

https://dspace.org/

Development of an HIV-1 intergrase enzyme strand transfer assay

Fish, Muhammad Qasim

2012-01-30
http://hdl.handle.net/10539/11197

Downloaded from a DSpace repository. https://dspace.org/

Development of an HIV-1 Integrase enzyme strand transfer assay

A dissertation submitted to the Faculty of Health Sciences, University of the Witwatersrand, in fulfillment of the requirements for the degree

of

Master of Science in Medicine

Declaration

I, Muhammad Qasim Fish declare that this dissertation is my own work. It is being submitted for the degree of Master of Science in the University of the Witwatersrand, Johannesburg. It has not been submitted before for any degree or examination at this or any other University.

……

……day of ……, 2010

Dedication

To my family

Publications and Conference Proceedings

Publications in DE accredited journals

Fish M.Q., Hewer R., Wallis C., Venter F., Stevens W. and Papathanasopoulos M.A. (2010). Baseline Integrase Inhibitor Resistance mutations amongst HIV-1 infected South African patients. *AIDS Research and Human Retroviruses*. 26:489-493.

Conference proceedings

Fish M.Q., Hewer R., Wallis C., Venter F., Stevens W. and Papathanasopoulos M.A. Baseline Integrase Inhibitor Resistance mutations amongst HIV-1 infected South African patients. 5th IAS Conference on HIV pathogenesis, treatment and prevention. 19-22 July, 2009. Cape Town, South Africa.

Abstract

The Human Immunodeficiency Virus type 1 (HIV-1) integrase is an essential enzyme required for viral replication. Integrase forms part of an ensemble of proteins known as the preintegration complex and functions by a two-step process. Firstly, the cleaving of the 3' ends of the viral cDNA genome, known as 3'-end processing. The second step is the insertion of these ends into host DNA by esterification, known as strand transfer. There is no human homologue to integrase which makes it an ideal drug target. However, the strand transfer inhibitor raltegravir is currently the only antiretroviral treatment available that inhibits integrase. The aims of this study were two-fold: firstly to characterise a cohort of South African patients so as to determine the viability of introducing raltegravir as a new treatment option, and secondly, to set up high-throughput integrase inhibitor screening assays (testing integrase enzymatic functionality). An HIV-1 subtype C specific RT-PCR and PCR assay was established for integrase genotyping using 51 integrase inhibitor-naïve patient plasma samples and 22 antiretroviral drug-naive primary viral isolates from South Africa. Seventy-one of the 73 samples were classified as HIV-1 subtype C and two samples were unique AC and CG recombinants in integrase. Amino acid sequence analysis revealed there were no primary mutations (Y143R/C/H, Q148H/R/K, and N155H/S) associated with reduced susceptibility to the integrase inhibitor raltegravir. However, one sample had the T97A mutation, three samples had the E157Q and V165I mutations, and the majority of samples contained the polymorphic mutation, V72I. The expected finding of no major integrase mutations conferring resistance to integrase inhibitors suggests that this new antiretroviral drug class will be effective in our region where HIV-1 subtype C predominates. However, the impact of E157Q and other naturally occurring polymorphisms warrants further phenotypic investigation. The integrase sequence of viral isolate, FV3, was closest to the consensus sequence, and thus chosen for preintegration

complex isolation for use in strand transfer assays. Isolation of preintegration complexes following FV3 infections of several cell lines was unsuccessful as determined by western blot analysis. Subsequently, the focus was changed to isolation of HIV-1 subtype B recombinant integrase and its functional evaluation. Expression of native integrase (IN_{wt}) and soluble integrase (IN_{sol}) was induced in *E. coli,* and both proteins were purified by nickel chelating chromatography. The purified recombinant proteins were used to develop three assays to test for strand transfer activity, of which two were successfully established. Furthermore, only IN_{sol} showed strand transfer activity in the high-throughput microtitre plate assay and scintillation proximity assay (SPA)-bead strand transfer assay. Activity of IN_{sol} was shown to be inhibited by the control compound, chicoric acid with an IC_{50} of 101.5nM in the high-throughput microtitre plate assay, whereas IN_{sol} activity as well as a dose response to chicoric acid with an IC_{50} of 248.5nM was recorded in the SPA-bead strand transfer assay. Visualization of radiolabelled enzymatic products of strand transfer by polyacrylamide gel electrophoresis of urea sequencing gels was unsuccessful. Overall, the high-throughput microtitre plate and SPA-bead strand transfer assays have been successfully established in our laboratories, and are available to screen compound libraries for potential antiretroviral drug candidates targeting integrase strand transfer.

Acknowledgements

My Gratitude goes to the All Wise, Knower of all things.

I thank my supervisor, Dr Maria Papathanasopoulos, especially for the great patience she has shown and all the guidance and help during the course of my studies.

I thank my co-supervisor Dr Raymond Hewer for the helpful discussions and great assistance during the course of the study.

Many thanks to all in the HIV pathogenesis lab, especially Nicole Cerutti, Naazneen Moolla and Mark Killick for all their immense assistance and many helpful suggestions during the study.

I thank Dr Salerwe Mosebi, for the great assistance with the functional assays, and insight on many practical aspects of the study.

I send thanks to Mintek, the Department of Molecular Medicine and Haematology and the Poliomyelitis Research Foundation for funding and granting me this opportunity.

To my parents, for all their love and support.

Table of contents

PAGE

DECLARATION....II

DEDICATION....III

PUBLICATIONS AND CONFERENCE PROCEEDINGS....IV

ABSTRACT....V

ACKNOWLEDGEMENTS....VII

LIST OF FIGURES....XI

LIST OF TABLES....XV

LIST OF ABBREVIATIONS....XVI

CHAPTER 1....1

1. INTRODUCTION....2

1.1. HIV: A General Background....2

1.2. HIV-1 antiretroviral drug targets, and highly active antiretroviral therapy (HAART)....4

1.3. The HIV-1 Integrase enzyme and Integration....6

1.4. The Preintegration Complex and Host Factors Influencing Integration....8

1.5. HIV-1 Integrase Inhibitors and resistance....11

1.6. HIV-1 Integrase Activity Assays....13

1.7. HIV-1 infection in South Africa....15

CHAPTER 2..... 18

2. MATERIALS AND METHODS..... 19

2.1. Patient Samples and Reagents used In This Study 19

2.2. Amplification and sequence analysis of the HIV-1 integrase.....21

2.2.1. Proviral DNA isolation from infected co-cultures 21

2.2.2. Viral RNA isolation from patient plasma.....21

2.2.3. Reverse Transcriptase (RT) Polymerase chain reaction (PCR).....22

2.2.4. Conventional PCR.....22

2.2.5. PCR Fragment Purification and Visualization 24

2.2.6. DNA Sequencing of HIV-1 *integrase*.....24

2.3. Mammalian Cell Culture.....26

2.4. HIV-1 Preintegration Complex Isolation 26

2.4.1. HIV-1 subtype C Infections.....26

2.4.2. Crude Extraction of HIV-1 PICs.....27

2.5. HIV-1 Recombinant Integrase Purification.....28

2.5.1. Preparation of competent *Escherichia (E.) coli* 28

2.5.2. Bacterial Transformations.....28

2.5.3. Gene Induction.....29

2.5.4. Bacterial Cell lysis and Fraction isolation 29

2.5.5. Nickel Chelating Column Chromatography.....30

2.5.6. Sodium Dodecyl Sulphate–Polyacrylamide Gel Electrophoresis (SDS-PAGE) and Western Blotting 31

2.6. Functional analysis of purified recombinant IN 32

2.6.1. High throughput microtitre plate HIV-1 integration assay 32

2.6.2. HIV-1 Integrase Urea-PAGE Based Functional Analysis.....34

2.6.3. HIV-1 Integrase Scintillation proximity assay Strand Transfer Assay 35

2.6.3.1. Preparation of SPA Bead complex 35

2.6.3.2. Radiolabelling of DNA.....36

2.6.3.3. SPA bead Strand Transfer Assay set up 37

2.6.4. Data analysis 37

CHAPTER 3 ..... **38**

3. RESULTS **39**

3.1. Amplification and sequencing of HIV-1subtype C integrase **39**

3.2. HIV-1 subtype C cell line infections, and detection of Preintegration Complexes **45**

3.3. HIV-1 subtype B recombinant Integrase Protein Isolation **45**

3.4. Recombinant IN and IN_{sol} activity **50**

3.4.1. High throughput microtitre plate HIV-1 integration assay 50

3.4.2. HIV-1 Integrase Urea-PAGE Based Functional Analysis 53

3.4.3. SPA bead Strand Transfer Assay 55

CHAPTER 4 ..... **57**

4. DISCUSSION **58**

CHAPTER 5 ..... **69**

5. REFERENCES **70**

CHAPTER 6 ..... **76**

6. SUPPLEMENTARY **77**

List of Figures

PAGE

Figure 1.1: Genetic organization of the HIV-1 genome (approximately 9.7kb). It contains 9 genes which code for at least 16 proteins, flanked by identical LTRs. The structural genes *gag, pol* and *env* encode for the Gag, Polymerase and Envelope glycoproteins, respectively. The Gag-Pol polyprotein is cleaved into the matrix (MA), capsid (CA), nucleocapsid (NC), P6, protease (PR), reverse transcriptase (RT) and integrase (IN). The regulatory genes *tat* and *rev* encode the Transcription transactivator Tat and regulatory factor Rev, respectively. The accessory genes *vif, vpr, vpu* and *nef* encode the viral infectivity factor, viral protein R, viral protein U and Negative factor, respectively. Adapted from (Sherman and Greene 2002). 3

Figure 1.2: Representative structure of the HIV-1 virion and associated structural and core proteins. Adapted from (Sherman and Greene 2002). 3

Figure 1.3: Typical life cycle of HIV-1 showing entry, reverse transcription, integration, transcription, translation, assembly and budding. Adapted from (Ciuffi and Bushman 2006). 5

Figure 1.4: Schematic of the domain structure of the HIV-1 Integrase (IN) enzyme. The functions and specific characterization of the domains are listed. Adapted from (McColl and Chen 2009). 7

Figure 1.5: Schematic of the HIV-1 integration reaction. The tetrameric integrase enzyme (IN) bound to viral cDNA generates a CA overhang in a process called 3'-end processing in the cytoplasm (i). The preintegration complex moves into the nucleus where the strand transfer reaction is initiated by binding Target DNA (ii), followed by nucleophilic attack and phosphodiester bond formation leaving a 5' DNA flap (iii), DNA repair via host mechanisms, resulting in an integrated provirus (iv). Adapted from (Dar *et al.* 2009). 9

Figure 1.6: Crystal structure of the prototype foamy virus IN active site in committed and drug-bound states showing IN without drug (i) and bound to RAL (ii) or ELV (iii). The top panels show protein and DNA cartoons, DNA bases and the side chains of indicated amino acids as sticks. The RAL and ELV drug atoms are colored: yellow, C; blue, N; red, O; grey, F; green, Cl. The bottom panels show the complex as a solvent-accessible surface, colored by atoms (light grey, C; red, O; blue, N). Grey spheres are Mn21 (i) or Mg21 (ii, iii) ions. Adapted from (Hare *et al.* 2010). 13

Figure 1.7: Schematic of a high throughput microtitre plate HIV-1 integration assay showing donor DNA immobilized on the well surface (i). After addition of IN, assembly and 3'-end processing is initiated (ii), followed by addition of labelled Target DNA leading to strand

transfer (iii). Successful strand transfer is detected by relevant detection methods. Picture drawn using CorelDraw version 12.0. 15

Figure 3.1: A representative 1% agarose gel showing *integrase* PCR products amplified from extracted proviral DNA. The names of the proviral isolates are indicated above the wells. A 1kb DNA size marker (New England Biolabs) is indicated on the left. The arrow indicates the approximately 1.3kb PCR amplicon. 40

Figure 3.2: A representative 1% agarose gel showing *integrase* RT-PCR products amplified from extracted viral RNA. The patient numbers are indicated above the wells. A 1kb DNA size marker (New England Biolabs) is indicated on the left. The arrow indicates the approximately 1.3kb PCR amplicon. 41

Figure 3.3: Phylogenetic relationships of the 73 newly characterized full-length integrase sequences with HIV-1 subtype reference sequences from the Los Alamos database (http://hiv-web.lanl.gov). The unique recombinant sequences are shown in red. Phylogenetic trees were constructed from nucleotide sequence alignments, using neighbor joining and the Kimura two-parameter distance matrix. The reliability of the branching order was estimated from 1000 bootstrap replicates, and only bootstrap values of 70% or higher are shown. 42

Figure 3.4: RIP 3.0 analysis of 09ZAP52 (top) and 09ZAP35 (bottom). The program compares the query sequence to a database of HIV-1 subtypes and selected recombinants and plots each nucleotide (x axis) to a percentage similarity (y axis) of all subtypes in the database. 09ZAP52 is shown to be a GC recombinant with part of its sequence not generally clear but possibly holding similarities to an AE recombinant around nucleotides 570 to 670. 09ZAP35 is shown to be an AC recombinant but also has high similarities to an AE recombinant between nucleotides 600 and 730. 43

