIV entry. Viruses. *4*, 309-324.

Dierynck,I., De,W.M., Gustin,E., Keuleers,I., Vandersmissen,J., Hallenberger,S., and Hertogs,K. (2007). Binding kinetics of darunavir to human immunodeficiency virus type 1 protease explain the potent antiviral activity and high genetic barrier. J. Virol. *81*, 13845-13851.

Dixon,S., McDonald,S., and Roberts,J. (2013). AIDS and economic growth in Africa: a panel data analysis. J Int Dev *13*, 411-426.

Dolling, D. I., Dunn, D., and The UK Collaborative Group on HIV Resistance. The increasing genetic diversity of HIV-1 in the UK, 2002-2010. 2013a.
Ref Type: Unpublished Work

Dolling,D.I., Dunn,D.T., Sutherland,K.A., Pillay,D., Mbisa,J.L., Parry,C.M., Post,F.A., Sabin,C.A., and Cane,P.A. (2013b). Low frequency of genotypic resistance in HIV-1-infected patients failing an atazanavir-containing regimen: a clinical cohort study. J. Antimicrob. Chemother.

Doms,R.W. and Peiper,S.C. (1997). Unwelcomed guests with master keys: how HIV uses chemokine receptors for cellular entry. Virology *235*, 179-190.

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

Doyon,L., Tremblay,S., Bourgon,L., Wardrop,E., and Cordingley,M.G. (2005). Selection and characterization of HIV-1 showing reduced susceptibility to the non-peptidic protease inhibitor tipranavir. Antiviral Res. *68*, 27-35.

Dragic,T., Trkola,A., Thompson,D.A., Cormier,E.G., Kajumo,F.A., Maxwell,E., Lin,S.W., Ying,W., Smith,S.O., Sakmar,T.P., and Moore,J.P. (2000). A binding pocket for a small molecule inhibitor of HIV-1 entry within the transmembrane helices of CCR5. Proc. Natl. Acad. Sci. U. S. A *97*, 5639-5644.

Dykes,C., Fox,K., Lloyd,A., Chiulli,M., Morse,E., and Demeter,L.M. (2001). Impact of clinical reverse transcriptase sequences on the replication capacity of HIV-1 drug-resistant mutants. Virology *285*, 193-203.

Dykes,C., Wu,H., Sims,M., Holden-Wiltse,J., and Demeter,L.M. (2010). Human immunodeficiency virus type 1 protease inhibitor drug-resistant mutants give discordant results when compared in single-cycle and multiple-cycle fitness assays. J. Clin. Microbiol. *48*, 4035-4043.

Engelman,A., and Cherepanov,P. (2012). The structural biology of HIV-1: mechanistic and therapeutic insights. *10*, 279-890.

Espeseth,A.S., Felock,P., Wolfe,A., Witmer,M., Grobler,J., Anthony,N., Egbertson,M., Melamed,J.Y., Young,S., Hamill,T., Cole,J.L., and Hazuda,D.J. (2000). HIV-1 integrase inhibitors that compete with the target DNA substrate define a unique strand transfer conformation for integrase. Proc. Natl. Acad. Sci. U. S. A *97*, 11244-11249.

Fassati,A. (2012). Multiple roles of the capsid protein in the early steps of HIV-1 infection. Virus Res. *170*, 15-24.

Fisher,R.D., Chung,H.Y., Zhai,Q., Robinson,H., Sundquist,W.I., and Hill,C.P. (2007). Structural and biochemical studies of ALIX/AIP1 and its role in retrovirus budding. Cell *128*, 841-852.

Flandre,P., Delaugerre,C., Ghosn,J., Chaix,M.L., Horban,A., Girard,P.M., Gladysz,A., Cohen-Codar,I., Van,P.N., Taburet,A.M., Rouzioux,C., and Delfraissy,J.F. (2009). Prognostic factors for virological response in antiretroviral therapy-naive patients in the MONARK Trial randomized to ritonavir-boosted lopinavir alone. Antivir. Ther. *14*, 93-97.

Fletcher,C.V. (2003). Enfuvirtide, a new drug for HIV infection. Lancet *361*, 1577-1578.

Franke,E.K., Yuan,H.E., and Luban,J. (1994). Specific incorporation of cyclophilin A into HIV-1 virions. Nature *372*, 359-362.

Fransen,S., Gupta,S., Danovich,R., Hazuda,D., Miller,M., Witmer,M., Petropoulos,C.J., and Huang,W. (2009). Loss of raltegravir susceptibility by human immunodeficiency virus type 1 is conferred via multiple nonoverlapping genetic pathways. J. Virol. *83*, 11440-11446.

Fun,A., Wensing,A.M., Verheyen,J., and Nijhuis,M. (2012). Human Immunodeficiency Virus Gag and protease: partners in resistance. Retrovirology. *9*, 63.

Gamble,T.R., Vajdos,F.F., Yoo,S., Worthylake,D.K., Houseweart,M., Sundquist,W.I., and Hill,C.P. (1996). Crystal structure of human cyclophilin A bound to the amino-terminal domain of HIV-1 capsid. Cell *87*, 1285-1294.

Ganser-Pornillos,B.K., Yeager,M., and Sundquist,W.I. (2008). The structural biology of HIV assembly. Curr. Opin. Struct. Biol. *18*, 203-217.

Garrus,J.E., von Schwelder,U.K., Pornillos,O.W., Morham,S.G., Zavitz,K.H., Wang,H.E., Wettstein,D.A., Stray,K.M., Cote,M., Rich,R.L., Myszka,D.G., and Sundquist,W.I. (2001). Tsg101 and the vacuolar protein sorting pathway are essential for HIV-1 budding. Cell. 107, 55-65.

Gatanaga,H., Suzuki,Y., Tsang,H., Yoshimura,K., Kavlick,M.F., Nagashima,K., Gorelick,R.J., Mardy,S., Tang,C., Summers,M.F., and Mitsuya,H. (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.

Geijtenbeek,T.B., Kwon,D.S., Torensma,R., van Vliet,S.J., van Duijnhoven,G.C., Middel,J., Cornelissen,I.L., Nottet,H.S., Kewalramani,V.N., Littman,D.R., Figdor,C.G., and van,K.Y. (2000). DC-SIGN, a dendritic cell-specific HIV-1-binding protein that enhances trans-infection of T cells. Cell *100*, 587-597.

Gheysen,D., Jacobs,E., de,F.F., Thiriart,C., Francotte,M., Thines,D., and De,W.M. (1989). Assembly and release of HIV-1 precursor Pr55gag virus-like particles from recombinant baculovirus-infected insect cells. Cell *59*, 103-112.

Ghosn,J., Delaugerre,C., Flandre,P., Galimand,J., Cohen-Codar,I., Raffi,F., Delfraissy,J.F., Rouzioux,C., and Chaix,M.L. (2011). Polymorphism in Gag gene cleavage sites of HIV-1 non-B subtype and virological outcome of a first-line lopinavir/ritonavir single drug regimen. PLoS. One. *6*, e24798.

Ghosn,J., Flandre,P., Cohen-Codar,I., Girard,P.M., Chaix,M.L., Raffi,F., Dellamonica,P., Ngovan,P., Norton,M., and Delfraissy,J.F. (2010). Long-term (96-week) follow-up of antiretroviral-naive HIV-infected patients treated with first-line lopinavir/ritonavir monotherapy in the MONARK trial. HIV. Med *11*, 137-142.

Gianotti,N. and Lazzarin,A. (2005). Sequencing antiretroviral drugs for long-lasting suppression of HIV replication. New Microbiol. *28*, 281-297.

Gilks,C.F., Walker,A.S., Dunn,D.T., Gibb,D.M., Kikaire,B., Reid,A., Musana,H., Mambule,I., Kasirye,R., Robertson,V., Ssali,F., Spyer,M., Pillay,D., Yirrell,D., and Kaleebu,P. (2012). Lopinavir/ritonavir monotherapy after 24 weeks of second-line antiretroviral therapy in Africa: a randomized controlled trial (SARA). Antivir. Ther. *17*, 1363-1373.

Goldstone,D.C., Ennis-Adeniran,V., Hedden,J.J., Groom,H.C., Rice,G.I., Christodoulou,E., Walker,P.A., Kelly,G., Haire,L.F., Yap,M.W., de Carvalho,L.P., Stoye,J.P., Crow,Y.J., Taylor,I.A., and Webb,M. (2011). HIV-1 restriction factor SAMHD1 is a deoxynucleoside triphosphate triphosphohydrolase. Nature *480*, 379-382.

Gong,Y.F., Robinson,B.S., Rose,R.E., Deminie,C., Spicer,T.P., Stock,D., Colonno,R.J., and Lin,P.F. (2000). In vitro resistance profile of the human immunodeficiency virus type 1 protease inhibitor BMS-232632. Antimicrob. Agents Chemother. *44*, 2319-2326.

Gottlinger,H.G., Dorfman,T., Sodroski,J.G., and Haseltine,W.A. (1991). Effects of mutations affecting the p6 gag protein on human immunodeficiency virus particle release. Proc Natl Acad Sci USA. 88, 3095-9.

Grobler,J.A., Stillmock,K., Hu,B., Witmer,M., Felock,P., Espeseth,A.S., Wolfe,A., Egbertson,M., Bourgeois,M., Melamed,J., Wai,J.S., Young,S., Vacca,J., and Hazuda,D.J. (2002). Diketo acid inhibitor mechanism and HIV-1 integrase: implications for metal binding in the active site of phosphotransferase enzymes. Proc. Natl. Acad. Sci. U. S. A *99*, 6661-6666.

Guay,L.A., Musoke,P., Fleming,T., Bagenda,D., Allen,M., Nakabiito,C., Sherman,J., Bakaki,P., Ducar,C., Deseyve,M., Emel,L., Mirochnick,M., Fowler,M.G., Mofenson,L., Miotti,P., Dransfield, K., Bray,D., Mmiro,F., and Jackson,J.B. (1999). Intrapartum and neonatal single-dose nevirapine compared with zidovudine for prevention of mother-to-child transmission of HIV-1 in Kampala, Uganda: HIVNET 012 randomised trial. Lancet *354,* 795-802.

Guenzel,C.A., Herate,C., and Benichou,S. (2014). HIV-1 Vpr-a still “enigmatic multitasker”. Front Microbiol. *5*, 127.

Guindon,S., Dufayard,J.F., Lefort,V., Anisimova,M., Hordijk,W., and Gascuel,O. (2010). New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. Syst. Biol *59*, 307-321.

Gulick,R.M., Mellors,J.W., Havlir,D., Eron,J.J., Gonzalez,C., McMahon,D., Richman,D.D., Valentine,F.T., Jonas,L., Meibohm,A., Emini,E.A., and Chodakewitz,J.A. (1997). Treatment with indinavir, zidovudine, and lamivudine in adults with human immunodeficiency virus infection and prior antiretroviral therapy. N. Engl. J. Med *337*, 734-739.

Gupta,R., Hill,A., Sawyer,A.W., and Pillay,D. (2008). Emergence of drug resistance in HIV type 1-infected patients after receipt of first-line highly active antiretroviral therapy: a systematic review of clinical trials. Clin. Infect. Dis. *47*, 712-722.

Gupta,R.K., Hue,S., Schaller,T., Verschoor,E., Pillay,D., and Towers,G.J. (2009). Mutation of a single residue renders human tetherin resistant to HIV-1 Vpu-mediated depletion. PLoS. Pathog. *5*, e1000443.

Gupta,R.K., Mlcochova,P., Pelchen-Matthews,A., Petit,S.J., Mattiuzzo,G., Pillay,D., Takeuchi,Y., Marsh,M., and Towers,G.J. (2009). Simian immunodeficiency virus envelope glycoprotein counteracts tetherin/BST-2?CD317 by intracellular sequestration. Proc Natl Acad Sci USA. *106*, 20889-94.

Gupta,R.K., Kohli,A., McCormick,A.L., Towers,G.J., Pillay,D., and Parry,C.M. (2010). Full-length HIV-1 Gag determines protease inhibitor susceptibility within in-vitro assays. AIDS *24*, 1651-1655.

Hahn,B.H., Shaw,G.M., Arya,S.K., Popovic,M., Gallo,R.C., and Wong-Staal,F. (1984). Molecular cloning and characterization of the HTLV-III virus associated with AIDS. Nature *312*, 166-169.

Hare,S., Gupta,S.S., Valkov,E., Engelman,A., and Cherepanov,P. (2010). Retroviral intasome assembly and inhibition of DNA strand transfer. Nature. *464*, 232-6.

