

Phenotypic effects and predictions of HIV-1 subtype C reverse transcriptase C-terminal domain mutations on reverse transcriptase inhibitors.

Nompumelelo Prudence Mkhwanazi


2018

A dissertation submitted in fulfillment of the requirement for the degree of Doctor of Philosophy in Virology in the School of laboratory medicine and medical science, College of Health Science, University of KwaZulu-Natal for PhD by research thesis

Declaration

I, Ms Nompumelelo Prudence Mkhwanazi, declare as follows:

1. That the work described in this thesis has not been submitted to UKZN or other tertiary institution for purposes of obtaining an academic qualification, whether by myself or any other party.
2. That my contribution to the project was as follows: Writing a reseach proposal, obtaining ethics approval, obtaining part of funding, conducting laboratory experiments and data analysis, writing up the findings of the project.
3. That the contributions of others to the project were as follows.
 - i. Sinaye Ngcapu provided the RNase H mutations from his Masters Thesis

4. Signed  Date 01 March 2019

Plagiarism Declaration

I, Nompumelelo Prudence Mkhwanazi declare that:

- i. The research reported in this dissertation, except where otherwise indicated, is my original work.
- ii. This dissertation has not been submitted for any degree or examination at any other university.
- iii. This dissertation does not contain other persons' data, pictures, graphs or other information, unless specifically acknowledged as being sourced from other persons.
- iv. This dissertation does not contain other persons' writing, unless specifically acknowledged as being sourced from other researchers. Where other written source have been quoted, then:
 - a. Their words have been re-written but the general information attributed to them has been referenced.
 - b. Where their exact words have been used, their writing has been placed inside quotation marks, and referenced.
- v. Where I have reproduced a publication of which I am an author , co-author or editor, I have indicated in detail which part of the publication was actually written by myself alone and have fully referenced such publications.
- vi. This dissertation does not contain texts, graphics or tables copied and pasted from internet, unless specifically acknowledged, and the source being in the dissertation and in the Reference sections.

Signed: _____ Date: _____

Ethical approval

This study obtained full ethical approval, from the Biomedical Research Ethics Committee of the Nelson R. Mandela School of Medicine, University of KwaZulu-Natal (ref: BE017/14). This study was part of the Protease Cleavage study conducted in Durban, South Africa, KwaZulu-Natal which was approved by BREC (BF068/08) and the informed patient consent was obtained.

Presentations

1. Part of this work was presented at the Keystone symposium (HIV vaccine X5) held in Fairmont Banff Springs, Banff, Alberta, Canada from 22 – 27 March 2015. A poster was presented titled: “*Development of phenotypic assay to evaluate the effect of RNase H mutations of HIV-1 subtype C on RTI.*”
2. Part of this work was also presented at the XXV International HIV Drug Resistance Workshop held in Boston, USA, from the 20 – 21 February, 2016. A poster was presented titled: “*Minimal Phenotypic Drug Susceptibility Effect of the E529D RNase H Mutation of HIV-1 Subtypes on the RTI*”.

Acknowledgements

I owe my deepest gratitude to my supervisor Dr Michelle Gordon for her understanding and great efforts to explain things. I am also grateful for the encouragement, sound advice, great ideas and good company which she has provided me with during my thesis-writing period.

I am deeply grateful to Professor Thumbi Ndung'u for allowing me to perform my laboratory experiments in the HIV Pathogenesis Programme Laboratory. His wide knowledge and important support throughout this work have been of great value to me.

I would like to thank a Mrs Keshni Hiramani, Ms Vanessa Naidoo, Avashna Singh and Nothemba Nontala for their research assistance in the HPP laboratory.

I am thankful to my friends Dr Bongwiwe Ndlovu and Dr Paradise Madlala for giving me emotional support during difficult time during thesis writing. I would also like to thank my colleagues in the laboratory Lungile Mkhize, Nothemba Nontala, Zenele Mncube for their loving support and for providing inspiring and fun working environment to learn and grow. A special thanks to Dr Catherine and Christina for their academic advice, friendship and support which they provided.

My special thanks are due to my two daughters (Noxolo and Nongcebo) for their understanding and loving support. A special thanks to my mother, father, brothers and sisters for their continuous support and for always believing in me.

Above all, I would like to thank my God Almighty for giving me strength to persevere and wisdom in making this project possible. I strongly believe that with God all things are possible.

Table of Contents

Declaration	ii
Plagiarism Declaration	iii
Ethical approval	iv
Presentations	v
Acknowledgements	vi
Abbreviations and Acronyms	xii
Abstract	xv
Chapter 1	1
Introduction and Literature review	1
1.1. Introduction	1
1.2. Literature review	3
1.2.1. The Human Immunodeficiency Virus-1	3
1.2.2. HIV-1 replication.....	8
1.2.3. HIV-1 Reverse Transcription.....	10
1.3. HIV-1 antiretroviral drugs/inhibitors	16
1.3.1. HIV-1 antiretroviral regimens used in South Africa.....	16
1.3.2. Mechanism of nucleoside and non-nucleoside reverse transcriptase inhibitors.....	17
1.3.3. HIV-1 drug resistance mutations	18
1.4. HIV-1 RT Connection and RNase H domain mutations	21
1.4.1. Effect of connection and RNase H mutations on drug susceptibility	22
1.4.2. Connection domain mutations and TAMs.....	22
1.4.3. HIV-1 Connection domain mutations and RNase H.....	24
1.4.4. HIV-1 RNase H mutations and TAMs	23
1.4.5. Connection and RNase H mutations in non-B subtypes	25
1.5. Principles of laboratory techniques	27
1.5.1. Site directed mutagenesis	27
1.5.2. Phenotypic drug susceptibility assay	28
1.5.3. Bayesian Network analysis	31
1.6. Rationale of the study	33
1.6.1. Hypothesis	34

1.6.2. Specific Aims	34
Chapter 2	35
Minimal phenotypic effects of RNase H domain mutations in HIV-1 subtype C on reverse transcriptase inhibitors	
2.1. Introduction	35
2.2. Material and methods	37
2.2.1. Antiretroviral drugs and cell lines.....	37
2.2.2. Construction of the RTsubCWT recombinant plasmid	38
2.2.3. Construction of viral mutants (RTsubCmut) using site-directed mutagenesis	39
2.2.4. Co-transfection.....	41
2.2.5. Replication capacity	42
2.2.6. Single cycle drug susceptibility assay.....	43
2.3. Results	44
2.3.1. Effects on viral replication capacity	44
2.3.1.1. T470S and T470P cause similar decrease in replication capacity	44
2.3.1.2. E529D impairs replication capacity	45
2.3.1.3. L517I+D67N double mutant rescues the impaired viral replication capacity caused by D67N.....	46
2.3.2. Effect on drug susceptibility	48
2.3.2.1. Minimal reduction in drug susceptibility induced by T470S and T470P viral mutants	48
2.3.2.2. D67N, L517I and the L517I+D67N double mutation decreases drug susceptibility to RTIs	49
2.3.2.3 E529D has a minimal effect on drug susceptibility.....	50
2.4. Discussion	52
Chapter 3	56
Identification of connection domains mutations of the HIV-1 C and their association with thymidine analog mutations using a Bayesian Network	56
3.1. Introduction	56
3.2. Material and Methods	58
3.2.1. Identification of HIV-1 connection domain mutations.....	58

3.2.2. Frequency of the HIV-1 RT connection domain mutations.....	58
3.2.3. Bayesian Network analysis	59
3.3. Results	59
3.3.1. Characterization and identification of drug resistance-associated mutations in HIV-1 RT connection domain and their correlation with thymidine analogue mutations.	59
3.4. Discussion	62
Chapter 4	66
The effects of HIV-1 C RNase H domain mutations on reverse transcriptase structure....	66
4.1. Introduction	66
4.2. Material and Methods	67
4.2.1. 3D Homology Modelling using Swissmodel.....	67
4.2.2. Visualization using Chimera (v1.8.1)	67
4.3. Results	67
4.3.1. Structural changes caused by the T470S and T470P mutations.....	69
4.3.2. Structural changes caused by the L517I and D67N double mutation	72
4.3.3. Structural changes caused by the E529D mutation.....	73
4.4. Discussion	76
Chapter 5	84
Discussion and conclusion	84
Chapter 6	85
References	85

Table of figures

Figure 1.1. Schematic presentation of HIV-1 virion indicating the approximate location of Gag proteins, the Env glycoproteins, and the Pol-encoded enzymes IN, RT, and PR (Freed, 1998).....	4
Figure 1.2. The organization of HIV-1 viral genome.....	8
Figure 1.3. Steps in the HIV-1 replication cycle..	10
Figure 1.4. HIV-1 Reverse transcription steps.	14
Figure 1.5 . RNase H activity with two modes.....	15
Figure 1.6. Mechanism of NNRTI drug resistance.....	21
Figure 2.1. Replication capacity measured in RLUs of the T470S and T470P mutants generated.....	45
Figure 2.2. Replication capacity of HIV-1 viral mutants generated by site directed mutagenesis.	46
Figure 2.3. Replication capacity (RLU) of the L517I and D67N mutants alone and in combination.	47.
Figure 2.4. Phenotypic drug susceptibility profile from T470S and T470P subtype C viral mutants to the NRTIs.....	48
Figure 2.5. Phenotypic drug susceptibility profile from D67N, L517I and L517I+D67N subtype C viral mutants to the RTIs.	49
Figure 2.6. Phenotypic drug susceptibility profile from E529D subtype B and subtype C viral mutants.	51
Figure 3.1. Comparison of HIV_1 RT protein sequences of subtype C from diferent regions.	61
Figure 3.2. Annotated Bayesian Network (BN) showing dependencies in emergence of resistance mutations, based on 500 bootstrap showing graphically the learned associations between nodes denoting amino acids of treatment experienced.An arrow represents a direct dependence.....	62
Figure 4.1. The structural model of HIV-1 RT (3lam.1.B.pdb) showing the two subunits (p66 and p51).	68
Figure 4.2. The superimposed structure of RNase H domain showing mutations T470S and T470P. ..	69
Figure 4.3. The superimposed structures of RNase H domain showing T470S and T470P mutations and RNase H active sites amino acids.....	70
Figure 4.4. The superimposed structure of of HIV-1 RNase H domain showing the L517I mutation. .	72
Figure 4.5. The superimposed structure of RNase H domain showing E529D mutation.	74

List of Tables

Table 1.1. Dosage of the HIV-1 antiretroviral regimens available in Southern Africa (Meintjies et al 2017).....	16
Table 1.2. CN and RNase domain mutations from different HIV-1 subtypes.....	25
Table 2.1. Primer sequences for the SDM.....	42
Table 3.1. Frequency (%) of CN mutations domain from HIV-1 subtype C infected patients (250) receiving antiretroviral therapy.....	60
Table 4.1. Distances between the RNase H mutations (T470S/P) and the nearby amino acids residues.....	70
Table 4.2. Distances between the RNase H active site amino acid residues and the RNase H domain mutations (T470S/P).....	71
Table 4.3. The distances between the TAM (D67N) and RNase H mutation (L517I).....	72

Abbreviations and Acronyms

ABC	-	Abacavir
AIDS	-	Acquired Immunodeficiency Syndrome
ARV	-	Antiretroviral
ATV	-	Atazanavir
AZT	-	Zidovudine
BN	-	Bayesian Network
CN	-	Connection domain
D4T	-	Stavudine
DEPC	-	Diethylpyrocarbonate
DDI	-	Didanosine
DLV	-	Delavirdine
DNA	-	Deoxyribonucleic acid
dNTP	-	Deoxynucleotide Triphosphate
DRV	-	Darunavir
DTT	-	Dithiothreitol

EFV	-	Efavirenz
ETR	-	Etravirine
FPV	-	Fosamprenavir
FTC	-	Emtricitabine
Gp41	-	Glycoprotein 41
Gp120	-	Glycoprotein 120
HAART	-	Highly Active Antiretroviral Therapy
HIV-1	-	Human Immunodeficiency Virus-1
IDV	-	Indinavir
IPTG	-	isopropyl β -D-1 -thiogalactopyranoside
LPV	-	Lopinavir
mRNA	-	messenger Ribonucleic Acid
NFV	-	Nelfinavir
NNRTI	-	Non-Nucleoside Reverse Transcriptase Inhibitor
NRTI	-	Nucleoside Reverse Transcriptase Inhibitor
NVP	-	Nevirapine
PCR	-	Polymerase Chain Reaction

PI	-	Protease Inhibitor
PR	-	Protease
RC	-	Replication capacity
RNA	-	Ribonucleic Acid
RNase H	-	Ribonuclease H
RT	-	Reverse Transcriptase
RTIs	-	Reverse Transcriptase Inhibitors
RTV	-	Ritonavir
SQV	-	Saquinavir
TAMs	-	Thymidine Analog Mutations
TDF	-	Tenofovir
TPV	-	Tipranavir
3TC	-	Lamivudine

Abstract

Background

Antiretroviral drug therapy has been shown to reduce the death of HIV-1 infected individuals. However, the emergence of HIV-1 drug resistance has hindered the success of HIV-1 treatment. Genotyping tests mainly concentrate on the polymerase domain of HIV-1 RT leaving the rest understudied. Recently, data from HIV-1 C suggested that drug resistance could be caused by mutations in the connection and RNase H domains either alone or in combination with mutations in the polymerase region. Here, the phenotypic effects of the specific RNase H domain mutations in HIV-1 subtype C on RTIs were investigated. The predictions of HIV-1 subtype C reverse transcriptase C-terminal domain mutations on reverse transcriptase inhibitors were also investigated.

Material and methods

Viral RNA was extracted from 500µl of plasma using the Viral RNA extraction kit (Qiagen,Germany) according to the manufacturer's instructions, and stored at -80°C until utilisation. The RNA was amplified using the Superscript III One-step RT-PCR system with Platinum Taq DNA polymerase (Invitrogen, Life Technologies Corporation, Carlsbad, CA, California). The HIV-1 RT amplicon was cloned into a TOPO vector using the TOPO TA cloning kit (Invitrogen, Life Technologies). Viral mutants were constructed using site-directed mutagenesis, introducing D67N in the polymerase domain and E529D, L517I, T470S, and T470P mutations in the RNase H domain. Viral replication capacity and drug susceptibilities were determined using the single cycle luciferase assay in TZM-bl cells.

To identify the HIV-1 connection domain (CN) mutations associated with drug resistance, HIV-1 subtype C sequences were downloaded from the Los Alamos and Stanford HIV- drug resistance

databases from drug naïve and treated-experienced patients. The presence of connection domain (CN) mutations were identified using REGA HIV-1 subtyping tools (Universiteit Leuven, Belgium). Bayesian Network (BN) analysis (B-course) was used to determine the association of connection domain (CN) mutations (condon 320-440) with other TAMs. The effect of RNase H domain mutations on the structure of reverse transcriptase was determined using Swiss Model and viewed in Chimera.

Results

The replication capacities of T470S, T470P, L517I, E529D RNase H domain mutations were lower than the HIV-1 subtype C wild type in the absence of drugs. The D67N mutation alone had a lower replication capacity compared with the wild-type. Combination of L517I and D67N showed a further decrease in replication capacity compared to the wild type in the absence of drugs. E529D mutation replication capacity was assessed in TZM-*bl* cell line in both HIV-1 subtype B and C. Although not statistically significant, both competent subtype B and C E529D mutant had a decreased growth and infectivity rate compared to their respective wildtypes. The RNase H domain mutation T470S showed a moderate level of resistance to NVP (10.2X), ETR (8.75X), d4T (5X) and no resistance to AZT and EFV. Interestingly, T470P showed a moderate level of resistance to NVP (6X) and no resistance to ETR and d4T, as well as EFV. It did however show a 5-fold change to AZT when compared to the wild-type virus. As expected, the thymidine analog mutation, D67N, showed a high level of resistance to AZT (103.3X), moderate level of resistance to d4T (6.2X) and low level of resistance to ETR (3.2X) and no resistance to NVP and EFV. The RNase H domain L517I mutation showed moderate level of resistance to AZT (5.2X), d4T (6.0X) and NVP (10.79X), and low level of resistance to ETR (3.50X). L517I mutation caused hypersusceptibility to EFV. The combination of RNase H and TAM (L517I+D67N) showed high

level of resistance to AZT (157.3X), moderate level of resistance to NVP (11.3X), low level of resistance to d4T (3.9X) and no resistance to ETR and EFV. Subtype C E529D mutation conferred up to 2-fold stavudine (d4T), 3-fold zidovudine (AZT) and nevirapine (NVP) resistance, respectively. These findings demonstrate that RNase H mutation E529D can confer mild resistance to nucleotide (AZT and d4T) and non-nucleotide (NVP) reverse transcriptase inhibitors.

New connection domain mutations identified were: D324G/N/P, T338S, I341F/L/V, M357R, E370D, M377T/L, A376S, I434M/L, A437V/I. E370D and A437V/I were directly associated with treatment in the BN, while N348I was only indirectly associated with treatment.

Discussion and Conclusions

Overall, the RNase H domain mutations impaired replication capacity in the absence of drugs, suggesting that they are acquired at a fitness cost (as with most drug resistance mutations). While T470S decreased drug susceptibility to ETR, it was shown to be hypersusceptible to EFV. Interestingly, T470S is very common in subtype C RTI treated patients and could indicate that a switch to the newer NNRTI might not be as beneficial as expected. The phenotypic data also suggests that the resistance pathways for T470S and T470P could be different; however further studies are required to investigate their mechanism of resistance.

The L517I mutation alone only minimally decreased drug susceptibility to both NRTI and NNRTIs. However it further decreased drug susceptibility to the NRTIs when in combination with D67N, compared to D67N alone. D67N is known to affect the NRTIs, and the combined effect of D67N and L517I on NVP, a NNRTI, was surprising. The decreased replication capacity in the L517I indicated that it has additive effect in fitness loss of the viral. Further site-directed mutagenesis studies are needed to understand the effect of these RNase H mutations alone and in

combination with other polymerase domain mutations or with connection domain mutations on nucleoside reverse transcriptase inhibitors.

Structural analysis showed that the T470S/P mutations cause an inward movement of the RNase H active site amino acids residues, which may have an affect on RNase H activity. There was no interaction between L517I and E529D with the RNase H active site amino acid residues observed. The observed interaction was between the amino acids that form part of the RNase H primer grip and these RNase H mutations. New HIV-1 subtype C connection domain mutations were identified and phenotypic studies are required to investigate their role in HIV-1 drug resistance.

In conclusion, this is the first study to show that T470S/P, L517I and E529D in HIV-1 subtype C affect NNRTI drug susceptibility. This provides further support for the monitoring of C-terminal domain mutations in relation to NNRTI, as well as NRTI drug resistance. In addition, these mutations need to be taken into account when designing newer NNRTI with the ability to retain activity against these mutants.

Chapter 1

Introduction and Literature review

1.1. Introduction

HIV-1 was first reported in 1981; initially it was associated with homosexuals, and later with drug users and finally with blood transfusion (CDC, 1981, 5th June; Heimer et al., 1992; Masur et al., 1981). The HIV-1 virus was first identified as lymphadenopathy associated virus (LAV), a T lymphotropic virus, in 1983 by Barre-Sinoussi and Montagnier (Barré-Sinoussi et al., 1983; Gallo et al., 1983). In addition, Gallo and his group claimed that this virus was similar in shape to other human T lymphotropic viruses (HTLVs) hence they named the virus as HTLV-III (Gallo et al., 1983). At the same time HTLV-III was reported to have been found in the tissue of AIDS patients in 1983 (Gallo et al., 1983; Gelmann et al., 1983). Soon after the isolation of HIV-1, CD4+ T cells were identified as the main target for HIV-1, these cells are critical to the immune system leading to the Acquired Immnodeficiency Syndrome (AIDS) (Hahn et al., 1984; Shaw et al., 1984).

The development of HIV-1 antiretroviral drugs has transformed HIV-1 from being a lethal infection into a controllable chronic infection (Ring et al., 1993; Tsibris and Hirsch, 2010). The first drug that was discovered to control HIV-1 replication was zidovudine (AZT) (Fischl et al., 1987; Ring et al., 1993). Since then, more than 25 HIV-1 antiretroviral drugs have been approved by the FDA. Despite these efforts, AIDS is still one of the leading health emergencies in the world, due to the development of HIV-1 drug resistance mutations (Schmit et al., 1996; Shafer et al., 1995).

The latest global report on the AIDS epidemic conducted by UNAIDS in 2016, showed that there are approximately 36.7 million of people living with HIV-1 globally;70% of them are in Sub-Saharan Africa (UNAIDS, 2016). Furthermore, more than 18.2 million people with HIV-1 are enrolled on antiretroviral therapy globally (UNAIDS, 2016). In South Africa, 7 million people are HIV-1 infected with more than 3 million people (approximately 48%) receiving antiretroviral drugs (UNAIDS, 2016).

The introduction of HIV-1 antiretroviral therapy has reduced the number of deaths due to HIV-1 infection and AIDS, however the emergence of HIV-1 drug resistance mutations has hindered the success of these drugs (Mas et al., 2002). HIV-1 drug resistance mutations are caused by non-adherence, insufficient drug level and transmission of the HIV-1 resistance strain. Therefore, the genotyping and phenotyping tests has been developed to detect HIV-1 drug resistance mutations. The standard genotyping test target the polymerase region of RT, leaving other parts of the untested. Drug resistance mutations have recently been identified in patients failing their first regimen in HIV-1 C in the C-terminus region of RT (Ngcapu et al., 2017). The current study therefore investigated the phenotypic effect of HIV-1 subtype C reverse transcriptase C- terminal domain mutations on the susceptibility to reverse transcriptase inhibitors.

1.2. Literature review

1.2.1. The Human Immunodeficiency Virus-1

Human immunodeficiency virus -1 (HIV-1) is a lentivirus that belongs to the family Retroviridae and contains a RNA genome. This virus was first identified as the causative agent for the acquired immunodeficiency syndrome over 35 years (Barré-Sinoussi et al., 1983; Gallo et al., 1984; Popovic et al., 1984; Schupbach et al., 1984). HIV-1 infects CD4+ T cell population of immune cells and is transmitted through sexual contact either homosexual or heterosexual. Other HIV-1 infections also occur through transfer of infected blood through needlestick or sharing needles among drug users. Untreated HIV-1 infection leads to depletion of CD4+ T lymphocytes leading to fatal immunodeficiency. There is no cure for HIV-1 infection and no successful vaccine that can prevent infection.

Human immunodeficiency virus-1 genome consists of nine genes which encode 19 proteins. These genes are *gag*, *pol*, *env*, *tat*, *rev*, *nef*, *vif*, *vpr*, and *vpu* (Freed, 1998; Freed, 2001). In addition, the HIV-1 genome has several open reading frames and has a refined system of differential RNA splicing to acquire nine different gene products which encodes 15 different proteins from its 9.2kb genome (Gibbs et al., 2003). The regulatory proteins of HIV are Tat and Rev which plays a role in viral infectivity. HIV also consist of accessory proteins Nef, Vif and Vpr that confer the ability on HIV to infect cells and produce new viral copies thereby causing the disease (Baur, 2004). The three primary HIV-1 translation products are initially made as polyprotein precursors, which are consequently processed by viral or cellular proteases into mature proteins (Figure 1.2).

The structure of HIV-1 is simply a spherical particle with a diameter between 100 and 180nm, and it is covered by a cell-derived lipid membrane containing surface proteins (Freed, 1998; McGovern et al., 2002). It has two copies of positive single stranded RNA that codes for nine genes, which is located inside the nucleocapsid, p7 (Freed, 1998). In addition, the membrane proteins situated on the surface are gp140 and gp41 and are called surface proteins (Chan et al., 1997). Some HIV-1 membrane proteins are captured from the host cell during viral budding and they give support. The matrix (MA) is located underneath the membrane proteins, consisting of trimers of matrix proteins (p17) (Figure 1.1).

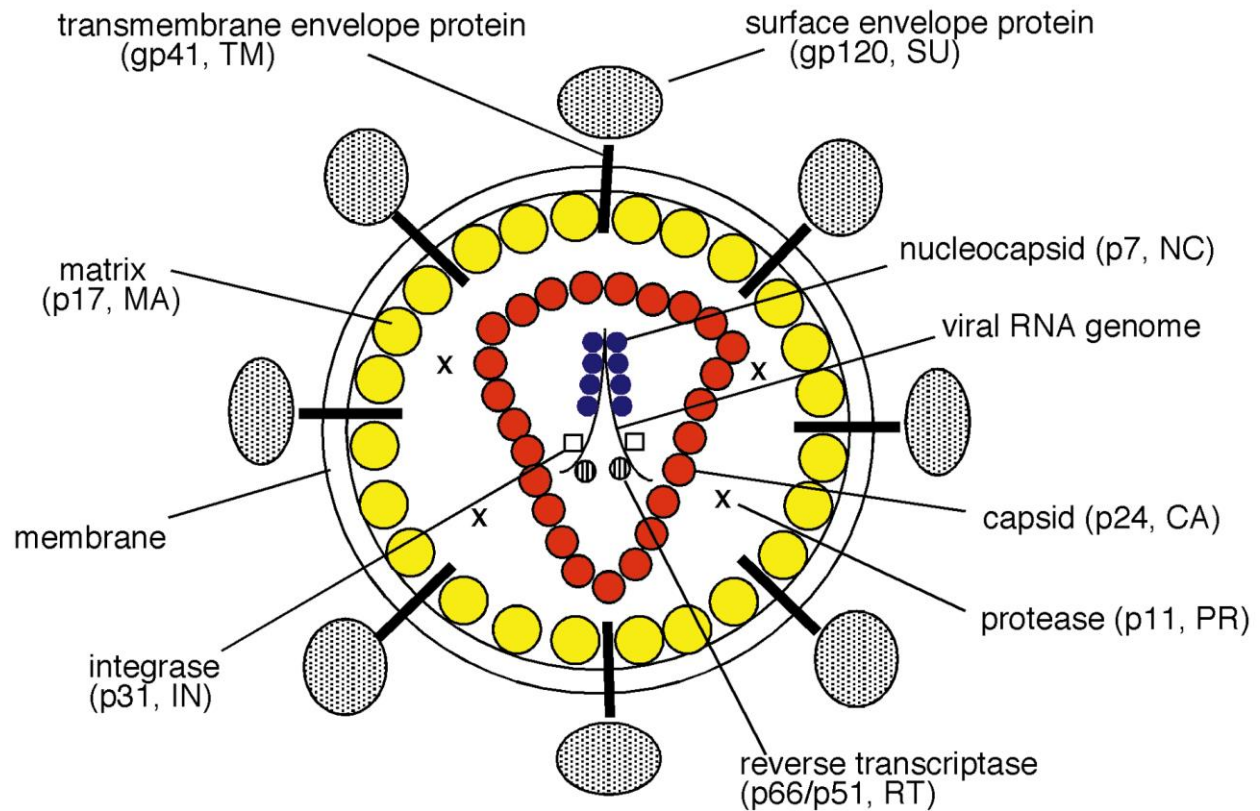


Figure 1.1. Schematic presentation of HIV-1 virion indicating the approximate location of Gag proteins, matrix, the Env glycoproteins, and the Pol-encoded enzymes IN, RT, and PR (Freed, 1998).

The matrix protein which is also known as p17, is the product of HIV *gag* gene. Env protein is located on the surface of the virus containing of three subunits called glycoproteins (gp41, gp120 and gp160) which assist the virus in attaching on the cell membrane (Chan et al., 1997) (Figure 1.1). The glycoprotein complex (gp160), cleaved by the cellular proteases to form glycoproteins (gp120 and gp41) (Chan et al., 1997; Huang and Jeang 1995). These glycoproteins fuse with the host cells and initiate HIV-1 replication (Chan et al., 1997) (Figure 1.2). They also promote the fusion of the infected cells with neighbouring uninfected cells forming syncytia (Chan and Kim, 1998). In addition, Env glycoproteins are the main target for entry and fusion inhibitors and some of these drugs are used in HIV-1 treatment e.g maraviroc (Figure 1.2).

Furthermore, HIV-1 *gag* gene encodes proteins for the infrastructure of the conical capsid that can be sliced by viral protease into structural proteins called, matrix (p17), nucleocapsid (p6 and p7) and capsid (p24) (Gibbs et al., 2003; Wainberg et al., 1988). Nucleocapsid protects the RNA from degradation by nucleases, whereas the viral protein p17 provides a protective matrix for the virion particles. Gag protein is one of the most important genes found in retroviruses and is required for assembly of viral proteins (Freed, 1998). Gag is made up of 1500 nucleotides and encodes four separate proteins which are building blocks of the viral core. It has diverse functions in the life cycle of the virus, including viral assembly (Goepfert et al., 2008) (Figure 1.1). In addition, Gag protein consists of a subunit called p24, which is a component of the virus that is relatively highly conserved. Immune response against p24 may give strong protection since the virus can only avoid these responses at significant fitness cost (Goepfert et al., 2008; Goulder and Watkins, 2008).

The HIV-1 Gag and Pol (Polymerase) proteins are located next to each other in the HIV-1 genome. The *pol* gene, which is translated as a combined *gag-pol* reading frame encodes for reverse transcriptase and integrase (Freed, 1998). The Pol reading frame codes for the polyprotein Pr160,

which consequently contains the four enzymes which are essential in viral replication: integrase (IN), protease (PR) and reverse transcriptase (RT) and RNase H (Freed, 2001; Miller et al., 1989) (Figure 1.2).

The regulatory protein, Tat (trans-activator of transcription) consists of 86 to 101 amino acids depending on the virus subtype. HIV-1 Tat protein accelerates the production of more HIV-1 virions (Henderson and Percipalle, 1997; Hiscott et al., 2001). HIV-1 strains that do not have Tat protein in their genome do not replicate. HIV-1 Tat protein induces chromatin remodeling and recruits elongation-competent transcriptional complexes on the viral LTR (Zheng et al., 2005). In addition, Tat inhibits LFA-1 mediated Ca^{2+} influx through the binding of the L-type Ca^{2+} channel and impairs NK cell cytotoxicity (Majumder et al, 2009).

Rev consists of 116 amino acids including nuclear localization sequence (NLS) and nuclear export sequence (NES) (Zheng et al., 2005). Rev regulates the expression of HIV proteins by controlling the export rate of mRNA (Pollard and Malim, 1998; Zheng et al., 2005) (Figure 1.2). It is also involved in shuttling RNA from the nucleus and the cytoplasm by binding to RRE RNA (Pollard and Malim, 1998). Vif disrupts the antiviral activity of APOBEC3G by targeting it for ubiquitination and cellular degradation. APOBEC3G is a cytidine deaminase enzyme that mutates viral nucleic acid (Zheng et al., 2005).

Negative factor (Nef) is an accessory protein which is synthesized early in the HIV life cycle and is important in disease progression (Collins et al., 1998; Lichterfeld et al., 2004). Nef influences the host's cellular machinery and thus allows infection, survival or replication of the pathogen (Page et al., 1997; Stumptner-Cuvelette et al., 2001). It also promotes the survival of infected cells by downregulating CD4⁺ T cells and HLA and allows infections by manipulating the host cellular

system by down modulating the expression of Major Histocompatibility complex-1 (MHC-1) and CD4+ T cells (Chaudhuri et al., 2007; Cohen et al., 1999). A more recent study reported that HIV-1 Nef protein is proposed to function as a viral suppressor of RNAi (Aqil et al., 2013).

HIV-1 Vpu protein is an accessory protein which consist of 81 amino acids (Deora and Ratner, 2001)(Figure 1.2). Vpu protein promotes the release of new virus particles from the infected cells. In addition, Vpu also plays a role in envelope maturation and downregulates CD4+ T cells (Pollard and Malim, 1998). It also contributes to HIV-1 induced CD4+ T cell receptor downregulation by mediating the proteosomal degradation of newly synthesized CD4 molecules in the endoplasmic reticulum.

Viral protein R (Vpr) is an accessory protein which consist of 96 amino acids and is found in HIV-1, HIV-2 and SIV (Bukrinsky and Adzhubei, 1999) (Figure 1.2). Vpr transport viral DNA into the nucleus as a component of pre-integration complex and also plays a role in cell cycle progression. It also regulates the apoptosis and transactivation of HIV LTR as well as host cell genes (Gibbs et al., 1995; Trono, 1995; Vodicka et al., 1998). Furthermore, Vpr dysregulates the immune system by disarming both adaptive and innate immune response, and it also affects of signal transduction (Goh et al., 1998; Le Rouzic and Benichou, 2005). It is normally found in sera and cerebrospinal fluid of AIDS patients indicating that it may exert biological functions in varied manners. Moreover, studies conducted in monkeys showed that *vpr* deletion in SIV results in low viral loads with rare progression to AIDS, suggesting a major role of *vpr* gene in disease progression in HIV infection (Gibbs et al., 1995; Hadian et al., 2009).