Figure 3.5: Distribution of variants amongst the 73 newly characterized full-length integrase sequences. The consensus subtype C sequence is shown at the top of each 30 amino acid sequence section, below which is the number of annotated sequences containing an unambiguous amino acid at the indicated position. Beneath are the variants with the superscript numbers indicating the amount of times that they occurred. Domains and important residues are highlighted. The N-terminal domain is highlighted in yellow, the catalytic core domain in white and the C-terminal domain in blue. The catalytic triad residues are indicated in red and relevant integrase inhibitor resistance mutations are in green. 44

Figure 3.6: OD readings of the p24 antigen ELISA results showing infection of MT2 cells and Supt1 with the FV3 viral isolate. OD readings of samples on day 0 and day 13 were taken. The key indicates Cell lines infected. Levels of p24 were above the upper limit of OD reading of 3 on day 13. 46

Figure 3.7: SDS-PAGE (A) and corresponding Western Blot (B) of induced BL21 pINSD.His cultures. Samples were taken at 1 hour intervals up to 3 hours after 1mM IPTG was added to the bacterial culture. Protein molecular weight markers (in kDa; kilodalton) are indicated on the left. The arrow indicates the expressed IN protein. 47

Figure 3.8: Western Blot analysis of induced *E. coli* BL21 pINSD.His bacterial samples. Lanes demarcated as Supernatant and Pellet were samples from *E. coli* BL21 pINSD.His that were not sonicated prior to centrifugation. Lanes demarcated as Lysate Supernatant and Lysate Pellet were samples taken after sonication and centrifugation. The Pellet sample indicates total protein (soluble and insoluble). The arrow indicates the expressed IN protein (approximately 34.5kDa). 48

Figure 3.9: SDS-PAGE Gel (A) and corresponding Western Blot (B) of IN protein purification by nickel chelating column chromatography. This sample was eluted with 350mM imidazole. Protein molecular weight markers (in kDa; kilodalton) are indicated on the left. The arrow indicates the expressed and purified IN protein. 48

Figure 3.10: SDS-PAGE (left) and corresponding Western Blot (right) of IN protein purified by nickel chelating column chromatography. This sample was eluted with 500mM imidazole. Protein molecular weight markers (in kDa; kilodalton) are indicated on the left. The arrow indicates the expressed and purified IN protein. 49

Figure 3.11: SDS-PAGE analysis of the purification of IN_{sol}. The enzyme is clearly noticed in the 24^{th} and 25^{th} fractions, however, contamination is noticed in these fractions. The protein eluted between fractions 23 and 60. The arrow indicates the purified IN protein. 50

Figure 3.12 : HIV-1 recombinant integrase activity from the high throughput microtitre plate HIV-1 integration assay, as measured in arbitrary fluorescence units (excitation 485nm and emission 535nm). A) Purified recombinant IN activity results using only a single Target DNA concentration (150nM). The blue, yellow and red bars indicate the fluorescence readings for the negative control (no-IN), test containing 1 µM of IN_{sol}, and test containing 1µM of IN_{wt}, respectively, showing that IN_{sol} is active. All tests were done in duplicate. B) Dose response of IN_{sol} at different ranges of Target DNA concentrations. Results were normalized to a negative control containing no IN at maximum Target DNA concentration (270nM), and show that IN_{sol} activity increases as the Target DNA concentration increases. All tests were done in duplicate. 51

Figure 3.13: HIV-1 recombinant integrase activity from the modified high throughput microtitre plate HIV-1 integration assay, as measured by absorbance (405nm). Results show a definitive difference between IN_{sol} and the blank control. Heights of the columns indicate an average mean of a triplicate. 52

Figure 3.14: HIV-1 recombinant integrase (IN) activity from the modified high throughput microtitre plate HIV-1 integration assay, as measured by absorbance (405nm) and

normalized to a positive control (%). As the concentration of the chicoric acid inhibitor increases, the recombinant IN activity decreases. Points indicate average mean of a triplicate. IC_{50} is calculated as 101.5nM on GraphPad Prism 5.1. 53

Figure 3.15: Representative autoradiography of the Urea-Gel based assay using Assay Buffer 2. Labelled Target DNA and Donor (32 mer) DNA were mixed with IN_{sol} from purification A (lane 1), IN_{sol} from purification B (lane 2), or a no IN enzyme control (lane 3). Labelled Donor DNA only and labelled Target DNA only controls are shown in lanes 4 and 5, respectively. 54

Figure 3.16: Representative autoradiography of the Urea-Gel based assay depicting use of Assay Buffer 1. Lanes are indicated. Labelled Target DNA and labelled Donor 1 (Table 2.2) DNA were mixed with IN_{sol} from purification A (lane 1), IN_{sol} from purification B (lane 2), or a no IN enzyme control (lane 3). Labelled Donor DNA only and labelled Target DNA only controls are shown in lanes 4 and 5, respectively. Donor DP (Table 2.2) concentration was 60nM and target concentration was 100nM. 54

Figure 3.17: SPA bead assay activity results. The bead assay results indicate that there is ST activity when comparing the IN_{sol} to the Blank control. 55

Figure 3.18: SPA bead assay with increasing inhibitor concentration. As the concentration of the chicoric acid inhibitor increases, the recombinant IN_{sol} activity decreases. Points indicate the average mean of a duplicate. IC_{50} is calculated as 248.5nM on GraphPad Prism 5.1. 56

List of Tables

PAGE

Table 2-1: Viral loads (RNA copies/ml) and gender of the 58 randomly selected patient plasma samples sent for routine viral load monitoring at the Charlotte Maxeke Johannesburg Hospital. 20

Table 2-2: Primers used throughout the course of this study. 23

List of Abbreviations

°C	Degrees Celsius
3'-P	3'-end processing
3TC	Lamivudine
AIDS	Acquired Immunodeficiency Syndrome
ART	Antiretroviral therapy
AZT	Zidovudine
BAF	Barrier-to-autointegration factor
bp	Base pair
CA	Capsid
CCD	Catalytic Core Domain
CCR5	CC Chemokine Receptor 5
CD4	Cluster of Differentiation number 4
cDNA	Complementary Deoxyribonucleic acid
CDR2	Second complementarity-determining region
CRF	Circulating recombinant form
CTD	C-terminal domain
CXCR4	CXC Chemokine Receptor 4
DKA	Diketo acids
DMSO	Dimethylsulfoxide
DNA	Deoxyribonucleic acid
dNTP	deoxy Nucleotide Triphosphate
EDTA	Ethylenediaminetetraacetic acid
EFV	Efavirenz
ELISA	Enzyme-linked immunosorbent assay
ELV	Elvitegravir
Env	Envelope glycoprotein
env	Envelope gene
FCS	Foetal calf serum

FTC	**Emtricitabine**
gp	**Glycoprotein**
HAART	**Highly active antiretroviral therapy**
HIV-1	**Human Immunodeficiency virus type 1**
HIV-2	**Human Immunodeficiency virus type 2**
HMGA1	**High mobility group protein A1**
IN	**Integrase**
INIs	**Integrase inhibitors**
kb	**Kilo base**
kDa	**Kilo Dalton**
LEDGF/p75	**Lens epithelium-derived growth factor/ protein 75**
LTR	**Long Terminal Repeat**
M	**Molar**
MA	**Matrix**
mAbs	**Monoclonal antibodies**
mg	**Milligram**
$MgCl_2$	**Magnesium Chloride**
ml	**Milliliter**
MW	**Molecular weight**
ng	**Nanogram**
nM	**Nanomolar**
NNRTI	**Non-nucleoside analogue reverse transcriptase inhibitor**
NRTI	**Nucleoside analogue reverse transcriptase inhibitor**
NTD	**N-terminal domain**
NVP	**Nevirapine**
OD	**Optical density**
PBS	**Phosphate Buffered Saline**
PCR	**Polymerase chain reaction**
PFV	**Prototype foamy virus**
PI	**Protease Inhibitor**

PIC	**Preintegration complex**
PR	**Protease**
RAL	**Raltegravir**
RPMI	**Roswell Park Memorial Institute Medium**
RT	**Reverse transcriptase**
RTI	**Reverse Transcriptase Inhibitor**
RT-PCR	**Reverse transcription polymerase chain reaction**
sol	**Soluble**
ST	**Strand transfer**
TDF	**Tenofovir**
TNF	**Tumour necrosis factor**
TRAF	**TNF receptor-associated factor**
TRIM	**Tripartite motif protein**
TTRAP	**TNF and TRAF receptor-associated protein**
UNAIDS	**The Joint United Nations Programme on HIV and AIDS**
Vpr	**Viral protein R**
Vpu	**Viral protein U**
wt	**Wild type**
µg	**Microgram**
µM	**Micromolar**

CHAPTER 1

Introduction

1. Introduction

1.1. HIV: A General Background

The Human Immunodeficiency Virus (HIV) belongs to the family *Retroviridae* within the subfamily/genus *Lentivirinae*. Infection with HIV ultimately leads to Acquired Immune Deficiency Syndrome (AIDS) in humans (Barre-Sinoussi *et al.* 1983; Gallo *et al.* 1984). By the end of 2008, there were 33.4 million people infected with HIV-1, with 67% living in sub-Saharan Africa (http:///www.unaids.org). Over 2.7 million people became HIV-1 positive during 2008, and there were over 2 million AIDS-related deaths in the same year. Although there are currently over 25 approved antiretroviral drugs, there is an ongoing need for the discovery and development of novel, cheaper and more accessible antiretroviral drugs, particularly for use in the developing world.

There are two types of HIV, namely HIV-1 and HIV-2. Phylogenetically, HIV-1 can be subdivided into 3 groups, the M (main), O (other) and N (non-M, non-O) group, where the M group is responsible for most HIV-1 infections globally (http://www.hiv.lanl.gov). Within HIV-1 group M there are 9 subtypes denoted with letters (A, B, C, D, F, G, H, J, K) according to their phylogenetic interpretation. Co-infection with more than one strain or subtype of HIV-1 could result in an inter-subtype recombination event and lead to a viral strain with characteristics of more than one subtype. An inter-subtype recombinant virus that gets transmitted and becomes one of the circulating strains in the HIV-1 epidemic is classified as a circulating recombinant form (CRF). To date, 48 CRF's have been described (http://www.hiv.lanl.gov).

The HIV-1 genome consists of 9 viral genes (Figure 1.1) that encode for at least 16 necessary proteins that aid in viral propagation. The structure of a typical HIV-1 virion is shown in Figure 1.2. HIV-1 has a characteristic life cycle as that of retroviruses from the *Lentivirinae* genus. HIV-1 begins its life cycle by interaction with a $CD4^+$ host T cell.

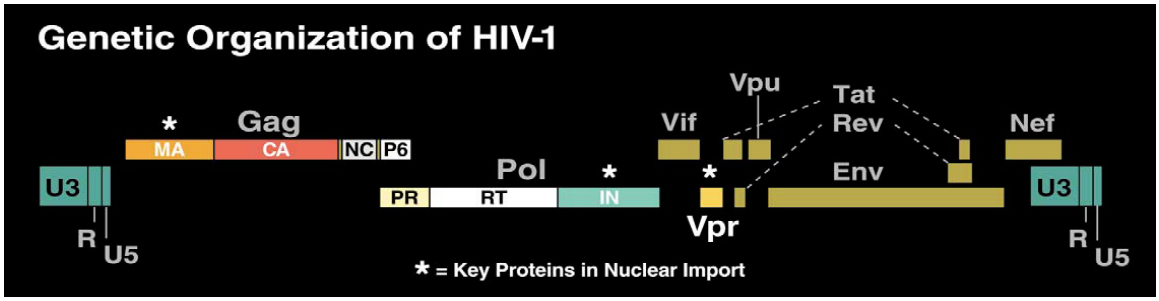


Figure 1.1: Genetic organization of the HIV-1 genome (approximately 9.7kb). It contains 9 genes which code for at least 16 proteins, flanked by identical LTRs. The structural genes *gag, pol* and *env* encode for the Gag, Polymerase and Envelope glycoproteins, respectively. The Gag-Pol polyprotein is cleaved into the matrix (MA), capsid (CA), nucleocapsid (NC), P6, protease (PR), reverse transcriptase (RT) and integrase (IN). The regulatory genes *tat* and *rev* encode the Transcription transactivator Tat and regulatory factor Rev, respectively. The accessory genes *vif, vpr, vpu* and *nef* encode the viral infectivity factor, viral protein R, viral protein U and Negative factor, respectively. Adapted from (Sherman and Greene 2002).

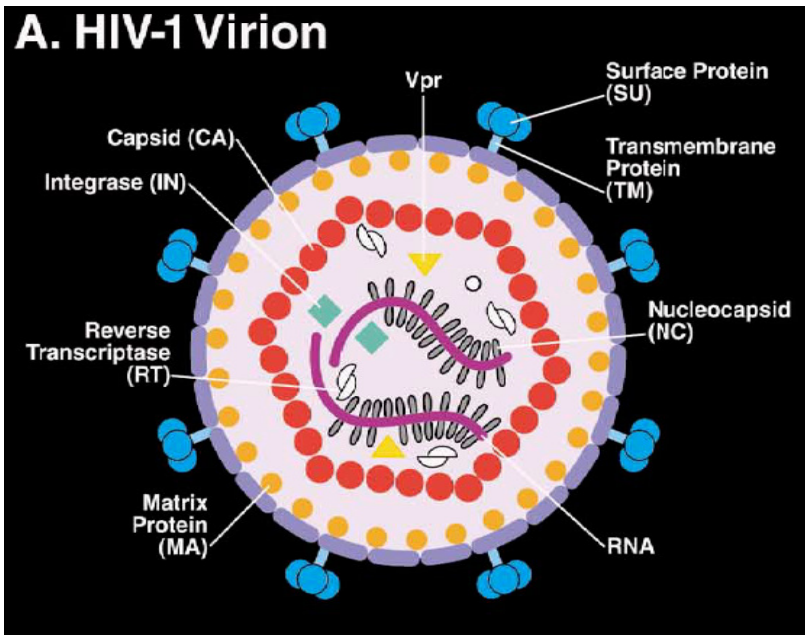


Figure 1.2: Representative structure of the HIV-1 virion and associated structural and core proteins. Adapted from (Sherman and Greene 2002).