Havlir,D.V., Marschner,I.C., Hirsch,M.S., Collier,A.C., Tebas,P., Bassett,R.L., Ioannidis,J.P., Holohan,M.K., Leavitt,R., Boone,G., and Richman,D.D. (1998). Maintenance antiretroviral

therapies in HIV infected patients with undetectable plasma HIV RNA after triple-drug therapy. AIDS Clinical Trials Group Study 343 Team. N. Engl. J. Med *339*, 1261-1268.

He,Y., Xiao,Y., Song,H., Liang,Q., Ju,D., Chen,X., Lu,H., Jing,W., Jiang,S., and Zhang,L. (2008). Design and evaluation of sifuvirtide, a novel HIV-1 fusion inhibitor. J. Biol Chem. *283*, 11126-11134.

Health Protection Agency. HIV in the United Kingdom: 2012 Report . 2012.
Ref Type: Report

Heider,D., Verheyen,J., and Hoffmann,D. (2010). Predicting Bevirimat resistance of HIV-1 from genotype. BMC. Bioinformatics. *11*, 37.

Hertogs,K., de Bethune,M.P., Miller,V., Ivens,T., Schel,P., Van,C.A., Van Den Eynde,C., Van,G., V, Azijn,H., Van,H.M., Peeters,F., Staszewski,S., Conant,M., Bloor,S., Kemp,S., Larder,B., and Pauwels,R. (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.

Hicks,C.B., Cahn,P., Cooper,D.A., Walmsley,S.L., Katlama,C., Clotet,B., Lazzarin,A., Johnson,M.A., Neubacher,D., Mayers,D., and Valdez,H. (2006). Durable efficacy of tipranavir-ritonavir in combination with an optimised background regimen of antiretroviral drugs for treatment-experienced HIV-1-infected patients at 48 weeks in the Randomized Evaluation of Strategic Intervention in multi-drug reSistant patients with Tipranavir (RESIST) studies: an analysis of combined data from two randomised open-label trials. Lancet *368*, 466-475.

Hijazi,K., Wang,Y., Scala,C., Jeffs,S., Longstaff,C., Stieh,D., Haggarty,B., Vanham,G., Schols,D., Balzarini,J., Jones,I.M,., Hoxie,J., Shattock,R., and Kelly,C.G. (2011). PLoS One. DC_SIGN increases the affinity of HIV-1 envelope glycoprotein interaction with CD4. *6,* e28307.

Hilditch,L., and Towers,G.J. (2014). A model for cofactor use during HIV-1 reverse transcription and nuclear entry. Curr Opin Virol. *4,* 32-6.

Hodcroft,E., Hadfield,J.D., Fearnhill,E., Philips,A., Dunn,D., O'Shea,S., Leigh Brown,A.J., UK HIV Drug Resistance Database, UK CHIC Study. (2014). Contribution of viral genotype to plasma viral set-point in HIV infection. PLoS Pathog. *10*, e1004112.

Hu,C., Saenz,D.T., Fadel,H.J., Walker,W., Peretz,M., and Poeschla,E.M. (2012). The HIV-1 central polypurine tract functions as a second line of defense against APOBEC3G/F. J. Virol. *84*, 11981-93

Hu,W.S. and Hughes,S.H. (2012). HIV-1 reverse transcription. Cold Spring Harb Perspect Med. *2*, a006882

Ibe,S., Yokomaku,Y., Shiino,T., Tanaka,R., Hattori,J., Fujisaki,S., Iwatani,Y., Mamiya,N., Utsumi,M., Kato,S., Hamaguchi,M., and Sugiura,W. (2010). HIV-2 CRF01_AB: first circulating recombinant form of HIV-2. J. Acquir. Immune. Defic. Syndr. *54*, 241-247.

Iglesias,C., Ringeard,M., Di Nunzio,F., Fernandez,J., Guadin,R., Souque,P., Charneau,P., and Arhel,N. (2011). Residual HIV-1 DNA Flap-independent nuclear import of cPPT/CTS double mutant viruses does not support spreading infection. Retrovirology, *8*, 92.

Jacobsen,H., Haenggi,M., Ott,M., Duncan,I.B., Andreoni,M., Vella,S., and Mous,J. (1996). Reduced sensitivity to saquinavir: an update on genotyping from phase I/II trials. Antiviral Res. *29*, 95-97.

Jayaweera,D. and Dilanchian,P. (2012). New therapeutic landscape of NNRTIs for treatment of HIV: a look at recent data. Expert. Opin. Pharmacother. *13*, 2601-2612.

Jinnopat,P., Isarangkura-na-ayuthaya,P., Utachee,P., Kitagawa,Y., de Silva,U.C., Siripanyaphinyo,U., Kameoka,Y., Tokunaga,K., Sawanpanyalert,P., Ikuta,K., Auwanit,W., and Kameoka,M. (2009). Impact of amino acid variations in Gag and protease of HIV type 1 CRF01_AE strains on drug susceptibility of virus to protease inhibitors. J. Acquir. Immune. Defic. Syndr. *52*, 320-328.

Johnson,M., Grinsztejn,B., Rodriguez,C., Coco,J., DeJesus,E., Lazzarin,A., Lichtenstein,K., Rightmire,A., Sankoh,S., and Wilber,R. (2005). Atazanavir plus ritonavir or saquinavir, and lopinavir/ritonavir in patients experiencing multiple virological failures. AIDS *19*, 685-694.

Johnson,V.A., Calvez,V., Gunthard,H.F., Paredes,R., Pillay,D., Shafer,R., Wensing,A.M., and Richman,D.D. (2011). 2011 update of the drug resistance mutations in HIV-1. Top. Antivir. Med *19*, 156-164.

Johnson,V.A., Calvez,V., Gunthard,H.F., Paredes,R., Pillay,D., Shafer,R.W., Wensing,A.M., and Richman,D.D. (2013). Update of the drug resistance mutations in HIV-1: March 2013. Top. Antivir. Med *21*, 6-14.

Johnston,E., Winters,M.A., Rhee,S.Y., Merigan,T.C., Schiffer,C.A., and Shafer,R.W. (2004). Association of a novel human immunodeficiency virus type 1 protease substrate cleft mutation, L23I, with protease inhibitor therapy and in vitro drug resistance. Antimicrob. Agents Chemother. *48*, 4864-4868.

Jolly,C., Kashefi,K., Hollinshead,M., and Sattentau,Q.J. (2004). HIV-1 cell to cell transfer across an Env-induced, actin-dependent synapse. J. Exp. Med *199*, 283-293.

Kagan,R.M., Shenderovich,M.D., Heseltine,P.N., and Ramnarayan,K. (2005). Structural analysis of an HIV-1 protease I47A mutant resistant to the protease inhibitor lopinavir. Protein Sci. *14*, 1870-1878.

Kameoka,M., Isarangkura-na-ayuthaya,P., Kameoka,Y., Sapsutthipas,S., Soonthornsata,B., Nakamura,S., Tokunaga,K., Sawanpanyalert,P., Ikuta,K., and Auwanit,W. (2010). The role of lysine residue at amino acid position 165 of human immunodeficiency virus type 1 CRF01_AE Gag in reducing viral drug susceptibility to protease inhibitors. Virology *405*, 129-138.

Karlstrom,O., Josephson,F., and Sonnerborg,A. (2007). Early virologic rebound in a pilot trial of ritonavir-boosted atazanavir as maintenance monotherapy. J. Acquir. Immune. Defic. Syndr. *44*, 417-422.

Karn,J., and Stoltzfus,C.M. (2012). Transcription and posttranscriptional regulation of HIV-1 gene expression. Cold Spring Harb Perspect Med. *2,* a006919.

Katlama,C., Clotet,B., Mills,A., Trottier,B., Molina,J.M., Grinsztejn,B., Towner,W., Haubrich,R., Nijs,S., Vingerhoets,J., Woodfall,B., and Witek,J. (2010a). Efficacy and safety of etravirine at week 96 in treatment-experienced HIV type-1-infected patients in the DUET-1 and DUET-2 trials. Antivir. Ther. *15*, 1045-1052.

Katlama,C., Esposito,R., Gatell,J.M., Goffard,J.C., Grinsztejn,B., Pozniak,A., Rockstroh,J., Stoehr,A., Vetter,N., Yeni,P., Parys,W., and Vangeneugden,T. (2007). Efficacy and safety of TMC114/ritonavir in treatment-experienced HIV patients: 24-week results of POWER 1. AIDS *21*, 395-402.

Katlama,C., Valantin,M.A., Algarte-Genin,M., Duvivier,C., Lambert-Niclot,S., Girard,P.M., Molina,J.M., Hoen,B., Pakianather,S., Peytavin,G., Marcelin,A.G., and Flandre,P. (2010b). Efficacy of darunavir/ritonavir maintenance monotherapy in patients with HIV-1 viral suppression: a randomized open-label, noninferiority trial, MONOI-ANRS 136. AIDS *24*, 2365-2374.

Keller,P.W., Adamson,C.S., Heymann,J.B., Freed,E.O., and Steven,A.C. (2011). HIV-1 maturation inhibitor bevirimat stabilizes the immature Gag lattice. J. Virol. *85*, 1420-1428.

Kempf,D.J., Isaacson,J.D., King,M.S., Brun,S.C., Xu,Y., Real,K., Bernstein,B.M., Japour,A.J., Sun,E., and Rode,R.A. (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.

Kempf,D.J., Marsh,K.C., Denissen,J.F., McDonald,E., Vasavanonda,S., Flentge,C.A., Green,B.E., Fino,L., Park,C.H., Kong,X.P., and . (1995). ABT-538 is a potent inhibitor of human immunodeficiency virus protease and has high oral bioavailability in humans. Proc. Natl. Acad. Sci. U. S. A *92*, 2484-2488.

King,M.S., Brun,S.C., and Kempf,D.J. (2005). Relationship between adherence and the development of resistance in antiretroviral-naive, HIV-1-infected patients receiving lopinavir/ritonavir or nelfinavir. J. Infect. Dis. *191*, 2046-2052.

King,N.M., Prabu-Jeyabalan,M., Bandaranayake,R.M., Nalam,M.N., Nalivaika,E.A., Ozen,A., Haliloglu,T., Yilmaz,N.K., and Schiffer,C.A. (2012). Extreme entropy-enthalpy compensation in a drug-resistant variant of HIV-1 protease. ACS Chem. Biol *7*, 1536-1546.

Kitahata,M.M., Gange,S.J., Abraham,A.G., Merriman,B., Saag,M.S., Justice,A.C., Hogg,R.S., Deeks,S.G., Eron,J.J., Brooks,J.T., Rourke,S.B., Gill,M.J., Bosch,R.J., Martin,J.N., Klein,M.B., Jacobson,L.P., Rodriguez,B., Sterling,T.R., Kirk,G.D., Napravnik,S., Rachlis,A.R., Calzavara,L.M., Horberg,M.A., Silverberg,M.J., Gebo,K.A., Goedert,J.J., Benson,C.A., Collier,A.C., Van Rompaey,S.E., Crane,H.M., McKaig,R.G., Lau,B., Freeman,A.M., and Moore,R.D. (2009). Effect of early versus deferred antiretroviral therapy for HIV on survival. N. Engl. J. Med *360*, 1815-1826.

Koh,Y., Amano,M., Towata,T., Danish,M., Leshchenko-Yashchuk,S., Das,D., Nakayama,M., Tojo,Y., Ghosh,A.K., and Mitsuya,H. (2010). In vitro selection of highly darunavir-resistant and replication-competent HIV-1 variants by using a mixture of clinical HIV-1 isolates resistant to multiple conventional protease inhibitors. J. Virol. *84*, 11961-11969.

Kohlstaedt,L.A., Wang,J., Friedman,J.M., Rice,P.A., and Steitz,T.A. (1992). Crystal structure at 3.5 A resolution of HIV-1 reverse transcriptase complexed with an inhibitor. Science *256*, 1783-1790.

Komarova,N.L., Anghelina,D., Voznesensky,I., Trinite,B., Levy,D.N., and Wodarz,D. (2013). Relative contribution of free-virus and synaptic transmission to the spread of HIV-1 through target cell populations. Biol Lett. *9*, 20121049.

Koning,F.A., Castro,H., Dunn,D., Tilston,P., Cane,P.A., and Mbisa,J.L. (2013). Subtype-specific differences in the development of accessory mutations associated with high-level resistance to HIV-1 nucleoside reverse transcriptase inhibitors. J. Antimicrob. Chemother. *68*, 1220-1236.