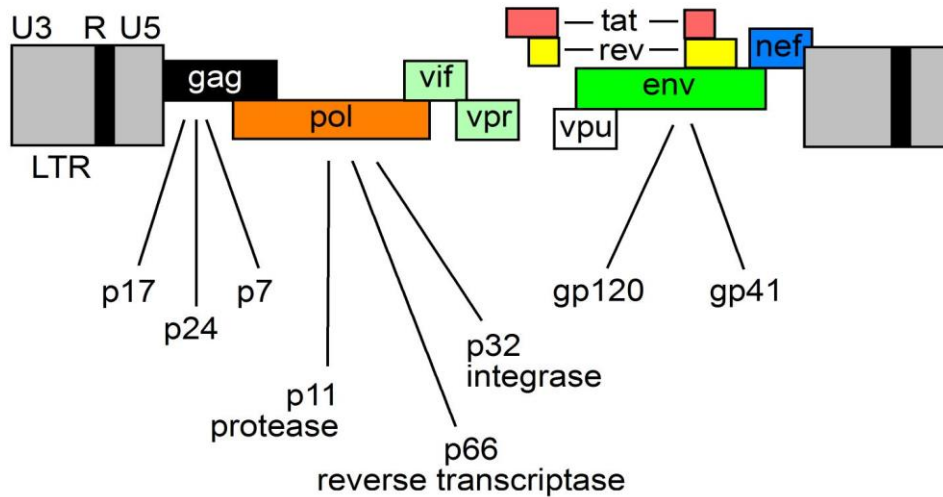


Figure 1.2. The organization of HIV-1 viral genome. (<https://www.hivbook.com/tag/structure-of-hiv-1/>).

1.2.2. HIV-1 replication

The HIV-1 life cycle begins when the gp120 of the envelope protein binds to CD4 T cell receptor and subsequently interacts with other coreceptors such CCR5 (chemokine receptor 5) and CXCR5 (C-X-C chemokine receptor type 4) (Chan and Kim, 1998; Freed, 1998; Freed, 2001). In addition, a subsequent conformational switch also permits the interaction with chemokine co-receptors CXCR4 and CCR5 on the cell surface, promoting the fusion of viral and cell membranes (Alkhatib, 2009; Berger et al., 1999; Bleul et al., 1997). Transmembrane gp41, induces the membrane fusion reaction that occurs between the lipid bilayer of the virion and the host cell plasma membrane. Consequently, the fusion results to the injection of the HIV RNA and various enzymes, including reverse transcriptase, integrase, ribonuclease and protease into host cells (Chen et al., 2002; Chen et al., 1998; Wyatt et al., 1998). Immediately after the release of the viral RNA into the host cell's cytoplasm, the genomic RNA is reverse transcribed into a linear, integration competent double-stranded DNA molecule (Figure 1.3). Furthermore, a ribonuclease enzyme (RNase H) cleaves the

ssRNA during the cDNA synthesis and simultaneously removes the polypurine tract. DNA polymerase facilitates the replication of cDNA into double stranded DNA (Freed, 2001). The microtubules then transport the viral DNA into the host cell's nucleus where the viral DNA is integrated into the host genome (Hiscott et al., 2001; Pollard and Malim, 1998) (Figure 1.3).

The integration of viral DNA into the host genome is catalyzed by integrase and enhanced by a viral nucleic acid called chaperone protein NC, which also protects viral DNA from degradation. After integration, the viral DNA is then transcribed into mRNA using the host cell machinery (Hiscott et al., 2001; Pollard and Malim, 1998) (Figure 1.3). Here, the integrated HIV-1 provirus acts as template for the transcription of viral messengers and genomic RNA, by the cellular RNA polymerase II. The viral promoter in the U3 part of the 5'-LTR requires activation by cellular transcription. After the transcription of mRNA, it is spliced and exported from the nucleus into the cytoplasm. The full length HIV-1 transcripts are transported to the cytoplasm either unspliced (genomic RNA, which also serves as the Gag and Gag-Pol mRNA), partially spliced (encoding Vif, Vpr, Vpu and Env) or full spliced (encoding Tat, Rev, Nef) (Hiscott et al., 2001; Pollard and Malim, 1998). The viral Rev protein links the incompleted spliced and unspliced viral RNA to the export machinery allowing unspliced RNA to exit the nucleus. In the cytoplasm, mRNA is then translated into viral protein and the enzymatic polyprotein precursors Gag and Gag-Pol, and these proteins are transported in different pathways to the plasma membrane (Anderson and Lever, 2006; Bell and Lever, 2013; Jacks et al., 1988). The immature virion buds off the host cell and acquires a phospholipid envelope. After budding, the viral protease is auto-activated and cleaves both Gag and Gag-pol precursors into functional individual proteins (Freed, 2001; Wyatt et al., 1998; Zack et al., 1990).

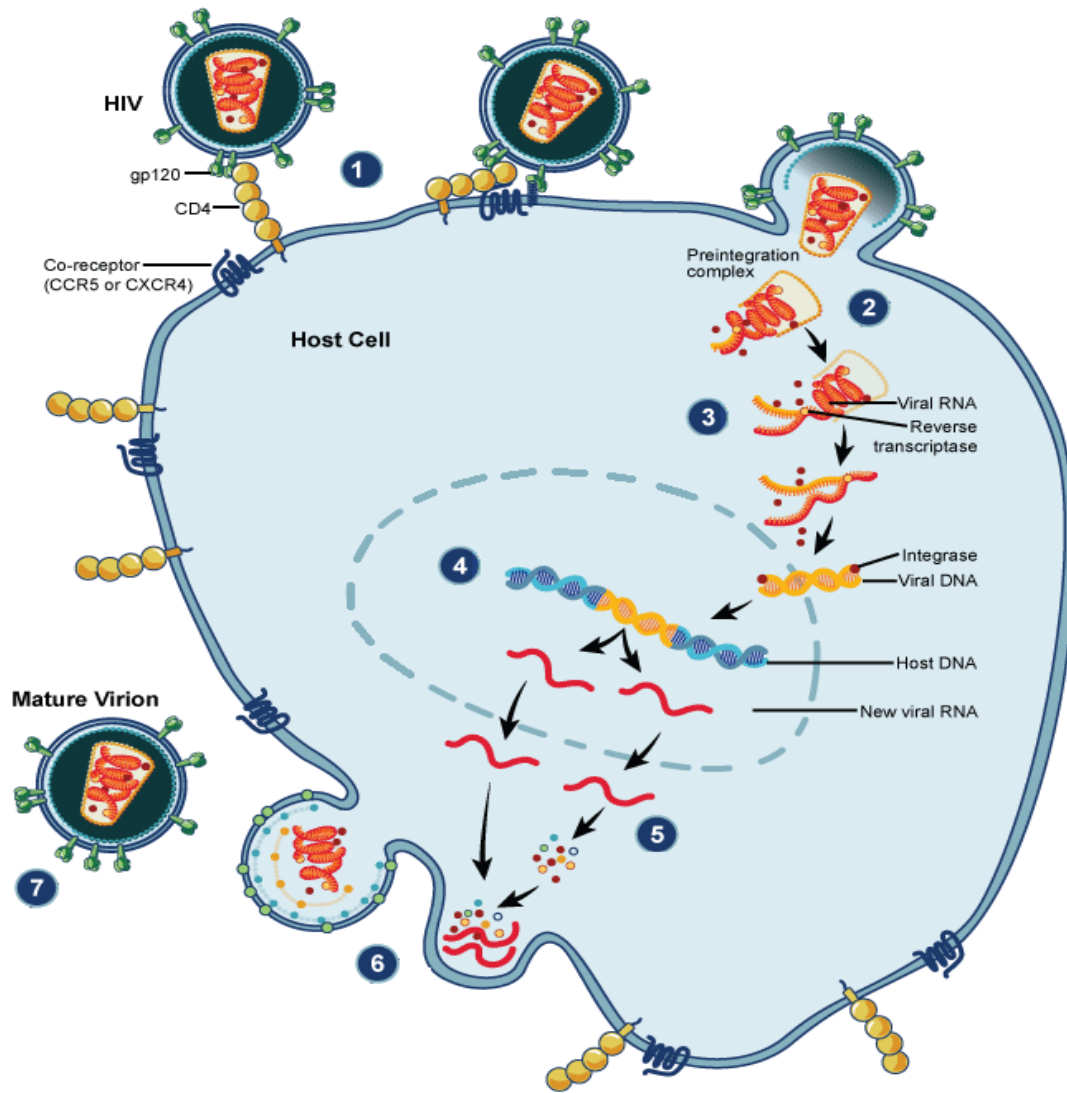


Figure 1.3. Steps in the HIV-1 replication cycle. 1. Fusion of the HIV cell to the host cell surface. 2. HIV RNA, reverse transcriptase, integrase, and other viral proteins enter the host cell. 3. Viral DNA is formed by reverse transcription. 4. Viral DNA is transported across the nucleus and integrates into the host DNA. 5. New viral RNA is used as genomic RNA and to make viral proteins. 6. New viral RNA and proteins move to cell surface and a new, immature, HIV virus forms. 7. The virus matures by protease releasing individual HIV proteins. (Source modified from : National Institute of Allergies and Infectious Diseases, Biology of HIV).

1.2.3. HIV-1 Reverse Transcription

Reverse transcription is an essential step in retroviral replication, it is facilitated by enzyme known as reverse transcriptase, which converts viral RNA into double stranded RNA (Baltimore, 1970b;

Menendez-Arias and Berkhout, 2008; Mizutani et al., 1970; Shafer et al., 2007). Once the viral capsid enters the host cell, reverse transcriptase initiates the process of reverse transcription. The viral capsid releases the single stranded (+) RNA genome from the attached viral protein and copies it into a complementary DNA (cDNA) molecule (Hu and Kuritzkes, 2011; Telesnitsky and Goff, 1997). Since reverse transcription is extremely error-prone, the mutations that occur during this process may cause drug resistance or allow the virus to evade the body's immune system (Freed, 2001; Zack et al., 1990) (Figure 1.4). Subsequently, HIV-1 reverse transcriptase derived from the Gag-Pol polyprotein is cleaved by viral proteases. The RT enzyme is a heterodimer and is composed of two subunits, which is p66 and p51 (Beilhartz and Götte, 2010). The larger subunit, consists of 560 amino acids while the smaller unit is made up of 440 amino acids. The p66 subunit has enzymatic activity, while p51 gives the structural support to p66. Moreover, the p66 subunit has three domains, which are: polymerase, RNase H and the connection domain. The connection domain connects the polymerase and RNase H domain (Ehteshami and Gotte, 2008; Hu and Kuritzkes, 2011; Shafer et al., 2007). Furthermore, the reverse transcriptase has RNA-dependent DNA polymerase, ribonuclease H and DNA-dependent DNA polymerase activities (Tu et al., 2010). Ribonuclease activity degrades the viral RNA during the synthesis of cDNA and polymerase facilitates the conversion of RNA into DNA (Götte, 2007).

Briefly, a specific cellular tRNA acts as a primer and hybridizes to a complementary part of the virus RNA genome called the primer binding site (PBS) (Beilhartz and Götte, 2010). In addition, the second strand DNA binds to the U5 (non-coding region) and R region (a direct repeat found at both ends of the RNA molecules) of the viral RNA synthesis would be primed by the polypurine tract (PPT), which is generated through the RNase H activity of RT (Beilhartz and Götte, 2010). Moreover, the DNA segment is extended from tRNA based on the sequence of the retroviral

genomic RNA. The viral R and U5 sequence are removed by RNase H and DNA hybridizes with the remaining R sequences at the 3' end (Hu and Kuritzkes, 2011; Sarafianos et al., 2002). Extension of the DNA strand from the 3' end occurs with removal of tRNA and PPT by RNase H (Figure 1.4). In addition, a second DNA strand is extended from the viral RNA and the PBS region of the second strand hybridizes with the PBS region of the first strand, resulting into extension of both DNA strands (Hu and Kuritzkes, 2011). Once the newly synthesized DNA strand is initiated from the viral RNA, another strand jump occurs where the PBS from the second strand hybridizes with the complementary PBS on the first strand. Both synthesized strands are extended further and can be incorporated into the host's genome by the enzyme integrase.

1.2.3.1. RNase H

RNase H belongs to the polynucleotidyl transferase superfamily. They are characterized by their catalytic sites that contain a central core composed of four highly conserved amino acids that are coordinated by two divalent Mg^{2+} cations (Esposito et al., 2007; Freed, 2001). The divalent metal ions are important for enzyme activity and they are observed in close proximity to the active site residues (De Clercq, 2002; Esnouf et al., 1995b; Esposito et al., 2007). RNase H is an endonuclease that specifically degrades the RNA portion of the RNA/DNA hybrid. Specifically, it degrades the (+) strand of the RNA genome while removing the (-) strand tRNA and polypurine tract (PPT) (Beilhartz and Götte, 2010). Consequently, RNase H activity can occur in two different modes, polymerase dependent and independent modes (Beilhartz and Götte, 2010; Ehteshami and Gotte, 2008) (Figure 1.5). The polymerase dependent mode is when the 3' end of the primer is bound to the polymerase site while the independent mode has no interaction between the 3' end the polymerase site (Beilhartz and Götte, 2010; Ehteshami et al., 2008a). Furthermore, polymerase dependent binding can occur in two distinct positions that are pre- and post-translocation. Pre-

translocation is where the N-site is occupied by the 3' primer terminus while the incoming nucleotides are blocked by the primer terminus and post translocation is when the 3' primer terminus occupies the P site leaving the N site unoccupied (Beilhartz and Götter, 2010). RNase H cleavage can occur at multiple positions on the template, which generate shorter fragments. In summary, the RNase H active site plays a role in tRNA removal which serve as a primer in (-) DNA synthesis and removal of polypurine tract (PPT) that serves as primer in (+) DNA synthesis (Ehteshami and Götter, 2008; Ehteshami et al., 2008b; Hu and Kuritzkes, 2011).

HIV-1 drug resistance mutations that occurred in the polymerase domain may affect the RNase H activity and polymerization efficiency. Recent studies report that W153L alone or in combination with K65R, M184I, K101E, K103N, E138K and Y181C impair enzyme processivity and polymerization efficiency (Guo et al., 2014; Sluis-Cremer, 2014; Tambuyzer et al., 2011; Xu et al., 2013). Similarly, recent biochemical studies report that the presence of N348I or M184V/N348I mutations decreased the replication capacity of viruses with E138K (Tambuyzer et al., 2011; Xu et al., 2013). Furthermore, M230L mutation cause resistance to EFV, NVP, ETR and was shown to be deficient in minus-strand DNA synthesis, both DNA- and RNA- dependent polymerase activities and RNase H activity (Xu et al., 2010).

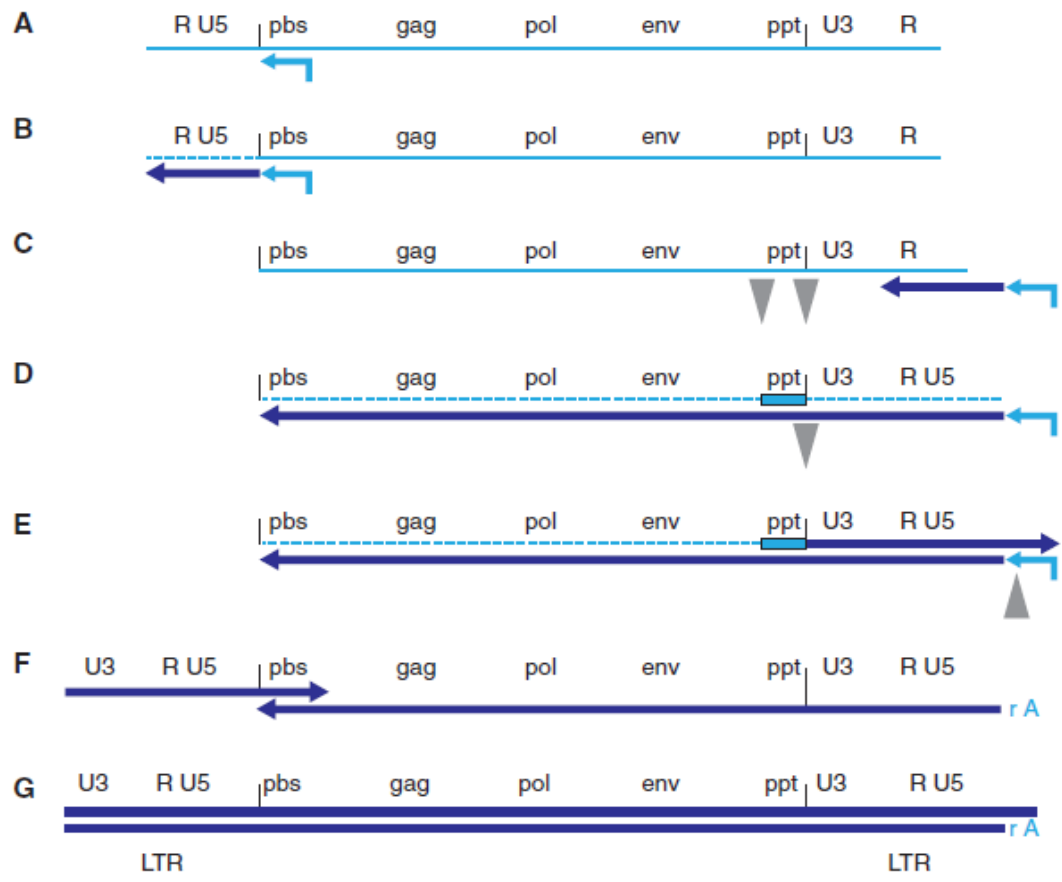


Figure 1.4. HIV-1 Reverse transcription steps. A. The RNA genome of retrovirus (light) with a tRNA primer base paired near the 5' end. B reverse transcription is initiated by RT generating (-) DNA (dark blue) and the RNase H activity of RT has degraded the RNA template (dashed line). C. Minus-strand transfer has occurred between the R sequences at both ends of the genome, allowing (-) DNA synthesis to continue. D. followed by degradation of tRNA. A polypurine tract (PPT) next to U3 is resistant to RNase H cleavage and serves as the primer for the synthesis of plus-strand. E. (+) DNA strand synthesis continues until the first 18 nucleotides of the tRNA are copied allowing the cleavage of RNA to remove the tRNA primer. The removal of the tRNA primer sets the stage for the second (+) transfer and the extension of the plus and minus strands leads to the synthesis of the complete double-stranded linear viral DNA. (Hu and Hughes, 2012).

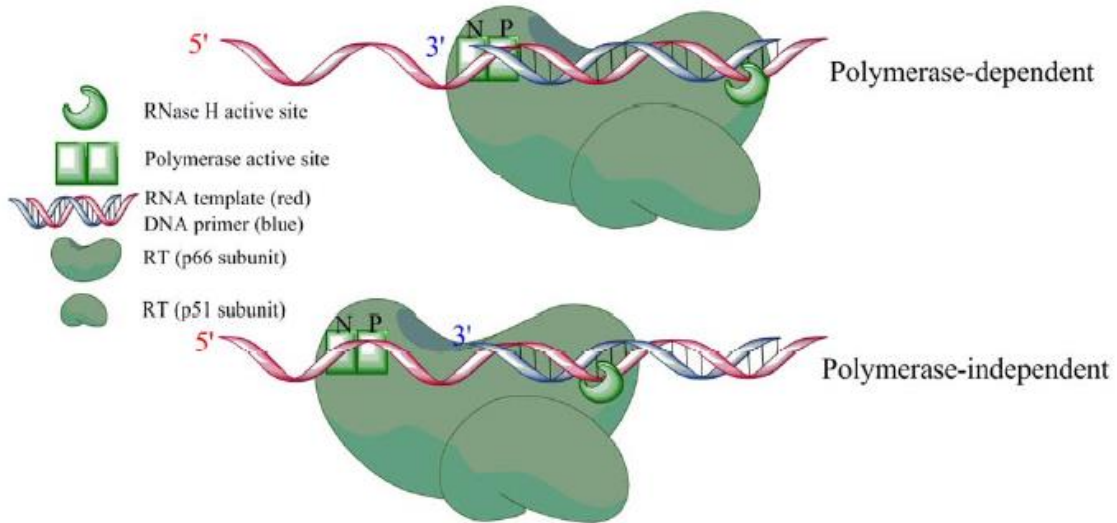


Figure 1.5 . RNase H activity with two modes. Polymerase-dependent mode is when the polymerase active site is in contact with the 3' primer terminus and polymerase independent mode is when the polymerase active site is not occupied by 3' end of the primer and RNase H cuts independently (Delviks-Frankenberry et al 2010).

1.3. HIV-1 antiretroviral drugs/inhibitors

The HIV-1 life cycle has several steps, which provides potential opportunities to develop therapeutic intervention (De Clercq, 2002; Esnouf et al., 1995b). Of the 25 HIV-1 drugs inhibitors that are approved by the FDA, currently 12 of those drugs target reverse transcription. The reverse transcription step is the most critical step of the HIV-1 replication and lasts for 10 hours of infection (Freed, 2001; Hu and Hughes, 2012). Reverse Transcriptase Inhibitors (RTIs) inhibit the polymerase activity of reverse transcription (Antikainen and Martin, 2005; Beilhartz and Götte, 2010). There are two classes of RTIs, the first are the nucleoside RTIs (NRTIs) which are: AZT, ddI, ZAL, d4T,3TC,ABC, FTC and TFV. They are all analogs of native nucleoside and nucleotide substrates (Arts and Hazuda, 2012; Jiang et al., 2011). For example, AZT is a thymidine analogue with an azido group instead of a hydroxyl group at the 3' position of the ribose. The second class are the non-nucleoside reverse transcriptase inhibitors (NNRTIs), which are: DLV, EFV, ETR, NVP and RPV. The proviral DNA integration step is also targeted by approved HIV-1 drug inhibitors.

1.3.1. HIV-1 antiretroviral regimens used in South Africa

There are different HIV-1 antiretroviral drugs available in Southern Africa that are used in suppressing HIV-1 replication (Meintjes et al., 2017; Menendez-Arias, 2002). In South Africa, the first line regimen of HIV-1 treatment has three options including TDF+emtracitabine (FTC) or 3TC + efavirenz (EFV), second option is TDF+emtracitabine (FTC) or 3TC +dolutegravir (DTG) and third option is TDF +emtracitabine (FTC) or (3TC) + rilpivirine (RPV) (Table 1.1). These options are used when the first line regimen fail to suppress HIV-1 replication as indicated by high viral loads and low CD4 counts. The genotyping test for HIV-1 drug resistance is performed and

hence switched to second line regimen which include; TDF+3TC or AZT +3TC and Ritonavir boosted or it can include Lopinavir/r or aluvir. Recently, third line regimen has been introduced in the events that second line fails.

Table 1.1. Dosage of the HIV-1 antiretroviral regimens available in Southern Africa (Meintjies et al 2017)

Generic name	Class of drug‡	Recommended dosage	Common or severe ADR§
Tenofovir (TDF)¶	NtRTI	300 mg daily	Renal failure , tubular wasting syndrome, reduced bone mineral density, nausea
Lamivudine (3TC)	NRTI	150 mg 12-hourly or 300 mg daily	Anaemia (pure red cell aplasia) (rare)
Emtricitabine (FTC)†	NRTI	200 mg daily	Palmar hyperpigmentation
Abacavir (ABC)	NRTI	300 mg 12-hourly or 600 mg daily	Hypersensitivity reaction
Zidovudine (AZT)¶	NRTI	300 mg 12-hourly	Anaemia, neutropenia , GI upset, headache, myopathy, hyperlactataemia or steatohepatitis (medium potential), lipodystrophy
Stavudine (d4T)¶	NRTI	30 mg 12-hourly	Peripheral neuropathy, lipodystrophy, hyperlactataemia or steatohepatitis (high potential), pancreatitis , dyslipidaemia
Didanosine (ddI)¶	NRTI	400 mg daily (250 mg daily if < 60 kg) taken on an empty stomach (only enteric-coated formulation available)	Peripheral neuropathy, pancreatitis , nausea, diarrhoea, hyperlactataemia or steatohepatitis (high potential)
Efavirenz (EFV)	NNRTI	600 mg at night (400 mg at night if < 40 kg)	Central nervous system symptoms (vivid dreams, problems with concentration, dizziness, confusion, mood disturbance, psychosis), rash, hepatitis , gynaecomastia
Nevirapine (NVP)	NNRTI	200 mg daily for 14 days, then 200 mg 12-hourly	Rash, hepatitis
Rilpivirine (RPV)	NNRTI	25 mg daily with food	Rash, hepatitis , central nervous system symptoms (all uncommon)
Etravirine (ETR)	NNRTI	200 mg 12-hourly	Rash, hepatitis (both uncommon)
Atazanavir (ATV)	PI	400 mg daily (only if ART-naive) or 300 mg with ritonavir 100 mg daily (preferable) with TDF; always 300/100 mg daily and with EFV 400/100 mg daily	Unconjugated hyperbilirubinaemia (visible jaundice in minority of patients), dyslipidaemia (low potential), renal stones (rare), hepatitis (uncommon)
Lopinavir/ritonavir (LPV/r)	Boosted PI	400/100 mg 12-hourly or 800/200 mg daily (only if PI-naive)	GI upset, dyslipidaemia, hepatitis
Darunavir (DRV)	PI	600 mg 12-hourly with 100 mg ritonavir 12-hourly or 800/100 mg daily (only if PI-naive)	GI upset, rash, dyslipidaemia, hepatitis (uncommon). Contains sulphonamide moiety (use with caution in patients with sulpha allergy)
Saquinavir (SQV) (rarely used)††	PI	1000 mg with 100 mg ritonavir; 12-hourly, or 1600 mg with 100 mg ritonavir daily (only if PI-naive); Take with a fatty meal, or up to 2 h after meal	GI disturbance (mild), hepatitis, hyperglycaemia, dyslipidaemia
Raltegravir (RAL)	InSTI	400 mg 12-hourly	Headache and other CNS side effects, GI upset, hepatitis and rash (rare), rhabdomyolysis (rare)
Dolutegravir (DTG)	InSTI	50 mg daily	Insomnia, headache and other CNS side effects, GI upset, hepatitis and rash (rare)
Maraviroc (MVC)	CCR5 blocker	150 mg, 300 mg or 600 mg 12-hourly (doses depend on concomitant medication and interactions)	Rash, hepatitis, fever, abdominal pain, cough, dizziness, musculoskeletal symptoms (all rare)

1.3.2. Mechanism of nucleoside and non-nucleoside reverse transcriptase inhibitors

Non nucleoside reverse transcriptase inhibitors (NNRTI) interact with reverse transcriptase of HIV-1 by binding to a single site on the p66 subunit, which is termed the NNRTI binding pocket (NNRTI-BP) (Esposito et al., 2007; Hsiou et al., 1996; Roquebert et al., 2007; Schultz and

Champoux, 2008). Previous studies have reported that NNRTI binding to the HIV-1 RT causes both short and long range distortions of the HIV-1 structure (Hsiou et al., 1996). Short range distortions include conformational changes of the amino acid residues or structural elements that form the NNRTI-BP. Similarly, the long range distortions involve a hinging movement of the p66 thumb subdomain that results in the displacement of the p66 thumb subdomain. This results in the displacement of the p66 connection and RNase H domains and p51 relative to the polymerase active site (de Bethune, 2010). Nucleoside Reverse Transcriptase Inhibitors (NRTIs) are analogues of naturally occurring dNTPs, which lack a 3'-hydroxyl group on the ribose sugar/pseudosugar (Schultz and Champoux, 2008; Sluis-Cremer et al., 2007). In order for NRTIs to perform their antiviral activity, they must be metabolically converted by host-cell kinases to their corresponding triphosphate form (de Bethune, 2010; Menendez-Arias, 2008; Menendez-Arias et al., 2011). These RT inhibitors compete with cellular triphosphates, which are natural substrates for the viral reverse transcriptase. They are also known as competitive inhibitors since they compete with viral RT, hence they act as chain terminators of nascent viral DNA synthesis (Arion and Parniak, 1999).

1.3.3. HIV-1 drug resistance mutations

The emergence of HIV-1 drug resistance mutations decrease the ability of antiviral drugs in inhibit viral replication, this has delayed the fight against AIDS. Some drug resistance mutations may be good for the virus, since they can assist virus to escape the immune system pressure thereby providing it with survival strategies. Similarly, almost all drug resistance mutations are harmful to the virus because they make the virus to create stops or changes in the structure of proteins/enzymes that are essential for viral replication. These changes in the protein structure may develop a viral strains that has better replicative capacity than the wild-type virus. Nonetheless,

antiretroviral drugs are proven to be effective agents in inhibiting the replication wild type strains of the HIV-1 virus. Once the resistance viral strain is harboured in the system the antiretroviral drug fail to inhibit the viral replication with resultant treatment failure.

Previous studies have reported on HIV-1 drug resistance mutations that are located in the first 300 amino acids of RT but not on the C- terminus of the RT (Esposito et al., 2007; Menendez-Arias, 2008; Michels et al., 2010; Pillay et al., 2008) (Figure 1.6). Nucleoside reverse transcriptase inhibitors acquire resistance mutations via different mechanisms. Most NRTI resistance mutations occur at or close to the drug-binding sites of the RT gene, resulting in increased discrimination by the RT. The second mechanism is via pyrophosphorolysis where the NRTI that terminates the chain extension is removed and enables the resumption of DNA synthesis (Julias et al., 2004; Kisic et al., 2008; Marconi et al., 2008)(Figure 1.6).

Zidovudine and stavudine enhance mutations by pyrophosphorolysis and these mutations commonly involved in excision of the terminating inhibitors are know as the thymidine analogue mutations (TAMS) (Hachiya et al., 2008b; Menendez-Arias, 2002, 2008, 2011). Pyrophosphorolysis is the primary mechanism of resistance to AZT and d4T, however it has been reported that these mutations confer reduced susceptibility to all NRTIs. Mutations conferred by the AZT and d4T are also known as thymidine analogue mutations (TAMs). Briefly, TAMS (M41L, D67N, K70R, L210W, T215F/Y and K219E/Q) are defined as non-polymorphic mutations selected by the thymidine analogues AZT and d4T (Bocket et al., 2004; Cases-González et al., 2007). The TAM mutational pathway occurs in two distinct but overlapping pathways: TAM-1 and TAM-2 (Cozzi-Lepri et al., 2005; Cunha et al., 2012; Novitsky et al., 2007). TAM-1 mutations (M41L, L210W, T215Y) interacts with the purine ring of the adenosine triphosphate, while TAM-2 mutations (D67N, K70R, and K219Q) have a larger effect on the proper orientation

of the PPi during the nucleophilic attack required for the excision reactions (Cunha et al., 2012; Mbisa et al., 2011; Menendez-Arias, 2008). TAM-1 mutations have a greater negative impact on virological responses to ABC, ddI or TDF containing regimen than TAM-2 mutations (Bocket et al., 2004).

Other NRTI-resistance associated mutations increase the ability of the enzyme to discriminate against the triphosphate derivate of the NRTI (Delviks-Frankenberry et al., 2010; Nikolenko et al., 2010; Radzio and Sluis-Cremer, 2008). For example, the M184V mutation has been shown to decrease the catalytic efficiency of incorporation of lamivudine-triphosphate (Ehteshami et al., 2008b; Götte, 2007; Gotte et al., 2000; Julias et al., 2004). These drug resistance mutations may affect the replication capacity of the virus and therefore influence how rapidly they become the dominant population in the quasispecies (Garcia-Lema et al., 2005; Santos and Soares, 2010).

Previously studies have shown that mutations which confer high level of resistance have a minimal effect on enzymatic activity, and hence viral fitness, emerges rapidly (Irina et al., 2012; Schultz and Champoux, 2008).

Non-Nucleoside Reverse Transcriptase inhibitors (NNRTIs) bind to amino acids located in a hydrophobic binding pocket within the reverse transcriptase (Hsiou et al., 2001; Tachedjian et al., 2001; Wang et al., 2004). Furthermore, this binding pocket is not part of the RT enzyme binding site, therefore the mechanism of resistance to NNRTI is straightforward. In addition, hydrophobic binding pocket does not exist in the absence of the drug, NNRTI open the pocket and blocks enzyme movement during the process of reverse transcription (Figure 1.6). Figure 1.6 illustrate that binding of NNRTI halt the DNA synthesis since the enzyme is not moving. Therefore, NNRTI mutations resistance involve amino acids that form hydrophobic binding pockets, these includes

L100I, Y181C, G190S/A and M230L (Bachelier et al., 2000; Gulick et al., 2004; Reuman et al., 2010; Richman et al., 1991; Shahriar et al., 2009).