The HIV-1 surface glycoprotein, gp120, sequentially interacts with CD4 and the co-receptors CCR5 and/or CXCR4 on the host cell membrane. This interaction exposes the viral gp41 fusion protein, and allows the HIV-1 p24 core to enter the cell either by endocytosis or by fusion of the gp41 with the cell membrane (Goto *et al.* 1997). Once entry is accomplished, the p24 core is uncoated and HIV-1 begins to reverse transcribe its plus strand RNA genome into cDNA by the action of the viral Reverse Transcriptase (RT) enzyme. The viral cDNA is then translocated into the nucleus (in the context of a preintegration complex; PIC) and is integrated into the host chromosome (Haseltine 1991). The integrated viral genome, or provirus, then undergoes transcription mediated with the aid of viral transcriptional factors and host transcriptional factors (Miller *et al.* 2000). Viral mRNA is then translated into polypeptide chains and proteolytically cleaved by viral or host proteases into functional proteins. Envelope glycoproteins are assembled on membrane-bound ribosomes and are transported and fused to the cell surface as membrane bound envelope glycoproteins (Kolegraff *et al.* 2006). The essential viral proteins and viral genomic RNA are transported to the cell membrane where the envelope glycoproteins are present, and a new, immature virion buds out of the cell membrane. Once the viral protease completes full cleavage of the Gag polyprotein, the mature HIV-1 particle is then capable of initiating a new round of infection. Figure 1.3 summarizes the typical life cycle of HIV-1 [for Reviews see (Haseltine 1991; Goto *et al.* 1997; Miller *et al.* 2000; Kolegraff *et al.* 2006)].

1.2. HIV-1 antiretroviral drug targets, and highly active antiretroviral therapy (HAART)

Each step of the HIV-1 life cycle represents a viable antiretroviral drug target, and is the subject of intensive international research. To date, four classes of antiretroviral drugs have been approved for human use by the Food and Drug Administration of the United States. These classes include inhibitors which target the viral RT, protease and integrase enzymes, and viral entry.

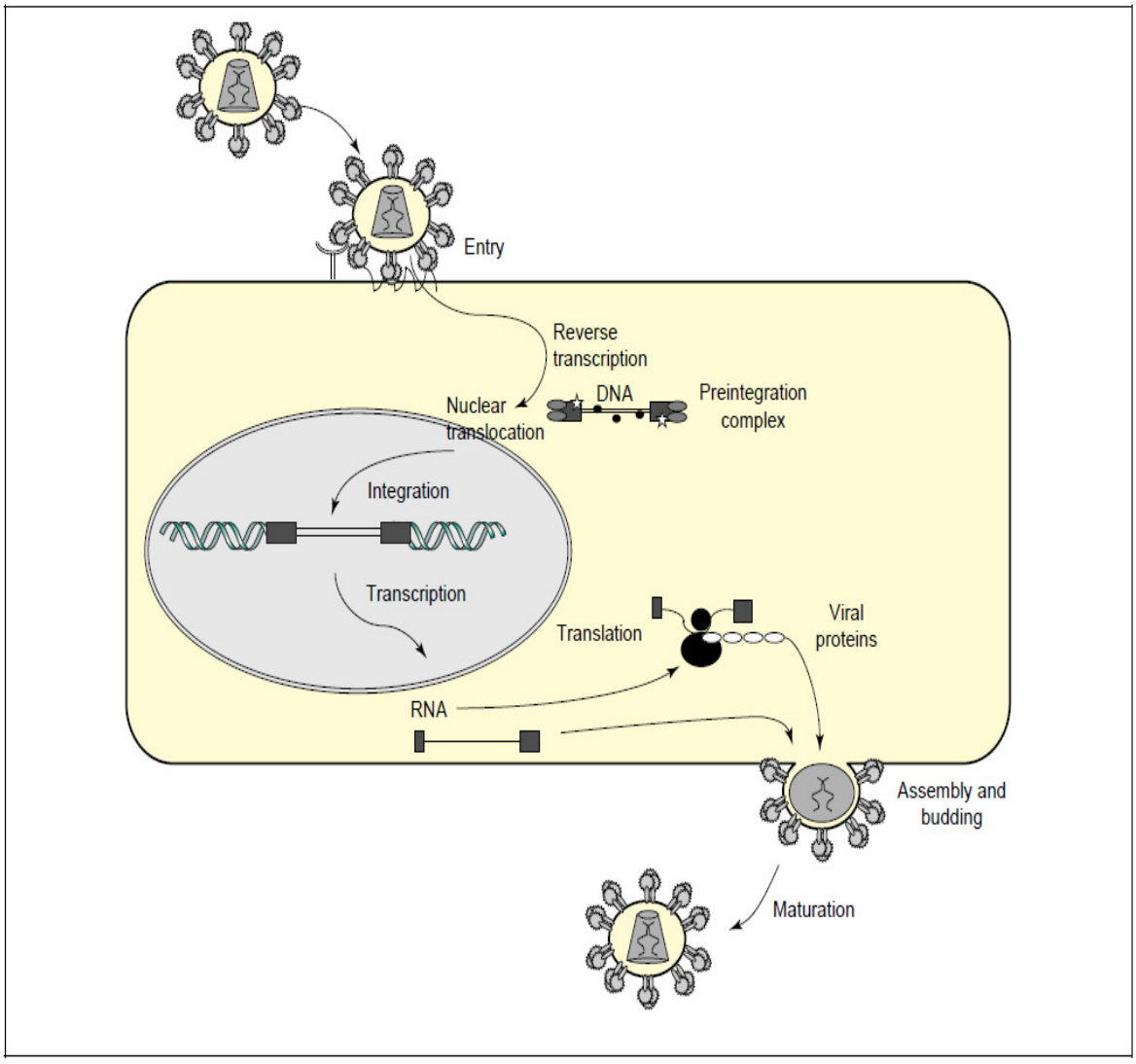


Figure 1.3: Typical life cycle of HIV-1 showing entry, reverse transcription, integration, transcription, translation, assembly and budding. Adapted from (Ciuffi and Bushman 2006).

Highly Active Antiretroviral Therapy (HAART) involves triple therapy consisting of a nucleoside analogue reverse transcriptase inhibitor (NRTI) backbone (consisting of two drugs) and either a non-nucleoside analogue reverse transcriptase inhibitor (NNRTI) or protease inhibitor. Since its introduction in 1996, HAART has significantly reduced morbidity worldwide, and prolonged the life-span of HIV-1 infected patients by preventing progression to Acquired Immunodeficiency Syndrome (Bally *et al.* 2004). However, a major downfall of HAART is that treatment failure does occur, and can be

attributed to patient non-adherence, poor response to antiretrovirals, toxicity to antiretroviral therapy or the development of HIV-1 drug resistance as a result of the error-prone RT enzyme during replication, or drug pressure (Perelson *et al.* 1996). The two other drug classes (integrase and entry inhibitors) are generally used as salvage therapy in patients that harbor multi-drug resistant virus.

1.3. The HIV-1 Integrase enzyme and Integration

During the HIV-1 replication cycle, there is a requirement that the viral genome, like other lentiviruses, be integrated into the host chromosome for successful propagation of the virus (Marchand *et al.* 2006). Integration occurs most efficiently after reverse transcription of HIV-1 RNA into cDNA (Iordanskiy *et al.* 2006), and is mediated by the viral integrase (IN) enzyme which results in a stable provirus that remains present for the lifetime of the infected cell. Mutations which destroy IN activity block viral replication (Dvorin *et al.* 2002), thus the integration step represents an important target for the development of new antiretroviral drugs.

The HIV-1 IN enzyme results from the proteolytic cleavage of the *gag-pol* precursor and belongs to the transposase family of DNA transferases (Mizuuchi 1992; Rice and Baker 2001). It has a molecular mass of 32 kDa (288 amino acids), and contains three distinct functional domains. The first 50 amino acids comprise the N-terminal domain (NTD) with zinc-binding properties (Bushman *et al.* 1993) which allows for increased catalytic activity and multimerisation of IN (Zheng *et al.* 1996). The Catalytic Core Domain (CCD) traverses residues 51–212 and contains the three highly conserved catalytically important residues Asp64, Asp116 and Glu152, also known as the DDE motif (Bushman *et al.* 1993). The third C-terminal domain (CTD) spanning from residues 213 to 288 is believed to be involved in DNA binding, oligomerization of IN *in vitro* (Jenkins *et al.* 1996) and interaction with RT (Dar *et al.* 2009). Therefore, all three domains are involved in oligomerization as well as specific functions within the complete IN enzyme framework

[for a review see (Marchand *et al.* 2006)]. A schematic of the IN domains is shown in Figure 1.4.

IN catalytic activity is accomplished by the tetrameric form of the protein *in vivo* and can be divided into a two step process. The 3'-end processing (3'-P) step is the first step, occurring within the cytosol. A dinucleotide adjacent to a phylogenetically conserved CA is hydrolytically cleaved off from the 3' end of each strand of the linear viral cDNA by IN (Sherman and Fyfe 1990; Bushman and Craigie 1991; Cherepanov *et al.* 1999). This results in two 3' hydroxyl groups that are used for a subsequent nucleophilic attack where the two viral DNA ends are inserted by IN into the host genome in a concerted fashion (Charvat *et al.* 2006) and this step is commonly referred to as strand transfer (ST). ST involves attack of two phosphodiester bonds on each strand of the cellular DNA by the 3'-hydroxyls at the ends of the viral DNA that results in the insertion of this DNA product within the host genome (Bushman and Craigie 1991; Cherepanov *et al.* 1999; Charvat *et al.* 2006). Integration seems to be favoured at chromosomal sites where active genes are present (Schröder *et al.* 2002), while integration is disfavoured in condensed heterochromatin (Carteau *et al.* 1998) and favoured in de-condensed chromatin regions (Albanese *et al.* 2008) suggesting that integration may be influenced by chromatin organization or directed to these sites by a molecular chaperone. Removal of the two unpaired nucleotides on the 5' viral DNA ends, also known as the DNA flap, and gap filling of the 5 base pair single stranded interfaces between the viral and host genomic DNA is then believed to be completed by the host DNA repair machinery [reviewed in (Marchand *et al.* 2006; Delelis *et al.* 2008)].

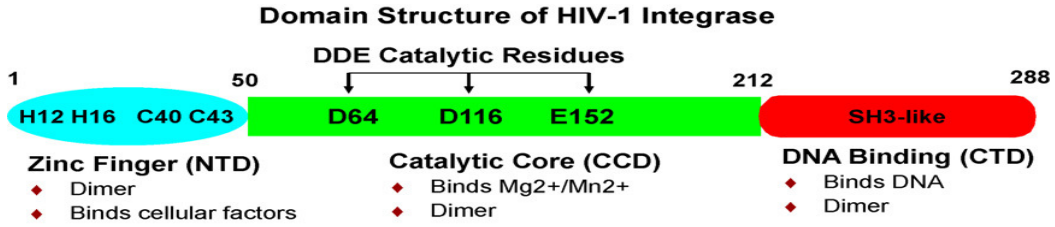


Figure 1.4: Schematic of the domain structure of the HIV-1 Integrase (IN) enzyme. The functions and specific characterization of the domains are listed. Adapted from (McColl and Chen 2009).

Completion of integration results in a fully functional provirus, which can then be used to initiate viral DNA transcription. IN also catalyses the disintegration of half site integration products of viral DNA out of the host DNA (Mazumder *et al.* 1994; Faust *et al.* 1996). IN functions while in dimeric form but is shown to exist as monomeric, dimeric, tetrameric, and high-order oligomeric states *in vitro* (Baranova *et al.* 2007) where interactions with oligonucleotides were shown to influence the IN oligomeric state (Baranova *et al.* 2007). The most accepted *in vivo* functional form of IN is its tetrameric state (Charvat *et al.* 2006) and more recently the crystal structure of tetrameric IN from prototype foamy virus, a retrovirus , in complex with viral DNA has been obtained (Hare *et al.* 2010), and may serve as a representative IN tetramer indicating the functional unit of IN *in vivo*. A schematic of the HIV-1 integration steps are shown in Figure 1.5.