Korber,B., Muldoon,M., Theiler,J., Gao,F., Gupta,R., Lapedes,A., Hahn,B.H., Wolinsky,S., and Bhattacharya,T. (2000). Timing the ancestor of the HIV-1 pandemic strains. Science *288*, 1789-1796.

Kosakovsky Pond,S.L. and Frost,S.D. (2005). Not so different after all: a comparison of methods for detecting amino acid sites under selection. Mol. Biol Evol. *22*, 1208-1222.

Koyanagi,Y., Miles,S., Mitsuyasu,R.T., Merrill,J.E., Vinters,H.V., and Chen,I.S. (1987). Dual infection of the central nervous system by AIDS viruses with distinct cellular tropisms. Science *236*, 819-822.

Krausslich,H.G., Ingraham,R.H., Skoog,M.T., Wimmer,E., Pallai,P.V., and Carter,C.A. (1989). Activity of purified biosynthetic proteinase of human immunodeficiency virus on natural substrates and synthetic peptides. Proc. Natl. Acad. Sci. U. S. A *86*, 807-811.

Kuiken,C., Foley,B., Leitner,T., Apetrei,C., Hahn,B., Mizrachi,I., Mullins,J., Rambaut,A., Wolinsky,S., and Korber,B. (2012). HIV Sequence Compendium. Theoretical Biology and Biophysics Group, Los Alamos National Laboratory).

Kutluay,S.B., and Bieniasz,P.D. (2010). Analysis of the initiating events in HIV-1 particle assembly and genome packaging. PLoS Pathog. *6*, e1001200.

Kuzembayeva,M., Dilley,K., Sardo,L., and Hu,W.S. (2014) Life of psi: how full-length HIV-1 RNAs become packaged genomes in viral particles. Virology. *454-455*, 362-70.

Lahouassa,H., Daddacha,W., Hofmann,H., Ayinde,D., Logue,E.C., Dragin,L., Bloch,N., Maudet,C., Bertrand,M., Graberg,T., Pancino,G., Priet,S., Canard,B., Laguette,N., Benkirane,M., Transy,C., Landau,N.R., Kim,B., and Margotting-Goguet,F. (2012). SAMHD1 restricts the replication of human immunodeficiency virus type 1 by depleting the intracellular pool of deoxynucleoside triphosphates. Nat Immunol. *12*, 223-8.

Lambert-Niclot,S., Masquelier,B., Cohen,C., I, Soulie,C., Delaugerre,C., Morand-Joubert,L., Charpentier,C., Ferre,V., Plantier,J.C., Montes,B., Carret,S., Perrot,V., Peytavin,G., Costagliola,D., Calvez,V., and Marcelin,A.G. (2012). Impact of lopinavir/ritonavir use on antiretroviral resistance in recent clinical practice. J. Antimicrob. Chemother. *67*, 2487-2493.

Landi,A., Iannucci,V., Nuffel,A.V., Meuwissen,P., and Verhasselt,B. (2011). One protein to rule them all: modulation of cell surface receptors and molecules by HIV Nef. Curr HIV Res. *9*, 496-504.

Larrouy,L., Chazallon,C., Landman,R., Capitant,C., Peytavin,G., Collin,G., Charpentier,C., Storto,A., Pialoux,G., Katlama,C., Girard,P.M., Yeni,P., Aboulker,J.P., Brun-Vezinet,F., and Descamps,D. (2010). Gag Mutations Can Impact Virological Response to Dual Boosted Protease Inhibitor Combinations in Antiretroviral Naive HIV Infected Patients (2 IP - ANRS 127 trial). Antimicrob. Agents Chemother.

Larrouy,L., Lambert-Niclot,S., Charpentier,C., Fourati,S., Visseaux,B., Soulie,C., Wirden,M., Katlama,C., Yeni,P., Brun-Vezinet,F., Calvez,V., Marcelin,A.G., and Descamps,D. (2011).

Positive impact of HIV-1 gag cleavage site mutations on the virological response to darunavir boosted with ritonavir. Antimicrob. Agents Chemother. *55*, 1754-1757.

Lataillade,M., Chiarella,J., Yang,R., DeGrosky,M., Uy,J., Seekins,D., Simen,B., St,J.E., Moreno,E., and Kozal,M. (2012). Virologic failures on initial boosted-PI regimen infrequently possess low-level variants with major PI resistance mutations by ultra-deep sequencing. PLoS. One. *7*, e30118.

Lataillade,M., Chiarella,J., Yang,R., Schnittman,S., Wirtz,V., Uy,J., Seekins,D., Krystal,M., Mancini,M., McGrath,D., Simen,B., Egholm,M., and Kozal,M. (2010). Prevalence and clinical significance of HIV drug resistance mutations by ultra-deep sequencing in antiretroviral-naive subjects in the CASTLE study. PLoS. One. *5*, e10952.

Latinovic,O., Kuruppu,J., Davis,C., Le,N., and Heredia,A. (2009). Pharmacotherapy of HIV-1 Infection: Focus on CCR5 Antagonist Maraviroc. Clin. Med Ther. *1*, 1497-1510.

Lehmann,M., Nikolic,D.S., and Piguet,V. (2011). How HIV-1 takes advantage of the cytoskeleton during replication and cell-to-cell transmission. Viruses. *3*, 1757-76.

Li,Y., Kappes,J.C., Conway,J.A., Price,R.W., Shaw,G.M., and Hahn,B.H. (1991). Molecular characterization of human immunodeficiency virus type 1 cloned directly from uncultured human brain tissue: identification of replication-competent and -defective viral genomes. J. Virol. *65*, 3973-3985.

Liao,F., Alkhatib,G., Peden,K.W., Sharma,G., Berger,E.A., and Farber,J.M. (1997). STRL33, A novel chemokine receptor-like protein, functions as a fusion cofactor for both macrophage-tropic and T cell line-tropic HIV-1. J. Exp. Med *185*, 2015-2023.

Lisovsky,I., Schader,S.M., Martinez-Cajas,J.L., Oliveira,M., Moisi,D., and Wainberg,M.A. (2010). HIV-1 protease codon 36 polymorphisms and differential development of resistance to nelfinavir, lopinavir, and atazanavir in different HIV-1 subtypes. Antimicrob. Agents Chemother. *54*, 2878-2885.

Little,S.J., Holte,S., Routy,J.P., Daar,E.S., Markowitz,M., Collier,A.C., Koup,R.A., Mellors,J.W., Connick,E., Conway,B., Kilby,M., Wang,L., Whitcomb,J.M., Hellmann,N.S., and Richman,D.D. (2002). Antiretroviral-drug resistance among patients recently infected with HIV. N. Engl. J. Med *347*, 385-394.

Liu,T.F. and Shafer,R.W. (2006). Web resources for HIV type 1 genotypic-resistance test interpretation. Clin. Infect. Dis. *42*, 1608-1618.

London HIV Consortium. Summary of ARV prescribing in London updated October 2011b. 2012. Ref Type: Online Source

Louis,J.M., Zhang,Y., Sayer,J.M., Wang,Y.F., Harrison,R.W., and Weber,I.T. (2011). The L76V drug resistance mutation decreases the dimer stability and rate of autoprocessing of HIV-1 protease by reducing internal hydrophobic contacts. Biochemistry *50*, 4786-4795.

Luca,A. (2006). The impact of resistance on viral fitness and its clinical implications. In Antiviral Resistance in Clinical Practice, A.M.Geretti, ed. Mediscipt).

Maddison,D.R., and Maddison,W.P. (2001). McClade. Sinauer Assoc. Sunderland, MA.

Maguire,M.F., Guinea,R., Griffin,P., Macmanus,S., Elston,R.C., Wolfram,J., Richards,N., Hanlon,M.H., Porter,D.J., Wrin,T., Parkin,N., Tisdale,M., Furfine,E., Petropoulos,C., Snowden,B.W., and Kleim,J.P. (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.

Malet,I., Roquebert,B., Dalban,C., Wirden,M., Amellal,B., Agher,R., Simon,A., Katlama,C., Costagliola,D., Calvez,V., and Marcelin,A.G. (2007). Association of Gag cleavage sites to protease mutations and to virological response in HIV-1 treated patients. J. Infect. *54*, 367-374.

Malim,M.H., and Bieniasz,P.D. (2012). HIV restriction factors and mechanisms of evasion. Cold Spring Harb Perspect Med. *2*, a006940.

Mammano,F., Ohagen,A., Hoglund,S., and Gottlinger,H.G. (1994). Role of the major homology region of human immunodeficiency virus type 1 in virion morphogenesis. J. Virol. *68*, 4927-4936.

Mammano,F., Petit,C., and Clavel,F. (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.

Marcelin,A.G., Flandre,P., de,M.C., Roquebert,B., Peytavin,G., Valer,L., Wirden,M., Abbas,S., Katlama,C., Soriano,V., and Calvez,V. (2007). Clinical validation of saquinavir/ritonavir genotypic resistance score in protease-inhibitor-experienced patients. Antivir. Ther. *12*, 247-252.

Martin,N., Welsch,S., Jolly,C., Briggs,J.A., Vaux,D., and Sattentau,Q.J. (2010). Virological synapse-mediated spread of human immunodeficiency virus type 1 between T cells is sensitive to entry inhibition. J. Virol. *84*, 3516-3527.

Martin, S. A., Cane, P. A, Pillay, D, and Mbisa, J. L. The emergence of HIV-1 integrase mutation Y143G upon re-initiation of raltegravir-containing therapy confers high-level resistance to raltegravir and improved viral fitness. Antivir.Ther. 17 Suppl 1. 2012.
Ref Type: Abstract

Martin-Serrano,J., Zang,T., and Bieniasz,P.D. (2001). HIV-1 and Ebola virus encode small peptide motifs that recruit Tsg101 to sites of particle assembly to facilitate egress. Nat. Med *7*, 1313-1319.

Martin-Serrano,J., and Neil,S.J. (2011). Host factors involved in retroviral budding and release. Nat Rev Microbiol. 9, 519-31.

Martinez-Picado,J., Savara,A.V., Sutton,L., and D'Aquila,R.T. (1999). Replicative fitness of protease inhibitor-resistant mutants of human immunodeficiency virus type 1. J. Virol. *73*, 3744-3752.

Mbisa,J.L., Delviks-Frankenberry,K.A., Thomas,J.A., Gorelick,R.J., and Pathak,V.K. (2009). Real-Time PCR Analysis of HIV-1 Replication Post-entry Events. In HIV Protocols, V.R.Prasad and G.V.Kalpana, eds. Humana Press), pp. 55-72.

Mbisa,J.L., Martin,S.A., and Cane,P.A. (2011). Patterns of resistance development with integrase inhibitors in HIV. Infect. Drug Resist. *4*, 65-76.

McKinnon,J.E., Delgado,R., Pulido,F., Shao,W., Arribas,J.R., and Mellors,J.W. (2011). Single genome sequencing of HIV-1 gag and protease resistance mutations at virologic failure during the OK04 trial of simplified versus standard maintenance therapy. Antivir. Ther. *16*, 725-732.

Mellors,J.W., Kingsley,L.A., Rinaldo,C.R., Jr., Todd,J.A., Hoo,B.S., Kokka,R.P., and Gupta,P. (1995). Quantitation of HIV-1 RNA in plasma predicts outcome after seroconversion. Ann. Intern. Med *122*, 573-579.

Meng,B. and Lever,A.M. (2013). Wrapping up the bad news: HIV assembly and release. Retrovirology. *10*, 5.

Menzo,S., Monachetti,A., Balotta,C., Corvasce,S., Rusconi,S., Paolucci,S., Baldanti,F., Bagnarelli,P., and Clementi,M. (2003). Processivity and drug-dependence of HIV-1 protease: determinants of viral fitness in variants resistant to protease inhibitors. AIDS *17*, 663-671.

Meynard,J.L., Vray,M., Morand-Joubert,L., Race,E., Descamps,D., Peytavin,G., Matheron,S., Lamotte,C., Guiramand,S., Costagliola,D., Brun-Vezinet,F., Clavel,F., and Girard,P.M. (2002). Phenotypic or genotypic resistance testing for choosing antiretroviral therapy after treatment failure: a randomized trial. AIDS *16*, 727-736.