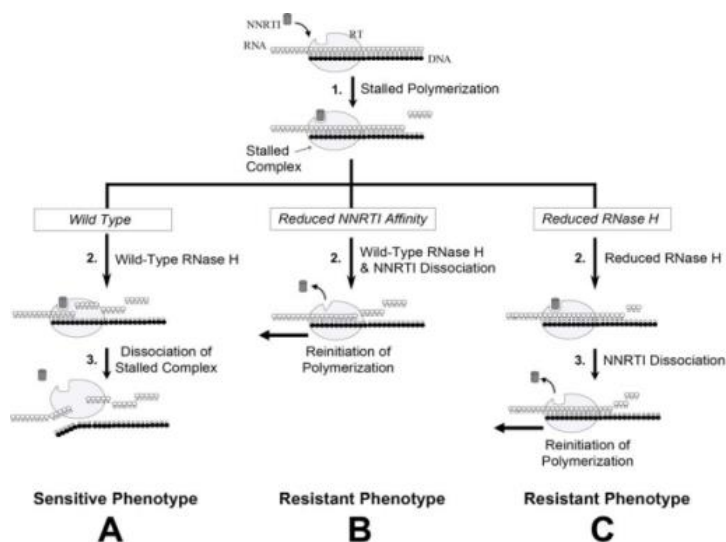


Figure 1.6. Mechanism of NNRTI drug resistance. In the case of drug sensitive virus, NNRTI bind to the hydrophobic pocket near the reverse transcriptase active site and block DNA polymerization, drug resistant virus NNRTI binding blocked and normal DNA polymerization (Nikolenko et al 2010).

1.4. HIV-1 RT Connection and RNase H domain mutations

As mentioned above, most HIV-1 drugs target RT, with well characterized RT mutations occur in the N terminal domain of RT while the C-terminal domain mutations are not well characterized. In addition, most of the HIV-1 drug resistance mutations have been well studied in subtype B, while non-B -subtypes remain understudied (Santos et al., 2008; Santos and Soares, 2010). Mutations implicated in RTI resistance in the CN and RNase H domains are shown in Table 1.1. E312Q, Q334D, G335C/D, N348I, M357R R358K, G359S, A360I/T/V, V365I, K366R, T369V, A371V, A376S, T377M, K390R, E399D, A400T, A435L, S468A,T470S, L484I, I506R, A508S, Q509L, L517I, Q524E, K527N, E529D, K530R and Q547K have a higher prevalence in ARV treated patients (Barral et al., 2016; Maiga et al., 2012; Ngcapu et al., 2017a; Saeng-aroon et al.,

2010; Santos et al., 2008; Schuckmann et al., 2010; von Wyl et al., 2010). In addition, N348I, A360V/T, T377M in the connection domain and D488E in RNase H were found to be selected independently of TAMS (Lengruber et al., 2011; Santos et al., 2008).

1.4.1. Effect of connection and RNase H mutations on drug susceptibility

Previous phenotypic studies have shown that N348I, T369V and A371V have reduced susceptibility to AZT in the presence of TAM-1 and/or TAM-2 pathway mutations, and also confer resistance to NVP (Gupta et al., 2011; Yap et al., 2007). Likewise, D488E and Q547K caused TAM-specific enhancement resistance to AZT while G359S replaced a AZT hypersusceptibility phenotype when combined with A371V (Delviks-Frankenberry et al., 2010; Lengruber et al., 2011). The G333D/E connection domain mutation causes resistance to both AZT and 3TC in subtype B viruses (Gallego et al., 2002). In contrast, this mutation has been reported in subtype C infected treatment naïve patients.

1.4.2. Connection domain mutations and TAMS

Connection domain (CN) mutation, N348I is observed with and without TAMS while A360V mutations appears to be associated with TAMS (Delviks-Frankenberry et al., 2008). Likewise, A360V mutation compensates for a TAM-mediated deficit in viral fitness and therefore appears later, after the emergence of TAMS. On the other hand, N348I mutation is able to recruit PPI as a substrate for the excision reaction and appears early, independent of TAMS (Nikolenko et al., 2005; Schuckmann et al., 2010; von Wyl et al., 2010). In addition, N348I together with M184V compensate for reduced NNRTI drug susceptibility observed in the single N348I mutant, with marginally improved replicative capacity (Gupta et al., 2011; Yap et al., 2007). Recently, studies

reported low viral replication capacity in viruses containing both E138K and N348I, while N348I enhanced resistance to ETR and RPV (Xu et al., 2014a) (Table 1.1).

1.4.3. HIV-1 Connection domain mutations and RNase H

Previous studies show that CN domain mutations are associated with high AZT resistance suggesting that C-terminus domain mutations should be considered in genotypic and phenotypic drug resistance testing (Julias et al., 2004; Julias et al., 2003; Nikolenko et al., 2007; Nikolenko et al., 2005). Several CN domain mutations are associated with RT template switching suggesting that they can change the predicted balance between nucleotide excision and RT template RNA degradation. The CN domain plays a role in positioning the incoming nucleic acid suggesting that CN mutations might affect the RNase H primer grip, resulting in inefficient and unspecific template cleavage (Nikolenko et al., 2007).

1.4.4. HIV-1 RNase H mutations and TAMs

Previous studies report that RNase H domain mutations can significantly contribute to drug resistance in combination with NRTI-resistance mutations in RT (Nikolenko et al., 2005). For instance, Q509L was shown to increase resistance to AZT in combination with the TAMs: D67N, K70R and T215F (Brehm et al., 2007). Brehm *et al* (2008) showed that RNase H mutations increased the AZT-monophosphate excision activity of RT on RNA/DNA template/primers (T/Ps) but not DNA/DNA T/Ps. Furthermore, Nikolenko *et al* (2005) proved that H539N and D549N RNase H mutations significantly increase AZT resistance alone and in combination with TAMs (Nikolenko et al., 2005; Roquebert and Marcelin, 2008). The latter findings provide evidence that RNase H domain mutations affects HIV-1 drug resistance, nonetheless more investigations are required to understand how they confer resistances in all the subtypes.

1.4.5. Connection domain and RNase H mutations in non-B subtypes

There are limited studies on non-B subtypes, however a recent study reported D404N as a novel NRTI and NNRTI-associated mutation found in HIV-1 subtype CRF08_BC treated isolates (Zhang et al., 2015b). The study show that D404N mutation conferred low-level resistance to NVP, EFV, RPV and AZT. In addition, Brazilian studies where HIV-1 subtype C is dominant, characterized some CN domain mutations: N348I (14%), M357R (5.9%), T369I/V (2%), A371V (12%), A376S (5%), A400T (10%) (Barral et al., 2016; Delviks-Frankenberry et al., 2008; Ngcapu et al., 2017a; Santos et al., 2008) (Table 1.1). Some of the CN domain mutations were characterized in HIV-1 subtype CRF02_AG were G333D/E, G335C/D, N348I, A360I/V, V365I, A371V, A376S and E399G (Maiga et al., 2012).

The Brazilian HIV-1C cohort identified two novel C-terminus mutations M357R and E529D (Barral et al., 2016). Previous data from South African HIV-1 subtype C cohort identified RNase H mutations: E438GKR, L517ISV, K527GENQR, E529DK and Q547HKR as treatment-related mutations (Ngcapu et al., 2017a). Interestingly, the E529D RNase H mutation was found in isolates that did not develop any classical NRTI mutations suggesting that E529D could be a primary/novel NRTI drug resistance mutation in HIV-1 subtype C (Ngcapu, 2012; Ngcapu et al., 2017a) (Table 1.2).

Table 1.2. CN and RNase domain mutations from different HIV-1 subtypes

CN and RNase H mutation	HIV-1 subtype	Primary or compensatory mutation	References
I326V	B	Primary	(Santos et al., 2008)
G333D/E	B and CRF02_AG	Primary and Secondary	(Gallego et al., 2002; Kemp et al., 1998; Maiga et al., 2012)
Q334D	B and C	Secondary mutation	(Delviks-Frankenberry et al., 2013; Hachiya et al., 2008a; Hachiya et al., 2009)
G335C/D	B and CRF01_AE	Primary	(Hachiya et al., 2009; Maiga et al., 2012; Nikolenko et al., 2007; Tanuma et al., 2010)
N348I	B, CRF01_AE and CRF02_AG	Primary and Secondary	(Gupta et al., 2011; Maiga et al., 2012; McCormick et al., 2011; Nikolenko et al., 2007; Tanuma et al., 2010; von Wyl et al., 2010; Yap et al., 2007)
R356K	B	Primary	(Maiga et al., 2012; Santos et al., 2008; von Wyl et al., 2010)
M357R	C	Primary	(Barral et al., 2016)
R358K	B	Primary	(Lengruher et al., 2011; Santos et al., 2008; von Wyl et al., 2010)
G359S	B	Primary	(Lengruher et al., 2011; Santos et al., 2008)
A360T/V	B, C and CRF02_AG	Secondary	(Brehm et al., 2007; Brehm et al., 2012a; Kemp et al., 1998; Maiga et al., 2012; Nikolenko et al., 2007; Santos et al., 2008; von Wyl et al., 2010; Yap et al., 2007)
V365I	B and CRF02_AG	Secondary	(Brehm et al., 2007; Hachiya et al., 2009; Kemp et al., 1998; Maiga et al., 2012; Nikolenko et al., 2007; Yap et al., 2007)
K366R	B	Primary	(Santos et al., 2008)Santos et al 2008
T369I/V	B and C	Primary and Secondary	(Gupta et al., 2011; Santos et al., 2008; Santos et al., 2011)
A371V	B, C and CRF02_AG	Primary And Secondary	(Brehm et al., 2007; Delviks-Frankenberry et al., 2013; Hachiya et al., 2009; Lengruher et al., 2011; Maiga et al., 2012; Santos et al., 2008; Santos et al., 2011 ; Tanuma et al., 2010; von Wyl et al., 2010)
A376S	B,C, CRF01_AE and CRF02_AG	Secondary	(Brehm et al., 2007; Delviks-Frankenberry et al., 2013; Hachiya et al., 2009; Maiga et al., 2012; Nikolenko et al., 2007; Paredes et al., 2011; Santos et al., 2011 ; Tanuma et al., 2010)
T377M	B	Primary	(Lengruher et al., 2011)
K390R	B	Primary	(Santos et al., 2008)
E399D	B and CRF02_AG	Primary	(Maiga et al., 2012; Michels et al., 2010)
A400T	B and C, CRF_AE	Primary and Secondary	(Delviks-Frankenberry et al., 2013; Lengruher et al., 2011; Santos et al., 2008; Tanuma et al., 2010)
D404N	CRF01_BC	Primary	(Zhang et al., 2015b)
A435L	C	Primary	(Ngcapu et al., 2017a)
K451R	B	Primary	(Lengruher et al., 2011)
Q457K	B	Primary	(Dau et al., 2010; Santos et al., 2008; Santos et al., 2011)
S468A	C	Primary	(Ngcapu et al., 2017a)
L469T/I/M/H	B	Primary	(Roquebert et al., 2007)
T470S	B and C	Primary	(Ngcapu et al., 2017a; Roquebert et al., 2007)
T470P	B	Primary	(Roquebert et al., 2007)
L484I	C	Primary	(Ngcapu et al., 2017a)
D488E	B and C	Primary	(Lengruher et al., 2011; Santos et al., 2011)
I506G	B	Primary	(Santos et al., 2008; Santos et al., 2011)
A508S	C	Primary	(Ngcapu et al., 2017a)
Q509L	B	Primary and secondary	(Brehm et al., 2007; Hachiya et al., 2009)
K512R	B	Primary	(Lengruher et al., 2011)
L517I	C	Primary	(Ngcapu et al., 2017a)
Q524E	C	Primary	(Ngcapu et al., 2017a)
K527N	B	Primary	(Santos et al., 2008)
E529D	C	Primary	(Barral et al., 2016; Ngcapu et al., 2017a)
K530R	B	Primary	(Santos et al., 2008)
H539N	B	Primary and Secondary	(Nikolenko et al., 2005)
Q547K	B	Primary	(Santos et al., 2011)
D549N	B	Primary and Secondary	(Nikolenko et al., 2005)
A55T/L/K	B	Primary	(Roquebert et al., 2007)
K558R/G/E	B	Primary	(Roquebert et al., 2007)

1.5. Principles of laboratory techniques

1.5.1. Site directed mutagenesis

Site directed mutagenesis (SDM) is a molecular biology technique used to make specific and intentional changes to the DNA sequence of a gene and any gene product (Antikainen and Martin, 2005; Ling and B.H., 1997; Shortle et al., 1981). It is also known as site-specific mutagenesis or oligonucleotide-directed mutagenesis since one nucleotide can move to a specific site and change the amino acid sequence. This technique is used to investigate the structure and biological activity of DNA, RNA and protein molecules and for protein engineering. It is a very important laboratory technique for introducing a mutation into DNA sequence. Previous study has attempted SDM using radiation or chemical mutagens, however the early attempts were non-specific, generating random mutations (Shortle et al., 1981). The present SDM technique used is the QuickChange (Stratagene; USA; La Jolla), where they developed their own mutagenesis kit and primer design tool.

Site directed mutagenesis requires the synthesis of a short DNA primer containing the desired mutation. This primer is complimentary to the template DNA around the mutation site so that it can hybridize with the DNA of the gene of interest. The introduced mutation may be a single or multiple base change (point mutation), deletion or insertion (Zheng et al., 2004). DNA polymerase is used to extend the single stranded primer and *Dpn* I digests the methylated parental DNA strand. The PCR products are directly transformed into host cells for example XL10 gold competent cell, colonies are picked and screened using restriction digestion. Introduction of the desired mutation is confirmed by DNA sequencing (Edelheit et al., 2009). In this study, the QuickChange Site directed mutagenesis from Stratagene (USA, La Jolla) was used, and was implemented according to the manufacturer's guidelines.

1.5.2. Phenotypic drug susceptibility assay

Phenotypic resistance assays assess the susceptibility of clinical HIV isolates to antiretroviral drugs (Haubrich, 2004). It measures the amount of drugs required to suppress 50% of the viral growth, known as the IC_{50} (Pattery et al., 2012; Petropoulos et al., 2000). The drug susceptibility is presented as the IC_{50} of the patient's virus compared to that of the reference/wild type virus. The ratio of these IC_{50} is called the fold change.

There are numerous drug susceptibility assays that have been developed. Older phenotypic drug susceptibility assays had several limitations; they required fresh healthy donor PBMCs and were labour intensive and time-consuming (Japour et al., 1993). It consisted of two steps: the first involved the co-cultivation of HIV-1 infected PBMC with seronegative PHA stimulated PBMCs to obtain a HIV-1 stock. Secondly, the virus stock was titrated for viral infectivity using serial dilutions of the drug in donor PBMCs (Japour et al., 1993; Kellam and Larder, 1994). This assay took a minimum of two weeks to complete and measured real time drug susceptibility.

The recombinant virus assay is based on direct amplification of the patient's RNA from plasma and inserting it into a deleted backbone vector (Petropoulos et al., 2000). In this assay, drug susceptibility of protease and RT inhibitors is measured using a reporter gene, such a luciferase indicator gene and GFP reporter gene. Since the vectors are replicative defective, luciferase activity can be measured following a single round of replication once the luciferase is incorporated in the vector, indicating in infectivity (Petropoulos et al., 2000; Puertas et al., 2012). More recently, studies have used TZM-bl reporter cell lines that contain luciferase (Hertogs et al., 1998; Petropoulos et al., 2000). This assay also has the capacity to incorporate multiple cycles of

infection (Puertas et al., 2012). Furthermore, in comparison with other traditional phenotypic drug susceptibility assays, it is rapid and affordable.

There are currently two commercially available phenotypic drug susceptibility assays; Antivirogram (Monogram) and PhenoSense (Virologic). The PhenoSense assay is a single-cycle assays that uses luciferase as a reporter gene (De Luca, 2006). The Antivirogram phenotypic drug susceptibility assay is another commercial assay used to determine drug susceptibility (Monogram Bioscience). Similarly, this assay also uses a resistant test vector with deleted proteins (PR and RT). The deleted proteins are replaced by the patients' derived genes. In this assay, the recombinant viruses are produced by nucleofection. The Antivirogram and PhenoSense assays both use polymerase chain reaction (PCR) to amplify the patients' proteins (PR or RT) from the patients' plasma. Furthermore a standardized virus inoculum infects 293 T cells in PhenoSense assay, while MT4 cells are infected in the Antivirogram. In PhenoSense assay, virus production is monitored by a luciferase gene that emits light in proportion with to the number of virions after single round of HIV-1 replication (Zhang et al., 2004; Zhang et al., 2005). In contrast, the Antivirogram assay measures the intensity of a tetrazolium dye that produces color when reduced by mitochondrial enzyme in proportion with viable cells (MTT assay).

A recent study by Weng *et al* (2016), developed a new phenotypic assay where two reporter genes were used ZsGreen and luciferase. In their protocol PR and the part of RT amplified from a patient isolate was packaged into a modified lentivirus carrying dual reporter genes ZsGreen and luciferase (Weng et al., 2016). The half replacement with presence of the ZsGreen reporter, decreased the cost of the Bright-Glo luciferase reagent, making the assay cheaper when larger samples are used (Weng et al., 2016). In this study, the phenotypic assay that was used was derived from the neutralization assay developed by Montefiori et al, (2009); briefly a TZM-bl cell line with the

luciferase gene was and the expression of luciferase was determine by the Bright-Glo reagent and measured on a Glomax luminometer (Montefiori, 2009).

1.5.3. 3D homology modelling

Homology modelling is also known as the comparative modelling of proteins; this refers to the construction of an atomic-resolution model of the target protein from its amino acid sequence and experimental three dimensional structure of a related homologous protein. It is also a fast tool for drug discovery and is used to get a rough idea where the alpha carbons of the key residues sit in the folded protein (Martz, 2001). The structure of proteins are usually determined using the X-ray diffraction and NMR distance measurement, X-ray diffraction is time-consuming while NMR looks at isolated proteins rather than protein complexes and is also time consuming (Cavasotto and Phatak, 2009; Chandonia and Brenner, 2005). Due to the complication of the above techniques, more methods were developed which are fast and easy to performed.

Homology modelling was used in structural biology to generate a reliable 3D model of a protein. There are steps that are followed in other to created a good 3D model structure. Firstly, identifying the template; single or multiple sequence alignment. Secondly, model building for the target based on the 3D structure of the template. Thirdly, model refinement and anlysis of alignment, gap deletion and additions. The last step in generation of the model is to do the model validation. There are several computer based software that are used to build the model. These are Swiss Model, Modeller, Modbase, 3D Crunch, PrisM, Composer and Concen (Cavasotto and Phatak, 2009; Li et al., 1997; Sali and Blundell, 1993; Sutcliffe et al., 1987; Yang and Honig, 1999). Modeller combines the sequences and structure into a complete alignment which can be examined using molecular graphics programmes and edited (Floudas et al., 2006; Sali and Blundell, 1993).

SwissModel, which is accessible on a webserver, accepts the sequence to be modelled and then delivers the model by electronic mail. It follows the standardized protocol of homology identification, sequence alignment, determining the core backbone and modelling loops and side chains (Hillisch et al., 2004). In addition, PriSM uses an alignment to build a composite template by selecting each secondary structure from the the most appropriate template. This method can be used for loop modelling and side chain dihedrals which are derived from either template or predicted structure based on main chain torsion angles and neural network algorithm (Yang and Honig, 1999). Similarly, COMPOSER uses a multiple template structure for building homology models. CONCEN creates models by developing distance restraints from the template structure and the sequence alignment of the target and template for atoms. The atoms include all backbone atoms and side-chain atoms of the same chemical type and hybridization state (Bower et al., 1997; Sutcliffe et al., 1987).

In this study, Swiss Model has been used to build the protein structures shown because it is accessible and easily to used. The structures were viewed in Chimera. Chimera is a highly extensible program for interactive visualization and analysis of molecular structure and related data, including density maps.

1.5.4. Bayesian Network analysis

A Bayesian Network (BN) is a graphical model that describes statistical correlation between multiple variables that hold in the domain (Deforche et al., 2008). The model encodes dependencies among all variables, and can be used to understand casual relationships and hence can be used to gain understanding about the problem domain and predict the consequences of intervention (Hackerman, 1996). A BN is learned from data by searching for the most credible network structure that explains casual and cause-effective relationships from data using a

minimum number of arcs (Deforche et al., 2006). Furthermore, it can elucidate a role for mutations selected during treatment, to identify resistance pathways and to investigate influences of background polymorphic positions (Deforche et al., 2007; Theys et al., 2010). Previous studies have shown that the association between mutations may reflect epistatic fitness interactions between mutations and therefore can be used to model a fitness landscape (Deforche et al., 2008; Theys et al., 2010). In addition, the directed acyclic structures of a BN are also capable of representing relationships between the variables through direct conditional or unconditional dependencies (Deforche et al., 2008; Myllymaki et al., 2002). Dependencies are represented by the presence of an arc from one variable to another, showing dependencies between all variables in the data (Myllymaki et al., 2002).

After the most credible BN network structure has been found, assessments of the network robustness are searched using bootstrapping (Deforche et al., 2008; Deforche et al., 2008c). One hundred replicates of non-parametric bootstraps are performed to derive network robustness by the presence or absence of a particular arc (Deforche et al., 2006). The existence and thickness of arcs showing a direct influence amongst the corresponding variables are relative to bootstrap values and their importances are colored according to the arc weight. For example, black arc weighed $\geq 10^9$, purple weighed $\geq 10^6$, green weighed $\geq 10^3$, and blue weighed $\geq 10^3$ while grey weighed ≥ 1 (Myllymaki et al., 2002; Pearl, 1988). The black arc indicates direct influence between resistance mutations, while associations between background polymorphisms are shown in green arcs and the blue arcs indicate an influence from background polymorphisms on drug resistance associated mutations.

1.6. Rationale of the study

HIV-1 reverse transcriptase enzyme is a multifunctional enzyme that contains RNA and DNA dependent polymerase activity, together with RNase H activity (Baltimore, 1970b; Delviks-Frankenberry et al., 2010; Ilina et al., 2012; Telesnitsky and Goff, 1997). It is composed of two subunits, a p51 subunit 440 amino acids long (the polymerase domain) and a p66 subunit which in addition to the polymerase domain has another 120 amino acids on the C-terminal end (the RNase H domain). Most studies on reverse transcriptase inhibitor drug resistance mutations focus on the N-terminal domain of RT (Brehm et al., 2007), while the C-terminal domain has not been fully characterized in the context of drug resistance mutations. Subtype B studies noted that mutations in the CN and RNase H domains are associated with high-level resistance to the NRTI (Brehm et al., 2007). Recently, our group has shown that isolates that did not develop any of the classical NRTI mutations in the RT domain harbored RNase H T470S, L517ISV and E529D mutations in RNase H. Additionally, BN analysis also suggested that E529D was a primary resistance mutation (Ngcapu, 2012). We therefore investigated the phenotypic effects of mutations in the connection and RNase H region of patient viral isolates, on susceptibility to NRTI and NNRTIs. In this study, we investigated the effect of E529D and T470S/P on N/NRTI susceptibility. We also identified novel mutations in CN domain and test their effect on N/NRTI susceptibility. Because the TAMs have also been implicated in CN and RNase H resistance, the effect of connection and RNase H mutations, in conjunction with TAMs, was also investigated. Secondly, BN was used to predict the possible pathway of HIV-1 subtype C reverse transcriptase mutations occurred concurrent with C-terminal mutations.

1.6.1. Hypothesis

We therefore hypothesized the following:

1. The E529D and T470S/P mutations in RNase H are primary drug resistance mutations in HIV-1 subtype C.
2. HIV-1 Connection and RNase H domains mutations, in combination with RT mutations affect viral fitness and drug susceptibility in HIV-1 subtype C.
3. Connection and RNase H domains mutations of HIV-1 subtype C affect the 3D conformation of the RT.

1.6.2. Specific Aims

1. To determine if the E529D mutation in RNase H is a primary drug resistance mutation in HIV-1 subtype C using site directed mutagenesis.
2. To identify novel resistance mutations in the connection domain and test the effect of these mutations on drug susceptibility and viral fitness.
3. To determine the effect of RNase H domain mutations in combination with the D67N thymidine analogue mutation (TAM) on drug susceptibility and viral fitness.
4. To perform 3D homology modeling of RNase H mutants.

Chapter 2

Minimal phenotypic effects of RNase H domain mutations in HIV-1 subtype C on reverse transcriptase inhibitors

2.1. Introduction

HIV-1 reverse transcriptase (RT) catalyses the conversion of viral single stranded RNA into double stranded DNA that integrates into the host genome (Baltimore, 1970a; Rhee et al., 2005; Spence et al., 1995b). In addition to RNA and DNA dependent polymerase activity, RT also has ribonuclease H (RNase H) activity, where RNase H removes the tRNA primer and polypurine tract (PPT) during reverse transcription (Beilhartz and Götte, 2010; Menendez-Arias and Berkhout, 2008). The native structure of HIV-1 RT consists of 66 kDa (p66) and 51 kDa (p51). The p66 subunit consists of the polymerase and RNase H domains linked by the connection domain (Baltimore, 1970a; Ilina et al., 2012; Sarafianos et al., 2009; Schultz and Champoux, 2008); the p51 subunit is catalytically inactive but plays a structural role (Schultz and Champoux 2008; Beilhartz and Götte 2010). Over half of the currently available FDA-approved drugs against HIV infection target the DNA polymerase activity of the RT enzyme, called the RT inhibitors (Ilina et al., 2012). These RT inhibitors are further classified into nucleoside and non-nucleoside RT inhibitors (NRTIs and NNRTIs) (Ilina et al., 2012; Pennings, 2013; Rhee et al., 2005), based on the way in which they inhibit HIV-1 replication. Nucleotide Reverse Transcriptase Inhibitors bind

at the nucleotide binding site and upon incorporation, inhibit the elongation of the viral DNA chain due to the absence of the 3'OH group, resulting in termination of DNA synthesis (Ilina et al., 2012; Iyidogan and Anderson, 2014). Non-Nucleotide Reverse Transcriptase Inhibitors are non-competitive inhibitors that induce a conformational change around the active site upon binding to the hydrophobic NNRTI binding pocket (NBP), where they block the chemical reaction step of DNA polymerization (Rittinger et al., 1995; Spence et al., 1995a; Xia et al., 2007). Moreover, NNRTIs distort the catalytic site shifting the three key aspartic acids at positions 110, 185 and 186 and restrict the flexibility of a key loop between positions 183 and 186. Mutations clustered around the hydrophobic site are associated with decrease susceptibility to the NNRTIs (Das and Arnold, 2013b; Delviks-Frankenberry et al., 2010; Esnouf et al., 1995a; Tambuyzer et al., 2009).

In spite of the successful use of the RTIs, the development of resistance has created a major hurdle in the long term use of these drugs (Paredes et al., 2011). Previous studies have identified two mechanisms of NRTI drug resistance. The first mechanism involves discriminatory mutations that weaken the binding affinity of the NRTI, allowing the dNTPs to bind, while the second mechanism involves nucleotide excision, resulting in the reversal of polymerization and restoration of DNA synthesis (Arion et al., 1998; Iyidogan and Anderson, 2014; Meyer et al., 1998). Previous studies reported that mutations in the RT CN and RNase H domains modulate HIV-1 susceptibility to RT inhibitors by 10 to 50 fold (Delviks-Frankenberry et al., 2008; Ehteshami et al., 2008a; Hachiya et al., 2008a; Nikolenko et al., 2007; Yap et al., 2007; Zelina et al., 2008).

Key studies have reported that RNase H mutations also dramatically increase resistance in the presence of the TAMs (Brehm et al., 2007; Brehm et al., 2012b). While H539N and D549N RNase H mutations increased AZT resistance 180-fold and reduced d4T susceptibility by 10-fold, when found together with D67N-K70 -T215Y-K219Q, AZT and d4T resistance increased

by 1,250-fold and 12,5 fold respectively (Brehm et al., 2007; Nikolenko et al., 2010; Nikolenko et al., 2005).

Most drug resistance studies have been performed in HIV-1 subtype B isolates, with limited data from subtype C, particularly for RNase H. Recently, two novel mutations in connection and RNase H (M357R and E529D respectively) were identified in a HIV-1 subtype C Brazilian cohort (Barral et al., 2016). Previous data from our group identified E529D and to a lesser extent L517I, as treatment-related RNase H mutations from a local subtype C cohort (Ngcapu et al., 2017b). The current study investigated the phenotypic effects of these RNase H mutations in the context of HIV-1 subtype C.

2.2. Material and methods

2.2.1. Antiretroviral drugs and cell lines

All antiretroviral drugs, including zidovudine (AZT), stavudine (d4T), nevirapine (NVP), etravirine (ETR) and efavirenz (EFV) as well as a pNL4.3 infectious molecular clone, were obtained from the NIH AIDS research and reference reagents program (<https://www.aidsreagent.org/index.cfm>). The HEK293T cell line was obtained from the American Type Culture Collection (ATCC). The TZM-bl cell line was a gift from Eric Hunter, Emory University, Atlanta, USA (Center for AIDS Research). The pNL4.3 RT deleted backbone was a gift from Mark and Zabrina Brumme (Simon Fraser University, British Columbia, Canada). TZM-bl and 293T cell lines were maintained in Dulbecco's Modified Eagle's Medium with L-glutamine, sodium pyruvate, glucose, pyridoxine and 25 mM HEPES (4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid) (Gibco BRL Life Technologies) containing 10% heat-inactivated fetal bovine serum (FBS) and 50 µg gentamicin/ml in vented T-75 culture flasks (Corning Costar).

Cells were incubated at 37°C in a humidified 5% CO₂. Unless otherwise specified, all incubations were carried out under these conditions. Cell monolayers were split 1:10 at confluence by treatment with 0.25% trypsin, 1 mM Ethylenediaminetetraacetic acid (EDTA) (Invitrogen) as described (Montefiori, 2009).

2.2.2. Construction of the RTsubCWT recombinant plasmid

A wild-type recombinant plasmid carrying the whole RT gene from a subtype C wild-type virus isolated from a stored plasma sample (PKE4N) was constructed. The sample was obtained from a patient enrolled in the Tropism study (BREC310807/41 hs), where samples were collected with informed consent from patients attending the ARV clinic at King Edward VIII Hospital, KwaZulu-Natal, South Africa. This sample was chosen because the RT sequence most closely matched the subtype C RT consensus sequence. Ethics approval to use this sample was obtained from the UKZN Biomedical Ethics Committee (BREC approval number: BE017/14).

Viral RNA was extracted using the Qiagen Viral RNA Mini kit (Qiagen, Netherlands), according to the manufacturer's instructions. The RNA was amplified using the Superscript III One-step RT-PCR system with Platinum Taq DNA polymerase (Invitrogen, Life Technologies Corporation, Carlsbad, CA, California) and the following primers: Forward primer: 5' TCCCTCAGATCACTATTTGGCA 3' and reverse primer: 5'CAGTCTACTTGTCCATGCATGGCTTC 3'. A second round of PCR was performed using the Q5 High fidelity DNA polymerase (New England Biolabs; Ipswich, MA) and the following primers: P66_chimericrecomb_F:

5'CTGCGGACATAAAGCTATAGGTACAGTATTAGTAGGACCTACACCTGTCAACATA ATTGGAAGAAATCTGTTGACTCAGATTGGCTGCACTTTAAATTTT 3' and

P66_chimericrecomb_R:

5'TTGCTACTACAGGTGGTAGGTTAAAATCACTAGCCATTGCTCTCCAATTACTGTGATATTCTCATGTTCTTCTTGGGCCTTATCTATTCCATCTAAAAAT3'. The PCR reaction was performed as follows: 98°C for 3 minutes, followed by 18 cycles of 98°C for 10 seconds, 55 °C for 30 seconds and 72°C for 5 minutes. The presence of the 2000bp RT PCR product was confirmed by running 5µl of each product on a 1% agarose gel. Products were purified using the QIAquick gel extraction kit (Qiagen, Netherlands) as per the manufacturer's instructions. The purified DNA was then sequenced using the Big Dye terminator sequencing kit v3.1 (Applied Biosystems, Foster City; CA) and run on an ABI Prism 3130 XL automated DNA sequencer (Applied Biosystems, Foster City; CA). The sequences were analyzed using Sequencher v5.4.6 (Gene Codes Corporation, Ann Arbor, MI, USA).

The RT amplicon was cloned into a TOPO vector using the TOPO TA cloning kit version 2.1 (Invitrogen, Carlsbad) according to the manufacturer's instructions. Colonies were screened for the insert using colony PCR and verified by restriction digest with Eco R1 (New England Biolab, Ipswich, MA). The DNA from positive clones was extracted using the GeneJet plasmid miniprep kit (Thermo scientific, Lithuania, EU).