1.4. The Preintegration Complex and Host Factors Influencing Integration

The IN enzyme functions in cells within the context of high molecular weight preintegration complexes (PICs) that are derived from the viral core and formed and mainly matured in the cytoplasm after reverse transcription of the viral RNA (Iordanskiy *et al.* 2006). The proteins forming the PIC are an assortment of both viral and host protein interactions. The cytoplasmic PIC creates an ideal environment for efficient catalysis of 3'- processing by protecting the cDNA from degradation by cellular nucleases. After 3'-processing, due to its karyophilic properties, IN aids the viral matrix (MA) and viral protein R (Vpr) in the nuclear localization of the PIC (Bukrinsky *et al.* 1993; Gulizia *et al.* 1994; Devroe *et al.* 2003), where movement into the nucleus is an energy dependent pathway (Bukrinsky *et al.* 1992). There is a possibility that IN allows for the movement across the nuclear membrane, (Gallay *et al.* 1997); however the nuclear import ability of IN has not been confirmed (Petit *et al.* 2000). Vpr has also been shown to be involved in nucleocytoplasmic shuttling across the nuclear membrane (Sherman *et al.* 2001) and may play a vital role in nuclear entry of the PIC. The PIC then assists in the

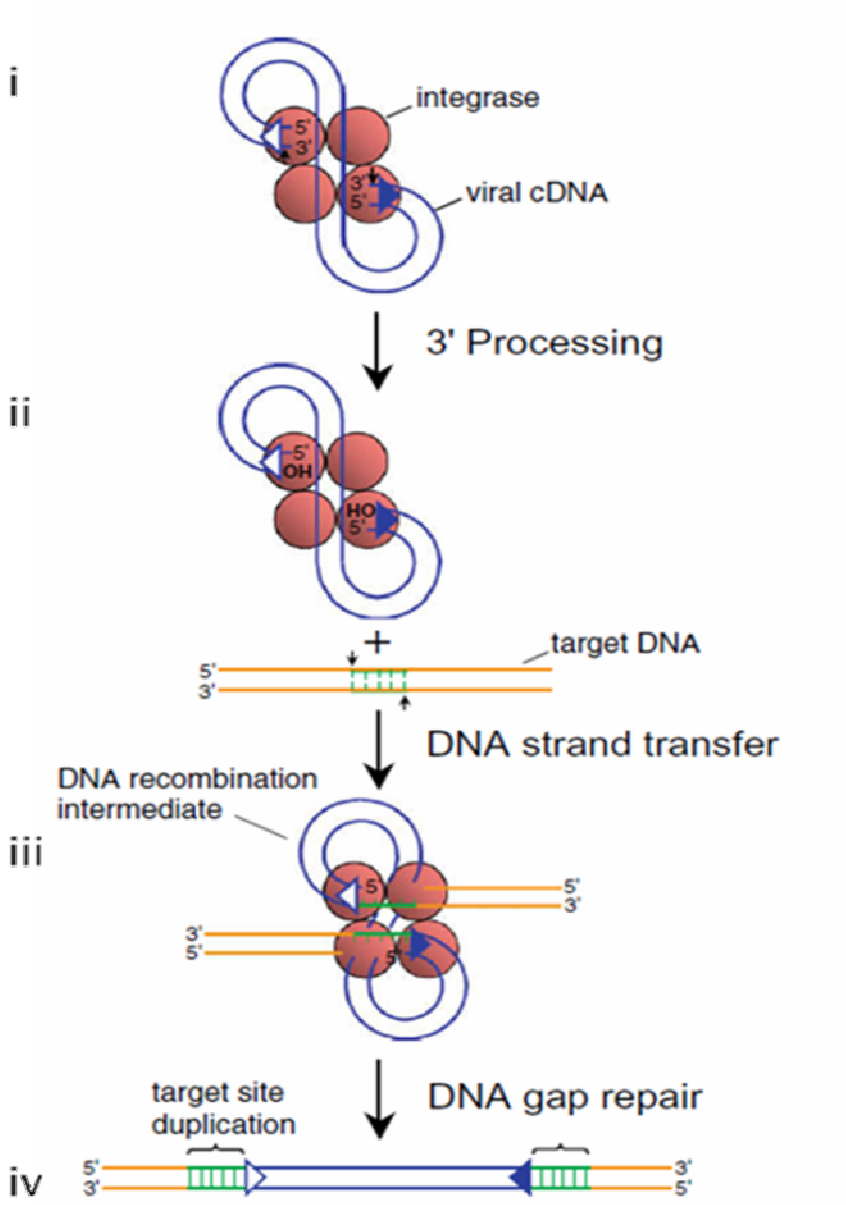


Figure 1.5: Schematic of the HIV-1 integration reaction. The tetrameric integrase enzyme (IN) bound to viral cDNA generates a CA overhang in a process called 3'-end processing in the cytoplasm (i). The preintegration complex moves into the nucleus where the strand transfer reaction is initiated by binding Target DNA (ii), followed by nucleophilic attack and phosphodiester bond formation leaving a 5' DNA flap (iii), DNA repair via host mechanisms, resulting in an integrated provirus (iv). Adapted from (Dar *et al.* 2009).

DNA ST activity especially by recruitment of the cellular cofactor lens epithelium-derived growth factor LEDGF/p75.

The precise composition of HIV-1 PICs is not known and probably changes during maturation from reverse transcription within the cytoplasm to integration within the nucleus. In addition to the viral factors HIV-1 cDNA, IN, Vpr, MA, RT and nucleocapsid (NC) proteins (Bukrinsky *et al.* 1993), known cellular host proteins which make up part of the PIC include the barrier-to-autointegration factor (BAF), high mobility group protein A1 (HMGA1), lens epithelial derived growth factor or protein p75 (LEDGF/p75), and others [reviewed in (Van Maele *et al.* 2006)]. Interestingly, components present in PICs, namely viral proteins, together with cellular factors are essential for the efficient integration of HIV-1 cDNA. In fact it has been shown that salt-stripped PICs lose their functionality, but can be returned to a functional state by addition of mammalian cell lysate and cytoplasmic extracts (Brooun *et al.* 2001).

LEDGF/p75 has been shown to directly interact with IN in the nucleus and allow for chromosomal tethering and directed integration of viral cDNA at LEDGF/p75 specific chromosomal sites [(Ciuffi *et al.* 2005; Marshall *et al.* 2007) and reviewed in (Ciuffi and Bushman 2006; Engelman and Cherepanov 2008)]. In addition, LEDGF/p75 may also aid in stabilizing the tetramer and prevent higher oligomerization (Michel *et al.* 2009). The first host factor to be shown to interact with IN was identified as the integrase interactor 1 (Kalpana *et al.* 1994) and has been implicated in post integration activities by shuttling the Gag-Pol precursor polypeptide by interaction with integrase in this context (Yung *et al.* 2001). BAF appears to play a role in coordinating integration of viral cDNA into the host chromosome and preventing integration of the viral DNA into itself and HMGA1 is implicated in enhancing transcription of the HIV provirus. [for review see (Van Maele *et al.* 2006)]. Other host proteins implicated in the PIC include TNF (tumour necrosis factor), TRAF (TNF receptor-associated factor) and receptor-associated protein

(TTRAP) (Zhang *et al.* 2009) and TRIM (tripartite motif protein) proteins in primates (Anderson *et al.* 2006).

These host cell factors may provide additional novel antiviral drug targets or impact on the ability of a putative IN inhibitor to successfully prevent integration.

1.5. HIV-1 Integrase Inhibitors and resistance

HIV-1 IN has no human homologue, making it an attractive target for the development of antiretroviral therapeutic agents with little adverse side-effects. Integration of HIV-1 has been extensively studied since the start of the pandemic, and great advances have been made in the development of novel IN inhibitors (INIs), especially by the Hazuda group from the Merck Research Laboratories and their development of a novel IN activity assay (Hazuda *et al.* 1994) which will be discussed later in this chapter. Although there are relatively few clinically viable INIs to date, it must be taken into account that the IN-integration pathway is that of an intricate nature, spanning from the cytosol and ending in the nucleus with multiple proteins involved in the process. The IN protein is insoluble mainly due to its hydrophobic nature (Jenkins *et al.* 1995; Jenkins *et al.* 1996) and this characteristic of HIV-1 IN has hampered the realization of a complete crystal structure, which would have served well in rational design of INIs as it did for RT and protease (PR) enzymes. Nevertheless, due to the pioneering work of the Merck team and other groups, progress in the field identified compounds from the diketo acids (DKA) group, Naphthyridine, Equesitin, β-hydroxy keto acids group and carboxylate group [reviewed in (McColl and Chen 2009)]. These inhibitors either affect 3'-P and ST activity or ST alone. The DKA group of INIs represents the most convincing, biologically validated inhibitors of IN. Raltegravir (RAL, MK-0518, Isentress), a compound of the DKA group, was the first INI to receive FDA approval for clinical use and has also been allowed for use as a first-line treatment in highly active antiretroviral therapy (HAART) (www.fda.gov). Elvitegravir (ELV), a monoketo acid, is currently in Phase III human clinical trials (Sato *et al.* 2009), and a new experimental quad pill regime of ELV along

with RT inhibitors Tenofovir, Emtricitabine and a cobicistat booster has just begun Phase III clinical trials (http://www.hivandhepatitis.com/recent/2010/0416_2010_c.html).

Structural and modelling studies have elucidated the mode of action of these compounds. Truncated IN has been crystallized and valuable structural data obtained (Cai *et al.* 1997; Goldgur *et al.* 1998; Chen *et al.* 2000; Wang *et al.* 2001; Cherepanov *et al.* 2005; Baranova *et al.* 2007), but the wild type enzyme is yet to be fully envisioned in its true form. Michel *et al* (2009) have obtained the outer shell of the functional tetrameric form of wild type IN in complex with LEDGF/p75 in the form of a cryo-negatively stained body, and used computer-based analysis and fitting of previously defined crystal structures of domains to determine a possible structural orientation of IN. This would be the closest, thus far, in obtaining a true structure of the HIV-1 IN functional unit. As mentioned previously, Hare *et al* (2010) have determined the structure of the IN tetramer of prototype foamy virus (PFV), which bound HIV-1 IN strand transfer inhibitors, RAL and ELV (Figure 1.6). Although the IN in their study is from a different virus, results can be extrapolated to HIV-1 and suggest that RAL and ELV interact with IN's active site only after 3'-P when the IN undergoes a conformational change (Baranova *et al.* 2007). RAL and ELV can then chelate essential metal ions (Mg^{2+} and Mn^{2+}) that bind to the IN active site. Furthermore, they are hydrophobic allowing them to interact with the hydrophobic pocket of the IN active site. They also interact with specific amino acids of the CCD, namely the DDE motif, and displace viral DNA out of the active site (Hare *et al.* 2010). The CCD domain and its associated DDE motif are highly conserved regions (Rhee *et al.* 2008) and the mechanism of action of INIs shown in this PFV virus may be highly related to actual method of action in wild type IN.

Many resistance mutations to these INIs, especially RAL and ELV have been documented. Antiretroviral drug resistance mutations are caused by RT fidelity or, rather, the lack thereof. The main mutational pathways associated with RAL and ELV resistance include signature mutations at integrase positions 143, 148 and 155. The

three major mutation pathways, Y143, Q148 and N155 are usually accompanied by T97A, G140S and E92Q, respectively. To date, there are at least 30 substitutions that have been specifically associated with the development of resistance to RAL and/or ELV (http://hivdb.stanford.edu).

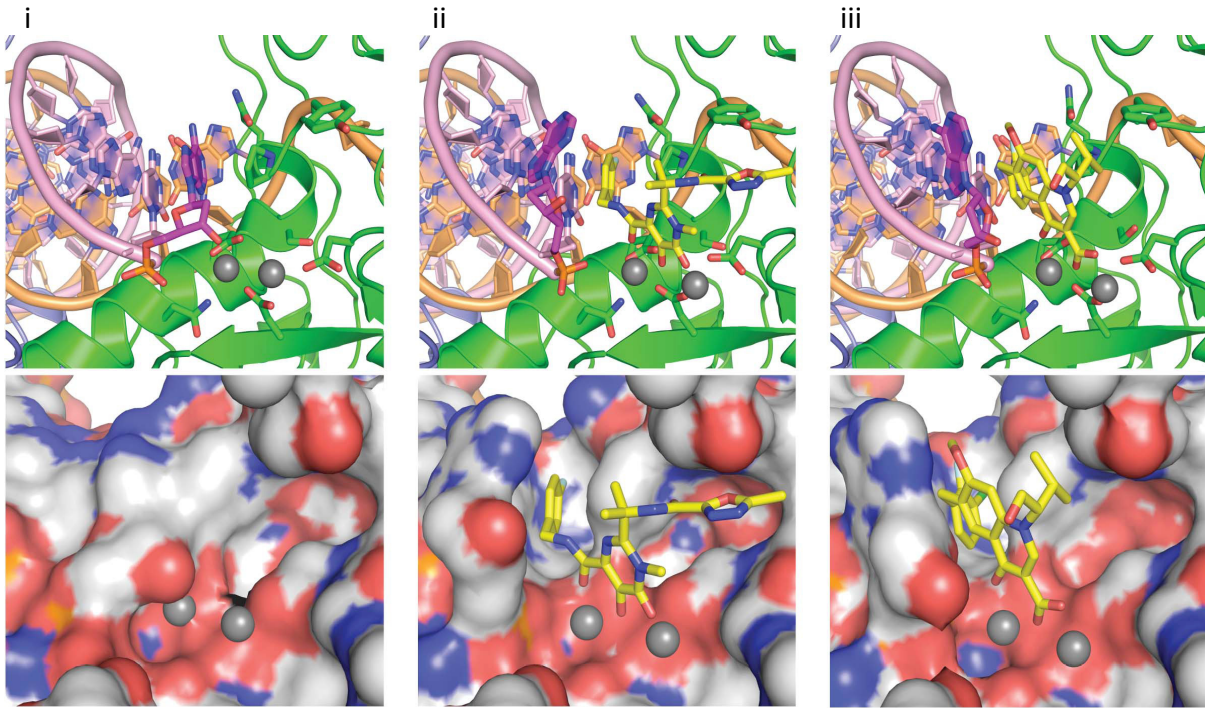


Figure 1.6: Crystal structure of the prototype foamy virus IN active site in committed and drug-bound states showing IN without drug (i) and bound to RAL (ii) or ELV (iii). The top panels show protein and DNA cartoons, DNA bases and the side chains of indicated amino acids as sticks. The RAL and ELV drug atoms are colored: yellow, C; blue, N; red, O; grey, F; green, Cl. The bottom panels show the complex as a solvent-accessible surface, colored by atoms (light grey, C; red, O; blue, N). Grey spheres are Mn21 (i) or Mg21 (ii, iii) ions. Adapted from (Hare *et al.* 2010).