Mills,A.M., Nelson,M., Jayaweera,D., Ruxrungtham,K., Cassetti,I., Girard,P.M., Workman,C., Dierynck,I., Sekar,V., Abeele,C.V., and Lavreys,L. (2009). Once-daily darunavir/ritonavir vs. lopinavir/ritonavir in treatment-naive, HIV-1-infected patients: 96-week analysis. AIDS *23*, 1679-1688.

Molina,J.M., Andrade-Villanueva,J., Echevarria,J., Chetchotisakd,P., Corral,J., David,N., Moyle,G., Mancini,M., Percival,L., Yang,R., Thiry,A., and McGrath,D. (2008). Once-daily atazanavir/ritonavir versus twice-daily lopinavir/ritonavir, each in combination with tenofovir and emtricitabine, for management of antiretroviral-naive HIV-1-infected patients: 48 week efficacy and safety results of the CASTLE study. Lancet *372*, 646-655.

Molla,A., Korneyeva,M., Gao,Q., Vasavanonda,S., Schipper,P.J., Mo,H.M., Markowitz,M., Chernyavskiy,T., Niu,P., Lyons,N., Hsu,A., Granneman,G.R., Ho,D.D., Boucher,C.A., Leonard,J.M., Norbeck,D.W., and Kempf,D.J. (1996). Ordered accumulation of mutations in HIV protease confers resistance to ritonavir. Nat. Med. *2*, 760-766.

Moore,M.D. and Hu,W.S. (2009). HIV-1 RNA dimerization: It takes two to tango. AIDS Rev. *11*, 91-102.

Moulard,M. and Decroly,E. (2000). Maturation of HIV envelope glycoprotein precursors by cellular endoproteases. Biochim. Biophys. Acta *1469*, 121-132.

Mugyenyi,P., Walker,A.S., Hakim,J., Munderi,P., Gibb,D.M., Kityo,C., Reid,A., Grosskurth,H., Darbyshire,J.H., Ssali,F., Bray,D., Katabira,E., Babiker,A.G., Gilks,C.F., Grosskurth,H., Munderi,P., Kabuye,G., Nsibambi,D., Kasirye,R., Zalwango,E., Nakazibwe,M., Kikaire,B., Nassuna,G., Massa,R., Fadhiru,K., Namyalo,M., Zalwango,A., Generous,L., Khauka,P., Rutikarayo,N., Nakahima,W., Mugisha,A., Todd,J., Levin,J., Muyingo,S., Ruberantwari,A., Kaleebu,P., Yirrell,D., Ndembi,N., Lyagoba,F., Hughes,P., Aber,M., Lara,A.M., Foster,S., Amurwon,J., Wakholi,B.N., Whitworth,J., Wangati,K., Amuron,B., Kajungu,D., Nakiyingi,J., Omony,W., Fadhiru,K., Nsibambi,D., Khauka,P., Mugyenyi,P., Kityo,C., Ssali,F., Tumukunde,D., Otim,T., Kabanda,J., Musana,H., Akao,J., Kyomugisha,H., Byamukama,A., Sabiiti,J., Komugyena,J., Wavamunno,P., Mukiibi,S., Drasiku,A., Byaruhanga,R., Labeja,O., Katundu,P., Tugume,S., Awio,P., Namazzi,A., Bakeinyaga,G.T., Katabira,H., Abaine,D., Tukamushaba,J.,

Anywar,W., Ojiambo,W., Angweng,E., Murungi,S., Haguma,W., Atwiine,S., Kigozi,J., Namale,L., Mukose,A., Mulindwa,G., Atwiine,D., Muhwezi,A., Nimwesiga,E., Barungi,G., Takubwa,J., Murungi,S., Mwebesa,D., Kagina,G., Mulindwa,M., Ahimbisibwe,F., Mwesigwa,P., Akuma,S., Zawedde,C., Nyiraguhirwa,D., Tumusiime,C., Bagaya,L., Namara,W., Kigozi,J., Karungi,J., Kankunda,R., Enzama,R., Latif,A., Hakim,J., Robertson,V., Reid,A., Chidziva,E., Bulaya-Tembo,R., Musoro,G., Taziwa,F., Chimbetete,C., Chakonza,L., Mawora,A., Muvirimi,C., Tinago,G., Svovanapasis,P., Simango,M., Chirema,O., Machingura,J., Mutsai,S., Phiri,M., Bafana,T., Chirara,M., Muchabaiwa,L., Muzambi,M., Mutowo,J., Chivhunga,T., Chigwedere,E., Pascoe,M., Warambwa,C., Zengeza,E., Mapinge,F., Makota,S., Jamu,A., Ngorima,N., Chirairo,H., Chitsungo,S., Chimanzi,J., Maweni,C., Warara,R., Matongo,M., Mudzingwa,S., Jangano,M., Moyo,K., Vere,L., Mdege,N., Machingura,I., Katabira,E., Ronald,A., Kambungu,A., Lutwama,F., Mambule,I., Nanfuka,A., Walusimbi,J., Nabankema,E., Nalumenya,R., Namuli,T., Kulume,R., Namata,I., Nyachwo,L., Florence,A., Kusiima,A., Lubwama,E., Nairuba,R., Oketta,F., Buluma,E., Waita,R., Ojiambo,H., Sadik,F., Wanyama,J., Nabongo,P., Oyugi,J., Sematala,F., Muganzi,A., Twijukye,C., Byakwaga,H., Ochai,R., Muhweezi,D., Coutinho,A., Etukoit,B., Gilks,C., Boocock,K., Puddephatt,C., Grundy,C., Bohannon,J., Winogron,D., Gibb,D.M., Burke,A., Bray,D., Babiker,A., Walker,A.S., Wilkes,H., Rauchenberger,M., Sheehan,S., Spencer-Drake,C., Taylor,K., Spyer,M., Ferrier,A., Naidoo,B., Dunn,D., Goodall,R., Darbyshire,J.H., Peto,L., Nanfuka,R., Mufuka-Kapuya,C., Kaleebu,P., Pillay,D., Robertson,V., Yirrell,D., Tugume,S., Chirara,M., Katundu,P., Ndembi,N., Lyagoba,F., Dunn,D., Goodall,R., McCormick,A., Lara,A.M., Foster,S., Amurwon,J., Wakholi,B.N., Kigozi,J., Muchabaiwa,L., Muzambi,M., Weller,I., Babiker,A., Bahendeka,S., Bassett,M., Wapakhabulo,A.C., Darbyshire,J.H., Gazzard,B., Gilks,C., Grosskurth,H., Hakim,J., Latif,A., Mapuchere,C., Mugurungi,O., Mugyenyi,P., Burke,C., Jones,S., Newland,C., Pearce,G., Rahim,S., Rooney,J., Smith,M., and Snowden,W. (2010). Routine versus clinically driven laboratory monitoring of HIV antiretroviral therapy in Africa (DART): a randomised non-inferiority trial. Lancet *375*, 123-131.

Muller,B., Anders,M., Akiyama,H., Welsch,S., Glass,B., Nikovics,K., Clavel,F., Tervo,H.M., Keppler,O.T., and Krausslich,H.G. (2009). HIV-1 Gag processing intermediates trans-dominantly interfere with HIV-1 infectivity. J. Biol. Chem. *284*, 29692-29703.

Muriaux,D., Mirro,J., Harvin,D., and Rein,A. (2001). RNA is a structural element in retrovirus particles. Proc. Natl. Acad. Sci. U. S. A *98*, 5246-5251.

Myint,L., Matsuda,M., Matsuda,Z., Yokomaku,Y., Chiba,T., Okano,A., Yamada,K., and Sugiura,W. (2004). Gag non-cleavage site mutations contribute to full recovery of viral fitness in protease inhibitor-resistant human immunodeficiency virus type 1. Antimicrob. Agents Chemother. *48*, 444-452.

Naldini,L., Blomer,U., Gallay,P., Ory,D., Mulligan,R., Gage,F.H., Verma,I.M., and Trono,D. (1996). In vivo gene delivery and stable transduction of nondividing cells by a lentiviral vector. Science *272*, 263-267.

Neil,S.J.D, Zang,T., and Bieniasz,P.D. (2008). Tetherin inhibits retrovirus release and is antagonized by HIV-1 Vpu. Nature. *451*, 425-431.

Neil,S. and Bieniasz,P. (2009). Human immunodeficiency virus, restriction factors, and interferon. J. Interferon Cytokine Res. *29*, 569-580.

Nijhuis,M., Schuurman,R., de,J.D., Erickson,J., Gustchina,E., Albert,J., Schipper,P., Gulnik,S., and Boucher,C.A. (1999). Increased fitness of drug resistant HIV-1 protease as a result of acquisition of compensatory mutations during suboptimal therapy. AIDS *13*, 2349-2359.

Nijhuis,M., van Maarseveen,N.M., and Boucher,C.A. (2007a). HIV protease resistance and viral fitness. Curr. Opin. HIV. AIDS *2*, 108-115.

Nijhuis,M., van Maarseveen,N.M., Lastere,S., Schipper,P., Coakley,E., Glass,B., Rovenska,M., de,J.D., Chappey,C., Goedegebuure,I.W., Heilek-Snyder,G., Dulude,D., Cammack,N., Brakier-Gingras,L., Konvalinka,J., Parkin,N., Krausslich,H.G., Brun-Vezinet,F., and Boucher,C.A. (2007b). A novel substrate-based HIV-1 protease inhibitor drug resistance mechanism. PLoS. Med. *4*, e36.

Notermans,D.W., Goudsmit,J., Danner,S.A., de,W.F., Perelson,A.S., and Mittler,J. (1998). Rate of HIV-1 decline following antiretroviral therapy is related to viral load at baseline and drug regimen. AIDS *12*, 1483-1490.

Onafuwa-Nuga,A. and Telesnitsky,A. (2009). The remarkable frequency of human immunodeficiency virus type 1 genetic recombination. Microbiol. Mol. Biol Rev. *73*, 451-80, Table.

Ono,A. (2010). HIV-1 assembly at the plasma membrane. Vaccine *28 Suppl 2*, B55-B59.

Ono,A., Ablan,S.D., Lockett,S.J., Nagashima,K., and Freed,E.O. (2004). Phosphatidylinositol (4,5) bisphosphate regulates HIV-1 Gag targeting to the plasma membrane. Proc. Natl. Acad. Sci. U. S. A *101*, 14889-14894.

Ono,A. and Freed,E.O. (2004). Cell-type-dependent targeting of human immunodeficiency virus type 1 assembly to the plasma membrane and the multivesicular body. J. Virol. *78*, 1552-1563.

Ott,D.E. (2009). Purification of HIV-1 virions by subtilisin in digestion or CD45 immunoaffinity depletion for biochemical studies. Methods Mol Biol. 485, 15-25.

Palmer,S., Kearney,M., Maldarelli,F., Halvas,E.K., Bixby,C.J., Bazmi,H., Rock,D., Falloon,J., Davey,R.T., Jr., Dewar,R.L., Metcalf,J.A., Hammer,S., Mellors,J.W., and Coffin,J.M. (2005). Multiple, linked human immunodeficiency virus type 1 drug resistance mutations in treatment-experienced patients are missed by standard genotype analysis. J. Clin. Microbiol. *43*, 406-413.

Pao,D., Andrady,U., Clarke,J., Dean,G., Drake,S., Fisher,M., Green,T., Kumar,S., Murphy,M., Tang,A., Taylor,S., White,D., Underhill,G., Pillay,D., and Cane,P. (2004). Long-term persistence of primary genotypic resistance after HIV-1 seroconversion. J. Acquir. Immune. Defic. Syndr. *37*, 1570-1573.

Parienti,J.J., Barrail-Tran,A., Duval,X., Nembot,G., Descamps,D., Vigan,M., Vrijens,B., Panhard,X., Taburet,A.M., Mentre,F., and Goujard,C. (2013). Adherence profiles and therapeutic responses of treatment-naive HIV-infected patients starting boosted atazanavir-based therapy in the ANRS 134-COPHAR 3 trial. Antimicrob. Agents Chemother. *57*, 2265-2271.

Parkin,N.T., Hellmann,N.S., Whitcomb,J.M., Kiss,L., Chappey,C., and Petropoulos,C.J. (2004). Natural variation of drug susceptibility in wild-type human immunodeficiency virus type 1. Antimicrob. Agents Chemother. *48*, 437-443.

Parry,C.M., Kohli,A., Boinett,C.J., Towers,G.J., McCormick,A.L., and Pillay,D. (2009). Gag determinants of fitness and drug susceptibility in protease inhibitor-resistant human immunodeficiency virus type 1. J. Virol. *83*, 9094-9101.