2.2.3. Construction of viral mutants (RTsubCmut) using site-directed mutagenesis

The following individual mutations: D67N, T470S, T470P, L517I, E529D and L517I+D67N combination were introduced into the TOPO plasmid containing the RT amplified product (RTsubCWT), using the Quick-Change II XL site directed mutagenesis kit (Stratagene, La Jolla, CA). Briefly, two complimentary oligonucleotides containing the desired mutations were designed

using the QuickChange mutagenic primer designer tool and are shown in Table 2.1(<https://www.chem.agilent.com/store/primerDesignProgram.jsp>). The underlined codon denotes the single nucleotide substitutions that were introduced into the plasmid. The control and sample reactions were prepared as follows: 5µl of 10x reaction buffer, 2µl (10 ng) of pWhitescript 4.5 kb control plasmid and 2 µl for Plasmid DNA (TOPO+ RT), 1µl of oligonucleotide primer (anti-sense) and 1 µl oligonucleotide primer (sense), 1 µl of dNTPs, 1µl of *PfuUltra* HF DNA polymerase was added. Finally, distilled water was added to a final volume of 50µl. Each reaction was cycled using the following PCR conditions: 95 °C for 1 minute; 95 °C for 50 seconds; 60 °C for 30 seconds; 68 °C for 5 minutes for 18 cycles. Following 68 °C for 7 minutes and the PCR product was placed on ice for 2 minutes to cool the reaction. The amplification products were digested with 1µl *Dpn* I restriction enzyme for 1 hour at 37 °C. Following the digestion of parental DNA strand, 1 µl *Dpn* I-treated DNA was transformed from each control and sample reaction into separate 50 µl aliquots of XL1- Gold ultracompetent cells according to the manufacturer's recommendations. The transformation of PCR products were done according to the manufacturer's recommendations. Table 2.1 shows the mutagenic primers that were used for the SDM reaction. The underlined codon denotes the single nucleotide substitutions that were introduced into the plasmid. Briefly, the positive clones were picked and cultured in the LB broth media with ampicilin overnight. Following the overnight incubation, DNA isolation was performed by GeneJet plasmid miniprep kit and confirmed by enzyme restriction digestion (Thermo Fisher Scientific, Inc). Positive clones were confirmed for the presence of the mutation by DNA sequencing. The sequencing reaction was performed using 0.4µl Pre-mixed Big-Dye terminator v3.1 (Applied Biosystems), 2µl 5x sequencing buffer, 4.6µl sterile DNase free water, 1µl DNA (20ng/µl) and 2µl of 3.2 pmol primers were added to the reaction. The final volume of the

sequencing reaction was 10µl. The reaction was placed in the thermocycler using the following cycling conditions: 96 °C for 1 minute; 96 °C for 10 seconds; 50 °C for 5 seconds; 60 °C for 4 minutes, sequencing reaction was done for 25 cycle. The sequences were run on a 3130XL automated DNA sequencer (Applied Biosystems).

Table 2.1. Primer sequences for the SDM

Name of Primer	Sequence of mutagenic primers
g199a_D67N sense	5'-TTTTCTCCATTTAGTACTGTTTTTTTTCTTTATGGCA <u>AACTAC</u> TGG AGT ATT G-3'
g199a_D67N antisense	5'-CAA TAC TCC AGT ATT TGC CAT AAA GAA AAA AAA CAG TAC TAA ATG GAG AAA A-3'
c1409_T470S antisense	5'-GATTTGTTGTGTCACCTTAGGGTGACAACCTTTTGTCTTCCT-3'
c1409g_T470S sense	5'-AGGAAGACAAAAAGTTGTCACCCTA <u>AGT</u> GACACAACAAATC-3'
a1408c_T470P sense	5'-GACAAAAAGTTGTCCCCCTA <u>CCGG</u> GACACAACAAATCAG-3'
a1408c_T470P antisense	5'-CTGATTTGTTGTGTCGGTAGGGGGACAACCTTTTGTGC-3'
t1549a_L517I sense	5'-GCT CTA TTA TTT GAC TGA CTA TCT CTG ATT CAC TCT TAT CTG G-3'
t1549a_L517I antisense	5'-CCA GAT AAG AGT GAA TCA GAG ATA GTC AGT CAA ATA ATA GAG C-3'
a267t_E529D antisense	5'-CATGACAGGTAGACCCTATCCTTTTT <u>TAT</u> TAGTTGTTCTATTATTGGT-3'
a267t_E529D sense	5'-ACCAAATAATAGAACAACATAAAAAAGG <u>ATA</u> GGGTCTACCTGTCATG-3'

2.2.4. Co-transfection

The co-transfection was performed as previously described by Xu et al 2014. Briefly, 16 µg of pNL4.3ΔRT and 7.5µg of the different RTsubCmut recombinant plasmids (D67N, T470S, T470P, L517I, L517+D67N, E529D) were cotransfected into HEK293 T cells using Fugene 6 (Invitrogen) according to the manufacturer's protocol. The pNL4.3 laboratory strain and RTsubCWT were used as controls. Cultures were incubated at 37°C, 5% CO₂, and harvested after 48 hours. Supernatants were clarified by centrifugation at 1700 rpm for 10 minutes followed by filtration through a

0.45µm pore-size filter and stored at -80°C. The presence of the introduced mutations was confirmed by sequencing.

The TCID₅₀ of the viruses were tested by infecting T_{ZM}-*bl* cells using a 10-fold dilution of each virus stock and was incubated at 37°C, 5% CO₂. After 48 hours, 100µl of culture was removed from the plates and replaced with 100µl of Bright-Glo luciferase reagent (Promega; Madison; WI; USA). After a two minute incubation at room temperature, 150µl of the mixture was transferred into a black plate (Costar; Corning Incorporated). The relative luminescence units (RLU) were measured using the Glomax luminometer (Modulus microplates). A standard amount of virus (50000 RLU; MOI=0.05) was used in the replication and drug susceptibility assays.

2.2.5. Replication capacity

The replication capacity of the recombinant viruses (RT_{sub}Cmut) and RT_{sub}CWT were evaluated in duplicate (and on repeated on a different day) in a non-competitive infectivity assay using T_{ZM}-*bl* cells (Asahchop et al., 2011; Azijn et al., 2010a). Briefly, 50µl of the titred viral stocks (MOI=0.05) were added in duplicate in 96-well flat bottom plates containing 10000 T_{ZM} *bl* cells/well in 100 µl of Dulbecco modified Eagle medium (DMEM; Invitrogen; Carlsband; CA) supplemented with 10% fetal bovine serum (Gibco), 1% gentamicin, 1% L-glutamine and HEPES. Viruses and T_{ZM}- *bl* cells were incubated at 37°C in 5% CO₂ for 48 hours, after which 100µl of culture was removed from the plates and the luciferase activity was measured as described in 2.2.3 above. The level of viral replication was expressed as a percentage of the RLU with reference to the respective wild type virus.

2.2.6. Single cycle drug susceptibility assay

The phenotypic effect of the RNase H mutations (T470S, T470P, L517I, E529D) and RT D67N on AZT, d4T, NVP, ETR and EFV drug susceptibility was measured using TZM-*bl* cells as described previously (Asahchop et al., 2013; Schader et al., 2012; Xu et al., 2014b; Xu et al., 2010). These drugs were selected because the drug resistance mutations that were discovered in patients which were administering were two NRTI either stavudine (d4T) or Zidovudine (AZT) plus lamivudine (3TC) and NNRTI, either efavirenz (EFV) or Nevirapine. Efavirenz was given to those patients previously took NVP during the onset of labor pains during delivery, preventing mother to child transmission of HIV-1. While ETR is the new generation of NNRTI, were chosen because we wanted to check whether these mutations were not associated with the newer drugs. However, d4T was faced out due to its toxicity and the new first line regimen was introduced. The current ARV regimen for a treatment-naive patient generally consists of two NRTIs, usually abacavir/lamivudine (ABC/3TC) or either tenofovir alafenamide/emtricitabine (TAF/FTC) or tenofovir disoproxil fumarate (TDF)/FTC, plus a drug from one of three drug classes: an INSTI, an NNRTI (EFV), or a boosted PI.

All assays were performed in duplicate and repeated on a different day. In short, the RTIs were serially diluted (1 in 10 dilution) in 96-well plates in 100µl of DMEM media supplemented as described above (2.2.5). Wild type or mutant virus (50µl) (MOI= 0.05) was added immediately after dilution of the drug. After a one hour incubation, the TZM-*bl* cells (10,000 cells per well) were added and incubated for 48 hrs. The luciferase activity was measured as described in section 2.2.5 above. Percent inhibition of each drug was calculated as follows: (% inhibition = $1 - \frac{RLU_{(virus\ test)}}{RLU_{(cell\ control)}/RLU_{(virus\ control)} - RLU_{(cell\ control)}} \times 100$). The 50% effective concentration (EC₅₀) was calculated using GraphPad Prism v5 (GraphPad Software, San Diego, CA). Drug

susceptibility was expressed as the fold change in EC_{50} of mutant virus compared with the wild type virus. Drug resistance levels were described as: moderate-level resistance (4- to 20-fold change), low-level resistance (2-to 4-fold change) and susceptible (0- to 2-fold change). This method has been previously used to describe NRTI and NNRTI drug resistance levels (Basson et al., 2015; Wu et al., 2012; Zhang et al., 2015a).

2.3. Results

2.3.1. Effects on viral replication capacity

2.3.1.1. T470S and T470P cause similar decrease in replication capacity

In Figure 2.1, the replication capacity of both T470S and T470P RNase H were compared to RTsubCWT. The effect on replication capacity was assessed by infecting TZM-bl cell with the mutant (T470S and T470P) and wild-type viruses. The results show that the replication capacity of the T470P mutant (39948.0 RLUs) was lower than the T470S mutant virus (83581.1 RLUs), although this difference was not significant. Both had replication capacities lower than the wild type C virus; however only the T470P mutant significantly affected replication ($p=0.0249$), while the T470S mutant was not significantly different to the wild-type ($p=0.0765$).

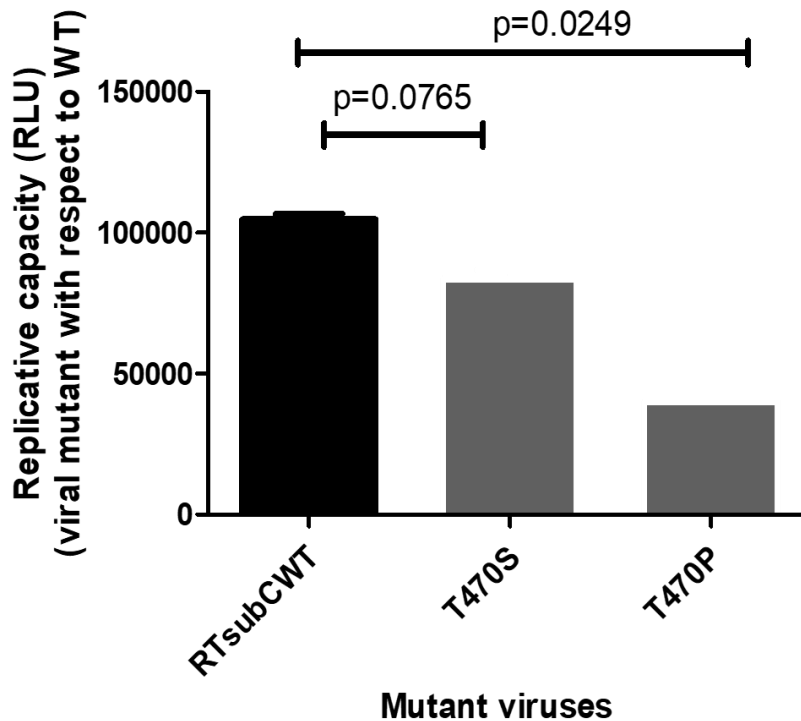


Figure 2.1. Replication capacity measured in RLUs of the T470S and T470P mutants generated. Relative luminescence units were shown in y-axis and the mutant viruses were shown in x-axis. All mutants had a lower replication capacity than the WT virus (104659 RLU). T470P had the lowest replication capacity (39948 RLU), and the T470S (83581 RLU).

2.3.1.2. E529D impairs replication capacity

Figure 2.2 shows that the replication capacity of the competent E529D mutants in both subtype B and C was significantly reduced compared to the wild-type viruses, with subtype C E529D mutant showing the lowest replication capacity among the competent mutants. The relative luminescence units (RLUs) of subtype C E529D mutant virus were lower compared to the RTsubCWT virus (65736.6 versus 70651 RLUs, $p < 0.0001$). Similarly, the subtype B E529D mutant virus had a lower RLUs compared to the subtype B wildtype virus (138664 versus 239782 RLUs, $p = 0.002$). There was no significant difference in RLUs of wild-type viruses, with subtype B virus showing the highest RLUs (239782) compared to 138664 RLUs found in subtype C virus ($p = 0.0001$). Mutation E529D significantly reduced the replication capacity of the subtype C mutant virus

compared to subtype B mutant virus (70651 versus 138664 RLUs, $p=0.0014$). There were significant differences between the wild types (both subtype B and C) and the RNase H E529D.

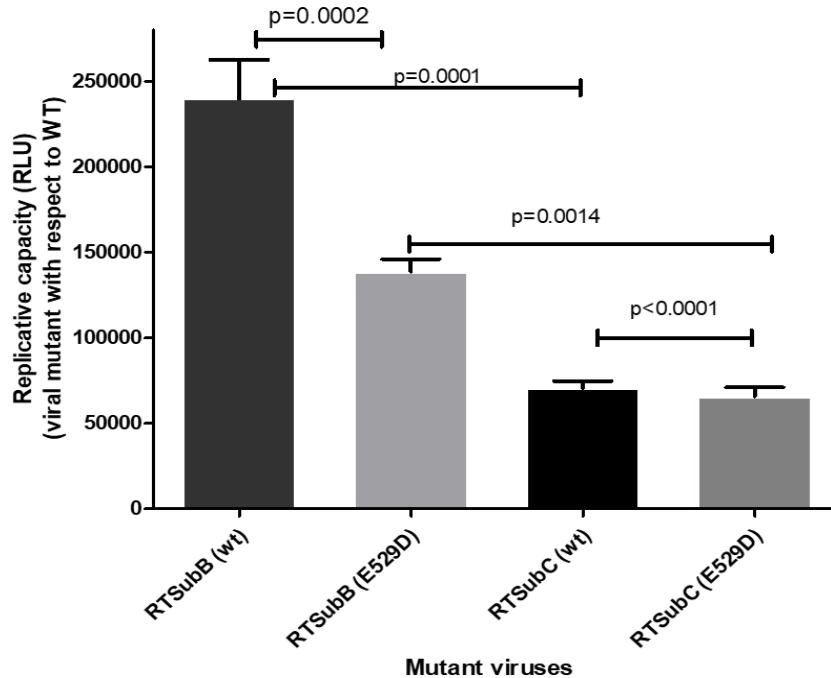


Figure 2.2. Replication capacity of HIV-1 viral mutants generated by site directed mutagenesis. The RC of RTSubB (NL4.3) wild type and mutant virus (RTSubB(E529D) was compared ($p=0.0002$). RTsubC wild type and RTSubC (E529D) (65736.6 RLU) mutant in the absence of drug. The Relative Luminiscence Units (RLU) are shown on the y-axis and viral mutants on the x-axis. Luciferase activity was measured at 48h post infection as an indication of viral replication.

2.3.1.3. L517I+D67N double mutant rescues the impaired viral replication capacity caused by D67N.

As shown in figure 2.3, the replication capacity of the virus containing the L517I viral mutant was 39831.60 RLU compared to RTSub-C wildtype virus (106659 RLU), and L517I+D67N viral mutant (10070 RLU) compared to the RTsubCWT (106659 RLU). The D67N viral mutant had a

more marked effect on replication capacity (51673.1 RLU). The replication capacity of the double mutation, L517I+D67N (10070 RLU,) was lower than the D67N single mutant, suggesting that it has additive effect on the fitness loss caused by D67N. There was significant difference in replication of wild type and L517I viral mutant ($p=0.0003$) also significant difference between the TAM (D67N) and L517I ($p=0.0004$).

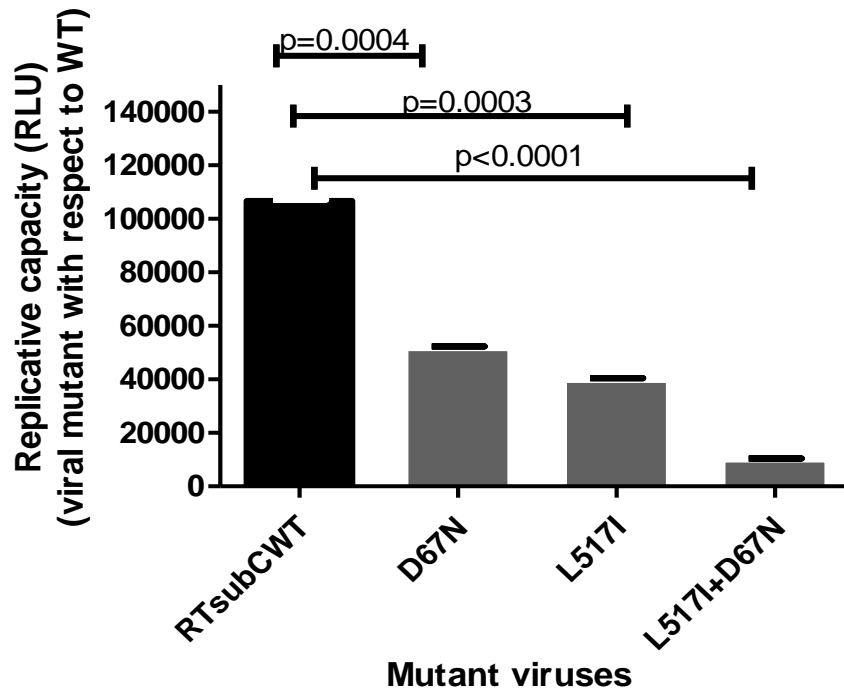


Figure 2.3. Replication capacity (RLU) of the L517I and D67N mutants alone and in combination. Viral mutants were shown on the x-axis and the replication capacity in RLUs were shown on the y-axis. All mutants had lower replication capacities than the wildtype virus: D67N was 51673.1 RLU, L517I was 39831.60 RLU and L517I+D67N was 10070 RLU, compared to HIV-1RTSubC wild type virus (106659 RLU).

2.3.2. Effect on drug susceptibility

2.3.2.1. Minimal reduction in drug susceptibility induced by T470S and T470P viral mutants

Figure 2.4 showed the phenotypic drug susceptibility analysis of T470S/P viral mutants. Drug susceptibilities were tested for d4T, AZT, ETR, EFV and NVP. The viral mutant with T470S had no effect on AZT (1-fold) and EFV (0.067-fold) susceptibility; however there was moderate resistance to NVP (10-fold), d4T (5-fold) and ETR (9-fold). Interestingly, in viruses with T470P, there was moderate resistance to AZT (5.09-fold) and NVP (6-fold); no effect was seen EFV, d4T and ETR (Figure 2.4).

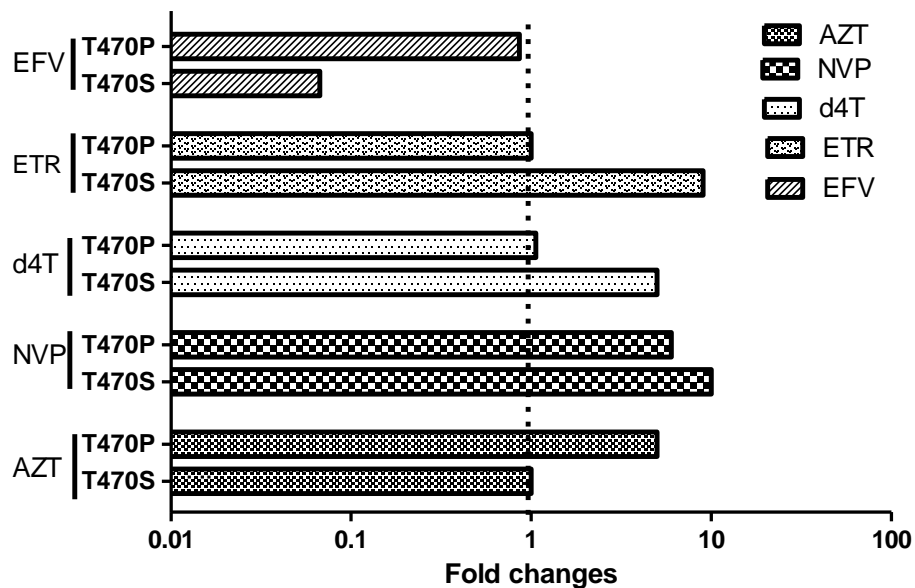


Figure 2.6. Phenotypic drug susceptibility profile from T470S and T470P subtype C viral mutants to the NRTIs: The fold changes for T470S viral mutant for AZT (1X), d4T (5X); and NNRTIs: NVP (10X), EFV (0.067X) and ETR

(9X). The The fold changes of the T470P viral mutants for AZT (5X), d4T (1.06X), and NNRTIs: NVP (6X), EFV (0.86X) and ETR (1X). The fold changes in IC₅₀ for virus mutants compared with WT are shown. The fold changes are shown on the x-axis and mutant viruses and drugs are shown on the y-axis, the bar graphs indicated the level of resistance for each drug.

2.3.2.2. D67N, L517I and the L517I+D67N double mutation decreases drug susceptibility to RTIs

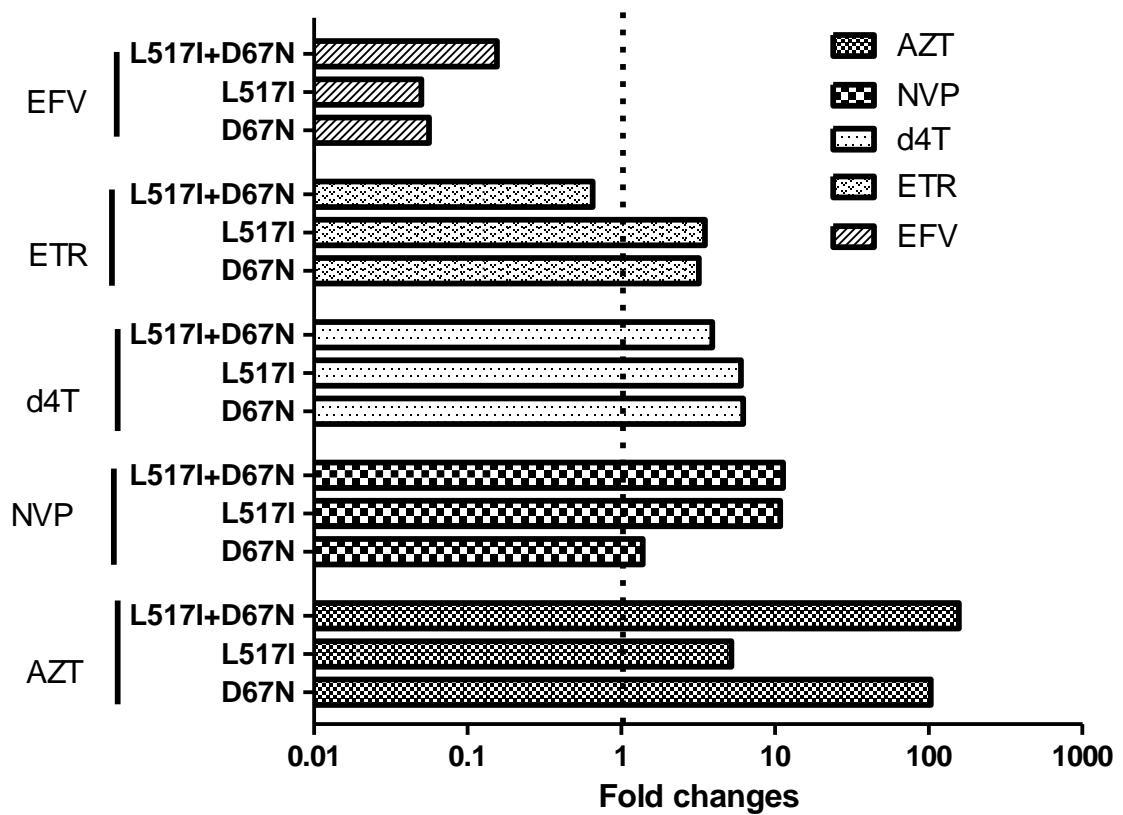


Figure 2.5. Phenotypic drug susceptibility profile from D67N, L517I and L517I+D67N subtype C viral mutants to the RTIs. The fold changes for D67N viral mutant for AZT (103X), d4T (6.2X); and NNRTIs: NVP (1.3X), EFV (0.056X) and ETR (3.2X). The fold changes of the L517I viral mutant for AZT (5.2X), d4T (6X), and NNRTIs: NVP (10.79X), EFV (0.05X) and ETR (3.5X). The fold changes of the L517I+D67N viral mutant for AZT (157X), d4T (3.9X), and NNRTIs: NVP (11.30X), EFV (0.155X) and ETR (0.65X) The fold changes in IC₅₀ for virus mutants compared with WT are shown. The fold changes are shown on the x-axis and mutant viruses and drugs are shown on the y-axis, the bar graphs indicated the level of resistance for each drug.

Figure 2.5 showed that the D67N mutation conferred low level resistance to ETR (3.29-fold change in EC_{50}), d4T (6.2-fold) and high level resistance to AZT (103-fold) but did not affect the susceptibility to NVP (1.03-fold) and EFV (0.056-fold) more sensitive. Viral mutants with the L517I RNase H mutation showed low level resistance to ETR (3.5-fold) and moderate resistance to NVP (10.79-fold), d4T (6-fold), AZT (5.2-fold) and EFV (0.05X) was more sensitive. Mutants with the L517I+D67N double mutant showed high level resistance to AZT (151.3-fold), but moderate resistance to d4T (3.9-fold). The resistance to NVP increase marginally to 11.3-fold; however there was more sensitive ETR (0.65-fold) and EFV (0.155-fold). The RNase H viral mutants exhibit greater phenotypic susceptibility to the NNRTI, EFV and ETR.

2.3.2.3. E529D has a minimal effect on drug susceptibility

Figure 2.6 shows that the subtype C E529D mutation had a minimal effect on drug susceptibility, conferring low level resistance to d4T (2-fold), AZT (3-fold) and NVP (3-fold). The subtype B E529D mutant gave a similar results, with low level resistance to d4T (5-fold), AZT (3-fold) and NVP (2-fold).

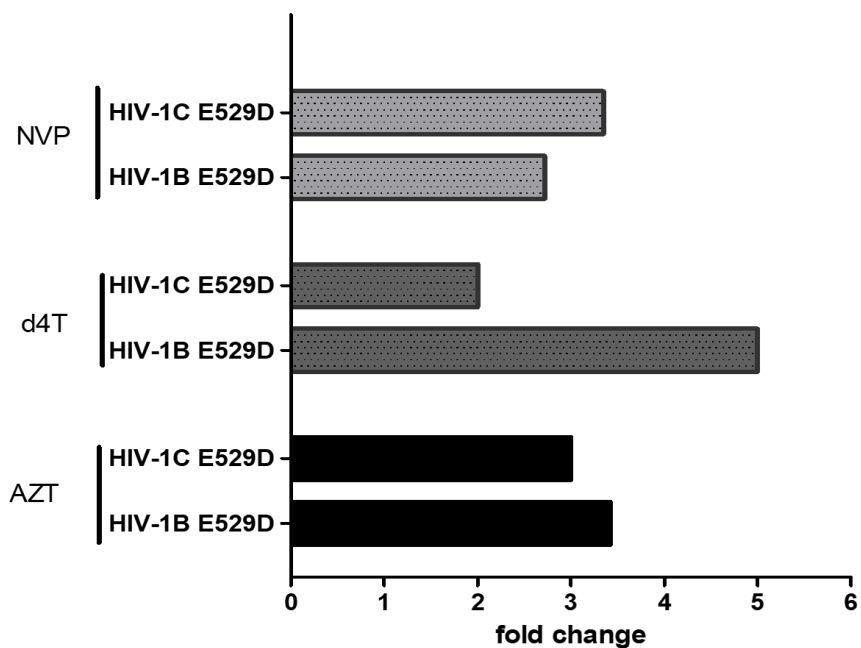


Figure 2.6. Phenotypic drug susceptibility profile from E529D subtype B and subtype C viral mutants. The fold changes for E529D subtype B viral mutant to the NRTIs: AZT (3-fold), d4T (5X); and NNRTIs: NVP (2-fold). The fold changes for E529D subtype C viral mutant to the NRTIs: AZT (3-fold), d4T (2X); and NNRTIs: NVP (3-fold). The fold changes in IC_{50} for virus mutants compared with WT are shown.

2.4. Discussion

The role and effects of HIV-1 subtype C RNase H domain mutations in contributing to HIV-1 treatment failure have not been fully characterized. It has been described, to some extent, in HIV-1 subtype B sequences from developed countries (Ehteshami and Gotte, 2008; Nikolenko et al., 2007; Nikolenko et al., 2004; Palaniappan et al., 1997; Roquebert et al., 2007; Santos et al., 2008). Our group previously identified several HIV-1 subtype C RNase H mutations that could be involved in drug resistance by comparing naïve and treated sequences (Ngcapu, 2012; Ngcapu et al., 2017b). They identified an increase in frequency in the treated group for E438GKR (6.72% vs 0%), T470S (39.3% vs 26.1%), L517I (14.3% vs 4.6%), K527GENQR (41.04% vs 26.12%), E529D (20.5% vs 2.3%) and Q547HKR (5.22% vs 0%). While the E529D mutation was the only mutation that remained significant ($p < 0.05$) after correction for multiple comparisons, direct links were also seen in their BN analysis connecting E529D and L517I with exposure to treatment. This suggested that these two mutations could possibly cause resistance on their own, without the presence of the well-known mutations in the polymerase domain, in particular the TAMs. In this study, we found that they decreased replication capacity and minimally affected drug susceptibility to both the NRTIs and NNRTIs.

The effect of T470S on replication capacity and drug resistance was compared to the already described T470P mutation (Roquebert and Marcelin, 2008; Roquebert et al., 2007b). Only T470S and not T470P was found in the Bayesian network by Ngcapu et al (2017), although it did not show a direct connection to treatment exposure. This study found that T470S showed a similar replication capacity to T470P, although their overall replication capacities were lower than the

wild type virus. This reduction in replication capacity could be as a result of impaired RNase H activity, although this was not tested in this study (Pingen et al., 2014).

The RNase H mutation T470P, a mutation known to be associated with NRTI drug resistance, decreased susceptibility to AZT by 5-fold. It was also shown that it affects NVP susceptibility (6-fold), which has not been previously shown. Although the T470S viral mutant showed no resistance to AZT, it caused minimal resistance to d4T (5-fold). It had a more marked reduction in NNRTI susceptibility, including NVP (10-fold) and ETR (9-fold). This is the first report that this variant is also associated with drug resistance. It is important to note that ETR is a next generation NNRTI that is meant to suppress viral replication in isolates that are already resistant to NVP (Azijn et al., 2010b) and could have implications for future therapy options for these patients. However, T470S was hyper susceptible to EFV. Efavirenz hypersusceptibility has also been associated with mutations at position 215, 208 and 118 (Haubrich et al., 2002; Whitcomb et al., 2002).

Codon 470 is close to the RNase H primer grip (codon 473 to 476; 501 and 505) and active site (D443, E473, D478, D549). Mutations in the primer grip and active site may reduce RNase H activity by indirectly modifying the interaction between the C-terminal of RT and its nucleic acid substrate (Julias et al., 2003; Nikolenko et al., 2005; Roquebert et al., 2007). It is possible that T470S has a different mechanism of drug resistance to the P variant, as it appears to affect NNRTI susceptibility more than the NRTIs. The NNRTIs bind to the NNRTI hydrophobic binding pocket located near the RT active site, suggesting that T470 indirectly changes the conformation of the binding pocket (Nikolenko et al., 2010; Nikolenko et al., 2005; Schultz and Champoux, 2008). Further analysis and understanding of the mechanism of resistance of both variants warrants further investigation.

Single mutants, D67N and L517I reduced the viral replication capacity in the absence of drugs when compared with wild type virus. The reduced replication capacity shown by D67N is consistent with a previous study that showed that this mutation comes at a fitness cost (Garcia-Lerma et al., 2004). The D67N+L517I double mutant was further reduce replication capacity when compared with single mutant suggesting that the L517I mutation may have the additive effect for the fitness loss of the D67N mutation. Other studies have suggested that additional mutations can modify viral fitness and also emerge under the selective pressure of drug treatment (Buckheit Jr, 2005). It was also previously suggested that some drug resistant viruses isolated during primary infection possess unique adaptive changes that allows for both increased viral replication capacity and resistance to one or more classes of antiretroviral drugs (Simon et al., 2003).

Previous studies have associated the D67N mutation with low level resistance to AZT (1-fold) and d4T (1.2-fold change) (Garcia-Lerma et al 2004) in the absence of additional TAMs. In the same study, the virus containing D67N and K219Q/E caused resistance to AZT (5.8-fold). In contrast, this study showed high level resistances to AZT (100-fold) and moderate resistance to d4T (10-fold) in the D67N single mutants. The presence of the L517I mutation increased resistance to AZT (157.3-fold), suggesting a role similar to that of the TAMs. This is in agreement with the study shown by Brehm et al (2007) showing that RNase H mutation (Q509L) causes a 1.3-fold change to AZT susceptibility; however when combined with the TAMs, AZT resistance increased up 50-fold. Like 470S, increased susceptibility to EFV was observed in D67N, L517I and L517I+D67N viral mutants.