1.6. HIV-1 Integrase Activity Assays

Assays testing 3'-P generally show cleavage of the terminal 3' CA of an oligonucleotide mimicking the viral LTR sequence (Sherman and Fyfe 1990). Visualization of these results are gel-based (denaturing polyacrylamide or agarose) making them time consuming and not conducive to high throughput screening.

There are several published strand transfer assay methodologies which utilize recombinant integrase in either its wild type or "soluble" forms (Hazuda *et al.* 1994; Chow 1997; Bally *et al.* 2004) or make use of authentic PICs isolated from the cytoplasm of infected cells (Farnet *et al.* 1996; Dar *et al.* 2009). These approaches serve to test the integration of a sequence-specific viral DNA strand into a target host mimic so as to determine enzyme activity dynamics, allowing for the identification of possible inhibitors or elucidation of enzymatic activity. The PICs are prime candidates for use in these assays, since those that are isolated from acutely-infected cells can integrate their endogenous cDNA into an added Target DNA *in vitro*. Purified HIV-1 PICs are capable of incorporating 15 to 95% of the target viral DNA into an exogenously supplied DNA target as full-site integration products (Farnet and Haseltine 1990). However, due to the difficulty in isolating functional PICs, as well as developing a suitable detection method, it has been challenging to construct a high-throughput assay using these complexes. However, the Engelman group has been significantly successful in isolation of and development of PIC assays and their applications (Dar *et al.* 2009).

By contrast, recombinant IN generally tends to catalyze integration of only one DNA end into a given target (half-site integration), and only under specific conditions, namely in its tetrameric form, can it then allow for concerted ST of two viral DNA ends into a Target DNA *in vitro* (Sinha and Grandgenett 2005; Bar-Magen *et al.* 2009). However, even with these limitations, assays using recombinant IN are still the preferred method of testing ST, mainly due to ease, reproducibility and high throughput.

The basic ideology of the ST assay utilizing recombinant IN, summarized in Figure 1.7, is the immobilization of a donor DNA that mimics a 3'-processed HIV-1 LTR. This takes advantage of the highly specific interaction of IN with the donor DNA (Hobaika *et al.* 2009; Hare *et al.* 2010). Thereafter, IN is bound to the immobilized donor, and this binding is referred to as assembly. After assembly, labelled Target DNA that would

mimic the host integration site is added. The IN would then integrate the donor into the target, hence immobilizing the target. Hereafter, the labelled target is detected by a suitable detection method to determine integration. The Hazuda group has been the forerunner in development of the recombinant IN protein assays (Hazuda *et al.* 1994) and many modifications have subsequently been made to increase the efficiency and effectiveness of this type of assay.

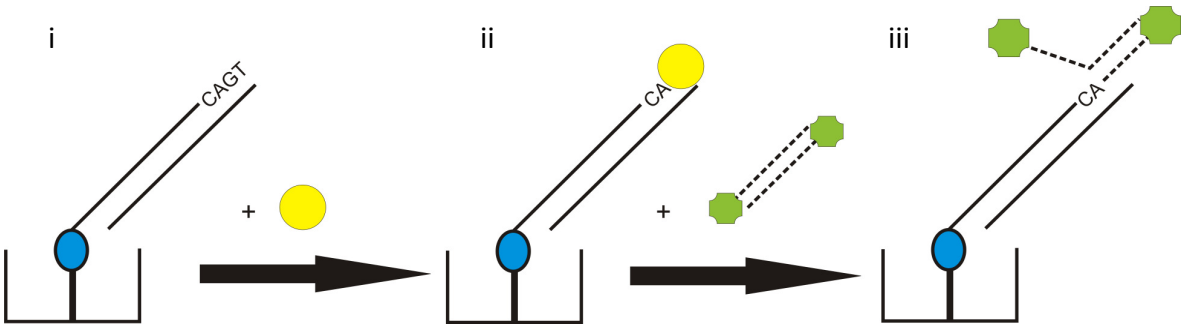


Figure 1.7: Schematic of a high throughput microtitre plate HIV-1 integration assay showing donor DNA immobilized on the well surface (i). After addition of IN, assembly and 3'-end processing is initiated (ii), followed by addition of labelled Target DNA leading to strand transfer (iii). Successful strand transfer is detected by relevant detection methods. Picture drawn using CorelDraw version 12.0.

1.7. HIV-1 infection in South Africa

South Africa is home to over 5.7 million people living with HIV-1, which is the largest infected population from a single country (http//:www.unaids.org). Most infections in South Africa are a result of HIV-1 subtype C transmission. There have been mild improvements in terms of prevalence and mortality since 2005 due to improvements in government policy; however, these statistics are nevertheless a daunting reality (http//:www.unaids.org). Access to antiretroviral drugs has escalated dramatically in South Africa since implementation of the government antiretroviral drug roll-out program in April 2004. Currently, there are over 1 million South Africans receiving HAART. The new first-line regimen for South African adults and adolescents effective in 2010, consists of Tenofovir (TDF), Lamivudine (3TC)/Emtricitabine (FTC) and Efavirenz

(EFV)/Nevirapine (NVP). In TB co-infected patients, EFV is preferred and NVP is administered to women of a child-bearing age who are not on reliable contraception. Patients on the previous d4T containing first-line regimen (d4T, 3TC, EFV/NVP) that exhibit no side effects remain on this regimen, and patients with contraindications for TDF are placed on Zidovudine (AZT), 3TC and EFV/NVP. The new second-line is comprised of TDF, 3TC/FTC and Alluvia (lopinavir/ritonovir) if the patient failed a d4T or AZT-based first-line regimen, and AZT, 3TC and Alluvia if they failed on a TDF-based first-line regimen. (Department of Health, 2010; http://www.sanac.org.za/documents/2010%20ART%20Guideline-Short.pdf). However, these regimens carry significant toxicities, requiring drug substitutions. In addition, an increasing number of patients require alternative regimens as they fail first and second-line regimens, due to non-adherence, toxicity or development of drug resistance. Thus, the effectiveness of integrase inhibitors, such as RAL, needs to be investigated as alternative antiretroviral drug treatment options are required. In addition, a concerted effort to screen for novel INIs against HIV-1 subtype C needs to be undertaken.

Several groups have published genotypic assays for the amplification and sequencing of *integrase* from diverse HIV-1 group M subtypes (Van Laethem *et al.* 2008; Eshleman *et al.* 2009; Van Baelen *et al.* 2009); however, none describe a subtype C specific *integrase* genotyping assay. In addition, there is no research group in South Africa screening for novel INIs against HIV-1 subtype C. Thus, the overall aim of the study was to establish an HIV-1 subtype C specific *integrase* genotyping assay and strand transfer assays to screen for novel IN inhibitors.

The specific objectives were to:

1. Establish an HIV-1 subtype C specific *integrase* genotyping assay, to determine baseline resistance mutations to integrase inhibitors in the HIV-1 infected South African population.
2. To select a representative viral isolate for isolation of preintegration complexes.

3. To establish strand transfer assays using HIV-1 subtype C isolated PICs, or subtype B purified recombinant IN that can be used to screen for novel IN inhibitors.

CHAPTER 2

Materials and Methods

2. Materials and Methods

2.1. Patient Samples and Reagents used In This Study

A total of 80 HIV-1 samples from patients attending the Charlotte Maxeke Johannesburg Hospital (Parktown, Johannesburg, South Africa) were available for purposes of this study. These included 58 randomly selected patient plasma samples sent for routine viral load monitoring (Table 2.1; sufficient patient sample available at the time of sample collection) and 22 primary HIV-1 isolates from antiretroviral drug naïve AIDS patients isolated during 2005 (Connell *et al.* 2008).

The following reagents were obtained through the NIH AIDS Research and Reference Reagent Program, Division of AIDS, NIAID, NIH: Sup-T1 from Dr. James Hoxie; CEM-SS (Cat# 776) from Dr. Peter L. Nara; CEM-GFP from Dr. Jacques Corbeil and MT-2 from Dr. Douglas Richman; HIV-1 NL4-3 Integrase Protein (F185H/C280S) from Dr. Robert Craigie as a positive control; HIV-1 IN Monoclonal Antibody (8G4) from Dr. Dag E. Helland, used as a primary antibody for western blot detection; Dr. Robert Craigie: pINSD.His, and is a expression vector that codes for integrase (IN) with a Hexahistidine motif grown in BL21 (DE3) *E. coli*, and pINSD.His(Sol) which contains the HIV-1NL4-3 IN coding sequence from pINSD.His with modifications to introduce the amino acid substitutions Phe 185→Lys and Cys 280→Ser. Anti-mouse IgG antibodies with horseradish peroxidase (HRP)-linked secondary antibodies (GE Healthcare, Waukesha, WI, US) were used for western blot detection. Chicoric acid (Sigma, St. Louis, MO, USA) was used as a control inhibitor in the modified microtitre plate assay.

Table 2-1: Viral loads (RNA copies/ml) and gender of the 58 randomly selected patient plasma samples sent for routine viral load monitoring at the Charlotte Maxeke Johannesburg Hospital.

Patient Number	Viral Load	Gender	Patient Number	Viral Load	Gender
1	180 000	F	30	1.7 million	F
2	270 000	F	31	9 300	M
3	12 000	F	32	43 000	M
4	18 000	F	33	140 000	M
5	6 400	F	34	3 800	F
6	9 100	M	35	28 000	M
7	32 000	M	36	12 000	F
8	1 500	F	37	6 000	F
9	11 000	F	38	43 000	F
10	6 100	M	39	18 000	F
11	25 000	F	40	3 800	M
12	4 300	M	41	88 000	F
13	3 900	M	42	5 300	F
14	4 800	F	43	43 000	F
15	43 000	M	44	200 000	F
16	310 000	M	45	21 000	M
17	21 000	M	46	930 000	F
18	5 000	F	47	960 000	F
19	29 000	F	48	56 000	M
20	34 000	F	49	120 000	M
21	19 000	F	50	4 600	F
22	47 000	M	51	84 000	F
23	220 000	M	52	52 000	M
24	160 000	F	53	140 000	M
25	260 000	F	54	540 000	M
26	17 000	F	55	40 000	M
27	59 000	M	56	310 000	F
28	> 3 million	F	57	78 000	M
29	29 000	F	58	870 000	F

2.2. Amplification and sequence analysis of the HIV-1 integrase

2.2.1. Proviral DNA isolation from infected co-cultures

Proviral DNA from the 22 viral isolates was extracted from p24 antigen-positive peripheral blood mononuclear cell (PBMC) cultures with a High Pure PCR Template Preparation Kit (Roche, Frankfurt, Germany) according to the manufacturers' instructions, using all buffers supplied in the kit. Briefly, PBMC cell cultures were pelleted and resuspended in 200µl of phosphate buffered saline (PBS). Binding buffer and proteinase K were added in volumes of 200µl and 40µl, respectively. The mixture was then incubated at 70˚C for ten minutes. Isopropanol (Merk, Darmstadt, Germany) was added in a volume of 100µl, mixed and was aliquoted to the reservoir of a High Filter Tube fitted to a collection tube. The sample volume was then centrifuged at 8000 *x g* for 1 minute (as were all subsequent spins unless otherwise stated), allowing it to pass through and bind the filter. Flow through was discarded and 500µl of Inhibitor Removal Buffer was added and centrifuged. Flow through was again discarded and 500µl of Wash Buffer was added to the filter tube and centrifuged and flow through discarded. This wash step was repeated once and after discarding flow through the filter was centrifuged at 13000 *x g* for 10 seconds to remove residual wash buffer and ethanol. The filter was then attached to a microtube and 200µl of Elution buffer at 70˚C was added. The filter was centrifuged and the flow through containing proviral DNA was collected.

2.2.2. Viral RNA isolation from patient plasma

Viral RNA was extracted from 58 patient plasma samples using the automated Roche MagNa Pure LC analyzer and the MagNA Pure LC Total Nucleic Acid Isolation Kit (Roche), according to the manufacturer's instructions. Fifty microliters of patient serum was added to 200µl of working solution from the kit, and extracted viral RNA was eluted in 50µl.

2.2.3. Reverse Transcriptase (RT) Polymerase chain reaction (PCR)

The full length *integrase* gene was RT-PCR amplified into cDNA from the extracted viral RNA samples using the reverse primer SQ9RC(3) (Table 2.2). Primers used in the RT PCR reaction, conventional PCR and sequencing reactions were all obtained from Integrated DNA Technologies and are listed in Table 2.2.