Parry,C.M., Kolli,M., Myers,R.E., Cane,P.A., Schiffer,C., and Pillay,D. (2011). Three residues in HIV-1 matrix contribute to protease inhibitor susceptibility and replication capacity. Antimicrob. Agents Chemother. *55*, 1106-1113.

Partaledis,J.A., Yamaguchi,K., Tisdale,M., Blair,E.E., Falcione,C., Maschera,B., Myers,R.E., Pazhanisamy,S., Futer,O., Cullinan,A.B., and . (1995). In vitro selection and characterization of human immunodeficiency virus type 1 (HIV-1) isolates with reduced sensitivity to hydroxyethylamino sulfonamide inhibitors of HIV-1 aspartyl protease. J. Virol. *69*, 5228-5235.

Patick,A.K., Rose,R., Greytok,J., Bechtold,C.M., Hermsmeier,M.A., Chen,P.T., Barrish,J.C., Zahler,R., Colonno,R.J., and Lin,P.F. (1995). Characterization of a human immunodeficiency virus type 1 variant with reduced sensitivity to an aminodiol protease inhibitor. J. Virol. *69*, 2148-2152.

Paton,N., Kityo,C., Hoppe,A., Hakim,J., van Oosterhout,J., Siika,A., Mwaba,P., Kambugu,A.,Easterbrook,P., Boles,J., Walker,S., Mugyenyi,P., EARNEST Trial Group. (2013). A pragmatic randomised controlled strategy trial of three second-line treatment options for use in public health rollout programme settings: the Europe-Africa Research Network for Evaluation of Second-line Therapy (EARNEST) Trial. 7th IAS Conference. Abstract WELBB02.

Pellegrin,I., Breilh,D., Montestruc,F., Caumont,A., Garrigue,I., Morlat,P., Le,C.C., Saux,M.C., Fleury,H.J., and Pellegrin,J.L. (2002). Virologic response to nelfinavir-based regimens: pharmacokinetics and drug resistance mutations (VIRAPHAR study). AIDS *16*, 1331-1340.

Pellegrin,I., Breilh,D., Ragnaud,J.M., Boucher,S., Neau,D., Fleury,H., Schrive,M.H., Saux,M.C., Pellegrin,J.L., Lazaro,E., and Vray,M. (2006). Virological responses to atazanavir-ritonavir-based regimens: resistance-substitutions score and pharmacokinetic parameters (Reyaphar study). Antivir. Ther. *11*, 421-429.

Pellegrin,I., Wittkop,L., Joubert,L.M., Neau,D., Bollens,D., Bonarek,M., Girard,P.M., Fleury,H., Winters,B., Saux,M.C., Pellegrin,J.L., Thiebaut,R., and Breilh,D. (2008). Virological response to darunavir/ritonavir-based regimens in antiretroviral-experienced patients (PREDIZISTA study). Antivir. Ther. *13*, 271-279.

Perez Valero,I., Gonzalez-Baeza,A., and Montest Ramirez,M.L. (2014). Central nervous system penetration and effectiveness of darunavir/ritonavir monotherapy. AIDS Rev. *16*, 101-8

Perno,C.F., Cozzi-Lepri,A., Balotta,C., Forbici,F., Violin,M., Bertoli,A., Facchi,G., Pezzotti,P., Cadeo,G., Tositti,G., Pasquinucci,S., Pauluzzi,S., Scalzini,A., Salassa,B., Vincenti,A., Phillips,A.N., Dianzani,F., Appice,A., Angarano,G., Monno,L., Ippolito,G., Moroni,M., and d' Arminio,M.A. (2001). Secondary mutations in the protease region of human immunodeficiency virus and virologic failure in drug-naive patients treated with protease inhibitor-based therapy. J. Infect. Dis. *184*, 983-991.

Petropoulos,C.J., Parkin,N.T., Limoli,K.L., Lie,Y.S., Wrin,T., Huang,W., Tian,H., Smith,D., Winslow,G.A., Capon,D.J., and Whitcomb,J.M. (2000). A novel phenotypic drug susceptibility assay for human immunodeficiency virus type 1. Antimicrob. Agents Chemother. *44*, 920-928.

Pettit,S.C., Henderson,G.J., Schiffer,C.A., and Swanstrom,R. (2002). Replacement of the P1 amino acid of human immunodeficiency virus type 1 Gag processing sites can inhibit or enhance the rate of cleavage by the viral protease. J. Virol. *76*, 10226-10233.

Pettit,S.C., Lindquist,J.N., Kaplan,A.H., and Swanstrom,R. (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.

Pettit,S.C., Moody,M.D., Wehbie,R.S., Kaplan,A.H., Nantermet,P.V., Klein,C.A., and Swanstrom,R. (1994). The p2 domain of human immunodeficiency virus type 1 Gag regulates sequential proteolytic processing and is required to produce fully infectious virions. J. Virol. *68*, 8017-8027.

Pettit,S.C., Simsic,J., Loeb,D.D., Everitt,L., Hutchison,C.A., III, and Swanstrom,R. (1991). Analysis of retroviral protease cleavage sites reveals two types of cleavage sites and the structural requirements of the P1 amino acid. J. Biol Chem. *266*, 14539-14547.

Phillips,A.N., Leen,C., Wilson,A., Anderson,J., Dunn,D., Schwenk,A., Orkin,C., Hill,T., Fisher,M., Walsh,J., Pillay,D., Bansi,L., Gazzard,B., Easterbrook,P., Gilson,R., Johnson,M., and Sabin,C.A. (2007). Risk of extensive virological failure to the three original antiretroviral drug classes over long-term follow-up from the start of therapy in patients with HIV infection: an observational cohort study. Lancet *370*, 1923-1928.

Plantier,J.C., Leoz,M., Dickerson,J.E., De,O.F., Cordonnier,F., Lemee,V., Damond,F., Robertson,D.L., and Simon,F. (2009). A new human immunodeficiency virus derived from gorillas. Nat. Med *15*, 871-872.

Pornillos,O., Ganser-Pornillos,B.K., and Yeager,M. (2011). Atomic-level modelling of the HIV capsid. Nature *469*, 424-427.

Prabu-Jeyabalan,M., Nalivaika,E., and Schiffer,C.A. (2002). Substrate shape determines specificity of recognition for HIV-1 protease: analysis of crystal structures of six substrate complexes. Structure. *10*, 369-381.

Prabu-Jeyabalan,M., Nalivaika,E.A., King,N.M., and Schiffer,C.A. (2003). Viability of a drug-resistant human immunodeficiency virus type 1 protease variant: structural insights for better antiviral therapy. J. Virol. *77*, 1306-1315.

Prabu-Jeyabalan,M., Nalivaika,E.A., King,N.M., and Schiffer,C.A. (2004). Structural basis for coevolution of a human immunodeficiency virus type 1 nucleocapsid-p1 cleavage site with a V82A drug-resistant mutation in viral protease. J. Virol. *78*, 12446-12454.

Price,M.N., Dehal,P.S., and Arkin,A.P. (2010). FastTree 2 – Approximately Maximum-Likelihood trees for large alignments. PLoS One, *5*, e9490.

Pulido,F., Delgado,R., Perez-Valero,I., Gonzalez-Garcia,J., Miralles,P., Arranz,A., Hernando,A., and Arribas,J.R. (2008). Long-term (4 years) efficacy of lopinavir/ritonavir monotherapy for maintenance of HIV suppression. J. Antimicrob. Chemother. *61*, 1359-1361.

Qari,S.H., Respess,R., Weinstock,H., Beltrami,E.M., Hertogs,K., Larder,B.A., Petropoulos,C.J., Hellmann,N., and Heneine,W. (2002). Comparative analysis of two commercial phenotypic assays for drug susceptibility testing of human immunodeficiency virus type 1. J. Clin. Microbiol. *40*, 31-35.

Quinones-Mateu,M. and Arts,E.J. (2001). HIV-1 Fitness: Implications for Drug Resistance, Disease Progression, and Global Epidemic Evolution. Theoretical Biology and Biophysics Group, Los Alamos National Library), pp. 134-170.

Rabi,S.A., Laird,G.M., Durand,C.M., Laskey,S., Shan,L., Bailey,J.R., Chioma,S., Moore,R.D., and Siliciano,R. (2013). Multi-step inhibition explains HIV-1 protease inhibitor pharmacodynamics and resistance. J Clin Invest.*123,* 3848–3860.

Rasaiyaah,J., Tan,C.P., Fletcher,A.J., Price,A.J., Blondeau,C., Hilditch,L., Jacques,D.A., Selwood,D.L., James,L.C., Noursadeghi,M., and Towers,G.J. (2013). HIV-1 evades innate immune recognition through specific cofactor recruitment. *503*, 402-5.

Ray,N., Blackburn,L.A., and Doms,R.W. (2009). HR-2 mutations in human immunodeficiency virus type 1 gp41 restore fusion kinetics delayed by HR-1 mutations that cause clinical resistance to enfuvirtide. J. Virol. *83*, 2989-2995.

Reijers,M.H., Weverling,G.J., Jurriaans,S., Wit,F.W., Weigel,H.M., Ten Kate,R.W., Mulder,J.W., Frissen,P.H., van,L.R., Reiss,P., Schuitemaker,H., de,W.F., and Lange,J.M. (1998). Maintenance therapy after quadruple induction therapy in HIV-1 infected individuals: Amsterdam Duration of Antiretroviral Medication (ADAM) study. Lancet *352*, 185-190.

Rhee,S.Y., Gonzales,M.J., Kantor,R., Betts,B.J., Ravela,J., and Shafer,R.W. (2003). Human immunodeficiency virus reverse transcriptase and protease sequence database. Nucleic Acids Res. *31*, 298-303.

Rhee,S.Y., Taylor,J., Wadhera,G., Ben-Hur,A., Brutlag,D.L., and Shafer,R.W. (2006). Genotypic predictors of human immunodeficiency virus type 1 drug resistance. Proc. Natl. Acad. Sci. U. S. A *103*, 17355-17360.

Riddler,S.A., Haubrich,R., DiRienzo,A.G., Peeples,L., Powderly,W.G., Klingman,K.L., Garren,K.W., George,T., Rooney,J.F., Brizz,B., Lalloo,U.G., Murphy,R.L., Swindells,S., Havlir,D., and Mellors,J.W. (2008). Class-sparing regimens for initial treatment of HIV-1 infection. N. Engl. J. Med *358*, 2095-2106.

Rodriguez-French,A., Boghossian,J., Gray,G.E., Nadler,J.P., Quinones,A.R., Sepulveda,G.E., Millard,J.M., and Wannamaker,P.G. (2004). The NEAT study: a 48-week open-label study to compare the antiviral efficacy and safety of GW433908 versus nelfinavir in antiretroviral therapy-naive HIV-1-infected patients. J. Acquir. Immune. Defic. Syndr. *35*, 22-32.

Saag,M.S., Tebas,P., Sension,M., Conant,M., Myers,R., Chapman,S.K., Anderson,R., and Clendeninn,N. (2001). Randomized, double-blind comparison of two nelfinavir doses plus nucleosides in HIV-infected patients (Agouron study 511). AIDS *15*, 1971-1978.

Sadler,B.M., Gillotin,C., Lou,Y., and Stein,D.S. (2001). Pharmacokinetic and pharmacodynamic study of the human immunodeficiency virus protease inhibitor amprenavir after multiple oral dosing. Antimicrob. Agents Chemother. *45*, 30-37.

Santos,J.R., Llibre,J.M., Imaz,A., Domingo,P., Iribarren,J.A., Marino,A., Miralles,C., Galindo,M.J., Ornelas,A., Moreno,S., Schapiro,J.M., and Clotet,B. (2012). Mutations in the protease gene associated with virological failure to lopinavir/ritonavir-containing regimens. J. Antimicrob. Chemother. *67*, 1462-1469.

Sarafianos,S.G., Marchand,B., Das,K., Himmel,D.M., Parniak,M.A., Hughes,S.H., and Arnold,E. (2009). Structure and function of HIV-1 reverse transcriptase: molecular mechanisms of polymerization and inhibition. J. Mol. Biol. *385*, 693-713.

Sato,M., Motomura,T., Aramaki,H., Matsuda,T., Yamashita,M., Ito,Y., Kawakami,H., Matsuzaki,Y., Watanabe,W., Yamataka,K., Ikeda,S., Kodama,E., Matsuoka,M., and Shinkai,H. (2006). Novel HIV-1 integrase inhibitors derived from quinolone antibiotics. J. Med Chem. *49*, 1506-1508.