The current study showed that the newly described E529D mutation conferred low-level resistance to AZT, d4T and NVP in both subtype B and C recombinant viruses (Barral et al., 2016). Previous

studies have suggested that fold changes above 2 can be regarded as resistance associated mutations, supporting the role of E529D in the development of drug resistance (Basson et al., 2015; Wu et al., 2012; Zhang et al., 2015b). Similar fold changes were observed for other RNase H mutations, D549N and H539N that conferred 2.4-fold and 4-fold resistance to d4T when in combination with thymidine analog mutations, respectively. These mutations also influenced resistance to NVP, which binds to RT with a lower affinity (Armstrong et al., 2011b; Nikolenko et al., 2005). While these mutations did not cause resistance to EFV, a second generation NNRTI, they did influence resistance to NVP, which binds to RT with a lower affinity (Armstrong et al., 2011a; Nikolenko et al., 2005). As only low-level resistance was observed, it is possible that this mutation acts in combination with other polymerase mutations and is not a primary resistance mutation. It is also important to point out that E529D was found in 21% of treated HIV-1 subtype C infected patients from Brazil (Barral et al., 2016) and in 20.5% of the South African subtype C treated isolates (Ngcapu et al., 2017a) thus the selection of E529D may depend on the balance between the level of drug resistance and the impact of the mutation on the viral fitness.

Taken together, this study was able to directly link the presence of RNase H mutations (E529D, L517I, T470S/P) with antiretroviral drug resistance. In this study we did not study the mechanism of how RNase H mutations caused mutations; however other studies that have investigated this have shown that RNase H mutations allow more time for NRTI excision (Nikolenko et al., 2005). As the mutations investigated in this study have largely been found to affect the NNRTIs, this suggests that RNase H mutations allow the NNRTIs to dissociate from the binding pocket allowing normal polymerization to continue. The way in which this occurs needs to be further investigated. In conclusion, the results from this study strongly support the inclusion of the C-terminal of the RT gene in genotyping tests; however larger association studies and more drug susceptibility assays are required to further investigate the effects of these mutations.

Chapter 3

Identification of connection domains mutations of the HIV-1 C and their association with thymidine analog mutations using a Bayesian Network

3.1. Introduction

As stated in the previous chapter, HIV-1 RT C-terminal domain mutations have been implicated in NRTI and NNRTI drug resistance (Brehm et al., 2007; Dau et al., 2010; Santos et al., 2008; Schuckmann et al., 2010). In 2008, Santos and colleagues identified connection domain mutations in 510 thumb and CN sequences from public sequence databases and correlated this with their treatment status. These mutations, R358K, G359S, A360V, K366R, A371V, K390R and A400T, were associated with AZT monotherapy failure. In another study, 345 treatment experienced patients enrolled in the OPTIMA trial treated with ABC, d4T, TNF or AZT harboured the following CN mutations: G333D/E, G335D, N348I, V365I, A371V, A376S (Dau et al., 2010). The G333D/E mutation has been found in patients treated with AZT-3TC dual therapy in subtype B (Kemp et al., 1998; Menendez-Arias et al., 2011). It has been established that G333D allows the enzyme to effectively discriminate between the normal substrate dCTP and the 3TC-TP inhibitor, restoring ATP-mediated excision of AZT-MP in the presence of M184V (Kemp et al., 1998; Larder and Kemp, 1989; Zelina et al., 2008).

The association of CN domain mutations with the TAMS has also been confirmed in vitro, where they showed that CN mutations (G335C/D, N348I, A360I/V, V365I and A376S), caused up to a

20-fold change to AZT susceptibility (Delviks-Frankenberry et al., 2008; Ehteshami et al., 2008a; Gupta et al., 2011; Hachiya et al., 2008a; Michels et al., 2010; Muniz et al., 2014; von Wyl et al., 2010; Yap et al., 2007). In addition, a previous study showed that double mutations in the CN domain (T369I and N348I) was associated with reduced ETR susceptibility (1.7X) in the presence of K101P or K103R/V179D; these mutations also reduced replication capacity (Gupta et al., 2011).

Some CN domain mutations have also been characterized in other non-B subtypes, although these studies are fewer than subtype B studies (Ehteshami et al., 2008a; Santos et al., 2011). The CN domain mutation, A400T was reported in CRF01_AE in patients failing AZT and harboring TAMS (Delviks-Frankenberry et al., 2009). Similarly, N348I and E399D were observed in a cohort of CRF01_AE patients failing d4T, 3TC and NVP (Delviks-Frankenberry et al., 2013; Saeng-aroon et al., 2010). Interestingly, known CN domain resistance mutations also occur as naturally occurring polymorphisms in non-B subtypes. In a HIV-1 subtype CRF02_AG infected treatment naïve cohort from Mali and West Africa, CN polymorphisms identified were: G335D (76%), A371V (63%), and E399D (11%) (Haidara et al., 2010; Maiga et al., 2012). Studies from subtype C have shown that both B and C CN mutants exhibit similar levels of AZT resistance; these mutations are Q334D, G335D, N348I, A371V, A376S and A400T (Delviks-Frankenberry et al., 2013; Muniz et al., 2014). It was also reported that HIV-1 subtype C CN mutations enhance resistance to other NRTI and NNRTI in the context of the TAMS (Delviks-Frankenberry et al., 2008).

Furthermore, a study from Southern Brazil showed that T369I and A376S, also known to cause resistance to the NNRTIs in subtype B, occurred in 10% of subtype C infected treatment naïve patients, suggesting that drug therapy is compromised in patients containing these mutations

(Santos et al., 2011). Evidence from *in vitro* experiments in non-B subtypes confirmed that N348I and A376S together with the TAMS caused NRTI resistance (Tanuma et al., 2010).

In the current study, the frequencies of HIV-1 connection domain in subtype C infected RTI treated patients were determined and the dependency of these mutations on exposure to RTI treatment was investigated using a BN.

3.2. Material and Methods

3.2.1. Identification of HIV-1 connection domain mutations.

One thousand and two HIV-1 RT subtype C sequences with connection and RNase H domains were retrieved from the Los Alamos (<http://hiv-web.lanl.gov>) and Stanford HIV drug resistance databases (<http://hivdb.stanford.edu>). Seven hundred and fifty three (753) of these sequences were from patients not on antiretroviral drugs therapy and 249 were on antiretroviral therapy. Duplicate sequences were checked using the Elimidupes tool at the Los Alamos database, and sequences were aligned by codon alignment (ClustalW).

3.2.2. Frequency of the HIV-1 RT connection domain mutations.

Naïve and treated sequences were analysed using the REGADB sequence analysis tools (<http://regatools.med.kuleuven.be/sequencetool/sequencetool.wt>). The frequency of the connection domain mutations was calculated in excel. Statistically significant differences between the frequencies of mutations in the naïve versus treated group were calculated using the Fischer'exact test in GraphPad Prism v5.1 and adjusted for multiple comparisons using the Bonferroni test. A p-value of <0.05 was considered significant.

3.2.3. Bayesian Network analysis

The BN was constructed using b-course (b-course.hiit.fi/obc/). Briefly, the dataset was saved as a text delimited file and uploaded. The dataset was analysed using the Dependency model; this model is used to find the dependencies between all variables in the data. Due to the large amount of computational power required for the analysis, the number of variable that could be analysed was limited. Therefore only CN mutations that were statistically significantly higher in the treated group, or showed a >5% difference between the naïve and treated patients, were included in the analysis. The association between the TAMS and CN domain mutations were investigated in the network.

3.3. Results

3.3.1. Characterization and identification of drug resistance-associated mutations in HIV-1 RT connection domain and their correlation with thymidine analogue mutations.

As shown in Figure 3.2, the following mutations had a higher frequency in sequences from treated vs naive isolates: D324E, S338T, I341V, N348I, M357R, T359N/S/H/G/A, A360P/S/T, K366R, E370D/K, A376 I/S/T/V, M377V/RQK/M/L/I, V380I,T400A/I/M/R/V, I434L/M and A437I/V. These did not remain significant after Bonferroni correction. Table 3.1 shows the percentage differences between the naïve and treated patients including the p-values. Mutations were considered for the BN analysis if the difference was $\geq 5\%$, or were known to be associated with resistance.

Table 3.1. Frequency (%) of CN mutations domain from HIV-1 subtype C infected patients (250) receiving antiretroviral therapy. The following mutations were identified D324G/N/P, T338S, N348I, M357R, G359N/S/H/G/A, A360P/S/T, K366R, E370D/K, M377A/G/I, V380I and T400A.

Mutations in the CN domain	Frequency of Subtype C naïve patients(n=749)	frequency of subtypeC treated patients (n=249)	%differences	p value
D324 E	11,44	29,7	18,26	P > 0.05
S338 T	3,170	6,820	3,650	P > 0.05
I341 V	1,190	2,400	1,210	P > 0.05
N348I	0,0000	2,400	2,400	P > 0.05
M357R	1,170	8,830	7,660	P > 0.05
T359 N/S/H/G/A	7,820	18,86	11,04	P > 0.05
A360P/S/T	3,840	5,610	1,770	P > 0.05
K366R	6,838	12,22	5,382	P > 0.05
E370D/K	2,380	4,400	2,020	P > 0.05
A376 I/S/T/V	11,10	14,82	3,719	P > 0.05
M377V/RQK/M/L/I	17,97	34,93	16,96	P > 0.05
V380I	0,4000	2,400	2,000	P > 0.05
T400A/I/M/R/V	2,991	11,40	8,409	P > 0.05
I434L/M	1,300	5,200	3,900	P > 0.05
A437I/V	2,510	44,80	42,29	P > 0.05

The associations between the presence of natural polymorphism, the emergence of CN mutations and HIV-1 treatment experience were explored using the BN. Figure 3.3 shows that some of the known CN mutations did not show a relationship with treatment experience (eRTI) although a direct robust interaction was observed between treatment experience and some of the TAMS: K70E/R and M41L. The E370D, M377T/L, A437V/I and A400T mutations were the only CN mutations to be directly associated with treatment experience. N348I was only connected to treatment exposure via the M41L TAM. K70E showed a strong association with T359N and A437V/I. Furthermore, there was a moderate association (bootstrap support of 60%) between A437V/I and the following mutations: M357R, M377T/L and I/I434M. Similarly, 65% bootstrap support was found for as M357R and T400A K366R and M41L as well A376S and T215F.

The HIV-1 subtype C connection domain sequences from different regions were aligned to discriminate the wild type and drug resistance mutations. Figure 3.1 show the consensus sequences of HIV-1 CN domain from different regions where HIV-1 subtype C is dominant. From the HIV-1 CN mutations analysed: D324E some in India have it as wild type. H334Q is the wild type wild in India and I341V is wild type in India and Zimbabwe. In Ethiopia, M357R is wild type where as in Brazil is reported as drug resistant mutation. In Botswana, A376S/T is in wild type while in other subtype C are drug resistance mutations. In addition, in other Brazilians K366R and V380I are seen in treated naïve individuals.

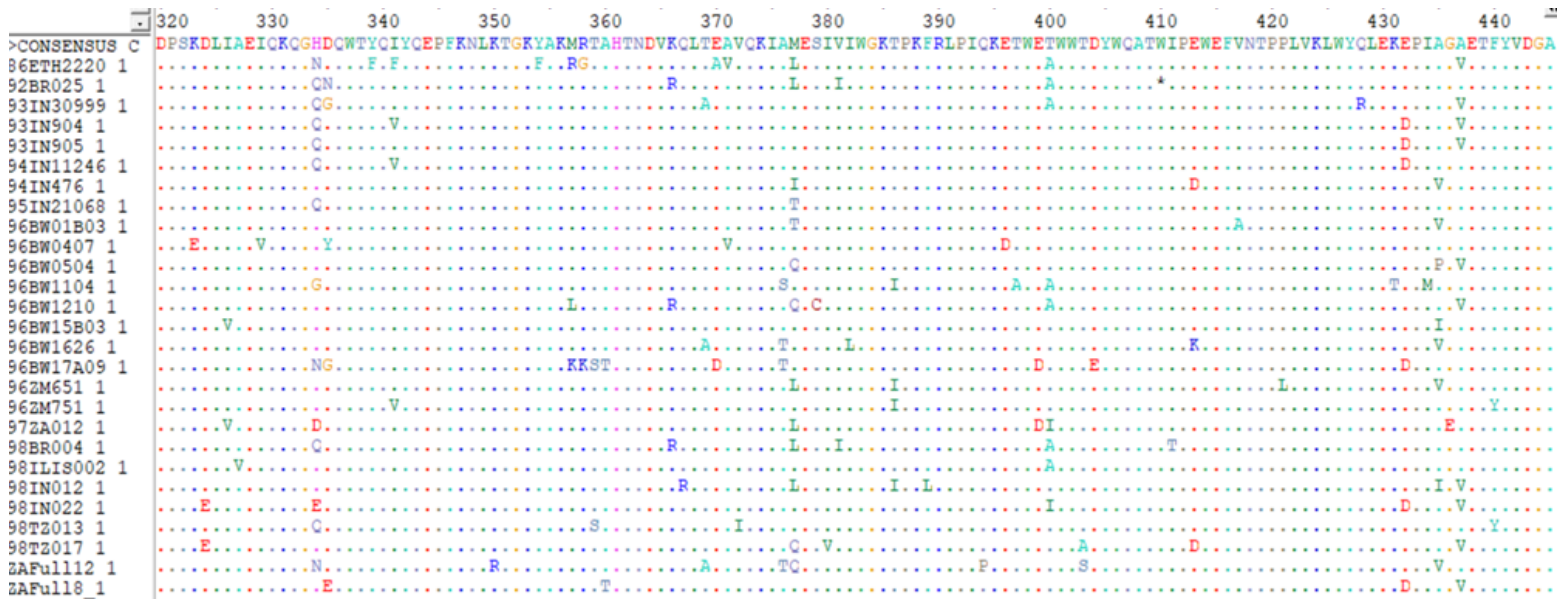


Figure 3.1. Comparison of HIV_1 RT protein sequences of subtype C from different regions. There were sequences from Brazil (2), Botswana (5), India (5), Ethiopia (1), Israel (1), South Africa (5), Tanzania (4), Zambia (2) and Kenya (1) were collected from Los Alamos HIV database.

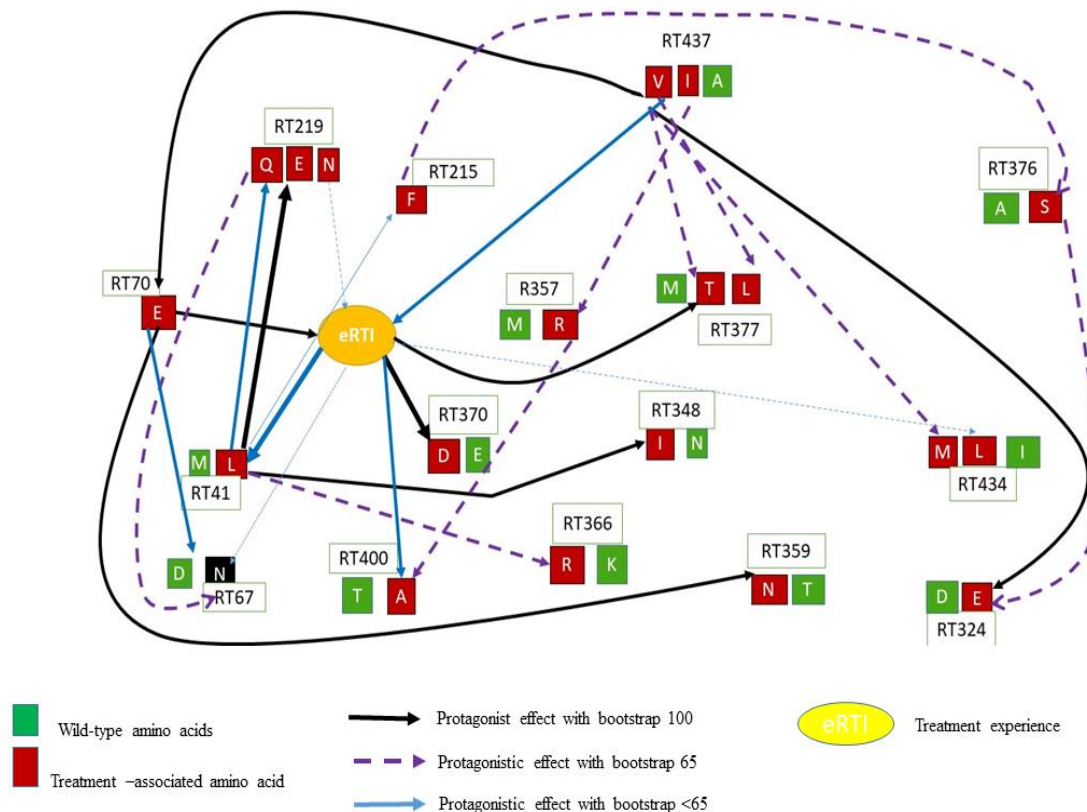


Figure 3.2. Annotated Bayesian Network (BN) showing dependencies in emergence of resistance mutations, based on 500 bootstrap showing graphically the learned associations between nodes denoting amino acids of treatment experienced. An arrow represents a direct dependence.

3.4. Discussion

The role and the effects of CN domain mutations in contributing to HIV-1 treatment failure have not been fully investigated, particularly in the context of subtype C. Although a recent study has been published including HIV-1 subtype C data from Brazil (Barral et al., 2016; Delviks-Frankenberry et al., 2013). Most available data that describes CN domain mutations are from subtype B sequences from developed countries, and are largely located between codon 320-400 in RT; the remaining portion has not been well characterized (Lengruber et al., 2011). In this study, treatment related mutations were identified by comparing downloaded HIV-1 CN domain

mutations in subtype C RTI treatment experienced sequences to those from subtype C treatment naïve sequences.

In the CN domain, the following mutations were increased in the treated group: D324E, S338T, I341V, N348I, M357R; T359N, A360PST, K366R, E370DK, A376S, M377TL, V380I, T400AMRV, I434LM and A437IV. Many of these mutations have already been described in HIV-1 subtype B, and to a lesser extent in non-B subtypes (Saeng-aroon et al., 2010; Santos et al., 2008; Santos et al., 2011 ; Tanuma et al., 2010). M357R, A360T/V and K366R were identified in subtype B isolates from Brazil (Barral et al., 2016; Delviks-Frankenberry et al., 2010; Santos et al., 2008). N348I and A376S were identified in subtype B isolates from the OPTIMA study (Brehm et al., 2012a; Dau et al., 2010; Paredes et al., 2011). While T400AMRV has been reported to contribute to AZT resistance in HIV-1 subtype CRF_AE, this study found to occurred in 11.40 % of the naïve and treated groups (Delviks-Frankenberry et al., 2013; Delviks-Frankenberry et al., 2009). Therefore those that have not been previously described to our knowledge are: D324E, S338T, I341V, T359N, E370DK, M377TL, V380I, I434LM and A437IV.

Interestingly, other known mutations such as G335D, Q334D, A360V/I, V365I and A371V, were not associated with treatment in this study (Delviks-Frankenberry et al., 2013; Radzio et al., 2010; Santos et al., 2008). Differences seen with the Brazilian study could be because they obtained their subtype C sequences from isolates from Brazil, as well as sequences from the Los Alamos database (Barral et al., 2016). Brazillian (and Indian) subtype C isolates are known to be different to subtype C isolates from other parts of the world, especially Sub-saharan Africa (Bello et al., 2008; de Oliveira et al., 2010; Novitsky et al., 2010).

The resistance pathways to NRTI were further investigated using BN learning (b-course) where the presence of CN domain mutations were directly or indirectly associated with NRTI treatment. The BN analysis show that there was a strong association between E370D, M377T, and A437I (with a lower bootstrap support) with RTI treatment. This suggests that these mutations could cause resistance without concomitant mutations in the polymerase domain. Similar results have been shown for mutations in RNase H (Ngcapu et al, 2107). Their role would still need to be confirmed using *in vitro* assays.

In the network, A437I was also associated with other CN mutations, some of which resulted in a “dead end”, such as D324E and I434M. Others appeared to go via a pathway of mutations to resistance, such as via M377T/L before the link to treatment exposure and via M357R to T400A before linking to treatment exposure. The M41L and K70E mutations were the only TAMs directly associated with HIV-1 treatment in the network. The other TAMs were connected via various pathways to treatment exposure: eRT> K70E>K219E> D67N and eRT>M41L>T215F). There were also associations between the TAMs and CN mutations: M41L and K366R; T215F and A376S; K70E and A437I and M357R; K70E and M357R. This suggests that CN mutations are also involved in the stepwise accumulation of mutations in the development of resistance to the thymidine analogues. None of these interactions have been described before and requires further investigation, including phenotypic studies.

Previous studies have reported that N348I acquisition in HIV-1 RT occurs early in therapy before the acquisition of TAMs, allowing the virus to select both TAMs and other mutations that are antagonistic to toward TAMs (Radzio et al., 2010). In contrast, the current Bayesian network analysis showed that N348I is not directly associated with treatment exposure in subtype C, but is connected via the M41L TAM. A previous study by von Wyl et al (2010) reported a pathway

between M184V and N348I which was restricted to exposure to both 3TC and AZT. The study speculated that the emergence of selected mutations follows a hierarchical order after drug: L74V>M41L>A371V>N348I (von Wyl et al., 2010). Presently, there is no study that has further investigated the interaction of M41L and N348I; however there are studies that have investigated the effect of TAMS and A371V in subtype B (Brehm et al., 2007; Nikolenko et al., 2007), where they found that the level of resistance to AZT increases from 11-fold to as much as 536-fold over wild type RT. It has been reported that N348I CN domain mutation reduce replication capacity in subtype B (Gupta et al., 2011), and increases resistance to AZT through both RNase H dependent and independent mechanisms.

In conclusion, this study identified novel HIV-1 subtype C connection domain mutations that are strongly associated with the TAMs. A limitation of the study is that the number of variables that could be analyzed simultaneously were limited and therefore some interactions may have been missed.

Chapter 4

The effects of HIV-1 C RNase H domain mutations on reverse transcriptase structure

4.1. Introduction

Biochemical studies have demonstrated that C-terminal domain mutations in HIV-1 RT indirectly increases AZT resistance by decreasing the secondary RNase H cleavages. In addition previous studies reported that mutations in the RNase H domain affect the initiation of DNA synthesis and specificity of RNase H cleavage demonstrating the interdependence of polymerase and RNase H activities during reverse transcription (Cristofaro et al., 2002; Julias et al., 2002; Mizrahi et al., 1994). It was previously reported using molecular modelling studies, that N348I decreases RNase H activity via an altered interaction with the RNA template (Radzio and Sluis-Cremer, 2011). In addition, it has been shown that any alteration in the structure of the RT enzyme may affect the binding affinity of the substrate, resulting in incorrect reverse transcription (Das and Arnold, 2013b).

The knowledge of the 3D structure of proteins, therefore, provides invaluable insights into the molecular basis of their functions. Previously, the 3D structure of a protein had to be determined experimentally, and this was done using X-ray crystallography or NMR spectroscopy; however these methods are very time consuming (Floudas et al., 2006). Homology modelling is therefore a useful tool for identifying where the alpha carbons of the key residues sit in the folded protein (Martz, 2001) and can provide information on the structure-function relationship of the protein (Vyas et al., 2012). Here, we investigate the structural effects caused by RNase H domain

mutations by comparing molecular models of the HIV-1 RT wild type and the mutants and how the RNase H active site affected.

4.2. Material and Methods

4.2.1. 3D Homology Modelling using Swissmodel

To investigate the structural changes induced by RNase H domain mutations, 3D homology models of the HIV-1 subtype C mutants were constructed. Sequence from a single patient (pKE4N) from King Edward Hospital was used to introduce the RNase H mutations (T470/P, L517I and E529D). Amino acid sequences were uploaded onto the SWISS-MODEL server (<http://swissmodel.expasy.org/>). A template of HIV-1 group M subtype B (3lam.1.B.pdb) was searched, the best matched query sequence with the highest percentage identity was used to build models. The model of the RNase H wildtype C sequence and with E529D as well were compared. The template was selected and the built-in algorithm on the SWISS MODEL server was used. Structures were viewed and analysed in Chimera v1.12 as described below (Pettersen et al., 2004).

4.2.2. Visualization using Chimera (v1.8.1)

Modelled structures of the wild type and mutant sequences containing the investigated mutations (T470S, T470P, L517I and E529D) were superimposed in Chimera. Firstly, the atomic structure of the mutated residues from the RNase H mutations model were visualized and compared to the unmutated model or wild type. Secondly, to understand the conformational changes that occurs between the RNase H mutations residues, the distance between the residues of each protein chain was measured. The distance between residues in the direct vicinity ($>10\text{\AA}$) of RNase H mutation were measured for all the mutations. The flexibility of the protein sidechains in the wildtype and

mutant structures were evaluated, as well as the affect on the neighbouring known RNase H mutations and RNase H active site.

4.3. Results

In figure 4.1, the three main domains of reverse transcriptase are shown: polymerase domain (blue), connection domain (green) and RNase H (cyan). The two subunits of the HIV-1 RT (p66 and p51) are also shown. The wild type residues where the RNase H domain mutations occurred were shown in black.

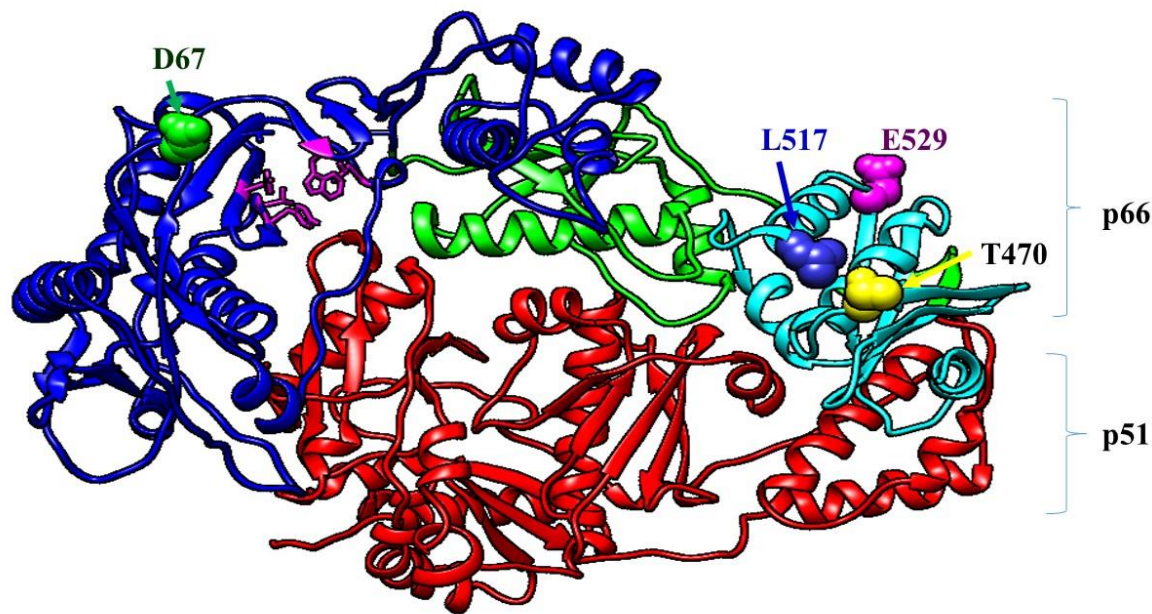


Figure 4.1. The structural model of HIV-1 RT (3lam.1.B.pdb) showing the two subunits (p66 and p51). Red represents the p51 and p66 was shown: polymerase domain (blue), connection domain (green), and RNase H (cyan). The amino acid residues where mutations occurred were shown, D67N (green), T470S/P (yellow), L517 (blue), E529D (Magenta).

4.3.1. Structural changes caused by the T470S and T470P mutations

Figure 4.2 shows the superimposed structures of the modelled mutant and wild type RNase H domain sequences. There was a slight change in T470S when compared with the wild type T470 causing a slight movement in the neighbouring amino acid residues. No significant changes identified in the presence of the T470S and T470P except the change observed in K451.

A

B

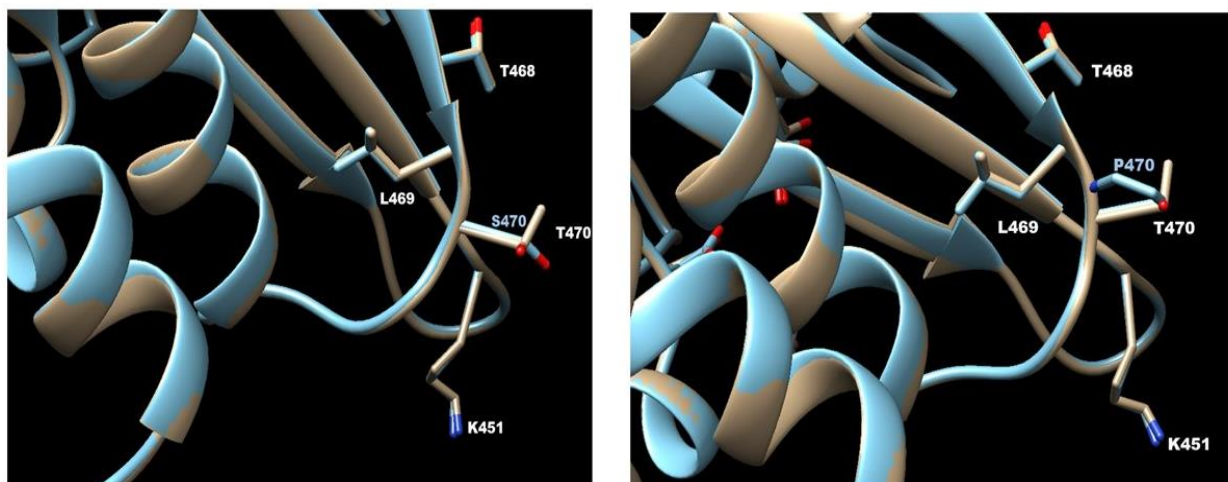


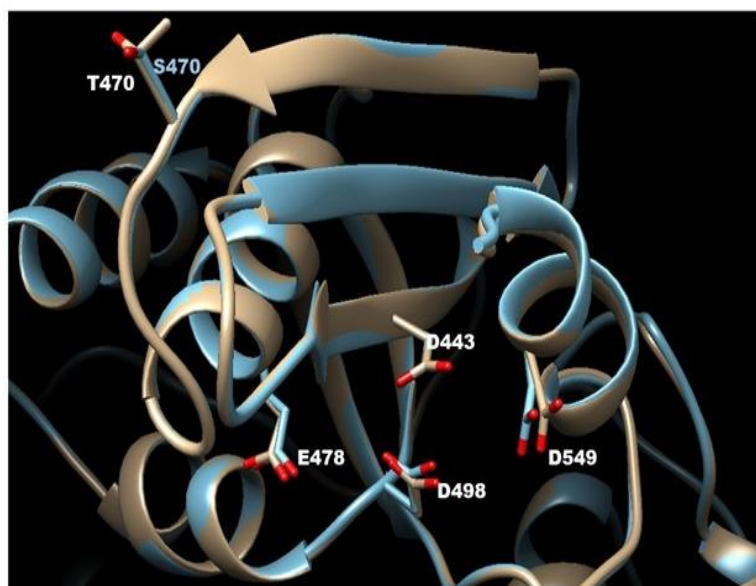
Figure 4.2. The superimposed structure of RNase H domain showing mutations T470S and T470P. The wild type structure is shown in brown and mutant structure shown in blue. A) T470S and B) T470P are shown. The amino acid residues located near these mutations were K451, T468 and L469 (at the distance less than 10 Å).

No structural changes between the wild type (T470) and mutant amino acid residue (T470P) observed (Figure 4.2B). The amino acid residues located near the T470S and T470P in a distance less than 10 Å were identified (Table 4.1). These were K451, L452, T468, L469, D471 and T472; some of these amino acid residues were known HIV-1 subtype B (K451) and C (T468, L469) RNase H domain mutations. The distances between the RNase H active site and the T470S/P mutations were measured (Table 4.2). These mutations were located far away from the RNase H active site.

Table 4.1. Distances between the RNase H mutations (T470S/P) and the nearby amino acids residues.