The Expand Reverse Transcriptase (RT) kit (Roche), was used as per the manufacturer's instructions with all reagents supplied in the kit. Briefly, volumes of 10.5µl were set up consisting of 1µg of total extracted RNA and 25pmol of primer oligonucleotide (SQ9RC(3)) made up in distilled water. DNA primers and RNA template strands were initially denatured at 65˚C for 10 minutes using the Gene Amp PCR System 9700 (Applied Biosystems, Foster City, CA, USA) and immediately placed on ice. The following reagents were then added ; reaction buffer was added to a 1X final concentration, dithiothreitol (DTT) to a final concentration of 10mM, dNTPs to a final concentration of 1mM each, 20 units (U) of RNase inhibitor and 50 U of RT. The RT PCR was initiated by incubation at 43˚C for 60 minutes on the Gene Amp PCR System 9700 (Applied Biosystems) and then was stopped by placing the reaction mixture on ice.

2.2.4. Conventional PCR

The Expand High Fidelity^{PLUS} PCR system (Roche) was used for all PCR reactions. The PCR reaction was carried out as per manufacturer's instructions and was shown to be optimal using buffer containing 1.5mM $MgCl_2$ and 25pmol of each primer. Forward primer SQ7F(2) and reverse primer SQ9RC(3) were used for the PCR reaction (Table 2.2). Template DNA was either 5µl of isolated proviral DNA (see section 2.1 above) or RT-PCR product (see section 2.3 above). The total volume of the reaction was 50µl. The thermal cycling was carried out as per thermal profile II of the manufacturer's instructions on the

Table 2-2: Primers used throughout the course of this study.

Primer name	Size (Mer)	Primer sequence	Experimental Use
SQ9RC(3)	24	5'-TTGGATGTCTGCTTTCATAATGAT-3'	PCR / RT PCR
SQ7F(2)	27	5'-AGGAGCAGAAACTTTCTATGTAGATGG-3'	PCR
QFIN1F	20	5'-ATGGGTACCAGCACACAAAG-3'	Sequencing
QFIN2F	20	5'-TAGCAGGAAGATGGCCAGTC-3'	Sequencing
QFIN1R	26	5'-CTCTCCTTGAAATATACATATGGTGC-3'	Sequencing
QFIN2R	22	5'-CTACTACTCCTTGACTTTGGGG-3'	Sequencing
Donor 1 (D1)	32	5'-ACCCTTTTAGTCAGTGTGGAAAATCTCTAGCA-3'	Microplate Assay/Gel Assay
Donor 2[a] (D2)	31	5'-ACTGCTAGAGATTTTCCACACTGACTAAAAG-3'	Microplate Assay/Gel Assay
Target 1[b]	20	5'-TGACCAAGGGCTAATTCACT-3'	Microplate assay /Gel assay
Target 2[b]	20	5'-AGTGAATTAGCCCTTGGTCA-3'	Microplate Assay / Gel assay
Donor[a] unprocessed (DU)	21	5'-GTGTGGAAAATCTCTAGCAGT-3'	Gel Assay
Donor[a] Processed (DP)	19	5'-GTGTGGAAAATCTCTAGCA-3'	Gel Assay
Donor Bottom 1 (B1)	21	5'-ACTGCTAGAGATTTTCCACAC-3'	Gel Assay (all assays)/ complementary strand
Donor Bottom 2 (B2)	19	5'-TGCTAGAGATTTTCCACAC-3'	Gel Assay (only used in figure SX) complementary strand

[a]Donor B2 DNA was either labelled with a 5' biotin or unlabeled at the 5'.
[b]Target DNA was either labelled with a 3' FITC label or unlabeled at the 3'.

Gene Amp PCR System 9700 (Applied Biosystems). Specifically, the cycles were set at 94˚C for initial denaturation. Thereafter, a ten times cycle of 94˚C for 30 seconds, 57˚C for 30 seconds for annealing and followed by 72˚C for one minute and 30 seconds for elongation. Following this was a 25 times cycle of 94˚C for 30 seconds, 57˚C for 30 seconds, followed by 72˚C for one minute and 30 seconds with an increasing increment

of ten seconds per cycle. This was followed by a final elongation step of 72˚C for seven minutes. Samples were stored at 4˚C until further use.

2.2.5. PCR Fragment Purification and Visualization

PCR fragments were purified using the high pure PCR cleanup kit (Roche) as per the manufacturer's instructions, with all buffers as supplied in the kit. Briefly; the 50µl PCR reaction volumes from section 2.2.4 was purified by addition of 500µl of binding buffer, mixed and added to a High Pure Filter tube. The filter was connected to a collection tube and centrifuged at 13000 *x g* for one minute. All subsequent centrifugation steps were of this profile unless otherwise stated. Flow through was discarded and 500µl of wash buffer supplemented with ethanol was added to the filter and centrifuged. The flow through was discarded and 200µl of wash buffer was added and centrifuged. Flow through was discarded and the filter was centrifuged for 30 seconds to dry and to remove excess wash buffer. Elution buffer was finally added in a volume of 50µl and the filter was attached to a clean microtube, centrifuged and elution, containing the PCR DNA product, was collected and stored at -20˚C. Fragments were separated by gel electrophoresis using 1% agarose gels stained with 5µg/ml ethidium bromide (Biorad, Hercules, CA, USA) and visualized on a Chemidoc (Biorad) system.

2.2.6. DNA Sequencing of HIV-1 *integrase*

Sequencing was achieved using the BigDye® Terminator v3.1 Cycle Sequencing Kit (Applied Biosystems). The sequencing reaction was prepared in a total volume of 20µl and was carried out as per the manufacturer's instructions. Primers QFIN1F, QFIN2F, QFIN1R and QFIN2R were used at 3.2pmol per reaction with 1µl of purified PCR product, reaction buffer and reaction mix (both supplied by the manufacturer). Thermal cycling was carried out on the Gene Amp PCR System 9700 (Applied Biosystems); with an initial single denaturation at 96˚C for 1 minute. Thereafter, in a 25 cycle format, each cycle begins with a ten second incubation at 96˚C; followed by a five second incubation at

50˚C for annealing and finally a four minute incubation at 60˚C for elongation. Thereafter, sequencing reactions are stored at 4˚C until further usage.

The sequencing reactions were then cleaned by adding 80µl of 80% isopropanol (Merck) to 20µl of the sequencing reaction. The reaction was thoroughly mixed and left to incubate for 15 minutes in the dark. This was then centrifuged at 2000 x *g* for 45 minutes on a model 5810 R Centrifuge with A-4-62 rotor (Eppendorf, Hamburg, Germany). The supernatant was immediately discarded by inverting the reaction plate and centrifuging in the inverted position at 700 x *g* for 1 minute. Ten microliters of HiDi formamide (Applied Biosystems) was then added to each sequencing reaction and the plates were stored at -20°C until sequencing. Sequencing was performed on the ABI Prism 3100- and 3730- *Avant* Genetic Analyzer (Applied Biosystems). Sequence data analysis was performed using the Sequencing Analysis V3.3 programs (Applied Biosystems) and assembled using Sequencher V4.5 (Genecodes, Ann Arbor, MI).

For subtyping, a multiple alignment of the full length *integrase* region with references from HIV-1 subtypes A to K, CRF01_AE and CRF02_AG (http://hiv-web.lanl.gov) was generated in Clustal X (http://www.clustal.org). Aligned sequences were converted to MEGA V3.0 format, and used in phylogenetic and molecular evolutionary analyses. A phylogeny reconstruction of each *integrase* gene was performed by Neighbor-Joining using the Kimura two-parameter distance matrix. The stability of the nodes was assessed by bootstrap analysis (1000 replicates) and bootstrap values greater than 70% were considered significant. Nucleotide sequences were converted to amino acid sequences, and compared to the 2004 consensus C sequence available from the Los Alamos website (http://hiv-web.lanl.gov). The amino acid sequence alignment was extensively analyzed on GeneDoc version 2.7 (http://www.nrbsc.org/gfx/genedoc) for the presence of primary mutations, non-polymorphic and polymorphic mutations associated with resistance to known integrase inhibitors (http://hivdb.stanford.edu).

The *integrase* sequences were submitted to GenBank using Sequin V5.35 (http://www.ncbi.nlm.nih.gov).

2.3. Mammalian Cell Culture

The SupT1, MT-2, CEM-SS and CEM-GFP suspension cell lines were maintained in RPMI media (Sigma) supplemented with 10% Foetal Bovine Serum (Gibco), 1% Penicillin and Streptomycin antibiotics (Sigma) and 1% L-Glutamine (Sigma). All cell lines were incubated at 37°C/ 5% CO_2 unless otherwise stated. Cell lines were split as they became confluent at a dilution of 1:10 to 1:20. Cell stocks were made by resuspending confluent cells at a concentration of 10^6 cells/ml in freezing mix comprised of 90% Foetal Calf Serum (Gibco) and 10% dimethylsulfoxide (DMSO) (Sigma). Cells were frozen in a nest box at -80°C overnight, and stored in liquid nitrogen.

2.4. HIV-1 Preintegration Complex Isolation

2.4.1. HIV-1 subtype C Infections

HIV-1 infections were done using a cell-line adapted isolate 05ZAFV3 available in our laboratory (previously propagated in a U87 cell line). SupT1, MT-2, CEM-SS and CEM-GFP cells were cultured as explained above until confluent and counted with a haemocytometer. For infections; $3x10^6$ cells were seeded in 5ml culture media and 150µl of HIV-1 FV3 containing supernatant was added, and culture was incubated overnight. There after the culture media was changed by pelleting the cells followed by resuspension in 5ml of fresh media. Cell culture was then incubated at 37˚C/5% CO_2 until usage in co-infections.

The Murex HIV Antigen mAb (Abbott, Illinois, USA) p24 ELISA detection assay was used to determine viral infection, as per the manufacturer's instructions. Briefly,

culture supernatant samples were diluted into 100µl volumes and added to the anti-p24 wells. Conjugate solution 1 was added in a 100µl volume and incubated at 37˚C for one hour. The wells were washed on an ELX50 Auto Strip Washer (BioTek Instruments Winooski, VT, United States) with available wash buffer and 200µl of conjugate 2 was added to the wells. This was incubated at 37˚C for 30 minutes. The wells were once again washed and 200µl of the substrate was added and incubated at room temperature for 30 minutes. A 50µl volume of HCl (1M) was added to stop the reaction. Binding was then calculated on an ELX808 Elisa plate reader (BioTek Instruments) using a dual wavelength (450nm and 690nm) protocol.

2.4.2. Crude Extraction of HIV-1 PICs

The following method was adapted from Brooun *et al.* (2001), Briefly, once infection was established, as determined by p24 ELISA, usually on day ten, the infected cells were co-cultured with a ten-fold concentration of uninfected SupT1 cells, and incubated for a further seven hours. Cells were then pelleted at 160 x g for five minutes and the supernatant was discarded. Cells were washed with buffer A (10mM Tris-HCl (Sigma) pH7.4; 15mM KCl (Sigma); 5mM $MgCl_2$ (Sigma) and 20µg/ml aprotinin (Sigma)) and pelleted at 160 x g for five minutes. Supernatant was discarded and cells were resuspended in lysis buffer (buffer A with 0.025% digitonin (Sigma)) at a concentration of 2.0 x 10^7 cells/ml based on original cell counts. The solution was left to incubate at room temperature for ten minutes, thereafter, nuclei were pelleted at 1000 x g for three minutes. Aliquots of supernatant were carefully removed and sucrose (Sigma) was added to a final concentration of 8% for each aliquot. This was then snap frozen in liquid CO_2 and stored at - 80˚C.

2.5. HIV-1 Recombinant Integrase Purification

2.5.1. Preparation of competent *Escherichia (E.) coli*

E. coli BL21 pLysE cells (Novagen, Darmstadt, Germany) were made competent by standard protocols. Briefly, overnight cultures of BL21 pLysE cells were propagated to make competent cells. These were diluted 25 times in Luria Broth supplemented with 50µg/ml chloramphenicol (Calbiochem, Darmstadt, Germany) and incubated at 37°C, in a shaking incubator, until the OD reading at 600nm reached 0.4. Cells were incubated on ice for ten minutes and then centrifuged at 2000 rpm/4°C (Eppendorf) for 15 minutes. The supernatant was removed and discarded and the pellet was resuspended in 16 ml autoclaved transformation buffer (100mM $CaCl_2$ (Sigma); 10mM Pipes-HCl (Merk) pH 7; 15% glycerol (Melford Labs, Ipswich, US). This was incubated on ice for 30 minutes and then centrifuged at 1500 rpm/4°C for 15 minutes. Supernatant was once again discarded and the pellet was gently resuspended in 4ml transformation buffer (and kept on ice). Aliquots of 200µl were stored at -80°C.

2.5.2. Bacterial Transformations

Competent BL21 (DE3) pLyse cells were transformed with the pINSD.His recombinant plasmid (see section 2.1). Briefly, 1µl of pINSD.His was added to 100µl of thawed competent BL21 pLyse cells and incubated on ice for 30 minutes. The solution was then heat-shocked at 37°C for 90 sec and placed on ice for five minutes. The solution was then plated on agar plates supplemented with 100µg/ml ampicillin (Melford Labs) and 50µg/ml chloramphenicol and incubated at 37°C overnight. Colonies that showed ampicillin resistance were selected as transformed bacteria (BL21 pINSD.His) and used in subsequent analyses.