Sauter,D., Schindler,M., Specht,A., Landford,W.N., Munch,J., Kim,K.A., Votteler,J., Schubert,U., Bibollet-Ruche,F., Keele,B.F., Takehisa,J., Ogando,Y., Ochsenbauer,C., Kappes,J.C., Ayouba,A., Peeters,M., Learn,G.H., Shaw,G., Sharp,P.M., Bieniasz,P., Hahn,B.H., Hatziioannou,T., and Kirchhoff,F. (2009). Tetherin-driven adaptation of Vpu and Nef function and the evolution of pandemic and nonpandemic HIV-1 strains. Cell Host. Microbe *6*, 409-421.

Sauter,D., (2014). Counteraction of the multifunctional restriction factor tetherin. Front Microbiol. *5*, 163.

Schaller,T., Ocwieja,K.E., Rasaiyaah,J., Price,A.J., Brady,T.L., Roth,S.L., Hue,S., Fletcher,A.J., Lee,K., Kewalramani,V.N., Noursadeghi,M., Jenner,R.G., James,L.C., Bushman,F.D., and Towers,G.J. (2011). HIV-1 Capsid-Cyclophilin Interactions Determine Nuclear Import Pathway, Integration Targeting and Replication Efficiency. PLoS. Pathog. *7*, e1002439.

Schapiro,J.M., Winters,M.A., Stewart,F., Efron,B., Norris,J., Kozal,M.J., and Merigan,T.C. (1996). The effect of high-dose saquinavir on viral load and CD4+ T-cell counts in HIV-infected patients. Ann. Intern. Med *124*, 1039-1050.

Scheffler,K., Martin,D.P., and Seoighe,C. (2006). Robust inference of positive selection from recombining coding sequences. Bioinformatics. *22*, 2493-2499.

Scherrer,A.U., Ledergerber,B., von,W., V, Boni,J., Yerly,S., Klimkait,T., Cellerai,C., Furrer,H., Calmy,A., Cavassini,M., Elzi,L., Vernazza,P.L., Bernasconi,E., and Gunthard,H.F. (2012). Minor protease inhibitor mutations at baseline do not increase the risk for a virological failure in HIV-1 subtype B infected patients. PLoS. One. *7*, e37983.

Schinazi,R.F., Lloyd,R.M., Jr., Nguyen,M.H., Cannon,D.L., McMillan,A., Ilksoy,N., Chu,C.K., Liotta,D.C., Bazmi,H.Z., and Mellors,J.W. (1993). Characterization of human immunodeficiency viruses resistant to oxathiolane-cytosine nucleosides. Antimicrob. Agents Chemother. *37*, 875-881.

Schroder,A.R., Shinn,P., Chen,H., Berry,C., Ecker,J.R., and Bushman,F. (2002). HIV-1 integration in the human genome favors active genes and local hotspots. Cell *110*, 521-529.

Shafer,R.W. (2006). Rationale and uses of a public HIV drug-resistance database. J. Infect. Dis. *194 Suppl 1*, S51-S58.

Sharova,N., Wu,Y., Zhu,X., Stranska,R., Kaushik,R., Sharkey,M., and Stevenson,M. (2008). Primate lentiviral Vpx commandeers DDB1 to counteract a macrophage restriction. PLoS Pathog. *4,* e1000057.

Sheehy,A.M., Gaddis,N.C., and Malim,M.H. (2003). The antiretroviral enzyme APOBEC3G is degraded by the proteasome in response to HIV-1 Vif. Nat. Med *9*, 1404-1407.

Shibata,J., Sugiura,W., Ode,H., Iwatani,Y., Sato,H., Tsang,H., Matsuda,M., Hasegawa,N., Ren,F., and Tanaka,H. (2011). Within-host co-evolution of Gag P453L and protease D30N/N88D demonstrates virological advantage in a highly protease inhibitor-exposed HIV-1 case. Antiviral Res. *90*, 33-41.

Shimura,K. and Kodama,E.N. (2009). Elvitegravir: a new HIV integrase inhibitor. Antivir. Chem. Chemother. *20*, 79-85.

Shuter,J., Sarlo,J.A., Kanmaz,T.J., Rode,R.A., and Zingman,B.S. (2007). HIV-infected patients receiving lopinavir/ritonavir-based antiretroviral therapy achieve high rates of virologic suppression despite adherence rates less than 95%. J. Acquir. Immune. Defic. Syndr. *45*, 4-8.

Sigal,A., Kim,J.T., Balazs,A.B., Dekel,E., Mayo,A., Milo,R., and Baltimore,D. (2011). Cell-to-cell spread of HIV permits ongoing replication despite antiretroviral therapy. Nature *477*, 95-98.

Simm,M., Shahabuddin,M., Chao,W., Allan,J.S., and Volsky,D.J. (1995). Aberrant Gag protein composition of a human immunodeficiency virus type 1 vif mutant produced in primary lymphocytes. J. Virol. *69*, 4582-4586.

Sloan,R.D., and Wainberg,M.A. (2011). The role of unintegrated DNA in HIV Infections. Retrovirology. *8*, 52.

Smith,D.M., May,S.J., Tweeten,S., Drumright,L., Pacold,M.E., Kosakovsky Pond,S.L., Pesano,R.L., Lie,Y.S., Richman,D.D., Frost,S.D., Woelk,C.H., and Little,S.J. (2009). A public health model for the molecular surveillance of HIV transmission in San Diego, California. AIDS *23*, 225-232.

Smith,S.M. (2004). HIV CTL escape: at what cost? Retrovirology. *1*, 8.

Solas,C., Poizot-Martin,I., Drogoul,M.P., Ravaux,I., Dhiver,C., Lafeuillade,A., Allegre,T., Mokhtari,M., Moreau,J., Lepeu,G., Petit,N., Durand,A., and Lacarelle,B. (2004) Therapeutic drug monitoring of lopinavir/ritonavir given alone or with a non-nucleoside reverse transcriptase inhibitor. Br. J. Clin. Pharacol. *57,* 436-40.

Stein,D.S., Fish,D.G., Bilello,J.A., Preston,S.L., Martineau,G.L., and Drusano,G.L. (1996). A 24-week open-label phase I/II evaluation of the HIV protease inhibitor MK-639 (indinavir). AIDS *10*, 485-492.

Stoltzfus,C.M. (2009). Chapter 1. Regulation of HIV-1 alternative RNA splicing and its role in virus replication. Adv. Virus Res. *74*, 1-40.

Stremlau,M., Owens,C.M., Perron,M.J., Kiessling,M., Autissier,P., and Sodroski,J. (2004). The cytoplasmic body component TRI

M5alpha restricts HIV-1 infection in Old World monkeys. Nature *427*, 848-853.

Sundquist,W.I. and Krausslich,H.G. (2012). HIV-1 assembly, budding, and maturation. Cold Spring Harb. Perspect. Med *2*, a006924.

Swain,A.L., Miller,M.M., Green,J., Rich,D.H., Schneider,J., Kent,S.B., and Wlodawer,A. (1990). X-ray crystallographic structure of a complex between a synthetic protease of human immunodeficiency virus 1 and a substrate-based hydroxyethylamine inhibitor. Proc. Natl. Acad. Sci. U. S. A *87*, 8805-8809.

Swindells,S., DiRienzo,A.G., Wilkin,T., Fletcher,C.V., Margolis,D.M., Thal,G.D., Godfrey,C., Bastow,B., Ray,M.G., Wang,H., Coombs,R.W., McKinnon,J., and Mellors,J.W. (2006). Regimen simplification to atazanavir-ritonavir alone as maintenance antiretroviral therapy after sustained virologic suppression. JAMA *296*, 806-814.

Tamura,K., Peterson,D., Peterson,N., Stecher,G., Nei,M., and Kumar,S. (2011). MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. Mol. Biol Evol. *28*, 2731-2739.

Tantillo,C., Ding,J., Jacobo-Molina,A., Nanni,R.G., Boyer,P.L., Hughes,S.H., Pauwels,R., Andries,K., Janssen,P.A., and Arnold,E. (1994). Locations of anti-AIDS drug binding sites and resistance mutations in the three-dimensional structure of HIV-1 reverse transcriptase. Implications for mechanisms of drug inhibition and resistance. J. Mol. Biol *243*, 369-387.

Tilton,J.C. and Doms,R.W. (2010). Entry inhibitors in the treatment of HIV-1 infection. Antiviral Res. *85*, 91-100.

Tisdale,M., Myers,R.E., Maschera,B., Parry,N.R., Oliver,N.M., and Blair,E.D. (1995). Cross-resistance analysis of human immunodeficiency virus type 1 variants individually selected for resistance to five different protease inhibitors. Antimicrob. Agents Chemother. *39*, 1704-1710.

Titanji,B.K., Aasa-Chapman,M., Pillay,D., and Jolly,C. (2013). Protease inhibitors effectively block cell-to-cell spread of HIV-1 between T cells. Retrovirology. *10*, 161.

Towers,G.J. (2007). The control of viral infection by tripartite motif proteins and cyclophilin A. Retrovirology. *4*, 40.

Tsibris,A.M. and Hirsch,M.S. (2010). Antiretroviral therapy in the clinic. J. Virol. *84*, 5458-5464.

Turner, B.G. and Summers, M.F. (1999) Structural biology of HIV. J. Mol. Biol. *285,* 1-32.

Turner,S.R., Strohbach,J.W., Tommasi,R.A., Aristoff,P.A., Johnson,P.D., Skulnick,H.I., Dolak,L.A., Seest,E.P., Tomich,P.K., Bohanon,M.J., Horng,M.M., Lynn,J.C., Chong,K.T., Hinshaw,R.R., Watenpaugh,K.D., Janakiraman,M.N., and Thaisrivongs,S. (1998). Tipranavir (PNU-140690): a potent, orally bioavailable nonpeptidic HIV protease inhibitor of the 5,6-dihydro-4-hydroxy-2-pyrone sulfonamide class. J. Med Chem. *41*, 3467-3476.

Vallari,A., Bodelle,P., Ngansop,C., Makamche,F., Ndembi,N., Mbanya,D., Kaptue,L., Gurtler,L.G., McArthur,C.P., Devare,S.G., and Brennan,C.A. (2010). Four new HIV-1 group N isolates from Cameroon: Prevalence continues to be low. AIDS Res. Hum. Retroviruses *26*, 109-115.

Vallari,A., Holzmayer,V., Harris,B., Yamaguchi,J., Ngansop,C., Makamche,F., Mbanya,D., Kaptue,L., Ndembi,N., Gurtler,L., Devare,S., and Brennan,C.A. (2011). Confirmation of putative HIV-1 group P in Cameroon. J. Virol. *85*, 1403-1407.

van Heeswijk,R.P., Veldkamp,A., Mulder,J.W., Meenhorst,P.L., Lange,J.M., Beijnen,J.H., and Hoetelmans,R.M. (2001). Combination of protease inhibitors for the treatment of HIV-1-infected patients: a review of pharmacokinetics and clinical experience. Antivir. Ther. *6*, 201-229.

Van Heuverswyn,F., Li,Y., Neel,C., Bailes,E., Keele,B.F., Liu,W., Loul,S., Butel,C., Liegeois,F., Bienvenue,Y., Ngolle,E.M., Sharp,P.M., Shaw,G.M., Delaporte,E., Hahn,B.H., and Peeters,M. (2006). Human immunodeficiency viruses: SIV infection in wild gorillas. Nature *444*, 164.

Van Laethem,K., Schrooten,Y., Dedecker,S., Van,H.L., Deforche,K., Van,W.E., Van,R.M., and Vandamme,A.M. (2006). A genotypic assay for the amplification and sequencing of gag and protease from diverse human immunodeficiency virus type 1 group M subtypes. J. Virol. Methods *132*, 181-186.

van Maarseveen,N.M., Andersson,D., Lepsik,M., Fun,A., Schipper,P.J., de,J.D., Boucher,C.A., and Nijhuis,M. (2012). Modulation of HIV-1 Gag NC/p1 cleavage efficiency affects protease inhibitor resistance and viral replicative capacity. Retrovirology. *9*, 29.

van Maarseveen,N.M., de,J.D., Boucher,C.A., and Nijhuis,M. (2006). An increase in viral replicative capacity drives the evolution of protease inhibitor-resistant human immunodeficiency virus type 1 in the absence of drugs. J. Acquir. Immune. Defic. Syndr. *42*, 162-168.

van Opijnen,T., de,R.A., Boerlijst,M.C., and Berkhout,B. (2007). Adaptation of HIV-1 depends on the host-cell environment. PLoS. One. *2*, e271.