Zones at the distance of >5%	470T (wild type) Å	470S Distance in angstrom (Å)	470P (Distances in angstrom Å)	Differences in distances Å(470S)	Differences in distances Å (470P)
K451	7.580	7.649	10.187	-0.069	-2.607
L452	6.861	6.820	6.085	0.041	0.776
T468	7.422	6.837	6.422	0.585	0.585
L469	5.730	5.801	6.739	-0.071	-0.071
D471	5.188	5.019	7.259	0.169	-2.071
T472	6.729	6.802	6.838	-0.072	-0.109

A



B

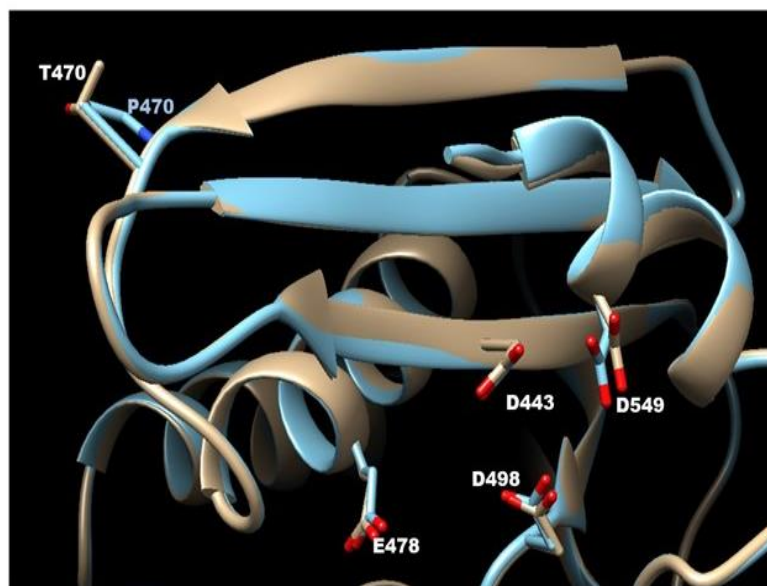


Figure 4.3. The superimposed structures of RNase H domain showing T470S and T470P mutations and RNase H active sites amino acids. The wild type structure is shown in brown and the mutant structure shown in blue. A. Showing the effect caused by T470S on the RNase H active site amino acids, D443, E478, D498 and D549 and B. show the effect caused by the T470P on the RNase H active site residues.

Table 4.2. Distances between the RNase H active site amino acid residues and the RNase H domain mutations (T470S/P)

RNase H active sites	470 T (Distances in Å)	470S (Distances in Å)	470P (Distances in Å)	Differences between active sites residues and 470S	Differences between active sites residues and 470P
D443	18.267	18.224	18.146	0.043	0.121
E478	17.146	16.3224	16.195	0.8236	0.951
D498	23.009	22.671	21.373	0.338	1.636
D549	21.151	20.967	21.448	0.184	-0.297

The changes in distance between the T470S and D443 was slightly decreased by 0.043 Å, while the distance between T470P and D443 further decreased by 0.121 Å. The changes in distances between these amino acids were not significant. The changes in distance between the T470S and E478 RNase H active site amino acid was 0.8236 Å while changes in distance in T470P and E478 was 0.951. The decrease in distance was pulling the binding pocket inward hence decreasing the active site. The RNase H active site amino acids D498 and D549 was located far away. The RNase H active site was compressed by the presence of RNase H domain mutations. The conformational changes in the active sites that affects with small movements which may affect resulting in improper binding of substrate.

4.3.2. Structural changes caused by the L517I and D67N double mutation

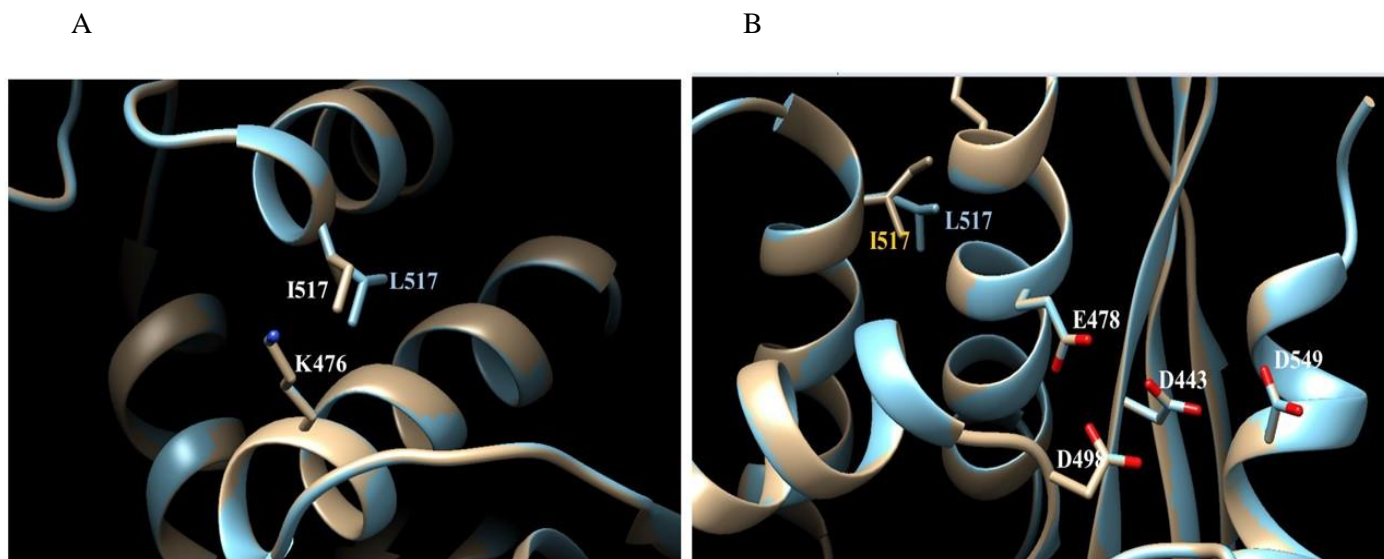


Figure 4.4. The superimposed structure of of HIV-1 RNase H domain showing the L517I mutation. A) L517 is shown in blue and I517 shown in brown, the neighbouring amino acid was also shown in brown with blue side chain. K476 forms part of RNase H primer grip located near the L517I and it was identified in a distance less 10 Å B) L517 shown in blue and I517 in brown, Rnase H active site amino acid residues were also shown. No changes observed in the RNase H active site amino acid in the presence of L517I.

Table 4.3. The distances between the TAM (D67N) and RNase H mutation (L517I).

Atom 1	Atom 2	Distances (Å)
L517	N67	63.067
L517	D67	66.220
I517	N67	64.782
I517	D67	66.162

Figure 4.4 (A and B) showed the amino acid residues located near the L517I in less 10 Å distance. The amino acids residues located around the L517I mutation were identified. These amino acids were T472, K476, L479, Q480, T483, S515, E516, V518, N519, Q520, I521 and I522. There were no known mutations identified around L517I in both HIV-1 subtype B and C. However it was located near the RNase H primer grip (K476) in a distance of 4.470Å. Table 4.3 showed the distances between the D67N and L517I which illustrating the structural effects caused by the presence of both polymerase and RNase H domain mutations. The distances between the 67N (mutant) and the L517 (wild type) was 63.093Å and distance between the D67 (wild type) and 517I (mutant) was 66.201Å. This suggests that polymerase domain and RNase H domain mutations may have a long range effect on the structure of RT. In addition, the distance between the I517 (mutant) and N67 (mutant) reduced to 64.782Å which pulls the RT structure inward. The distance between the I517 (mutant) and D67 (wild type) was 66.162 Å, are located far away from each other. This results suggest that D67N L517I mutations could affect the structure of RT by influencing the RT affinity.

4.3.3. Structural changes caused by the E529D mutation

In figure 4.5, the RNase H domain mutation (depicted in brown) and the wild type (depicted in blue) were shown. The D529 amino acid residue shifts the E529 slightly away from its original position. The amino acid residues located near E529D mutation were identified in a distance less than 10 Å. It was identified that E529D mutation was located near E430, L491, E492, K527, K528, K530 and V531 (Table 4.3) (Fig 4.5 A). The E529D mutation was identified to be located near the identified HIV-1 subtype B and C RNase H mutations L491P (4.160 Å, K527N (8.750Å), K530R (4.843Å) (Ngcapu et al., 2017a). One hydrogen bond was identified between the L491 and D529 and the distance is 3.011Å. In Figure 4.5 B show the E529D mutation and RNase H active

site amino acid residues, however no changes observed in these amino acid. This suggested that E529D mutation has no effect on the position of the RNase H active site amino acid residues.

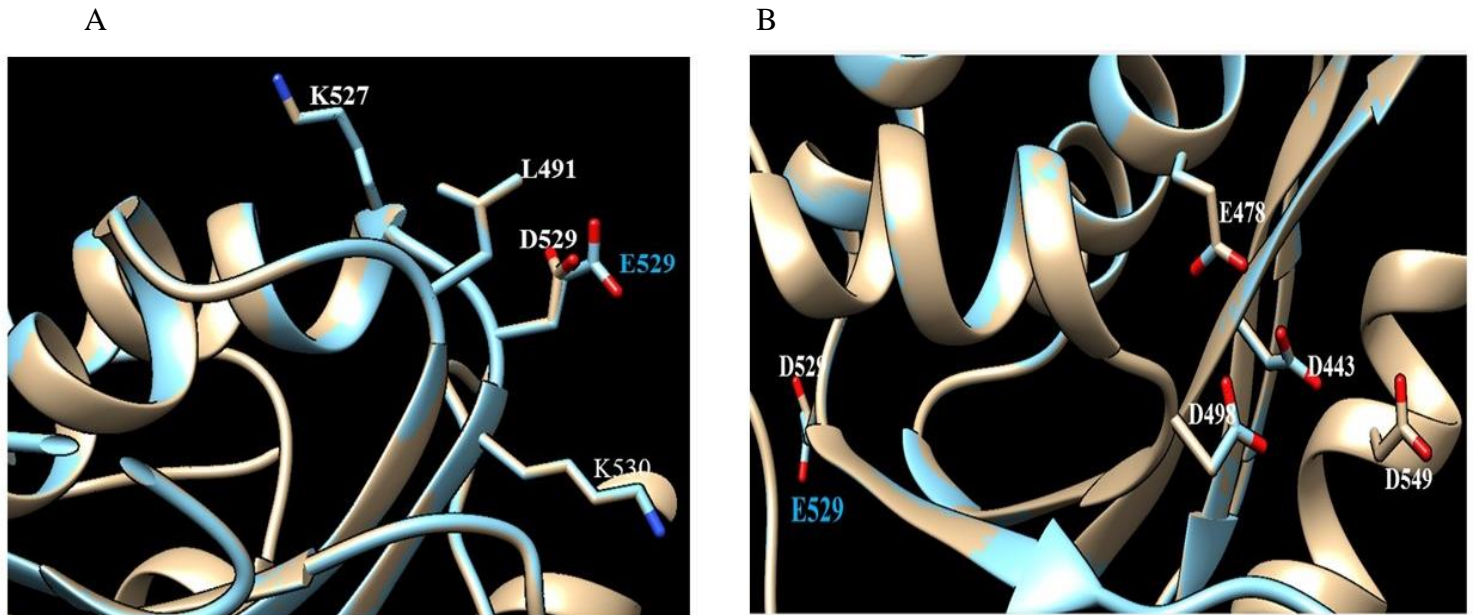


Figure 4.5. The superimposed structure of RNase H domain showing E529D mutation. A. E529 is the wild type amino acid residues depicted in blue and D529 is the mutant amino acid depicted in brown, L491, K527 and K530 are amino acids located near E529D. B. RNase H active site amino acid residues, D443, E478, D498, D549 were shown.

Table 4.4. Distances between the nearby amino acids residues and E529D RNase H domain mutation.

Zones at the distance of >5%	E529 (Distances in Å)	D529 (Distances in Å)	Differences between the E529 and D529
E430	10.089	9.802	0.287
L491	4.830	4.160	0.67
E492	6.713	5.569	1.144
K527	10.756	8.750	2.006
K528	8.201	5.520	2.681
K530	7.669	4.843	2.826
V531	8.645	8.529	0.116

4.4. Discussion

The structures represent structural effects caused by the RNase H domain mutations on the RNase H active site. The information obtained from these structural analysis will provide a basis on how RNase H domain mutations affects structure and critical function of reverse transcriptase. RNase H domain mutations were located far away from the RNase H active site residues, however the distances between the RNase H domain mutations and the RNase H active sites were changed. The changes in distances in the RNase H domain mutations and RNase H active site residues reduces the binding pocket of the substrate suggesting that these mutations affect the RNase H activity. However, these mutations affect position of the RNase active site amino acid residues, suggesting that RNase H domain mutations may have a long range effect which will affect the RNase H activity. Previous studies reported that the changes in the orientation of the active site may affect the binding of RNase H to the template, thereby affecting RNase H activity (Das et al., 2008; Davis et al., 2011). Furthermore previous studies have reported that despite the NNRTI binding site being located on the palm domain of the HIV-1 RT polymerase domain approximately 10-15Å from the polymerase active site, C- terminal domain mutations increase resistance to both NRTIs and NNRTIs (Delviks-Frankenberry et al., 2008; Delviks-Frankenberry et al., 2010; Nikolenko et al., 2004). It has been previously shown that RNase H domain mutations cause resistance by reducing template degradation and enhance AZT excision and increase template switching frequency. The mutation in the polymerase domain of the RT may interfere with RNase H activity through the long range effect, influencing the structure of the RNA:DNA hybrid substrate (Hang et al., 2007). This study showed that L517I together with D67N mutation may affect the structure of the RT which may affect the positions of the amino acids in the polymerase active site and NNRTI binding pocket.

The current study showed that the T470S/P RNase H domain mutations affect the RNase H active sites by reducing its size which affect RNase H activity, however further experimentation is required to investigate this speculation. Further site directed mutagenesis studies are required to confirm this speculation. These results are in agreement with previous study shown that AZT drug resistance mutations in HIV-1 RT can induce long range conformational changes (Ren et al., 1998). Even though the distances between the RNase H mutations and RNase H active sites were far apart from each, it may still cause conformational changes to RT. These results show that long range effect could represent a common mechanism for generating drug resistance in other systems.

The study indicates that the RNase H domain mutations may cause HIV-1 drug resistance in combination with polymerase domain and the mutations that occurred near the RNase H catalytic site. We investigated this possibility by identifying all the amino acid residues around the RNase H domain mutations. The structural analysis indicated that T470S was located closely to the published HIV-1 subtype B and C RNase H domain mutations K451, T468 and L469, suggesting that T470S may occur concurrently with these mutations. Previous studies showed that K451, T468, L469 are primary RNase H domain mutations (Lengruber et al., 2011; Ngcapu et al., 2017a; Roquebert et al., 2007). However these amino acid residues were also observed near T470P mutation more further apart compared to the T470S, the distances were located in less than 10 Å. In addition T470S/P mutations were located far away from the RNase H active site, however the superimposed images indicated that there is a shift in the RNase H active site of mutant. These findings suggest that these RNase H domain mutations (T470S/P) conferred HIV-1 drug resistance in combination with the previous identified mutations from other HIV-1 subtypes. Further functional studies are required to confirm these structural findings.

The structural effect of L517I mutation was located far away from the DNA polymerase region and NNRTI binding pocket. Therefore the amino acids residues located in a distance less than 10Å were identified, analysing any known mutations that may occur concurrently with RNase H domain mutations. Among these amino acid residues not known mutations were observed near L517I mutation. However, L517I mutation was located near the K476 which form part in RNase H primer grip. Previous studies reported that mutations within the RNase H primer grip region of the HIV-1 RT affect the RNase H activity and RNase H primer grip mutations enhances AZT resistance (Delviks-Frankenberry et al., 2007; Palaniappan et al., 1997). In addition, mutations in the RNase H primer grip decrease efficiency and accuracy of plus strand DNA strand, hence the RNase H primer grip alteration modifies cleavage specificity (Mbisa et al., 2005; Rausch et al., 2002). Furthermore, HIV-1 RNase H and DNA polymerase activities depend on proper positioning of the RT on the primer template. Hence the alteration within RT results in improper primer positioning causing the failure of the RT to perform DNA synthesis or RNase cleavage (Mbisa et al., 2005; Palaniappan et al., 1997). In addition, RNase H primer grip mutations may influence both polymerase and RNase H activities by changing RT affinity for template primer complex, and also they may affect the enzyme processing (Delviks-Frankenberry et al., 2007). These results shown that L517I may reduce the RT affinity for the template-primer complex since it occurred near the amino acid that form RNase H primer grip. This study may suggest that that L517I and K476 may affect the RNase H activity. More phenotypic studies may be required to investigate combination of L517I and K476 which form part of the RNase primer grip amino acid residue.

In conclusion, the T470S and T470P mutations alter the orientation of the RNase H active site amino acids which may affect the RNase H activity, however further experimentation are required to investigate these findings. RNase H domain mutations may also affect the functioning of RTI by altering the polymerase domain active site (Das and Arnold, 2013a; Das et al., 2016). New classes of HIV-1 drugs that target RNase H domain may be required in HIV-1 treatment in combination with the RT inhibitors. At present, there are no licenced RNase H inhibitors. The understanding of the structural changes caused by drug resistance mutations provide information on improve the HIV drugs on clinical management of HIV-1.

Chapter 5

General discussion and conclusion

In the study by Ngcapu et al (2017), their BN analysis suggested that certain mutations in the RNase H domain were related to drug resistance in HIV-1 subtype C. Of particular interest were the T470P/S, L517I and E529D mutations. The T470P mutation was already known as a resistance associated mutation in HIV-1 subtype B (Roquebert et al., 2007). However, it was the first time that the T470S mutation was described to be involved in RTI drug resistance. In addition, the E529D and L517I mutations were identified as potential “primary” drug resistance mutations. In this study, the effect of these mutations on replication capacity and drug susceptibility were investigated in order to confirm their role in ARV drug resistance. The study was further extended by identifying novel connection domain mutations and investigating their interactions using a BN.

All RNase H mutations lowered the replication capacity when compared to the wildtype subtype C virus, suggesting that that these mutations come at a fitness cost, and are only beneficial when ARVs are present. These findings are in agreement with previous studies that have shown that resistance mutations are often associated with decreased replication capacity in the absence of drugs (Asahchop et al., 2011; Back et al., 1996; De Luca, 2006; Deeks et al., 2005; Hu and Kuritzkes, 2011; Xu et al., 2014a). A decrease in RC has also been associated with decreased RT processivity (Chunduri et al., 2011; Nicastri et al., 2003).

In this study, the effect of L517I lowered RC further when in combination with D67N, therefore, suggesting that L517I mutation may have additive effect for the fitness loss of D67N viral mutant in the absence of drugs. Previous studies showed that the reduction in replication capacity from polymerase mutation together with the RNase H mutation was associated with significantly greater AZT resistance (Brehm et al., 2007; Nikolenko et al., 2005). The authors have argued that the additional mutations serve to impair the viral replication capacity emerge under the selective pressure of drug treatment (Buckheit Jr, 2005; Joly et al., 2004).

Our study clearly showed that T470S/P, L517I and E529D caused a reduction in drug susceptibility. Although this effect was minimal, with a range in fold change from 2X to 157.3X, there have been other studies that have shown a similar effect *in vivo* (Little et al., 1999; Sutherland et al., 2015). These studies demonstrated reduced ARV susceptibility of more than a 2.5 to 10 fold change. In a recent study, mutations in Gag reduced drug susceptibility to PIs by 4.7 to 9.6 fold (Sutherland et al., 2015).

It is also possible that the minimal effect seen is because the RNase H mutations has the reduced replication capacity and have a greater effect on drug susceptibility when in combination with mutations in other regions of the virus, such as the connection and polymerase domain. Previous studies have shown that RNase H mutations in combination with the TAMs cause a marked decrease in drug susceptibility (Brehm et al., 2007; Nikolenko et al., 2005). Nikolenko et al (2005) showed a 1250-fold increase in resistance to AZT when D549N occurred in combination with TAMs (Nikolenko et al., 2005). In this study, the double mutant showed a 157.30-fold, compared to a 103.3 fold change in D67N alone, and supports the role for L517I as having the additive effect

on the fitness loss. The greater resistance illustrate the trade-off between replication capacity and resistance.

New connection domain mutations identified using a Bayesian Network

The study also investigated the CN domain mutations seen in HIV-1 subtype C treated isolates. To our knowledge, there is only one other study that has described mutations in the connection region specific to HIV-1 subtype C (Barral et al., 2016). In our study, in addition to identifying mutations that were increased in frequency in treated isolates, the association of these mutations was determined using a BN. This method has been used successfully to identify novel mutations and pathways to resistance in HIV-1 (Deforche et al., 2008; Ngcapu et al., 2017a). Interestingly, some of HIV-1 subtype C CN mutations identified as treatment related mutations were similar to those CN mutations in other HIV-1 subtypes (Delviks-Frankenberry et al., 2013; Santos et al., 2008). In agreement with the study by Barral et al (2016), the M357R mutation was identified. However, there was no association with resistance found with the D404N mutations seen in the CRF08_BC strain from China (Zhang et al., 2015b). New CN domain mutations in HIV-1 subtype C, which have also not previously been reported in subtype B, were identified including D324E, S338T, I341V, K366R, E370D/K, M377V, V380I, I434L/M and A437I/V.

In the BN, some CN mutations were directly associated with HIV-1 treatment (E370D, M377T and A437I), while others were indirectly associated (D324E, N348I, M357R, T359N, K366R, A376S, I434M). Interestingly, a known HIV-1 CN mutation N348I was linked to M41L and not directly to exposure to treatment. This is similar to a study by von Wyl et al (2010) where N348I was indirectly associated with ddI exposure and required the development of a number of TAMs

(L74V and M41L) and an additional CN domain mutation (A371V) before it emerged. The connections were also among the TAMs involved in the TAM 1 pathway, which suggests that this may not be a pathway commonly seen in subtype C isolates.

The BN also showed some strong associations between the M41L, K70E, 215F and N348I, T359N, K366R, A376S, A437V. TAMs, such as T215I/F, have also been shown to come at a fitness cost (Brehm et al., 2007). This supports the role of CN domain mutations as compensatory mutations, compensating for the loss of fitness caused by the TAMs; however this requires further investigation using phenotypic assays.

The overall effects of the RNase H domain mutations were investigated using homology modelling. Here, the RNase H domain mutations were shown to affect the position of the RNase H active site. These changes may also affect the binding of the NNRTIs at the polymerase active site, despite these mutations being quite far away from it (Das et al., 2014). Hang et al (2007) showed that the NNRTI binding site, which is located in the palm domain of HIV-1 RT, 10-15Å from the polymerase active site, affects the binding of the substrate. Previously, it has shown that NNRTI binding induces an additional hinge movement of the p66 thumb to adopt a configuration that is even more extended than the DNA-bound RT structure (Das et al., 2008; Hsiou et al., 1996). The study also showed that L517I mutation is located near the RNase H primer grip suggesting that it will affect the RNase H activity. Furthermore, E529D mutation was located near the L491S, K527N and K530R which are known RNase H mutations suggesting that it may increase the effect of these mutations by altering the RNase H active site. It has been shown further that HIV-1 RT conformational changes and enhanced RNase H interactions of RNA/DNA facilitate RNase H cleavage (Das et al., 2014). Understanding the effects of structural changes caused by drug

resistance mutations may provide ideas in developing new HIV-1 drugs targeting different proteins in HIV-1.

In conclusion, this is the first study to show that T470S, L517I and E529D in the RNase H domain can affect drug susceptibility. In addition, this study has identified novel connection domain mutations in HIV-1 subtype C. Since these regions are not included in routine genotyping tests for HIV-1 drug resistance, it is recommended that newer genotype tests are developed that include the N-terminal half of RT, and that revised resistance-associated mutation lists include the RNase H mutations. In addition, these mutations need to be taken into account when developing next generation RTIs, as well as the newer classes of RNaseH inhibitors. More information on the structural implications of these mutations will facilitate drug design strategies. A limitation of the study is that while information about the single mutants were obtained, not all combinations of mutations were tested to fully understand the implications of these mutations in the context of the TAMS.

Chapter 6

References

1. Alkhatib, G., 2009. The biology of CCR5 and CXCR4. *Curr Opin HIV AIDS* 4, 96–103.
2. Anderson, E.C., Lever, A.M.L., 2006. Human Immunodeficiency Virus Type 1 Gag Polyprotein Modulates Its Own Translation. *J Virol.* 80, 10478–10486.
3. Antikainen, N.M., Martin, S.F., 2005. Altering protein specificity: techniques and applications. *Bioorg Med Chem* 13, 2701-2716.
4. Aqil, M., Naqvi, A.R., Bano, A.S., Jameel, S., 2013. The HIV-1 Nef Protein Binds Argonaute-2 and Functions as a Viral Suppressor of RNA Interference. *PLoS ONE* 8, e74472.
5. Arion, D., Kaushik, N., McCormick, S., Borkow, G., Parniak, M.A., 1998. Phenotypic mechanism of HIV-1 resistance to 3'-azido-3'-deoxythymidine (AZT): increased polymerization processivity and enhanced sensitivity to pyrophosphate of the mutant viral reverse transcriptase. *Biochemistry* 37, 15908-15917.
6. Arion, D., Parniak, M.A., 1999. HIV resistance to zidovudine: the role of pyrophosphorolysis. *Drug Resistance Updates* 2, 91-95.
7. Armstrong, K., Lee, T., Essex, M., 2011a. Replicative Fitness Costs of Nonnucleoside Reverse Transcriptase Inhibitor Drug Resistance Mutations on HIV Subtype C. *Antimicrobial Agents And Chemotherapy* 55, 2146-2153.
8. Armstrong, K., Lee, T., Essex, M., 2011b. Resistance Mutations on HIV Subtype C Reverse Transcriptase Inhibitor Drug Replicative Fitness Costs of Nonnucleoside. *Antimicrob. Agents Chemother.* 55, 2146.
9. Arts, E.J., Hazuda, D.J., 2012. HIV-1 Antiretroviral Drug Therapy. *Cold Spring Harb Perspect Med* 2, a007161.
10. Asahchop, E., Oliveira, M., Wainberg, M., Brenner, B., Moisi, D., d'Aquin Toni, T., Tremblay, C., 2011. Characterization of the E138K Resistance Mutations in

HIV-1 Reverse Transcriptase Conferroing Susceptibility to in B and Non-B HIV-1 subtypes. *Antimicrobial Agents and Chemotherapy* 55, 600-607.

11. Asahchop, E.L., Wainberg, M.A., Oliveira, M., Xu, H., Brenner, B.G., Moisi, D., Ibanescu, I.R., Tremblay, C., 2013. Distinct resistance patterns to etravirine and rilpivirine in viruses containing nonnucleoside reverse transcriptase inhibitor mutations at baseline. *AIDS* 27, 879-887.
12. Azijn, H., Tirry, I., Vingerhoets, J., de Be´thune, M., Kraus, G.B., K., Jochmans, D., Van Craenenbroeck, E., Picchio, G., Rimsky, L.T., 2010a. TMC278, a Next-Generation Nonnucleoside Reverse Transcriptase Inhibitor (NNRTI), Active against Wild-Type and NNRTI-Resistant HIV-1 *Antimicrobial agents and chemotherapy* 54, 718-727.
13. Azijn, H., Tirry, I., Vingerhoets, J., de Béthune, M., Kraus, G., Boven, K., Jochmans, D., Van Craenenbroeck, E., Picchio, G., Rimsky, L.T., 2010b. TMC278, a Next-Generation Nonnucleoside Reverse Transcriptase Inhibitor (NNRTI), Active against Wild-Type and NNRTI-Resistant HIV-1. *Antimicrob. Agents Chemother.* 54, 718-727
14. Bacheler, L.T., Anton, E.D., Kudish, P., Baker, D., Bunville, J., Krakowski, K., Bolling, L., Aujay, M., Wang, X.V., Ellis, D., Becker, M.F., Lasut, A.L., George, H.J., Spalding, D.R., Hollis, G., Abremski, K., 2000. Human immunodeficiency virus type 1 mutations selected in patients failing efavirenz combination therapy. *Antimicrob Agents Chemother* 44, 2475-2484.
15. Back, N.K., Nijhuis, M., Keulen, W., Boucher, C.A., Oude Essink, B.O., van Kuilenburg, A.B., van Gennip, A.H., Berkhout, B., 1996. Reduced replication of 3TC-resistant HIV-1 variants in primary cells due to a processivity defect of the reverse transcriptase enzyme. *EMBO J.* 15, 4040–4049.
16. Baltimore, D., 1970a. RNA-dependent DNA polymerase in virions of RNA tumour viruses. *Nature* 226, 1209–1211.
17. Baltimore, D., 1970b. RNA-dependent DNA polymerase in virions of RNA tumour viruses. *Nature* 226, 1209-1211.
18. Barral, M.F.M., Sousa, A.K.P., Santos, A.F., Abreu, C.M., Tanuri, A., Soares, M.A., C, F.t.B.C.f.t.S.o.H.-S., 2016. Identification of Novel Resistance-Related

Polymorphisms in HIV-1 Subtype C RT Connection and RNase H Domains from Patients Under Virological Failure in Brazil. *Aids Research And Human Retroviruses* 00, 1-7.

19. Barré-Sinoussi, F., Chermann, J.C., Rey, F., Nugeyre, M.T., Chamaret, S., Gruest, J., Dauguet, C., Axler-Blin, C., Vézinet-Brun, F., Rouzioux, C., Rozenbaum, W., Montagnier, L., 1983. Isolation of a T-lymphotropic retrovirus from a patient at risk for acquired immune deficiency syndrome (AIDS). *Science* 220, 868-871.
20. Basson, A.E., Rhee, S., Parry, C.M., El-Khatib, Z., Charalambous, S., De Oliveira, T., Pillay, D., Hoffmann, C., Katzenstein, D., Shafer, R.W., Morris, L., 2015. Impact of Drug Resistance-Associated Amino Acid Changes in HIV-1 Subtype C on Susceptibility to Newer Nonnucleoside Reverse Transcriptase Inhibitors. *Antimicrob Agents Chemother* 59, 960-971.
21. Baur, A., 2004. Functions of the HIV-1 Nef protein. *Curr Drug Targets Immune Endocr Metabol Disord* 4, 309-313.
22. Beilhartz, G.L., Götte, M., 2010. HIV-1 Ribonuclease H: Structure, Catalytic Mechanism and Inhibitors. *Viruses* 2, 900-926.
23. Bell, N.M., Lever, A.M., 2013. HIV Gag polyprotein: processing and early viral particle assembly. *Trends Microbiol.* 21, 136-144.
24. Bello, G., Passaes, C.P., Guimaraes, M.L., Lorete, R.S., Matos Almeida, S.E., Medeiros, R.M., Alencastro, P.R., Morgado, M.G., 2008. Origin and evolutionary history of HIV-1 subtype C in Brazil. *AIDS* 22, 1993-2000.
25. Berger, E.A., Murphy, P.M., Farber, J.M., 1999. Chemokine receptors as HIV-1 coreceptors: roles in viral entry, tropism, and disease. *Annu Rev Immunol* 17, 657-700.
26. Bleul, C.C., Wu, L., Hoxie, J.A., Springer, T.A., Mackay, C.R., 1997. The HIV coreceptors CXCR4 and CCR5 are differentially expressed and regulated on human T lymphocytes. *Proc Natl Acad Sci U S A* 94, 1925–1930.
27. Bocket, L., Yazdanpanah, Y., Ajana, F., Gerard, Y., Viget, N., Goffard, A., Alcaraz, I., Wattré, P., Mouton, Y., 2004. Thymidine analogue mutations in antiretroviral-naive HIV-1 patients on triple therapy including either zidovudine or stavudine. *J Antimicrob Chemother* 53, 89-94.

28. Bower, M.J., Cohen, F.E., Dunbrack, R.L.J., 1997. Prediction of protein side-chain rotamers from a backbone-dependent rotamer library: a new homology modeling tool. *J Mol Biol.* 267, 1268-1282.
29. Brehm, J.H., Koontz, D., Meteer, J.D., Pathak, V.K., Sluis-Cremer, N., Mellors, J.W., 2007. Selection of Mutations in the Connection and RNase H Domains of Human Immunodeficiency Virus Type 1 Reverse Transcriptase That Increase Resistance to 3-Azido-3-Dideoxythymidine. *Journal of Virology* 81, 7852-7859.
30. Brehm, J.H., Koontz, D.L., Wallis, C.L., Shutt, K.A., Sanne, I., Wood, R., McIntyre, J.A., Stevens, W.S., Sluis-Cremer, N., Mellors, J.W., 2012a. Frequent emergence of N348I in HIV-1 subtype C reverse transcriptase with failure of initial therapy reduces susceptibility to reverse-transcriptase inhibitors. *Clin Infect Dis* 55, 737-745.
31. Brehm, J.H., Scott, Y., Koontz, D.L., Perry, S., Hammer, S., Katzenstein, D., Mellors, J.W., Sluis-Cremer, N., 2012b. Zidovudine (AZT) monotherapy selects for the A360V mutation in the connection domain of HIV-1 reverse transcriptase. *PLoS One* 7, e31558.
32. Buckheit Jr, R.W., 2005. Understanding HIV resistance, fitness, replication capacity and compensation: targeting viral fitness as a therapeutic strategy. *Expert Opinion on Investigational Drugs* 13, 933-958.
33. Bukrinsky, M., Adzhubei, A., 1999. Viral protein R of HIV-1. *Reviews in medical virology* 9, 39-49.
34. Cases-González, C.E., Franco, S., Martínez, M.A., Menéndez-Arias, L., 2007. Mutational Patterns Associated with the 69 Insertion Complex in Multi-drug-resistant HIV-1 Reverse Transcriptase that Confer Increased Excision Activity and High-level Resistance to Zidovudine. *J. Mol. Biol.* 365, 298-309.
35. Cavasotto, C.N., Phatak, S.S., 2009. Homology modeling in drug discovery: current trends and applications. *Drug Discov Today* 14, 676-683.
36. CDC, 1981, 5th June. 'Pneumocystis Pneumonia MMWR, Los Angeles'.
37. Chan, D.C., Fass, D., Berger, J.M., Kim, P.S., 1997. Core structure of gp41 from the HIV envelope glycoprotein. *Cell* 89, 263-273.
38. Chan, D.C., Kim, P.S., 1998. HIV entry and its inhibition. *Cell* 93, 681-684.