2.5.3. Gene Induction

BL21 pINSD.His and pINSD.His(sol) cells were induced for gene expression. Firstly, Luria Broth supplemented with ampicillin (100µg/ml) and chloramphenicol (50µg/ml) was inoculated with BL21 pINSD.His and incubated overnight at 37°C shaking. This was then diluted 50 times and incubated at 37°C and allowed to reach an OD between 0.5 – 0.8 at 600nm. Isopropyl β-D-1-thiogalactopyranoside (Calbiochem) was then added to the culture at a 1mM concentration. The culture was then incubated at 37°C shaking for an optimal time period of three hours. Induced BL21 pINSD.His bacteria were then harvested by centrifugation at 4000 rpm (Eppendorf) for 30 minutes and resuspended in binding buffer (1M NaCl (Sigma); 20mM Tris (Sigma) pH 7.9; 5mM imidazole (Sigma)) and stored at -20°C.

2.5.4. Bacterial Cell lysis and Fraction isolation

Induced BL21 pINSD.His cells were lysed by sonication. Sonicator (Bandelin) settings were 30 seconds for six cycles at 75-80 Hz, and this was repeated five times. The sonicate was then centrifuged at 40 000 x g on an Optima L-80 XP Ultra Centrifuge (Beckman, Brea, CA, USA) set at 4˚C. Supernatants were collected as the soluble fraction and depending on usage was stored on ice or at -20˚C.

Induced BL21 pINSD.His(sol) cells were resuspended in lysis buffer (1M NaCl, 20mM HEPES, 5mM imidazole, 0.25% CHAPS (Sigma), 1mM PMSF (Sigma), 35 U of DNAase1 (Sigma)) and incubated on ice for 30 minutes. Cells were then lysed by homogenizing five times for 1 minute using a homogenizer (Bandelin), followed by sonication, three times in a sonicator bath (Bandelin) for 1 minute. Finally, the lysate was centrifuged at 15 000 x g for 30 minutes on a model 5810 R Centrifuge with A-4-62 rotor (Eppendorf) and enzyme containing supernatant was collected and immediately used in subsequent purification.

2.5.5. Nickel Chelating Column Chromatography

The following method for protein purification was used for wild type IN (IN_{wt}) proteins only which were induced from pINSD.His. Nickel-bound resin was prepared by incubation of 0.1M $NiSO_4$ (Sigma) to sepharose beads on ice for 15 minutes and then equilibrated with binding buffer for five minutes on ice. Resin and protein fractions were then mixed and incubated in a 50ml tube (Nunc), shaking at 4˚C overnight. The mixture was then added to a Luer-Lock, Non-jacketed Liquid Chromatography Column (Sigma) and resin was allowed to pack by gravity while a flow through sample was collected. Two hundred millilitres of two wash buffers each (1M NaCl (Sigma); 50mM Sodium Phosphate Buffer (Sigma) pH 8; 50mM imidazole (Sigma)) and (1M NaCl (Sigma); 50mM Sodium Phosphate Buffer (Sigma) pH 8; 100mM imidazole (Sigma)) were sequentially allowed to flow through the column using the Biologic LP System (Biorad) at a flow rate of 3.5ml/minute. The protein was then eluted with 25ml elution buffer (1M NaCl (Sigma); 50mM Sodium Phosphate Buffer (Sigma) pH 8; 500mM imidazole (Sigma)) at a flow rate of 1ml/minute. Eluate was dialyzed against buffer (1M NaCl (Sigma); 20mM HEPES (Sigma) pH 7.5; 0.1mM ethylenediaminetetraacetic acid (EDTA) (Sigma); 1mM DTT (Sigma); 10% glycerol (Melford Labs)) overnight at 4˚C. Dialysed eluate was then concentrated using 10kDa centricon tubes (Millipore, Billerica, MA, USA).

The following protocol was used for purification of the soluble mutant of IN (IN_{sol}) from induced bacteria transformed with pINSD.His(sol). A pre-packaged 5ml HisTrap HP Column (GE Healthcare) was attached to the ACTAPrime Plus (GE Healthcare) and equilibrated with binding buffer (1M NaCl (Sigma), 20mM HEPES (Sigma) pH 7, 5mM imidazole (Melford Labs)) and loaded with 10ml of sonicated supernatant containing His-tagged protein. After protein binding, the column was washed with 100ml of binding buffer at a flow rate 2ml/minute allowing for the pressure to be below the maximum of 0.3 mp. All flow rates were the same for subsequent washes unless otherwise stated. Thereafter, the column was washed with 200ml of wash

buffer 1 (1M NaCl (Sigma), 20mM HEPES (Sigma) pH 7, 60mM imidazole (Melford Labs)) followed by 200ml of wash buffer 2 (1M NaCl (Sigma), 20mM HEPES (Sigma) pH 7, 150mM imidazole). Nickel-bound proteins were eluted using a linear gradient of imidazole ranging from 150mM to 600mM at a flow rate of 1ml/minute. The gradient was set to 150 ml final volume. Elution was measured spectrophotometrically and was monitored with PrimeView (GE Healthcare). Fractions with protein were collected (fractions 23-60) and were pooled and concentrated using 3kDa centricon tubes (Millipore). Buffer exchange was achieved by passage through a p10 G25 Sephadex column (GE Healthcare).

2.5.6. Sodium Dodecyl Sulphate–Polyacrylamide Gel Electrophoresis (SDS-PAGE) and Western Blotting

The isolation of PICs and recombinant IN_{wt} and IN_{sol} protein expression was investigated by SDS-PAGE and Western blot analysis. SDS-PAGE gels of 12.5% were prepared and nitrocellulose (Amersham, Waukesha, WI, US) was blotted according to Hoefer guidelines. Briefly, the isolated PICs and recombinant protein samples were added to equal amounts of loading buffer (0.5M Tris-HCl pH 6.8, 5% glycerol, 2% SDS and 100mM DTT) and electrophoresed on 12.5% polyacrylamide gels at a current of 10 milliamperes (mA) initially until the prestained molecular weight marker (New England Biolabs, Ipswich, MA, US; Fermentas, Maryland, USA) entered the running gel, then current was increased to 30mA per gel. The gels were then equilibrated with transfer buffer and proteins were transferred to the nitrocellulose membrane (Amersham) for 50 minutes at 50mA using the TransBlot SD Semi-Dry Electrophoretic Transfer Cell (BioRad). Nitrocellulose membranes were blocked with 5% skim milk powder in 1 x Tris Buffered Saline (TBS) (50mM Tris (Sigma) pH 7.6, 150mM NaCl (Sigma)) for one hour. The membrane was then washed with 0.1% Tween-TBS three times for five minutes. Primary antibodies used were HIV-1 IN Monoclonal Antibody (8G4) and membranes were incubated at a concentration of 1:500 for one hour. The membrane was washed as before and anti-mouse horse radish peroxidase (HRP)-linked secondary antibody (GE

Healthcare) was incubated with the membrane in a 1:3000 concentration for one hour. The membrane was washed as before. Super Signal West Pico Chemiluminescent Substrate (Pierce, Rockfork, IL, USA) was prepared by addition of equal amounts of Super Signal West Pico Sable Peroxide to Super Signal West Pico Luminol/Enhancer Solution. The substrate was incubated with the membrane at room temperature and fluorescence was visualized on a Chemidoc system (Biorad).

2.6. Functional analysis of purified recombinant IN

Purified recombinant IN protein was used to set up and optimize three different assays to test for one-site integration activity using a high throughput microtitre plate and SPA HIV-1 integration assay, and full-site integration using a gel-based HIV-1 integration assay.

2.6.1. High throughput microtitre plate HIV-1 integration assay

Microplate analysis for recombinant IN_{wt} and IN_{sol} protein functionality was performed using the fluorescent properties of a FITC (fluorescein isothiocyanate) labelled primer bound to a biotinylated donor immobilized in streptavidin coated wells. Oligomeric primers used for this experiment are listed in Table 2.2. The oligomers and their complementary strands were annealed by heating to 95˚C for five minutes and allowed to slowly reach room temperature. This yielded donor: Donor 1/Donor 2(5'Biotin) and target: (3' FITC) Target 1/Target 2(3' FITC) pairs. Immobilizer streptavidin F96 Black (Nunc) plates were firstly primed as per manufacturer's instructions with the biotinylated donor. Briefly, wells used for the test were washed three times with 5 x SSCT buffer (750mM NaCl (Sigma), 75mM Sodium Citrate (Sigma), 0.05% Tween 20 (Sigma)). Biotinylated donor DNA was prepared in 5 x SSCT at a concentration of 50mM, and 100µl was added to each test well. The plates were incubated at room temperature for one hour with gentle agitation. Thereafter, wells were aspirated and washed with 2 x SSCT (150mM NaCl (Sigma),

15mM Sodium Citrate (Sigma), 0.05% Tween 20 (Sigma)) three times. Purified IN was diluted to a concentration of 1µM in Assay Buffer 1 (20mM HEPES (Sigma) pH 7.3, 75mM NaCl (Sigma), 10mM $MnCl_2$ (Sigma), 2µM $ZnCl_2$ (Sigma), 1% glycerol (Melford Labs)) or Assay Buffer 2 (25mM HEPES (Sigma) pH 7.5, 10mM $MgCl_2$ (Sigma), 5mM DTT (Sigma), 10% polyethylene glycol (PEG) 8000 (Sigma), 10% DMSO (Sigma), 20µM $ZnCl_2$ (Sigma), 75mM NaCl (Sigma)) and 100µl was added to each test well, while negative control wells contained only Assay Buffer. Plates were then incubated for 30 minutes at 37˚C with gentle agitation. Wells were then washed with Assay Buffer 1 or Assay Buffer 2 two times with 5 minute incubations per wash. FITC labelled Target DNA was then diluted in assay buffer to various concentrations to determine the optimal concentration, and 100µl was added to each test well including the negative control wells. The plates are incubated at 37˚C for one hour with gentle agitation. Wells were then washed three times with 300µl of 2 x SSC (150mM NaCl (Sigma), 15mM Sodium Citrate (Sigma)) with ten minute incubations per wash. Finally, wells were read on an Appliskan fluorometer (Thermo Electron Corp. Rockfork, IL, USA) with excitation wavelength of 485nm and emission wavelength of 535nm and data was recorded with SkanIt RE 2.3 software (Thermo Electron).

A modification to the abovementioned protocol was carried out to improve the recombinant IN_{sol} protein detection. The oligonucleotides used were the same as above. Donor substrate at a concentration of 140nM was diluted in a donor dilution buffer (25mM Tris-HCl (Sigma) pH 7.5, 150mM NaCl (Sigma), 0.05% Tween-20 (Sigma), 100µg/ml BSA (Sigma)) and added to Even Coat Streptavidin-coated microplates (R & D Systems, Minneapolis, MN, US). These were incubated at room temperature for 1 hour with gentle agitation. Unbound donor substrate was then removed by three washes with 300µl PBS. IN enzyme was then prepared in Assay Buffer 1 to a concentration of 1µM and 100µl added to each well and incubated at room temperature for 30 minutes with gentle agitation. The plates were washed

twice with Assay Buffer 1 with gentle agitation at room temperature for five minutes. Assay Buffer 1 was then added to each well in a volume of 90µl or 80µl and 10µl of 100% DMSO or Chicoric inhibitor (Sigma) and the plates were incubated at 37˚C for 30 minutes. Target substrate was diluted in Assay Buffer 1 to a concentration of 2.5µM and 10µl were added per well (final concentration of 250nM). Plates were incubated at 37˚C for one hour and thereafter washed three times with 2 x SSC with 10 minute incubations per wash at room temperature with gentle agitation. Monoclonal Alkaline Phosphatase conjugated anti-FITC antibody (Sigma) was prepared by 10 000 times dilution in Tween (0.05%)-TBS (T-TBS) and 200µl was added to the wells and incubated at room temperature for two hours with gentle agitation. Plates were then washed three times with TBS for ten minutes per wash at room temperature with gentle agitation. *para*-Nitrophenylphosphate (pNPP) substrate (Sigma) was added to each well in a volume of 200µl and incubated at 37˚C for one hour with gentle agitation. The reactions were then stopped by dilution of the 3µM NaOH stock to a final concentration of 50nM. Absorbance of the wells was measured at 405nm using the xMark Microplate Absorbance Spectrophotometer (Biorad).