Velazquez-Campoy,A., Todd,M.J., Vega,S., and Freire,E. (2001). Catalytic efficiency and vitality of HIV-1 proteases from African viral subtypes. Proc. Natl. Acad. Sci. U. S. A *98*, 6062-6067.

Vercauteren,J., Wensing,A.M., van de Vijver,D.A., Albert,J., Balotta,C., Hamouda,O., Kucherer,C., Struck,D., Schmit,J.C., Asjo,B., Bruckova,M., Camacho,R.J., Clotet,B., Coughlan,S., Grossman,Z., Horban,A., Korn,K., Kostrikis,L., Nielsen,C., Paraskevis,D., Poljak,M., Puchhammer-Stockl,E., Riva,C., Ruiz,L., Salminen,M., Schuurman,R., Sonnerborg,A., Stanekova,D., Stanojevic,M., Vandamme,A.M., and Boucher,C.A. (2009). Transmission of drug-resistant HIV-1 is stabilizing in Europe. J. Infect. Dis. *200*, 1503-1508.

Vergne,L., Stuyver,L., Van,H.M., Butel,C., Delaporte,E., and Peeters,M. (2006). Natural polymorphism in protease and reverse transcriptase genes and in vitro antiretroviral drug susceptibilities of non-B HIV-1 strains from treatment-naive patients. J. Clin. Virol. *36*, 43-49.

Verheyen,J., Litau,E., Sing,T., Daumer,M., Balduin,M., Oette,M., Fatkenheuer,G., Rockstroh,J.K., Schuldenzucker,U., Hoffmann,D., Pfister,H., and Kaiser,R. (2006). Compensatory mutations at the HIV cleavage sites p7/p1 and p1/p6-gag in therapy-naive and therapy-experienced patients. Antivir. Ther. *11*, 879-887.

Verheyen,J., Verhofstede,C., Knops,E., Vandekerckhove,L., Fun,A., Brunen,D., Dauwe,K., Wensing,A.M., Pfister,H., Kaiser,R., and Nijhuis,M. (2010). High prevalence of bevirimat resistance mutations in protease inhibitor-resistant HIV isolates. AIDS *24*, 669-673.

Vermeiren,H., Van,C.E., Alen,P., Bacheler,L., Picchio,G., and Lecocq,P. (2007). Prediction of HIV-1 drug susceptibility phenotype from the viral genotype using linear regression modeling. J. Virol. Methods *145*, 47-55.

Vernazza,P., Daneel,S., Schiffer,V., Decosterd,L., Fierz,W., Klimkait,T., Hoffmann,M., and Hirschel,B. (2007). The role of compartment penetration in PI-monotherapy: the Atazanavir-Ritonavir Monomaintenance (ATARITMO) Trial. AIDS *21*, 1309-1315.

Vora,S., Marcelin,A.G., Gunthard,H.F., Flandre,P., Hirsch,H.H., Masquelier,B., Zinkernagel,A., Peytavin,G., Calvez,V., Perrin,L., and Yerly,S. (2006). Clinical validation of atazanavir/ritonavir genotypic resistance score in protease inhibitor-experienced patients. AIDS *20*, 35-40.

Wain,L.V., Bailes,E., Bibollet-Ruche,F., Decker,J.M., Kelle,B.F., Van Heuverswyn,F., Li,Y., Takehisa,J., Ngole,E.M., Shaw,G.M,. Peeters,M., Hahn,B.H., and Sharp,P.M. (2007). Adaptation of HIV-1 to its human host. Mol Biol Evol. *24*, 1853-60.

Wainberg,M.A. and Albert,J. (2010). Can the further clinical development of bevirimat be justified? AIDS *24*, 773-774.

Wainberg,M.A., Miller,M.D., Quan,Y., Salomon,H., Mulato,A.S., Lamy,P.D., Margot,N.A., Anton,K.E., and Cherrington,J.M. (1999). In vitro selection and characterization of HIV-1 with reduced susceptibility to PMPA. Antivir. Ther. *4*, 87-94.

Walmsley,S., Bernstein,B., King,M., Arribas,J., Beall,G., Ruane,P., Johnson,M., Johnson,D., Lalonde,R., Japour,A., Brun,S., and Sun,E. (2002). Lopinavir-ritonavir versus nelfinavir for the initial treatment of HIV infection. N. Engl. J. Med *346*, 2039-2046.

Walmsley,S.L., Becker,M.I., Zhang,M., Humar,A., and Harrigan,P.R. (2001). Predictors of virological response in HIV-infected patients to salvage antiretroviral therapy that includes nelfinavir. Antivir. Ther. *6*, 47-54.

Wei,X., Decker,J.M., Liu,H., Zhang,Z., Arani,R.B., Kilby,J.M., Saag,M.S., Wu,X., Shaw,G.M., and Kappes,J.C. (2002). Emergence of resistant human immunodeficiency virus type 1 in patients receiving fusion inhibitor (T-20) monotherapy. Antimicrob. Agents Chemother. *46*, 1896-1905.

Wensing,A.M., van Maarseveen,N.M., and Nijhuis,M. (2010). Fifteen years of HIV Protease Inhibitors: raising the barrier to resistance. Antiviral Res. *85*, 59-74.

Westby,M., Lewis,M., Whitcomb,J., Youle,M., Pozniak,A.L., James,I.T., Jenkins,T.M., Perros,M., and van der Ryst,E. (2006). Emergence of CXCR4-using human immunodeficiency virus type 1 (HIV-1) variants in a minority of HIV-1-infected patients following treatment with the CCR5 antagonist maraviroc is from a pretreatment CXCR4-using virus reservoir. J. Virol. *80*, 4909-4920.

Westby,M., Smith-Burchnell,C., Mori,J., Lewis,M., Mosley,M., Stockdale,M., Dorr,P., Ciaramella,G., and Perros,M. (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.

WHO. Progress Report 2011: Global HIV/AIDS response. 2011.
Ref Type: Report

WHO: Core epidemiological slides; HIV/AIDS estimates. 2013.

Wilen,C.B,, Tilton,J.C., and Doms,R.W. (2012). HIV: cell binding and entry. Cold Spring Harb Perspect Med. *2*, a006866.

Wilkin,T.J., McKinnon,J.E., DiRienzo,A.G., Mollan,K., Fletcher,C.V., Margolis,D.M., Bastow,B., Thal,G., Woodward,W., Godfrey,C., Wiegand,A., Maldarelli,F., Palmer,S., Coffin,J.M., Mellors,J.W., and Swindells,S. (2009). Regimen simplification to atazanavir-ritonavir alone as maintenance antiretroviral therapy: final 48-week clinical and virologic outcomes. J. Infect. Dis. *199*, 866-871.

Williams,I., Churchill,D., Anderson,J., Boffito,M., Bower,M., Cairns,G., Cwynarski,K., Edwards,S., Fidler,S., Fisher,M., Freedman,A., Geretti,A.M., Gilleece,Y., Horne,R., Johnson,M., Khoo,S., Leen,C., Marshall,N., Nelson,M., Orkin,C., Paton,N., Phillips,A., Post,F., Pozniak,A., Sabin,C., Trevelion,R., Ustianowski,A., Walsh,J., Waters,L., Wilkins,E., Winston,A., and Youle,M. (2012). British HIV Association guidelines for the treatment of HIV-1-positive adults with antiretroviral therapy 2012. HIV. Med *13 Suppl 2*, 1-85.

Winters,B., Van,C.E., Van der Borght,K., Lecocq,P., Villacian,J., and Bacheler,L. (2009). Clinical cut-offs for HIV-1 phenotypic resistance estimates: update based on recent pivotal clinical trial data and a revised approach to viral mixtures. J. Virol. Methods *162*, 101-108.

World Health Organisation. Consolidated guidelines on the use of antiretroviral drugs for treating and preventing HIV infection. 2013.
Ref Type: Report

Wright,E., Temperton,N.J., Marston,D.A., McElhinney,L.M., Fooks,A.R., and Weiss,R.A. (2008). Investigating antibody neutralization of lyssaviruses using lentiviral pseudotypes: a cross-species comparison. J. Gen. Virol. *89*, 2204-2213.

Wu,J., Matunis,M.J., Kraemer,D., Blobel,G., and Coutavas,E. (1995). Nup358, a cytoplasmically exposed nucleoporin with peptide repeats, Ran-GTP binding sites, zinc fingers, a cyclophilin A homologous domain, and a leucine-rich region. J. Biol Chem. *270*, 14209-14213.

Wyatt,R., Kwong,P.D., Desjardins,E., Sweet,R.W., Robinson,J., Hendrickson,W.A., and Sodroski,J.G. (1998). The antigenic structure of the HIV gp120 envelope glycoprotein. Nature *393*, 705-711.

Yamashita,M., and Emerman,M., (2006). Retroviral infection of non-dividing cells: old and new perspectives. *344*, 88-93.

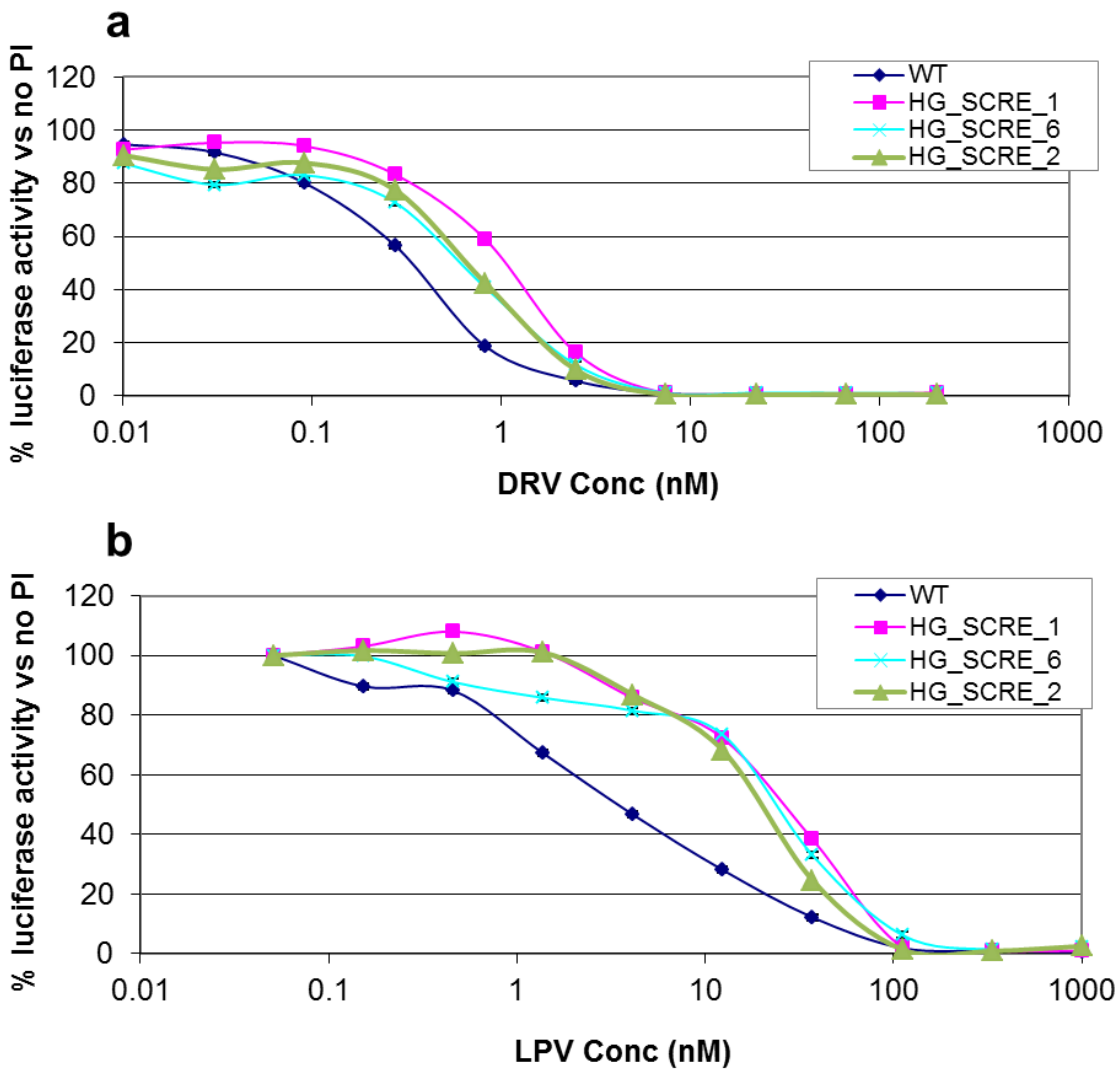
Yates,P.J., Hazen,R., St,C.M., Boone,L., Tisdale,M., and Elston,R.C. (2006). In vitro development of resistance to human immunodeficiency virus protease inhibitor GW640385. Antimicrob. Agents Chemother. *50*, 1092-1095.