39. Chandonia, J.M., Brenner, S.E., 2005. Implications of structural genomics target selection strategies: Pfam5000, whole genome, and random approaches. *Proteins* 58, 166-179.
40. Chaudhuri, R., Lindwasser, O.W., Smith, W.J., Hurley, J.H., Bonifacino, J.S., 2007. Downregulation of CD4 by Human Immunodeficiency Virus Type 1 Nef Is Dependent on Clathrin and Involves Direct Interaction of Nef with the AP2 Clathrin Adaptor. *J. Virol.* 81, 3877-3890
41. Chen, D., Wang, M., Zhou, S., Zhou, Q., 2002. HIV-1 Tat targets microtubules to induce apoptosis, a process promoted by the pro-apoptotic Bcl-2 relative Bim. *EMBO J.* 21, 6801–6810.
42. Chen, Z., Gettie, A., Ho, D.D., Marx, P.A., 1998. Primary SIVsm isolates use the CCR5 coreceptor from sooty mangabeys naturally infected in west Africa: a comparison of coreceptor usage of primary SIVsm, HIV-2, and SIVmac. *Virology* 246, 113-124.
43. Chunduri, H., Crumpacker, C., Sharma, P.L., 2011. Reverse transcriptase mutation K65N confers a decreased replication capacity to HIV-1 in comparison to K65R due to a decreased RT processivity. *Virology* 414, 34-41.
44. Cohen, G.B., Gandhi, R.T., Davis, D.M., Mandelboim, O., Chen, B.K., Strominger, J.L., Baltimore, D., 1999. The selective downregulation of class I major histocompatibility complex proteins by HIV-1 protects HIV-infected cells from NK cells. *Immunity* 10, 661-671.
45. Collins, K.L., Chen, B.K., Kalams, S.A., Walker, B.D., Baltimore, D., 1998. HIV-1 Nef protein protects infected primary cells against killing by cytotoxic T lymphocytes. *Nature* 391, 397-401.
46. Cozzi-Lepri, A., Ruiz, L., Loveday, C., Phillips, A.N., Clotet, B., Reiss, P., Ledergerber, r.B., Holkmann, C., Staszewski, S., Lundgren, J.D., Group., E.S., 2005. Thymidine analogue mutation profiles: factors associated with acquiring specific profiles and their impact on the virological response to therapy. *Antivir Ther.* 10, 791-80.

47. Cristofaro, J.V., Rausch, J.W., Le Grice, S.F., DeStefano, J.J., 2002. Mutations in the ribonuclease H active site of HIV-RT reveal a role for this site in stabilizing enzyme-primer-template binding. *Biochemistry* 41, 10968-10975.
48. Cunha, R., Abreu, C., Gonzalez, L., Nijhuis, M., de Jong, D., Aguiar, R., Afonso, A., Brindeiro, R., Tanuri, A., 2012. Differential *In Vitro* Kinetics of Drug resistance Mutation Acquisition in HIV-1 RT of subtype B and C. *Plos one* 7, e46622.
49. Das, K., Arnold, E., 2013a. HIV-1 reverse transcriptase and antiviral drug resistance. Part 1. *Current Opinion in Virology* 3, 111-118.
50. Das, K., Arnold, E., 2013b. HIV-1 reverse transcriptase and antiviral drug resistance. Part 2. *Current Opinion in Virology* 3, 119-128.
51. Das, K., Balzarini, J., Miller, M.T., Maguire, A.R., DeStefano, J.J., Arnold, E., 2016. Conformational States of HIV-1 Reverse Transcriptase for Nucleotide Incorporation vs Pyrophosphorolysis-Binding of Foscarnet. *ACS Chem Biol* 11, 2158-2164.
52. Das, K., Bauman, J.D., Clark, A.D., Jr., Frenkel, Y.V., Lewi, P.J., Shatkin, A.J., Hughes, S.H., Arnold, E., 2008. High-resolution structures of HIV-1 reverse transcriptase/TMC278 complexes: strategic flexibility explains potency against resistance mutations. *Proc Natl Acad Sci U S A* 105, 1466-1471.
53. Das, K., Martinez, S.E., Bandwar, R.P., Arnold, E., 2014. Structures of HIV-1 RT-RNA/DNA ternary complexes with dATP and nevirapine reveal conformational flexibility of RNA/DNA: insights into requirements for RNase H cleavage. *Nucleic Acids Res* 42, 8125-8137.
54. Dau, B., Ayers, D., Singer, J., Harrigan, P.R., Brown, S., Kyriakides, T., Cameron, D.W., Angus, B., Holodniy, M., 2010. Connection domain mutations in treatment-experienced patients in the OPTIMA trial. *J Acquir Immune Defic Syndr.* 54, 160-166.
55. Davis, C.A., Parniak, M.A., Hughes, S.H., 2011. The effects of RNase H inhibitors and nevirapine on the susceptibility of HIV-1 to AZT and 3TC. *Virology* 419, 64-71.

56. de Bethune, M., 2010. Non-nucleoside reverse transcriptase inhibitors (NNRTIs), their discovery, development, and use in the treatment of HIV-1 infection: A review of the last 20 years (1989-2009). *Antiviral Research* 85, 75-90.
57. De Clercq, E., 2002. Strategies in the design of antiviral drugs. *Nat. Rev. Drug Discov*, 13-25.
58. De Luca, A., 2006. The impact of resistance on viral fitness and its clinical implications, in: Geretti AM, e. (Ed.), *Antiretroviral Resistance in Clinical Practice*. Mediscript,, London.
59. de Oliveira, T., Pillay, D., Gifford, R.J., Resistance, U.K.C.G.o.H.D., 2010. The HIV-1 subtype C epidemic in South America is linked to the United Kingdom. *PLoS One* 5, e9311.
60. Deeks, S.G., Hoh, R., Neilands, T.B., Liegler, T., Aweeka, F., Petropoulos, C.J., Grant, R.M., Martin, J.N., 2005. Interruption of treatment with individual therapeutic drug classes in adults with multidrug-resistant HIV-1 infection. *J Infect Dis.* 192, 1537-1544.
61. Deforche, K., Camacho, R., Grossman, Z., Silander, T., Soares, M.A., Moreau, Y., Shafer, R.W., Van Laethem, K., Carvalho, A.P., Wynhoven, B., Cane, P., Snoeck, J., Clarke, J., Sirivichayakul, S., Ariyoshi, K., Holguin, A., Rudich, H., Rodrigues, R., Bouzas, M.B., Cahn, P., Brigido, L.F., Soriano, V., Sugiura, W., Phanuphak, P., Morris, L., Weber, J., Pillay, D., Tanuri, A., Harrigan, P.R., Shapiro, J.M., Katzenstein, D.A., Kantor, R., Vandamme, A.M., Abate-Shen, C., 2007. Bayesian network analysis of resistance pathways against HIV-1 protease inhibitors. *Infection, Genetics and Evolution* 7, 382–390.
62. Deforche, K., Camacho, R.J., Grossman, Z., Soares, M.A., van Laethem, K., Katzenstein, D.A., Harrigan, P.R., Kantor, R., Shafer, R., Vandamme, A., workinggroup, o.b.o.t.n.-B., 2008. Bayesian network analyses of resistance pathways against efavirenz and nevirapine. *AIDS* 22, 2107-2115.
63. Deforche, K., Cozzi-Lepri, A., Theys, K., Clotet, B., Camacho, R., Kjaer, J., Van Laethem, K., Phillips, A.N., Moreau, Y., 2008c. Modelled in vivo HIV fitness under drug selective pressure and estimated genetic barrier towards resistance are predictive for virological response. *Antivir Ther.* 13, 399-407.

64. Deforche, K., Silander, T., Camacho, R.J., Grossman, Z., Soares, M.A., Van Laethem, K., Kantor, R., Moreau, Y., Vandamme, A., 2006. Analysis of HIV-1 pol sequence using Bayesian network: implications for drug resistance. *Bioinformatics* 22, 2975-2979.
65. Delviks-Frankenberry, K., Nikolenko, G., Boyer, P., Hughes, S., Coffin, J., Jere, A., Pathak, V., 2008. HIV-1 reverse transcriptase connection subdomain mutations reduce template RNA degradation and enhance AZT excision. *PNAS* 105, 10943-10948.
66. Delviks-Frankenberry, K.A., Lengruber, R.B., Santos, A.F., Silveira, J.M., Soares, M.A., Kearney, M.F., Maldarelli, F., Pathak, V.K., 2013. Connection subdomain mutations in HIV-1 subtype-C treatment-experienced patients enhance NRTI and NNRTI drug resistance. *Virology* 435, 433-441.
67. Delviks-Frankenberry, K.A., Nikolenko, G.N., Barr, R., Pathak, V.K., 2007. Mutations in Human Immunodeficiency Virus Type 1 RNase H Primer Grip Enhance 3-Azido-3-Deoxythymidine Resistance. *Journal Of Virology* 81, 6837-6845.
68. Delviks-Frankenberry, K.A., Nikolenko, G.N., Maldarelli, F., Hase, S., Takebe, Y., Pathak, V.K., 2009. Subtype-specific differences in the human immunodeficiency virus type 1 reverse transcriptase connection subdomain of CRF01_AE are associated with higher levels of resistance to 3'-azido-3'-deoxythymidine. *J Virol* 83, 8502-8513.
69. Delviks-Frankenberry, K.A., Nikolenko, G.N., Pathak, V.K., 2010. The "Connection" Between HIV Drug Resistance and RNase H. *Viruses* 2, 1476-1503.
70. Deora, A., Ratner, L., 2001. Viral protein U (Vpu)-mediated enhancement of human immunodeficiency virus type 1 particle release depends on the rate of cellular proliferation. *J Virol* 75, 6714-6718.
71. Edelheit, O., Hanukoglou, A., Hanukoglu, I., 2009. Simple and efficient site-directed mutagenesis using two single-primer reactions in parallel to generate mutants for protein structure-function studies. *BMC Biotechnology* 9, 1472-6750/1479/1461.

72. Ehteshami, M., Beilhartz, G.L., Scarth, B.J., Tchesnokov, E.P., McCormick, S., Wynhoven, B., Harrigan, P.R., Gotte, M., 2008a. Connection domain mutations N348I and A360V in HIV-1 reverse transcriptase enhance resistance to 3'-azido-3'-deoxythymidine through both RNase H-dependent and -independent mechanisms. *J Biol Chem* 283, 22222-22232.
73. Ehteshami, M., Gotte, M., 2008. Effects of mutations in the connection and RNase H domains of HIV-1 reverse transcriptase on drug susceptibility. *AIDS Rev* 10, 224-235.
74. Ehteshami, M., Scarth, B.J., Tchesnokov, E.P., Dash, C., Le Grice, S.F., Hallenberger, S., Jochmans, D., Gotte, M., 2008b. Mutations M184V and Y115F in HIV-1 reverse transcriptase discriminate against "nucleotide-competing reverse transcriptase inhibitors". *J Biol Chem* 283, 29904-29911.
75. Esnouf, R., Ren, J., Ross, C., Jones, Y., Stammers, D., Stuart, D., 1995a. Mechanism of inhibition of HIV-1 reverse transcriptase by non-nucleoside inhibitors. *Nat. Struct. Biol.* 2, 303-308.
76. Esnouf, R., Ren, J., Ross, C., Jones, Y., Stammers, D., Stuart, D., 1995b. Mechanism of inhibition of HIV-1 reverse transcriptase by nonnucleoside inhibitors. *Nat. Struct. Biol.* 2, 303–308.
77. Esposito, F., Zinzula, L., Tramontano, E., 2007. Antiviral drug discovery: HIV-1 RNase H, the next hit target. *Communicating Current Research and Educational Topics and Trends in Applied Microbiology*, 799-809.
78. Fischl, M.A., Richman, D.D., Grieco, M.H., Gottlieb, M.S., Volberding, P.A., Laskin, O.L., Leedom, J.M., Groopman, J.E., Mildvan, D., Schooley, R.T., et al., 1987. The efficacy of azidothymidine (AZT) in the treatment of patients with AIDS and AIDS-related complex. A double-blind, placebo-controlled trial. *N Engl J Med* 317, 185-191.
79. Floudas, C.A., Fung, H.K., McAllister, S.R., Mönnigmann, M., Rajgaria, R., 2006. Advances in protein structure prediction and de novo protein design: A review. *Chem Eng Sci.* 61, 966–988.
80. Freed, E.O., 1998. HIV-1 gag proteins: diverse functions in the virus life cycle. *Virology* 251, 1-15.

81. Freed, E.O., 2001. HIV-1 replication. *Somat Cell Mol Genet.* 26, 13-33.
82. Gallego, O., Corral, A., de Mendoza, C., Rodes, B., Soriano, V., 2002. Prevalence of G333D/E in naive and pretreated HIV-infected patients. *AIDS Res Hum Retroviruses.* 18, 857-860.
83. Gallo, R.C., Salahuddin, S.Z., Popovic, M., Shearer, G.M., Kaplan, M., Haynes, B.F., Palker, T.J., Redfield, R., Oleske, J., Safai, B., et al., 1984. Frequent detection and isolation of cytopathic retroviruses (HTLV-III) from patients with AIDS and at risk for AIDS. *Science* 224, 500-503.
84. Gallo, R.C., Sarin, P.S., Gelmann, E.P., Robert-Guroff, M., Richardson, E., Kalyanaraman, V.S., Mann, D., Sidhu, G.D., Stahl, R.E., Zolla-Pazner, S., Leibowitch, J., Popovic, M., 1983. Isolation of human T-cell leukemia virus in acquired immune deficiency syndrome (AIDS). *Science* 220, 865-867.
85. Garcia-Lema, J.G., (2005), 56, , doi:10.1093/jac/dki194, 2005, A.A.p.J., among, D.o.t.a.r.g., persons, n.d.H.-i., Garcı, J.G., , a-Lerma, 2005. Diversity of thymidine analogue resistance genotypes among newly diagnosed HIV-1-infected persons. *Journal of Antimicrobial Chemotherapy* 56, 265-269.
86. Garcia-Lerma, J.G., Marclnnes, H., Bennett, D., Weinstock, H., Heneine, W., 2004. Transmitted Human Immunodeficiency Virus Type 1 Carrying the D67N or K219Q/E Mutation Evolves Rapidly to Zidovudine Resistance in Vitro and Shows a High Replicative Fitness in the Presence of Zidovudine. *J.Virol.* 78, 7545-7552.
87. Gelmann, E.P., Popovic, M., Blayney, D., Masur, H., Sidhu, G.D., Stahl, R.E., Gallo, R.C., 1983. Proviral DNA of a retrovirus, human T-cell leukemia virus, in two patients with AIDS. *Science* 220, 862-865.
88. Gibbs, J.E., Rashid, T., Thomas, S.A., 2003. Effect of Transport Inhibitors and Additional Anti-HIV Drugs on the Movement of Lamivudine (3TC) across the Guinea Pig Brain Barriers. *The Journal Of Pharmacology And Experimental Therapeutics* 306, 1035–1104.
89. Gibbs, J.S., Lackner, A.A., Lang, S.M., Simon, M.A., Sehgal, P.K., Daniel, M.D., Desrosiers, R.C., 1995. Progression to AIDS in the absence of a gene for vpr or vpx. *J Virol* 69, 2378-2383.

90. Goepfert, P.A., Lumm, W., Farmer, P., Matthews, P., Prendergast, A., Carlson, J.M., Derdeyn, C.A., Tang, J., Kaslow, R.A., Bansal, A., Yusim, K., Heckerman, D., Mulenga, J., Allen, S., Goulder, P.J., Hunter, E., 2008. Transmission of HIV-1 Gag immune escape mutations is associated with reduced viral load in linked recipients. *J Exp Med* 205, 1009-1017.
91. Goh, W.C., Rogel, M.E., Kinsey, C.M., Michael, S.F., Fultz, P.N., Nowak, M.A., Hahn, B.H., Emerman, M., 1998. HIV-1 Vpr increases viral expression by manipulation of the cell cycle: a mechanism for selection of Vpr in vivo. *Nat Med* 4, 65-71.
92. Götte, M., 2007. Should We Include Connection Domain Mutations of HIV-1 Reverse Transcriptase in HIV Resistance Testing? *Plos one* 4, 1859-1861.
93. Gotte, M., Arion, D., Parniak, M.A., Wainberg, M.A., 2000. The M184V mutation in the reverse transcriptase of human immunodeficiency Virus type-1 impairs rescue of chain-terminated DNA synthesis. *J.Virol.* 74, 3579-3585.
94. Gulick, R.M., Ribaldo, H.J., Shikuma, C.M., Lustgarten, S., Squires, K.E., Meyer, W.A., 3rd, Acosta, E.P., Schackman, B.R., Pilcher, C.D., Murphy, R.L., Maher, W.E., Witt, M.D., Reichman, R.C., Snyder, S., Klingman, K.L., Kuritzkes, D.R., 2004. Triple-nucleoside regimens versus efavirenz-containing regimens for the initial treatment of HIV-1 infection. *N Engl J Med* 350, 1850-1861.
95. Guo, W., Li, H., Zhuang, D., Jiao, L., Liu, S., Li, L., Liu, Y., Gui, T., Jia, L., Li, J., 2014. Impact of Y181C and/or H221Y mutation patterns of HIV-1 reverse transcriptase on phenotypic resistance to available non-nucleoside and nucleoside inhibitors in China. *BMC Infectious Diseases* 14.
96. Gupta, S., Fransen, S., Paxinos, E.E., Stawiski, E., Huang, W., Petropoulos, C.J., 2011. Combinations of mutations in the connection domain of human immunodeficiency virus type 1 reverse transcriptase: assessing the impact on nucleoside and nonnucleoside reverse transcriptase inhibitor resistance. *Antimicrob Agents Chemother* 54, 1973-1980.
97. Hachiya, A., Kodama, E.N., Sarafianos, S.G., Schuckmann, M.M., Sakagami, Y., Matsuoka, M., Takiguchi, M., Gatanaga, H., Oka, S., 2008a. Amino Acid Mutation N348I in the Connection Subdomain of Human Immunodeficiency Virus Type 1

- Reverse Transcriptase Confers Multiclass Resistance to Nucleoside and Nonnucleoside Reverse Transcriptase Inhibitors. *Journal of Virology* 82, 3261-3270.
98. Hachiya, A., Kodama, E.N., Sarafianos, S.G., Schuckmann, M.M., Sakagami, Y., Matsuoka, M., Takiguchi, M., Gatanaga, H., Oka, S., 2008b. Amino acid mutation N348I in the connection subdomain of human immunodeficiency virus type 1 reverse transcriptase confers multiclass resistance to nucleoside and nonnucleoside reverse transcriptase inhibitors. *J Virol* 82, 3261-3270.
99. Hachiya, A., Shimane, K., Sarafianos, S.G., Kodama, E.N., Sakagami, Y., Negishi, F., Koizumi, H., Gatanaga, H., Matsuoka, M., Takiguchi, M., Oka, S., 2009. Clinical relevance of substitutions in the connection subdomain and RNase H domain of HIV-1 reverse transcriptase from a cohort of antiretroviral treatment-naive patients. *Antiviral Res* 82, 115-121.
100. Hackerman, D., 1996. *A Tutorial on Learning With Bayesian Network*, p. 57.
101. Hadian, K., Vincendeau, M., Mausbacher, N., Nagel, D., Hauck, S.M., Ueffing, M., Loyer, A., Werner, T., Wolff, H., Brack-Werner, R., 2009. Identification of a heterogeneous nuclear ribonucleoprotein-recognition region in the HIV Rev protein. *J Biol Chem* 284, 33384-33391.
102. Hahn, B.H., Shaw, G.M., Arya, S.K., Popovic, M., Gallo, R.C., Wong-Staal, F., 1984. Molecular cloning and characterization of the HTLV-III virus associated with AIDS. *Nature* 312, 166-169.
103. Haidara, A., Chamberland, A., Sylla, M., Aboubacrine, S.A., Cissé, M., Traore, H.A., Maiga, M.Y., Tounkara, A., Nguyen, V.K., Tremblay, C., .. A.I.T.A.R.e.A.d.l.O.A.G., 2010. High level of primary drug resistance in Mali. *HIV Med.* 11, 404-411.
104. Hang, J.Q., Li, Y., Yang, Y., Cammack, N., Mirzadegan, T., Klumpp, K., 2007. Substrate-dependent inhibition or stimulation of HIV RNase H activity by non-nucleoside reverse transcriptase inhibitors (NNRTIs). *Biochemical and Biophysical Research Communications* 352, 341-350.
105. Haubrich, R.H., 2004. Resistance and Replication capacity assays: Clinical utility and interpretation. *Topics in HIV medicine* 12, 52-56.

106. Haubrich, R.H., Kemper, C.A., Hellmann, N.S., Keiser, P.H., Witt, M.D., Forthal, D.N., Leedom, J., Leibowitz, M., Whitcomb, J.M., Richman, D., McCutchan, J.A., California Collaborative Treatment, G., 2002. The clinical relevance of non-nucleoside reverse transcriptase inhibitor hypersusceptibility: a prospective cohort analysis. *AIDS* 16, F33-40.
107. Heimer, R., Myers, S.S., Cadman, E.C., Kaplan, E.H., 1992. Detection by polymerase chain reaction of human immunodeficiency virus type 1 proviral DNA sequences in needles of injecting drug users. *J Infect Dis.* 165, 781-782.
108. Henderson, B.R., Percipalle, P., 1997. Interactions between HIV Rev and nuclear import and export factors: the Rev nuclear localisation signal mediates specific binding to human importin-beta. *J Mol Biol.* 274, 693-707.
109. Hertogs, K., de BE THUNE, M.P., Miller, V., Ivens, T., Schel, P., Van cauwenberge, A., Van den eynde, C., Van gerwen, V., Azijn, H., Van houtte, M., Peeters, F., Staszewski, S., Conant, M., Bloor, S., Kemp, S., Larder, B., 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. *Antimicrobial agents and chemotherapy* 42, 269-276.
110. Hillisch, A., Pineda, L.F., Hilgenfeld, R., 2004. Utility of homology models in the drug discovery process. *DDT* 9, 660-669.
111. Hiscott, J., Kwon, H., Génin, P., 2001. Hostile takeovers: viral appropriation of the NF-kappaB pathway. *J Clin Invest.* 107, 143-151.
112. Hsiou, Y., Ding, J., Das, K., Clark, A.D., Jr., Boyer, P.L., Lewi, P., Janssen, P.A., Kleim, J.P., Rosner, M., Hughes, S.H., Arnold, E., 2001. The Lys103Asn mutation of HIV-1 RT: a novel mechanism of drug resistance. *Journal of molecular biology* 309, 437-445.
113. Hsiou, Y., Ding, J., Das, K., Clark Jr., A.D., Hughes, S.H., Arnold, E., 1996. Structure of unliganded HIV-1 reverse transcriptase at 2.7Å resolution: implications of conformational changes for polymerization and inhibition mechanisms. *Structure* 4, 853-860.

114. Hu, W., Hughes, S.H., 2012 HIV-1 Reverse Transcription. *Cold Spring Harb Perspect Med.* 2, a006882.
115. Hu, Z., Kuritzkes, D., 2011. Interaction of Reverse Transcriptase (RT) Mutations Conferring Resistance to lamivudine and Etravirine: Effects on fitness and RT Activity of Human Immunodeficiency virus Type 1. *Journal of Virology* 85, 11309-11314.
116. Huang, L., Jeang, K., 1995. HIV Vpr: Roles in Viral Replication and Cellular Metabolism.
117. Ilina, T., Labarge, K., Sarafianos, S.G., Ishima, R., Parniak, M.A., 2012. Inhibitors of HIV-1 Reverse Transcriptase-Associated Ribonuclease H Activity. *Biology (Basel)* 1, 521-541.
118. Iyidogan, P., Anderson, K.S., 2014. Current Perspectives on HIV-1 Antiretroviral Drug Resistance. *Viruses* 6, 4095-4139.
119. Jacks, T., Power, M.D., Masiarz, F.R., Luciw, P.A., Barr, P.J., Varmus, H.E., 1988. Characterization of ribosomal frameshifting in HIV-1 gag-pol expression. *Nature* 331, 280-283.
120. Japour, A.J., Mayers, D.L., Johnson, V.A., Kuritzkes, D.R., Beckett, L.A., Arduino, J.M., Lane, J., Black, R.J., Reichelderfer, P.S., D'Aquila, R.T., Crumpacker, C.S., group, T.R.-s., group, T.A.c.T.g.v.c.r.w., 1993. Standardized peripheral blood mononuclear cell culture assay for determination of drug susceptibilities of clinical human immunodeficiency virus type 1 isolates. The RV-43 Study Group, the AIDS Clinical Trials Group Virology Committee Resistance Working Group. *Antimicrob Agents Chemother.* 37, 1095-1101.
121. Jiang, Y., Liu, X., De Clercq, E., 2011. New therapeutic approaches targeted at the late stages of the HIV-1 replication cycle. *Curr Med Chem* 18, 16-28.
122. Joly, V., Descamps, D., Peytavin, G., Touati, F., Mentre, F., Duval, X., Delarue, S., Yeni, P., Brun-Vezinet, F., 2004. Evolution of human immunodeficiency virus type 1 (HIV-1) resistance mutations in nonnucleoside reverse transcriptase inhibitors (NNRTIs) in HIV-1-infected patients switched to antiretroviral therapy without NNRTIs. *Antimicrob Agents Chemother* 48, 172-175.

123. Julias, J.G., Boyer, P.L., McWilliams, M.J., Alvord, W.G., Hughes, S.H., 2004. Mutations at position 184 of human immunodeficiency virus type-1 reverse transcriptase affect virus titer and viral DNA synthesis. *Virology* 322, 13-14.
124. Julias, J.G., McWilliams, M.J., Sarafianos, S.G., Alvord, W.G., Arnold, E., Hughes, S.H., 2003. Mutation of amino acids in the connection domain of human immunodeficiency virus type 1 reverse transcriptase that contact the template-primer affects RNase H activity. *J Virol* 77, 8548-8554.
125. Julias, J.G., McWilliams, M.J., Sarafianos, S.G., Arnold, E., Hughes, S.H., 2002. Mutations in the RNase H domain of HIV-1 reverse transcriptase affect the initiation of DNA synthesis and the specificity of RNase H cleavage in vivo. *Proc Natl Acad Sci U S A* 99, 9515-9520.
126. Kellam, P., Larder, B., 1994. Recombinant Virus Assay: a Rapid, Phenotypic Assay for assessment of drug susceptibility of human immunodeficiency Virus Type-1 isolates. *ANTIMICROBIAL AGENTS AND CHEMOTHERAPY* 38, 23-30.
127. Kemp, S.D., Shi, C., Bloor, S., Harrigan, P.R., Mellors, J.W., Larder, B.A., 1998. A Novel Polymorphism at Codon 333 of Human Immunodeficiency Virus Type 1 Reverse Transcriptase Can Facilitate Dual Resistance to Zidovudine and L-2',3'-Dideoxy-3'-Thiacytidine. *Journal of Virology* 72, 5093-5098.
128. Kusic, M., Mendieta, J., Puertas, M.C., Parera, M., Martinez, M.A., Martinez-Picado, J., Menendez-Arias, L., 2008. Mechanistic basis of zidovudine hypersusceptibility and lamivudine resistance conferred by the deletion of codon 69 in the HIV-1 reverse transcriptase coding region. *Journal of molecular biology* 382, 327-341.
129. Larder, B., Kemp, S., 1989. Multiple mutations in HIV-1 reverse transcriptase confer high-level resistance to zidovudine (AZT). *Science* 246, 1155-1158.
130. Le Rouzic, E., Benichou, S., 2005. The Vpr protein from HIV-1: distinct roles along the viral life cycle. *Retrovirology* 2, 11.
131. Lengruber, R.B., Delviks-Frankenberry, K.A., Nikolenko, G.N., Baumann, J., Santos, A.F., Pathak, V.K., Soares, M.A., 2011. Phenotypic characterization of drug resistance-associated mutations in HIV-1 RT connection and RNase H

- domains and their correlation with thymidine analogue mutations. *Journal of Antimicrobial Chemotherapy* 66, 702-708.
132. Li, H., Tejero, R., Monleon, D., Bassolino-Klimas, D., Abate-Shen, C., Bruccoleri, R.E., Montelione, G.T., 1997. Homology modeling using simulated annealing of restrained molecular dynamics and conformational search calculations with CONGEN: application in predicting the three-dimensional structure of murine homeodomain Msx-1. *Protein Sci.* 6, 956–970.
133. Lichterfeld, M., Yu, X.G., Waring, M.T., Mui, S.K., Johnston, M.N., Cohen, D., Addo, M.M., Zaunders, J., Alter, G., Pae, E., Strick, D., Allen, T.M., Rosenberg, E.S., Walker, B.D., Altfeld, M., 2004. HIV-1-specific cytotoxicity is preferentially mediated by a subset of CD8(+) T cells producing both interferon-gamma and tumor necrosis factor-alpha. *Blood* 104, 487-494.
134. Ling, M.M., B.H., R., 1997. Approaches to DNA mutagenesis:an overview. *Anal Biochem* 244, 157-178.
135. Little, S.J., Daar, E.S., D'Aquila, R.T., Keiser, P.H., Connick, E., Whitcomb, J.M., Hellmann, N.S., Petropoulos, C.J., Sutton, L., Pitt, J.A., Rosenberg, E.S., Koup, R.A., Walker, B.D., Richman, D.D., 1999. Reduced antiretroviral drug susceptibility among patients with primary HIV infection. *JAMA* 282, 1142-1149.
136. Maiga, A.I., Penugonda, S., Katile, D., Diallo, F., Fofana, D.B., Berzins, B., Maiga, M.Y., Sylla, A., Traore, H.A., Marcelin, A.G., Calvez, V., Tounkara, A., Bellosillo, N., Murphy, R., Taiwo, B., 2012. Connection domain mutations during antiretroviral treatment failure in Mali: frequencies and impact on reverse transcriptase inhibitor activity. *J Acquir Immune Defic Syndr* 61, 293-296.
137. Marconi, V.C., Sunpath, H., Lu, Z., Gordon, M., Koranteng-Apegyei, K., Hampton, J., Carpenter, S., Giddy, J., Ross, D., Holst, H., Losina, E., Walker, B.D., Kuritzkes, D.R., 2008. Prevalence of HIV-1 drug resistance after failure of a first highly active antiretroviral therapy regimen in KwaZulu Natal, South Africa. *Clin Infect Dis* 46, 1589-1597.
138. Martz, E., 2001. Homology modelling for Beginners. Free Software <http://www.umass.edu/molvis/workshop/homolmod.htm>.
139. Mas, A., Vazquez-Alvarez, B.M., Domingo, E., Menendez-Arias, L., 2002. Multidrug-resistant HIV-1 reverse transcriptase: involvement of ribonucleotide-

- dependent phosphorolysis in cross-resistance to nucleoside analogue inhibitors. *Journal of molecular biology* 323, 181-197.
140. Masur, H., Michelis, M.A., Greene, J.B., Onorato, I., Stouwe, R.A., Holzman, R.S., Wormser, G., Brettman, L., Lange, M., Murray, H.W., Cunningham-Rundles, S., 1981. An outbreak of community-acquired *Pneumocystis carinii* pneumonia: initial manifestation of cellular immune dysfunction. *N Engl J Med* 305.
141. Mbisa, J.L., Martin, S.A., Cane, P.A., 2011. Patterns of resistance development with integrase inhibitors in HIV. *Infect Drug Resist* 4, 65-76.
142. Mbisa, J.L., Nikolenko, G.N., Pathak, V.K., 2005. Mutations in the RNase H primer grip domain of murine leukemia virus reverse transcriptase decrease efficiency and accuracy of plus-strand DNA transfer. *J Virol* 79, 419-427.
143. McCormick, A.L., Parry, C.M., Crombe, A., Goodall, R.L., Gupta, R.K., Kaleebu, P., Kityo, C., Chirara, M., Towers, G.J., Pillay, D., 2011. Impact of the N348I Mutation in HIV-1 Reverse Transcriptase on Nonnucleoside Reverse Transcriptase Inhibitor Resistance in Non-Subtype B HIV-1. *Antimicrob Agents Chemother.* 55, 1806–1809.
144. McGovern, S.L., Caselli, E., Grigorieff, N., Shoichet, B.K., 2002. A common mechanism underlying promiscuous inhibitors from virtual and high-throughput screening. *J Med Chem.* 45, 1712-1722.
145. Meintjes, G., Maartens, G., 2014. Dosing of raltegravir when given with rifampicin. *Lancet Infect Dis* 14, 442-443.
146. Meintjes, G., Moorhouse, M.A., Carmona, S., Davies, N., Dlamini, S., van Vuuren, C., Manzini, T., Mathe, M., Moosa, Y., Nash, J., Nel, J., Pakade, Y., Woods, J., Van Zyl, G., Conradie, F., Venter, F., 2017. Adult antiretroviral therapy guidelines 2017. *South Afr J HIV Med* 18, 776.
147. Menendez-Arias, L., 2002. Targeting HIV: antiretroviral therapy and development of drug resistance. *Trends Pharmacol Sci* 23, 381-388.
148. Menendez-Arias, L., 2008. Mechanisms of resistance to nucleoside analogue inhibitors of HIV-1 reverse transcriptase. *Virus Res* 134, 124-146.