Zennou,V., Mammano,F., Paulous,S., Mathez,D., and Clavel,F. (1998). Loss of viral fitness associated with multiple Gag and Gag-Pol processing defects in human immunodeficiency virus type 1 variants selected for resistance to protease inhibitors in vivo. J. Virol. *72*, 3300-3306.

Zhai,Q., Landesman,M.B., Chung,H.Y., Dierkers,A., Jeffries,C.M., Trewhella,J., Hill,C.P., and Sundquist,W.I. (2011). Activation of the retroviral budding factor ALIX. J. Virol. *85*, 9222-9226.

Zhang,Y.M., Imamichi,H., Imamichi,T., Lane,H.C., Falloon,J., Vasudevachari,M.B., and Salzman,N.P. (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.

Appendix:



Appendix Figure 1. Representative titration curves for phenotypic PI susceptibility assay. Pseudotyped viruses were produced by co-transfection of 293T cells with an HIV Gag-Pol expression vector p8.9NSX+, VSV-g expression vector pMDG and pCSFLW. 18 hours post-transfection, cells were incubated with serial dilutions of PI and the infectivity of pseudovirions produced in the presence of each PI concentration was tested. Typical curves for the PIs DRV (a) and LPV (b) are shown, with the reference strain p8.9NSX+ shown in dark blue and the other three viruses generated from patient-derived Gag-protease (green, light blue and pink).

```
patient HG          MGARASILSG GKLDKWEKIR LRPGGKKQYK LKHLVWASRE LERFAVNPGL
patient SP          ......V..P .E..Q..... .......K.R .......... ..........
patient FRD         ......V... ....A..... .......K.R .......... .....L....
patient MD          ......V... ....S..... .......K.R .......... .....L....
patient BOY         ........T. ....Q..... .......... .......... .....I....
patient MF          ......V... ....S..... .......K.. .......... .....L....

patient HG          LESSEGCRQI LEHLQPSLKA GSEELRSLFN TVATLYCVHQ NIEVRDTKEA
patient SP          ..T.I..... ..Q.....QT ........Y. ....I..... K...K.....
patient FRD         ..TT...Q.L M.Q..SA..T ........Y. .I...W.... R..IK.....
patient MD          ..TA...Q.. M.Q..SA.GT .....K.... .....W...R R.DIK.....
patient BOY         ..TA...Q.. ..Q..ST..T .....K.... .I...W.... K.DI......
patient MF          ..NA...Q.V M.QF.ST..T .....K..Y. .......... R...K.....

patient HG          LDKIEEEQNK SRKKAQQAAA AADTGNRSQV SQNYPIVQNH QGQMVHQALS
patient SP          .......... .K.......? -?....SN.. .........L ........I.
patient FRD         V..V..I.D. ..Q.T..... .-TAATG.-S .........A ....T..SM.
patient MD          ...V..V... .KQ..... ?---?TG.-S ...F.....A ....T..SM.
patient BOY         ...V..A.K. .KQ.T..... ?--?ATG.-. .........A .......PM.
patient MF          ......I.K. ..QQT..... .-TGSSN... ......M..A ....TY....

patient HG          PRTLNAWVKV VEEKAFSPEV IPMFSALSEG ATPQDLNTML NTVGGHQAAM
patient SP          .......... I......... .......... .......... ..........
patient FRD         .......... I......... .......... .......M.. .I........
patient MD          .......... I......... .......... .......M.. .I........
patient BOY         .......... I......... ....T..... .......M.. .I........
patient MF          .......... .......... .......A.. .P.H...... ..........

patient HG          QILKEAINEE AAEWDRLHPV QAGPVAPGQL REPRGSDIAG TTSTLQEQIG
patient SP          .M...T.... .......... H...I....M .......... ..........
patient FRD         .M..DT.... ......V... H...IP...M .......... ..........
patient MD          .M..DT.... ......V... H...IP...M .......... ...N......
patient BOY         .M..D..... ......T... H...IP...M .......... ..........
patient MF          .M..DT.... ......Q..Q ....LP.... .......... .........R

patient HG          WMTHNPPIPV GEIYKRWIII GLNKIVRMYS PTSILDIKQG PKEPFRDYVD
patient SP          ...N...... .........L .......... .V.....R.. ..........
patient FRD         ...S...... .........VL .......... .......... ..........
patient MD          ...S...... .........VL .......... .......... ..........
patient BOY         ...S...V.. .........VL .......... .V........ ..........
patient MF          ...ST..V.. .........L .......... .V........ ..........

patient HG          RFYKTLRAEQ ATPEVKNWMT ETLLVQNANP DCKTILKALG PGASLEEMMT
patient SP          .......... .SQD...... .......... .......... .A.T......
patient FRD         ..F....... ..Q....... .......... ...A..R... ...T......
patient MD          ..F....... ..Q....... .......... ...S..R... ...T......
patient BOY         ..F....... ..Q....... ....I..... ...S..R... ...T......
patient MF          ..F.I..... ..Q....... .......... .......... .A.T......

patient HG          ACQGVGGPAH KARVLAEAMS QLTNS?*R?? DAERQF*EPK KDG*VLQLWQ
patient SP          ........S. .......... .A.S.?YH?? ......Q.S. ..S..F....
patient FRD         ........G. .......... .V?-TDQC?? .......R?? ENNK.F....
patient MD          ........G. .......... .V?-TVQH?? .......G?? ETNKMF....
patient BOY         S.......S. .......... .A?ATVQHSN .....L.G?? ENNK.F....
patient MF          ........G. .......... .T.GMAAA?N .......G.. .KY..F....
```

```
patient HG          RRAHSQKL*G T*EKGLLEMW KGGTPNERLH *ETG*FFRED LAFPQGEARE
patient SP          ..P.....Q. S......... .RRA.....Y ....Q..... ..........
patient FRD         ..TS....Q. SQ.E...... E.R......Y .K.......N ...Q.....K
patient MD          ..TP.K..Q. P......... E.R......Y .........N ...Q.....K
patient BOY         ..TP..E.Q. .......... ..R.S..... .........N ...Q.....K
patient MF          G.TS....Q. P.K.....V. ....S..... GK.......N ...Q.R....

patient HG          FSSEQTRANS PTRANSPTSR ELQVWGRDNN SPSEAGADRS ?--?TVS??S
patient SP          .......... ..?-----?. .......... .........Q ?--?...SF.
patient FRD         ......GT?- ----?..... ..WDG....- PLP..RTKGQ ?--?.I.SFN
patient MD          .....IGT?- ----?..... ..WD??...- LL...RTEGQ GGQG.I.SFN
patient BOY         .......TKT R.?-?..... DPGDG...S- LL....TGGQ ?--?.I.SL.
patient MF          .....A..?- ----?..SCG DPR.RR..S- LLP.T..EGE ?--?VI.??G

patient HG          FPQITLWQRP LVTIKIGGQL KEALLDTGAD DTVLEDMNLP GRWKPKMIGG
patient SP          .......... .......... .......... .....E.S.Q ..........
patient FRD         .......... ...VR..... I......... .....EI..L .K........
patient MD          .......... ...V..... I......... .....EI... .K........
patient BOY         .......... ...VRV.... I......... .....EI... ..........
patient MF          .......... ..KVQV..KI I......... .....EID.. .K........

patient HG          IGGFIKVRQY DQIPIEICGH KAVGTVLIGP TPVNIIGRNL LTQLGCTLNF
patient SP          .......... ..V....... ..I....... .......... ..........
patient FRD         .......... .........K ..I....V.. .........M ...I......
patient MD          .......... ...LL....K R.I....V.. .........M ...I......
patient BOY         .......... .........EK R.I....V.. .........M ...I......
patient MF          .......... ...MM..S.K ..I....V.. .........M ...I......
```

Appendix Figure 2. Alignment of Gag-protease consensus viral variants subjected to phenotyping from patients infected with subtype B viruses experiencing therapy failure (HG and SP) and achieving virologic suppression (FRD, MD, BOY, MF). In total, 78 variant amino acids are present in Gag and 16 in protease. Of these, three amino acid residues correlated with treatment outcome – Gag 34, 93 and 490 (HXB2 position 473).

Appendix Table S1. Amino acid positions within Gag previously associated with PI resistance or exposure that demonstrate intra-patient variability within patient #1403 (KON).

	Viral variants from patient #1403 (KON)																			
Amino Acid Change	**Screening**										**Failure**									
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
Y79F	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	F	Y	Y	F	Y
M200I	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	V	M
A374P/S	A	A	A	A	T	A	A	A	T	A	T	T	T	T	T	T	T	T	T	T
I376V	V	V	V	V	V	V	V	V	V	V	I	I	V	V	V	I	V	V	V	V
Q474L	P	P	P	P	Q	P	P	P	Q	P	P	P	P	P	P	P	P	P	P	P

Appendix Table S2. Amino acid positions within Gag previously associated with PI resistance or exposure that demonstrate intra-patient variability within patient #508 (SP).

	Viral variants from patient #508 (SP)																			
Reported Mutation	**Screening**										**Failure**									
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
V128A	V	V	V	V	V	V	A	V	V	V	V	V	V	V	V	V	V	V	V	V
V370A/M	A	V	A	A	A	A	A	V	A	A	A	A	A	A	A	A	A	A	A	A
T375N/S	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	A	T	T
R380K	R	R	R	R	R	R	R	R	R	R	R	K	R	R	R	R	R	R	R	R
S451N	S	S	S	S	S	S	S	S	S	S	S	S	N	S	S	S	S	S	S	S

Appendix Table S3. Amino acid positions within Gag previously associated with PI resistance or exposure that demonstrate intra-patient variability within patient #1404 (DIO).

	Viral variants from patient #1404 (DIO)																			
Reported Mutation	**Screening**										**Failure**									
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
Y79F	Y	Y	Y	F	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
E428G	E	E	E	E	K	E	E	E	E	E	E	E	E	E	E	E	E	E	E	E

Appendix Table S4. Amino acid positions within Gag previously associated with PI resistance or exposure that demonstrate intra-patient variability within patient #4201 (SO)

	Viral variants from patient #4201 (SO)																			
Reported Mutation	**Screening**										**Failure**									
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
G62R	K	N	N	N	N	N	N	N	N	K	K	K	K	K	K	K	K	K	K	K
L75R	L	I	I	L	L	I	L	I	I	L	L	L	L	L	L	L	L	L	L	L
Y79F	Y	F	F	F	F	F	F	F	F	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
A374P/S	T	A	T	A	T	A	T	T	T	T	T	T	T	T	T	T	T	T	T	T
G381S	S	G	G	G	S	G	G	G	G	S	G	G	G	G	G	G	G	G	G	G

Appendix Table S5. Amino acid positions within Gag previously associated with PI resistance or exposure that demonstrate intra-patient variability within patient (HG).

Reported Mutation	**Viral variants from patient #3204 (HG)**																			
	Screening										**Failure**									
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
E12K	K	Q	K	K	K	Q	Q	K	K	K	K	Q	K	K	K	K	K	K	K	K
R76K	R	R	R	R	K	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R
T81A	A	T	T	T	T	T	T	T	T	T	A	A	A	A	A	A	A	A	A	A
H219Q	Q	Q	Q	Q	Q	Q	Q	Q	Q	Q	Q	Q	Q	Q	Q	Q	Q	Q	Q	K
S373P	P	P	S	S	P	S	P	S	P	S	S	P	S	S	S	S	S	S	S	S
I376V	I	I	V	V	I	I	I	V	I	V	I	I	I	I	I	I	I	I	I	I
I389T	M	M	T	T	M	T	M	T	M	T	T	M	T	T	T	T	T	T	T	T
S451N	S	N	S	S	N	S	N	S	S	N	S	S	S	S	S	S	S	S	S	S