149. Menendez-Arias, L., 2011. A structural frame for understanding the role of thymidine analogue resistance mutations in resistance to zidovudine and other nucleoside analogues. *Antivir Ther* 16, 943-946.
150. Menendez-Arias, L., Berkhout, B., 2008. Retroviral reverse transcription. *Virus Res* 134, 1-3.
151. Menendez-Arias, L., Betancor, G., Matamoros, T., 2011. HIV-1 reverse transcriptase connection subdomain mutations involved in resistance to approved non-nucleoside inhibitors. *Antiviral Res* 92, 139-149.
152. Menendez-Arias, L., Tozser, J., 2008. HIV-1 protease inhibitors: effects on HIV-2 replication and resistance. *Trends Pharmacol Sci* 29, 42-49.
153. Merluzzi, V.J., Hargrave, K.G., Labadia, M., Grozinger, K., Skoog, M., Wu, J.C., Shih, C.K., Eckner, K., Hattox, S., Adams, J., Rosenthal, A.S., Faanes, R., Eckner, R.J., Koup, R.A., Sullivan, J.L., 1990. Inhibition of HIV-1 replication by a nonnucleoside reverse transcriptase inhibitor. *Science* 250, 1411–1413
154. Meyer, P.R., Matsuura, S.E., So, A.G., Scott, W.A., 1998. Unblocking of chain-terminated primer by HIV-1 reverse transcriptase through a nucleotide-dependent mechanism. *Proceedings of the National Academy of Sciences of the United States of America* 95, 13471-13476.
155. Michels, I., Staszewski, S., Gurtler, L., Nisius, G., Muller, A., Locher, L., Doerr, H.W., Sturmer, M., 2010. Mutations in the C-terminal region of the HIV-1 reverse transcriptase and their correlation with drug resistance associated mutations and antiviral treatment. *European Journal of medical research* 15, 415-421.
156. Miller, M., Jaskólski, M., Rao, J.K., Leis, J., Wlodawer, A., 1989. Crystal structure of a retroviral protease proves relationship to aspartic protease family. *Nature*. 337, 576-579.
157. Mizrahi, V., Brooksbank, R.L., Nkabinde, N.C., 1994. Mutagenesis of the conserved aspartic acid 443, glutamic acid 478, asparagine 494, and aspartic acid 498 residues in the ribonuclease H domain of p66/p51 human immunodeficiency virus type I reverse transcriptase. Expression and biochemical analysis. *J Biol Chem* 269, 19245-19249.

158. Mizutani, S., Boettiger, D., Temin, H.M., 1970. A DNA-dependent DNA polymerase and a DNA endonuclease in virions of Rous sarcoma virus. *Nature* 228, 424-427.
159. Montefiori, D., 2009. Measuring HIV Neutralization in a Luciferase Reporter Gene Assay. *Human Protocols: Second Edition* 485, 395-404.
160. Muniz, C.P., Soares, M.A., Santos, A.F., 2014. Early selection of resistance-associated mutations in HIV-1 RT C-terminal domains across different subtypes: role of the genetic barrier to resistance. *J Antimicrob Chemother* 69, 2741-2745.
161. Myllymaki, P., Silander, T., Tirri, H., Uronen, P., 2002. B-course: a web-based tutorial for Bayesian and casual data analysis. *Int J Artif Intell tools* 11, 396-387.
162. Ngcapu, S., 2012. Nucleoside Reverse Transcriptase Inhibitors-associated Mutations in the RNase H Region of HIV-1 isolates in South African Adults and Children failing Highly Active Antiretroviral Therapy. Department of Molecular Virology, Faculty of Health Science, Nelson R. Mandela School of Medicine at the University of KwaZulu-Natal., Durban.
163. Ngcapu, S., Theys, K., Libin, P., Marconi, V.C., Sunpath, H., Ndung'u, T., Gordon, M.L., 2017a. Characterization of Nucleoside Reverse Transcriptase Inhibitor-Associated Mutations in the RNase H Region of HIV-1 Subtype C Infected Individuals. *Viruses* 9, 1-10.
164. Nicastri, E., Sarmati, L., d'Ettorre, G., Palmisano, L., Parisi, S.G., Uccella, I., Rianda, A., Concia, E., V., V., Vella, S., Andreoni, M., 2003. Replication capacity, biological phenotype, and drug resistance of HIV strains isolated from patients failing antiretroviral therapy. *J Med Virol.* 69, 1-6.
165. Nikolenko, G.N., Delviks-Frankenberry, K.A., Palmer, S., Maldarelli, F., Fivash, M.J., Jr., Coffin, J.M., Pathak, V.K., 2007. Mutations in the connection domain of HIV-1 reverse transcriptase increase 3'-azido-3'-deoxythymidine resistance. *Proc Natl Acad Sci U S A* 104, 317-322.
166. Nikolenko, G.N., Delviks-Frankenberry, K.A., Pathak, V.K., 2010. A Novel Molecular Mechanism of Dual Resistance to Nucleoside and Nonnucleoside Reverse Transcriptase Inhibitors. *Journal of Virology* 84, 5238-5249.

167. Nikolenko, G.N., Palmer, S., Maldarelli, F., Mellors, J.W., Coffin, J.M., Pathak, V.K., 2005. Mechanism for nucleoside analog-mediated abrogation of HIV-1 replication: balance between RNase H activity and nucleotide excision. *Proc Natl Acad Sci U S A* 102, 2093-2098.
168. Nikolenko, G.N., Svarovskaia, E.S., Delviks, K.A.e.a., 2004. Antiretroviral drug resistance mutations in human immunodeficiency virus type 1 reverse transcriptase increase template-switching frequency. *J Virol* 78, 8761-8770.
169. Novitsky, V., Wang, R., Lagakos, S., Essex, M., 2010. HIV-1 Subtype C Phylogenetics in the Global Epidemic. *Viruses* 2, 33-54.
170. Novitsky, V., Wester, C.W., DeGruttola, V., Bussmann, H., Gaseitsiwe S, Thomas, A., Moyo, S., Musonda R, Van Widenfelt, E., Marlink, R.G., Essex, M., 2007. The reverse transcriptase 67N 70R 215Y genotype is the predominant TAM pathway associated with virologic failure among HIV type 1C-infected adults treated with ZDV/ddI-containing HAART in southern Africa. *AIDS Res Hum Retroviruses*. 23, 868-878.
171. Page, K.A., van Schooten, W.C., Feinberg, M.B., 1997. Human immunodeficiency virus type 1 Nef does not alter T-cell sensitivity to antigen-specific stimulation. *J Virol* 71, 3776-3787.
172. Palaniappan, C., Wisniewski, M., Jacques, P.S., Le Grice, S.F.J., Fay, P.J., Bambara, R.A., 1997. Mutations within the Primer Grip Region of HIV-1 Reverse Transcriptase Result in Loss of RNase H Function. *The Journal Of Biological Chemistry* 272, 11157-11164.
173. Paredes, R., Puertas, M., Bannister, W., Kisic, M., Cozzi-Lepri, A., Pou, C., o Bellido, R., Betancor, G., Bogner, J., Gargalianos, P., Ba´nhegyi, D., Clotet, B., Lundgren, J., Mene´ndez-Arias, L., Martinez-Picado, J., Group, T.E.S., 2011. A376S in the Connection Subdomain of HIV-1 Reverse Transcriptase Confers Increased Risk of Virological Failure to Nevirapine Therapy. *The Journal of Infectious Diseases* 204, 741-752.
174. Pattery, T., Verlinden, Y., De Wolf, H., Nauwelaers, D., Van Baelen, K., van Houtte, M., Mc Kenna, P., Villacian, J., 2012. Development and Performance of Conventional HIV-1 Phenotyping (Antivirogram) and Genotyped-Based

- Calculated Phenotyping Assay (virco TYPE HIV-1) on Protease and Reverse Transcriptase Genes to Evaluate Drg Resistance. *Intervirology* 55, 136-146.
175. Pearl, J., 1988. Probabilistic Reasoning in intelligent systems Networks of Plausible Inference Morgan Kaufmann, San Mateu, CA.
176. Pennings, P., 2013. HIV-1 drug resistance: problems and perspectives *Infectious Disease Reports* 5, 21-25.
177. Petropoulos, C.J., Parkin, N.T., Limoli, K.L., Lie, Y.S., Wrin, T., Huang, W., Tian, Y., Smith, D., Winslow, G.A., Capon, D.J., Whitcomb, J.M., 2000. A Novel Phenotypic Drug Susceptibility Assay for Human Immunodeficiency Virus Type 1. *Antimicrob Agents Chemother* 44, 920-928.
178. Pettersen, E.F., Goddard, T.D., Huang, C.C., Couch, G.S., Greenblatt, D.M., Meng, E.C., Ferrin, T.E., 2004. UCSF Chimera--a visualization system for exploratory research and analysis. *J Comput Chem* 25, 1605-1612.
179. Pillay, V., Ledwaba, J., Hunt, G., Rakgotho, M., Singh, B., Makubalo, L., Bennett, D.E., Puren, A., Morris, L., 2008. Antiretroviral drug resistance surveillance among drug-naive HIV-1-infected individuals in Gauteng Province, South Africa in 2002 and 2004. *Antivir Ther* 13 Suppl 2, 101-107.
180. Pinggen, M., Sarrami-Forooshani, R., Wensing, A.M., van Ham, P., Drewniak, A., Boucher, C.A., Geijtenbeek, T.B., Nijhuis, M., 2014. Diminished transmission of drug resistant HIV-1 variants with reduced replication capacity in a human transmission model. *Retrovirology* 11, 113.
181. Pollard, V.W., Malim, M.H., 1998. The HIV-1 Rev protein. *Annu Rev Microbiol* 52, 491-532.
182. Popovic, M., Sarngadharan, M.G., Read, E., Gallo, R.C., 1984. Detection, isolation, and continuous production of cytopathic retroviruses (HTLV-III) from patients with AIDS and pre-AIDS. *Science* 224, 497-500.
183. Puertas, M., Buzon, M., Ballester, M., Van Den Esde, P., Clotet, B., Prado, J., Martinez-Picado, J., 2012. Novel Two-Round Phenotypic Assay for Protease Inhibitor Susceptibility Testing of Recombinant and primary HIV-1 isolates. *Journal of Clinical Microbiology* 50, 3909-3916.

184. Radzio, J., Sluis-Cremer, N., 2008. Efavirenz Accelerates HIV-1 Reverse Transcriptase Ribonuclease H Cleavage, Leading to Diminished Zidovudine Excision. *Molecular Pharmacology* 73 601-606.
185. Radzio, J., Sluis-Cremer, N., 2011. Subunit-specific mutational analysis of residue N348 in HIV-1 reverse transcriptase. *Retrovirology* 8, 69.
186. Radzio, J., Yap, S., Tachedjian, G., Sluis-Cremer, N., 2010. N348I in Reverse Transcriptase Provides a Genetic Pathway for HIV-1 to Select TAMs and Mutations Antagonistic to TAMs. *AIDS* 24, 659-667.
187. Rausch, J.W., Lener, D., Miller, J.T., Julias, J.G., Hughes, S.H., Le Grice, S.F.J., 2002. Altering the RNase H Primer Grip of Human Immunodeficiency Virus Reverse Transcriptase Modifies Cleavage Specificity. *Biochemistry* 41, 4856-4865.
188. Ren, J., Esnouf, R.M., Hopkins, A.L., Jones, E.Y., Kirby, I., Keeling, J., Ross, C.K., Larder, B.A., Stuart, D.I., Stammers, D.K., 1998. 3'-Azido-3'-deoxythymidine drug resistance mutations in HIV-1 reverse transcriptase can induce long range conformational changes. *Proc Natl Acad Sci U S A* 95, 9518-9523.
189. Reuman, E.C., Rhee, S.Y., Holmes, S.P., Shafer, R.W., 2010. Constrained patterns of covariation and clustering of HIV-1 non-nucleoside reverse transcriptase inhibitor resistance mutations. *J Antimicrob Chemother* 65, 1477-1485.
190. Rhee, S., Fessel, W., Zolopa, A., Hurley, L., Liu, T., Taylor, J., Nguyen, D., Slome, S., Klein, D., Horberg, M., Flamm, J., Follansbee, S., Schapiro, J., Shafer, R., 2005. HIV-1 Protease and Reverse-Transcriptase Mutations: Correlations with Antiretroviral Therapy in Subtype B Isolates and Implications for Drug-Resistance Surveillance. *J Infect Dis.* 192, 456-465.
191. Richman, D., Shih, C.K., Lowy, I., Rose, J., Prodanovich, P., Goff, S., Griffin, J., 1991. Human immunodeficiency virus type 1 mutants resistant to nonnucleoside inhibitors of reverse transcriptase arise in tissue culture. *Proc Natl Acad Sci U S A* 88, 11241-11245.

192. Ring, C.S., Sun, E., McKerrow, J.H., Lee, G.K., Rosenthal, P.J., Kuntz, I.D., Cohen, F.E., 1993. Structure-based inhibitor design by using protein models for the development of antiparasitic agents. *Proc Natl Acad Sci U S A* 90, 3583-3587.
193. Rittinger, K., Divita, G., Goody, R.S., 1995. Human immunodeficiency virus reverse transcriptase substrate induced conformational changes and the mechanism of inhibition by nucleoside inhibitors. *Proceedings of the National Academy of Sciences of the United States of America* 92, 8046-8049.
194. Roquebert, B., Marcelin, A.G., 2008. The involvement of HIV-1 RNase H in resistance to nucleoside analogues. *J Antimicrob Chemother* 61, 973-975.
195. Roquebert, B., Wirden, M., Simon, A., Deval, J., Katlama, C., Calvez, V., Marcelin, A., 2007. Relationship Between Mutations in HIV-1 RNase H Domain and Nucleoside Reverse Transcriptase inhibitors Resistance Mutations in Naive and Pre-Treated HIV infected Patients. *Journal of medical Virology* 79, 207-211.
196. Saeng-aroon, S., Tsuchiyab, N., Auwanita, W., Isarangura Na Ayuthayaa, P., Pathipvanich, P., Sawanpanyalerta, P., Rojanawiwata, A., Kannagid, M., Ariyoshib, K., Sugiura, W., 2010. Drug-resistant mutation patterns in CRF01 AE cases that failed d4T + 3TC + nevirapine fixed-dosed, combination treatment: Follow-up study from the Lampang cohort. *Antiviral Research* 87, 22-29.
197. Sali, A., Blundell, T.L., 1993. Comparative protein modelling by satisfaction of spatial restraints. *J Mol Biol.* 234, 779-715.
198. Santos, A., Lengruher, R., Soares, E., Jere, A., Sprinz, E., Martinez, A., Silveira, J., Sion, F., Pathak, V., Soares, M., 2008. Conservation Patterns of HIV-1 RT Connection and RNase H Domains: Identification of New Mutations in NRTI-Treated Patients. *PLoS ONE* 3, e1781.
199. Santos, A., Silveira, J., Muniz, C., Tornatore, M., Goes, L., Mendoza-Sassi, R., Martinez, A., Tupinambas, U., Greco, D., Soares, M., 2011. Primary HIV-1 drug resistance in the C-terminal domains of viral reverse transcriptase among drug-naive patients from Southern Brazil. *Journal of Clinical Virology* 52, 373-376.

200. Santos, A., Soares, M., 2010. HIV Genetic Diversity and Drug Resistance. *Viruses* 2, 503-531.
201. Sarafianos, S.G., Clark, A.D., Jr., Das, K., Tuske, S., Birktoft, J.J., Ilangkumaran, P., Ramesha, A.R., Sayer, J.M., Jerina, D.M., Boyer, P.L., Hughes, S.H., Arnold, E., 2002. Structures of HIV-1 reverse transcriptase with pre- and post-translocation AZTMP-terminated DNA. *EMBO J* 21, 6614-6624.
202. Sarafianos, S.G., Marchand, B., Das, K., Himmel, D.M., Parniak, M.A., Hughes, S.H., Arnold, E., 2009. Structure and function of HIV-1 reverse transcriptase: molecular mechanisms of polymerization and inhibition. *Journal of molecular biology* 385, 693-713.
203. Schader, S., Oliveira, M., Ibanescu, R., Moisi, D., Colby-Germinario, S., Wainberg, M., 2012. In Vitro Resistance Profile of the Candidate HIV-1 Microbicide Drug Dapivirine. *Antimicrobial Agents and Chemotherapy* 56, 751-756.
204. Schmit, J.C., Cogniaux, J., Hermans, P., Van Vaeck, C., Sprecher, S., Van Remoortel, B., Witvrouw, M., Balzarini, J., Desmyter, J., De Clercq, E., Vandamme, A.M., 1996. Multiple drug resistance to nucleoside analogues and nonnucleoside reverse transcriptase inhibitors in an efficiently replicating human immunodeficiency virus type 1 patient strain. *J. Infect. Dis.* 174, 962-968.
205. Schuckmann, M.M., Marchand, B., Hachiya, A., Kodama, E.N., Kirby, J.R., Singh, K.P., Sarafianos, S.G., 2010. The N348I Mutation at the Connection Subdomain of HIV-1 Reverse Transcriptase Decrease Binding to Nevirapine. *J.Biol.Chem* 285, 38700-38709.
206. Schultz, S.J., Champoux, J.J., 2008. RNase H activity: structure, specificity, and function in reverse transcription. *Virus Res* 134, 86-103.
207. Schupbach, J., Popovic, M., Gilden, R.V., Gonda, M.A., Sarngadharan, M.G., Gallo, R.C., 1984. Serological analysis of a subgroup of human T-lymphotropic retroviruses (HTLV-III) associated with AIDS. *Science* 224, 503-505.

208. Shafer, R., Rhee, S., Pillay, D., Miller, V., Sandstrom, P., Schapiro, J., Kuritzkese, D., Bennett, D., 2007. HIV-1 protease and reverse transcriptase mutations for drug resistance surveillance. *AIDS* 21, 215-223.
209. Shafer, R.W., Iversen, A.K., Winters, M.A., Aguiniga, E., Katzenstein, D.A., Merigan, T.C., 1995. Drug resistance and heterogeneous long-term virologic responses of human immunodeficiency virus type 1-infected subjects to zidovudine and didanosine combination therapy. *J. Infect. Dis.* 172, 70-78.
210. Shahriar, R., Rhee, S.Y., Liu, T.F., Fessel, W.J., Scarsella, A., Towner, W., Holmes, S.P., Zolopa, A.R., Shafer, R.W., 2009. Nonpolymorphic human immunodeficiency virus type 1 protease and reverse transcriptase treatment-selected mutations. *Antimicrob Agents Chemother* 53, 4869-4878.
211. Shaw, G.M., Hahn, B.H., Arya, S.K., Groopman, J.E., Gallo, R.C., Wong-Staal, F., 1984. Molecular characterization of human T-cell leukemia (lymphotropic) virus type III in the acquired immune deficiency syndrome. *Science* 226, 1165-1171
212. Shortle, D., DiMaio, D., Nathans, D., 1981. Directed Mutagenesis. *Annual Review of Genetics* 15, 265-294.
213. Simon, V., Padte, N., Murray, D., Vanderhoeven, J., Wrin, T., Parkin, N., Di Mascio, M., Markowitz, M., 2003. Infectivity and Replication Capacity of Drug-Resistant Human Immunodeficiency Virus Type 1 Variants Isolated during Primary Infection. *Journal of Virology* 77, 7736-7745.
214. Sluis-Cremer, N., 2014. The Emerging Profile of Cross-Resistance among the Nonnucleoside HIV-1 Reverse Transcriptase Inhibitors. *Viruses* 6, 2960-2973.
215. Sluis-Cremer, N., Sheen, C.W., Zelina, S., Torres, P.S., Parikh, U.M., Mellors, J.W., 2007. Molecular mechanism by which the K70E mutation in human immunodeficiency virus type 1 reverse transcriptase confers resistance to nucleoside reverse transcriptase inhibitors. *Antimicrob Agents Chemother* 51, 48-53.
216. Spence, R.A., Kati, W.M., Anderson, K.S., Johnson, K.A., 1995a. Mechanism of inhibition of HIV-1 reverse transcriptase by nonnucleoside inhibitors. *Science* 267, 988-993.

217. Spence, R.A., Kati, W.M., Anderson, K.S., Johnson, K.A., 1995b. Mechanism of Inhibition of HIV-1 Reverse Transcriptase by Nonnucleoside Inhibitors. *Science* 267, 988-998.
218. Stumptner-Cuvelette, P., Morchoisne, S., Dugast, M., Le Gall, S., Raposo, G., Schwartz, O., Benaroch, P., 2001. HIV-1 Nef impairs MHC class II antigen presentation and surface expression. *Proc Natl Acad Sci U S A* 98, 12144-12149.
219. Sutcliffe, M.J., Hayes, F.R., Blundell, T.L., 1987. Knowledge based modelling of homologous proteins, Part II: Rules for the conformations of substituted sidechains. *Protein Eng* 1, 385-392.
220. Sutherland, K.A., Parry, C.M., McCormick, A., Kapaata, A., Lyagoba, F., Kaleebu, P., Gilks, C.F., Goodall, R., Spyer, M., Kityo, C., Pillay, D., Gupta, R.K., Group, D.V., 2015. Evidence for Reduced Drug Susceptibility without Emergence of Major Protease Mutations following Protease Inhibitor Monotherapy Failure in the SARA Trial. *PLoS One* 10, e0137834.
221. Tachedjian, G., Orlova, M., Sarafianos, S.G., Arnold, E., Goff, S.P., 2001. Nonnucleoside reverse transcriptase inhibitors are chemical enhancers of dimerization of the HIV type 1 reverse transcriptase. *Proc Natl Acad Sci U S A* 98, 7188-7193.
222. Tambuyzer, L., Azijn, H., Rimsky, L.T., Vingerhoets, J., Lecocq, P., Kraus, G., Picchio, G., de Béthune, M., 2009. Compilation and prevalence of mutations associated with resistance to non-nucleoside reverse transcriptase inhibitors. *Antiviral therapy* 14, 103-109.
223. Tambuyzer, L., Nijs, S., Daems, B., Picchio, G., Vingerhoets, J., 2011. Effect of Mutations at Position E138 in HIV-1 Reverse Transcriptase on Phenotypic Susceptibility and Virologic Response to Etravirine. *J Acquir Immune Defic Syndr* 58, 18-22.
224. Tanuma, J., Hachiya, A., Ishigaki, K., Gatanaga, H., Thi Minh Lien, T., Duc Hien, N., Van Kinh, N., Kaku, M., Oka, S., 2010. Impact of CRF01_AE-specific polymorphic mutations G335D and A371V in the connection subdomain of human immunodeficiency virus type 1 (HIV-1) reverse transcriptase (RT) on susceptibility to nucleoside RT inhibitors. *Microbes and Infection* 12, 1170-1177.

225. Telesnitsky, A., Goff, S., 1997. Retroviruses. Cold Spring Harbor Laboratory Press: Cold Spring Harbor NY, USA, 121-160.
226. Theys, K., Deforche, K., Libin, P., Camacho, R.J., Van Laethem, K., Vandamme, A.M., 2010. Resistance pathways of human immunodeficiency virus type 1 against the combination of zidovudine and lamivudine. *Journal of General Virology* 91, 1898–1908.
227. Trono, D., 1995. HIV accessory proteins: leading roles for the supporting cast. *Cell* 82, 189-192.
228. Tsibris, A.M.N., Hirsch, M.S., 2010. Antiretroviral Therapy in the Clinic. *J Virol.* 84, 5458–5464.
229. Tu, X., Das, K., Han, Q., Bauman, J.D., Clark Jr, A.D., Hou, X., Frenkel, Y.V., Gaffney, B.L., Jones, R.A., Boyer, P.L., Hughes, S.H., Sarafianos, S.G., Arnold, E., 2010. Structural basis of HIV-1 resistance to AZT by excision. *Nat Struct Mol Biol.* 17, 1202-1209.
230. UNAIDS, 2016. Global AIDS Update 2016.
231. Vodicka, M.A., Koepp, D.M., Silver, P.A., Emerman, M., 1998. HIV-1 Vpr interacts with the nuclear transport pathway to promote macrophage infection. *Genes Dev* 12, 175-185.
232. von Wyl, V., Ehteshami, M., Demeter, L., Burgisser, P., Nijhuis, M., Symons, J., Yerly, S., Boni, J., Klimkait, T., Schuurman, R., Ledergerber, B., Gottel, M., Günthard, H., Study, a.t.S.H.C., 2010. HIV-1 Reverse Transcriptase Connection Domain Mutations: Dynamics of Emergence and Implications for Success of Combination Antiretroviral Therapy. *Clinical Infectious Diseases* 51, 620-628.
233. Vyas, V.K., Ukawala, R.D., Ghate, M., Chintla, C., 2012. Homology Modeling a Fast Tool for Drug Discovery: Current Perspectives. *Indian J Pharm Sci* 74, 1-17.
234. Wainberg, M.A., Kendall, O., Gilmore, N., 1988. Vaccine and antiviral strategies against infections caused by human immunodeficiency virus. *CMAJ* 138, 797-807.

235. Wang, L.Z., Kenyon, G.L., Johnson, K.A., 2004. Novel mechanism of inhibition of HIV-1 reverse transcriptase by a new non-nucleoside analog, KM-1. *J Biol Chem* 279, 38424-38432.
236. Weng, Y., Zhang, L., Huang, J., Zhao, J., Luo, P., Bi, S., Yang, Z., Zhu, H., Allain, J.P., Li, C., 2016. A simple and cost-saving phenotypic drug susceptibility testing of HIV-1. *Sci Rep.* 19, :33559.
237. Whitcomb, J.M., Huang, W., Limoli, K., Paxinos, E., Wrin, T., Skowron, G., Deeks, S.G., Bates, M., Hellmann, N.S., Petropoulos, C.J., 2002. Hypersusceptibility to non-nucleoside reverse transcriptase inhibitors in HIV-1: clinical, phenotypic and genotypic correlates. *Aids* 16, F41-47.
238. Wu, H., Zhang, H.J., Zhang, X.M., Xu, H.F., Wang, M., Huang, J.D., Zheng, B.J., 2012. Identification of drug resistant mutations in HIV-1 CRF07_BC variants selected by nevirapine in vitro. *PLoS One* 7, e44333.
239. Wyatt, R., Kwong, P.D., Desjardins, E., Sweet, R.W., Robinson, J., Hendrickson, W.A., Sodroski, J.G., 1998. The antigenic structure of the HIV gp120 envelope glycoprotein. *Nature* 393, 705-711.
240. Xia, Q., Radzio, J., Anderson, K.S., Sluis-Cremer, N., 2007. Probing nonnucleoside inhibitor-induced active site distortion in HIV-1 reverse transcriptase by transient kinetic analyses. *Protein science : a publication of the Protein Society* 16, 1728-1737.
241. Xu, H., Colby-Germinario, S., Asahchop, E., Oliveira, M., McCallum, M., Schader, S., Han, Y., Quan, Y., Sarafianos, S., Wainberg, M., 2013. Effect of Mutations at Position E138 in HIV-1 Reverse Transcriptase and Their Interactions with the M184I Mutation on Defining Patterns of Resistance to Nonnucleoside Reverse Transcriptase Inhibitors Rilpivirine and Etravirine. *Antimicrobial Agents and Chemotherapy* 57, 3100-3109.
242. Xu, H., Colby-Germinario, S., Oliveira, M., Han, Y., Quan, Y., Zanichelli, V., Wainberg, M., 2014a. The Connection Domain Mutation N348I in HIV-1 Reverse Transcriptase Enhances Resistance to Etravirine and Rilpivirine but Restricts the Emergence of the E138K Resistance Mutation by Diminishing Viral Replication Capacity. *Journal of Virology* 88, 1536-1547.

243. Xu, H., Colby-Germinario, S.P., Oliveira, M., Rajotte, D., Bethell, R., Wainberg, M.A., 2014b. Effects of the W153L Substitution in HIV Reverse Transcriptase on Viral Replication and Drug Resistance to Multiple Categories of Reverse Transcriptase Inhibitors. *Antimicrobial agents and chemotherapy* 58 4515–4526.
244. Xu, H., Quan, Y., Schader, S., Oliveira, M., Bar-Magen, T., Wainberg, M., 2010. The M230L Nonnucleoside Reverse Transcriptase Inhibitor Resistance Mutation in HIV-1 Reverse Transcriptase Impairs Enzymatic Function and Viral Replicative Capacity. *ANTIMICROBIAL AGENTS AND CHEMOTHERAPY* 6, 2401-2408.
245. Yang, A.S., Honig, B., 1999. Sequence to structure alignment in comparative modelling using PriSM. *Proteins* 37, 66-72.
246. Yap, S.H., Sheen, C., Fahey, J., Zanin, M., Tyssen, D., Lima, V.D., Wynhoven, B., Kuiper, M., Sluis-Cremer, N., Harrigan, P.R., Tachedjian, G., 2007. N348I in the Connection Domain of HIV-1 Reverse Transcriptase confers Zidovudine and Nevirapine. *Plos One* 4, e335.
247. Zack, J.A., Arrigo, S.J., Weitsman, S.R., Go, A.S., Haislip, A., Chen, I.S., 1990. HIV-1 entry into quiescent primary lymphocytes: molecular analysis reveals a labile, latent viral structure. *Cell* 61, 213-222.
248. Zelina, S., Sheen, C., Radzio, J., Mellors, J.W., Sluis-Cremer, N., 2008. Mechanisms by Which the G333D Mutation in Human Immunodeficiency Virus Type 1 Reverse Transcriptase Facilitates Dual Resistance to Zidovudine and Lamivudine. *Antimicrobial Agents And Chemotherapy* 52, 157-163.
249. Zhang, H., Zhou, Y., Alcock, C., Kiefer, T., Monie, D., Siliciano, J., Li, Q., Pham, P., Cofrancesco, J., Persuad, D., Siliciano, R.F., 2004. Novel single-cell-level phenotypic assay for residual drug susceptibility and reduced replication capacity of drug-resistant human immunodeficiency virus type 1. *J. Virol.* 78, 1718-1729.
258. Zhang, J., Rhee, S.Y., Taylor, J., Shafer, R.W., 2005. Comparison of the precision and sensitivity of the Antivirogram and PhenoSense HIV drug susceptibility assays. *J Acquir Immune Defic Syndr* 38, 439-444.

250. Zhang, X., Wu, H., Zhang, Q., Lau, T.C., Chu, H., Chen, Z., Jin, D., Zheng, B., 2015a. A novel mutation, D404N, in the connection subdomain of reverse transcriptase of HIV-1 CRF08_BC subtype confers cross-resistance to NNRTI. *Journal of antimicrobial chemotherapy*, 1-10.
251. Zhang, X., Wu, H., Zhang, Q., Lau, T.C., Chu, H., Chen, Z., Jin, D., Zheng, B., 2015b. A novel mutation, D404N, in the connection subdomain of reverse transcriptase of HIV-1 CRF08_BC subtype confers cross-resistance to NNRTIs. *Journal of Antimicrobial Chemotherapy* 70, 1381-1390.
252. Zheng, L., Baumann, U., Reymond, J.L., 2004. An efficient one-step site-directed and site-saturation mutagenesis protocol. *Nucleic Acids Res* 32, e115.
253. Zheng, Y.H., Lovsin, N., Peterlin, B.M., 2005. Newly identified host factors modulate HIV replication. *Immunol Lett* 97, 225-234.