2.6.2. HIV-1 Integrase Urea-PAGE Based Functional Analysis

Gel-based IN functional analysis was undertaken utilizing the DNA separation properties of Urea-PAGE. Assay conditions were as follows; firstly, IN proteins were diluted in 50µl Assay Buffer 1 (20mM HEPES pH 7.3, 75mM NaCl, 10mM $MnCl_2$, 2µM $ZnCl_2$, 1% glycerol) or Assay Buffer 2 (25mM HEPES pH 7.5, 10mM $MgCl_2$, 5mM DTT, 10% PEG 8000, 10% DMSO, 20µM $ZnCl_2$, 75mM NaCl) to varying concentrations. Unlabelled or radiolabelled Donor DNA (Table 2.2, Donor 1, DU and DP) was diluted with 25µl of varying assay buffers to varying concentrations and added to the diluted IN proteins. The reaction was incubated for 30 minutes on ice to allow for IN/DNA complex formation. Radiolabelled Target DNA (Table 2.2, Target 1 and Target 2) was diluted in 25µl of varying assay buffers to varying concentrations and added to the reaction

mixture. The reaction mixture was incubated for 30 minutes on ice and thereafter for one hour at 37°C to allow for ST. The reactions were stopped by adding 20µl formamide loading buffer (95% formamide (Sigma), 0.025% bromophenol blue (Sigma), 0.5µM EDTA (Sigma)). Thereafter, reactions were denatured at 95°C for 5 minutes and placed on ice or stored at -20°C. Urea-PAGE Gels were prepared as 20% acrylamide gels (20% polyacrylamide (Merk), 8M Urea (Sigma), 89mM Tris (Sigma), 89mM Boric acid (Sigma), 2mM EDTA (Sigma)). Gels were pre-run at 200 volts for one hour in TBE (89mM Tris (Sigma), 89mM Boric acid (Sigma), 2mM EDTA (Sigma)) tank buffer heated to 70°C. Reaction samples were added to the wells and were electrophoresed at 150 volts for two hours (until the dye front exited the gel completely). Gels were then wrapped in cellophane paper and exposed to Kodak CL-Xposure X-ray film (Sigma) for 24 hours at -70°C. The X-ray film was then developed in the dark by five minute incubation in Kodak developer (Sigma), followed by a one minute wash in distilled water. Finally, the developed film was fixed in Kodak Fixer (Sigma) for three minutes and washed with distilled water for 1 minute and allowed to air dry.

2.6.3. HIV-1 Integrase Scintillation proximity assay Strand Transfer Assay

2.6.3.1. Preparation of SPA Bead complex

Further functional analysis was undertaken utilizing the scintillation properties of streptavidin-coated scintillation proximity assay (SPA) beads (GE Healthcare). SPA beads were prepared by resuspension in SPA buffer containing NaCl (27.8mM HEPES (Sigma) pH 7.8, 27.8mM $MgCl_2$ (Sigma), 111.1µg/ml BSA (Sigma), 5.56mM β-mercaptoethanol (Sigma), 50mM NaCl (Sigma)) to a concentration of 10mg/ml. Biotinylated donor DNA from Table 2.2 was added to the suspension at a final concentration of 360nM. The mixture was then rocked at room temperature for 30 minutes. Thereafter, the SPA beads are centrifuged at 870 x g for five minutes and the supernatant discarded. The pellet was resuspended in 500µl of SPA buffer (NaCl+) and centrifuged at 870 x g for five

minutes. The supernatant was discarded and the pellet was resuspended in 2.5ml of SPA buffer containing no NaCl. IN proteins were diluted in Integrase Control Buffer (50mM Tris-HCl (Sigma) pH 7.8, 1M NaCl (Sigma), 2mM β-mercaptoethanol (Sigma), 20% glycerol (Melford Labs)) to 4µM and added to the prepared SPA bead suspension to a final concentration of 0.5µM. This solution was then rocked at room temperature for 30 minutes and the resultant IN/Donor/SPA bead complex was stored at 4˚C for not more than two weeks.

2.6.3.2. Radiolabelling of DNA

Donor and Target DNA (Table 2.2) were radiolabelled using the DNA 5' End-Labeling System (Promega, Madison, WI, USA) as per the manufacturer's instructions. Briefly, 10pmol of the oligomer substrate (D1, DP, DU and either target) was added to a 50 µl reaction of 1 x calf intestinal alkaline phosphatase reaction buffer and 0.1 units of calf intestinal alkaline phosphatase. Reaction mixtures were incubated first at 37°C for 15 minutes and then at 56°C for 15 minutes. Thereafter, an additional 0.1 unit of calf intestinal alkaline phosphatase was added and the above incubations were repeated. Tris-EDTA-saturated phenol:chloroform:isoamyl alcohol (25:24:1) (Sigma) was added to the reactions, vortexed and centrifuged at 12 000 x *g* for two minutes, and the upper aqueous phase was aspirated and placed in a clean microtube. The phenol extraction was repeated and the upper aqueous phase was extracted to a fresh microtube and NaCl was added to a final concentration of 0.5M. Two volumes of ethanol were then added and the solution was vortexed and centrifuged at 12 000 x *g* for 15 minutes. The ethanol was aspirated and the pellet was allowed to air dry until no liquid ethanol was present. The pellets containing the dephosphorylated oligomers were then resuspended in 29µl of distilled water. All oligomer substrates were then phosphorylated with ^{32}P, and Target DNA for the SPA bead Stand Transfer assay was also labeled with ^{33}P in a 50µl reaction containing the 29µl dephosphorylated target substrate (10pmol), 1X T4 polynucleotide kinase reaction buffer, 100µCi of ^{32}P or 50µCi ^{33}P 3000Ci/mmol Easy Tides Adenosine 5'-triphosphate (Perkin Elmer Inc., Boston, MA, US) and 10 units of T4

polynucleotide kinase. The reactions were incubated at 37°C for 10 minutes, and then stopped with 2µl of 0.5M EDTA. The solutions were then heat inactivated for 1 minute at 78°C. Thereafter 10pmol of complementary strands were added. The solutions were heated to 95°C for five minutes and allowed to cool to room temperature to allow for annealing.

2.6.3.3. SPA bead Strand Transfer Assay set up

Reactions were prepared in a 96 well plate in 100µl reactions by diluting 50µl of SPA bead complexes (section 2.6.3.1) in 40µl of Bead assay buffer (25mM HEPES (Sigma) pH 7.8, 62.5mM NaCl (Sigma), 31.25mM $MgCl_2$ (Sigma), 100µg/ml BSA, 5mM β-mercaptoethanol (Sigma)) or 30µl of Bead assay buffer and 10µl of chicoric acid with pre-incubation with inhibitor for 1 hour. Radiolabelled ^{33}P-Target DNA was diluted to 1µM with Target Dilution buffer (25mM HEPES (Sigma) pH 7.8, 100µg/ml BSA (Sigma), 5mM β-mercaptoethanol (Sigma)) and 10µl was added to reactions at a final concentration of 100nM and incubated at 37°C for one hour. Reactions were stopped with 50µl of 3 x Quenching solution (6M CsCl (Sigma), 0.5M EDTA (Sigma)). The scintillation was measured using a TopCount Microplate Scintillation Counter (Perkin-Elmer Inc.)

2.6.4. Data analysis

All data and graphs for the functional assays were analyzed in GraphPad Prism version 5.1. Unpaired Students t-tests were used to compare individual columns while a Non-Linear curve fitting for (Log)inhibitor vs response was used to calculate IC_{50} values.

CHAPTER 3

Results

3. Results

3.1. Amplification and sequencing of HIV-1subtype C integrase

An RT-PCR and PCR protocol for the amplification of an approximately 1.3kb fragment encompassing the HIV-1 *integrase* was successfully established. RT-PCR and PCR amplification was successful for 73 of the 80 (91.3%) study samples, and included all of the 22 proviral DNA samples (Figure 3.1), and 51 of the 58 patient plasma samples (Figure 3.2).

Patient viral loads for the samples that were successfully amplified ranged from greater than 3 million to 1500 RNA copies/ml (Table 2.1). Of the 51 samples that were amplified, 41.4% were from male patients.

All 73 amplicons were successfully sequenced and analyzed. Nucleotide sequences of the 73 newly characterized full-length *integrase* genes ranged from 867 to 876 base pairs, with intact open reading frames confirming the presence of functional *integrase* genes in all patients and primary virus isolates. The *integrase* intrasubtype nucleotide diversity for all 73 newly characterized samples varied between 2.2% and 7.9 %.

Phylogenetic analysis of the full-length *integrase* nucleotide sequences, together with viruses from the major subtypes revealed that 71 sequences clustered within HIV-1 subtype C with a bootstrap value of 100%, while two samples (09ZAP35 and 09ZAP52) showed evidence of recombination (Figure 3.3). These two sequences were further analyzed using the RIP 3.0 Program (http://hiv-web.lanl.gov) (Figure 3.4), which showed that sample 09ZAP52 was a unique GC recombinant with a region at the 3'-end of the gene that was unassigned, since it showed similarity to an AE recombinant, a CY recombinant or an A subtype. 09ZAP35 was shown to be a unique AC recombinant.

All sequences were submitted to Genbank using the Sequin (version 9.5) program available from the national centre for biotechnology information website (http://www.ncbi.nlm.nih.gov) and are available under the accession numbers GQ872465 to GQ872538.

The nucleotide sequences were converted to amino acid sequences and an alignment of all 73 integrase sequences was compared and extensively analyzed. A distribution of the variants is shown in Figure 3.5. The proteins varied from 288 to 292 amino acids in length, primarily due to insertion of an unusual tetra-peptide insertion (QNME) in the C-terminus of the protein in 23 samples.

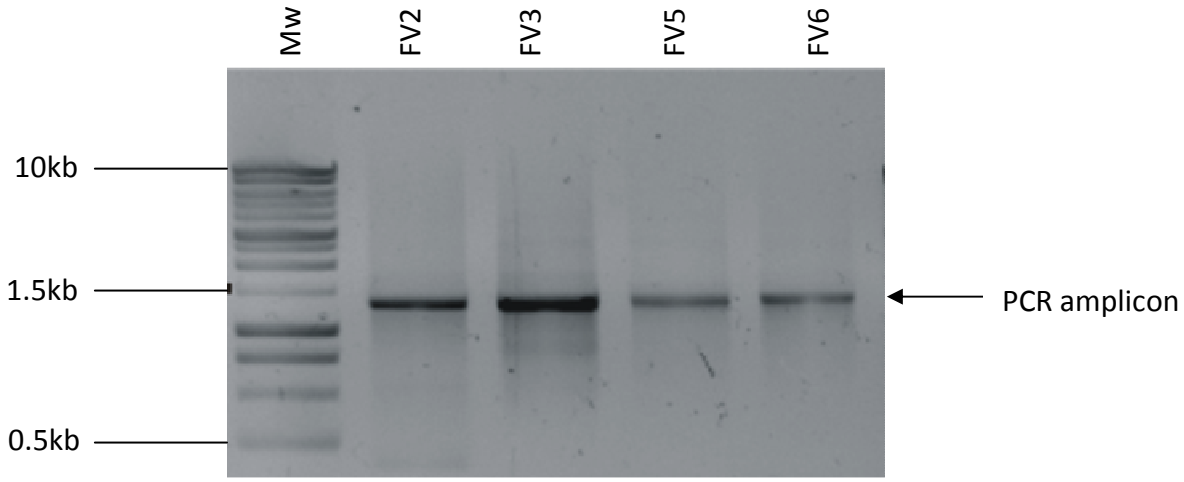


Figure 3.1: A representative 1% agarose gel showing *integrase* PCR products amplified from extracted proviral DNA. The names of the proviral isolates are indicated above the wells. A 1kb DNA size marker (New England Biolabs) is indicated on the left. The arrow indicates the approximately 1.3kb PCR amplicon.

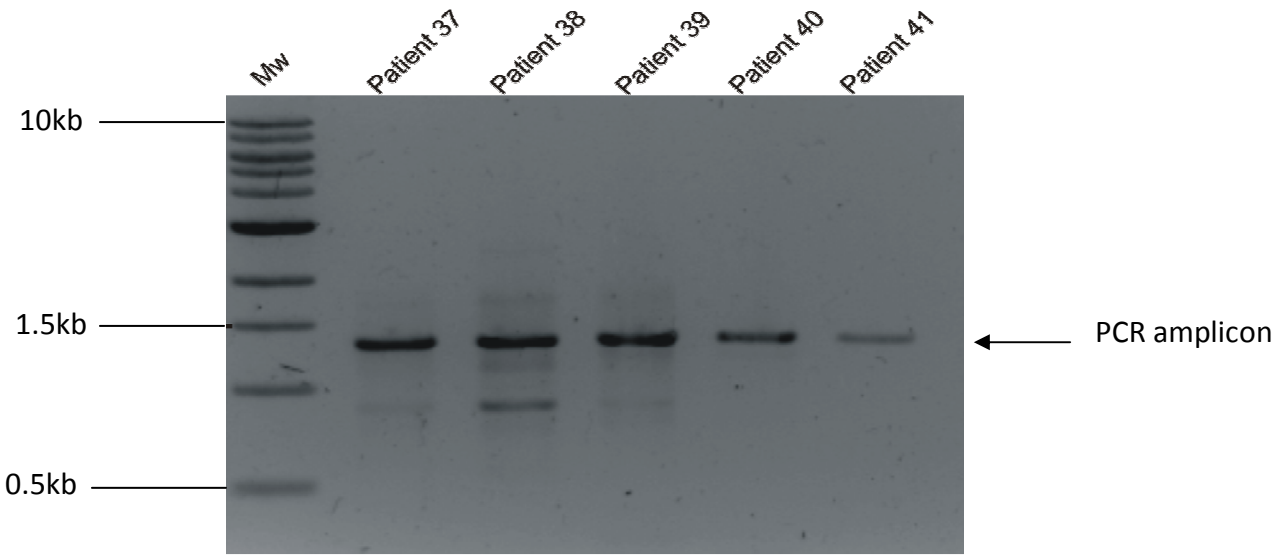


Figure 3.2: A representative 1% agarose gel showing *integrase* RT-PCR products amplified from extracted viral RNA. The patient numbers are indicated above the wells. A 1kb DNA size marker (New England Biolabs) is indicated on the left. The arrow indicates the approximately 1.3kb PCR amplicon.

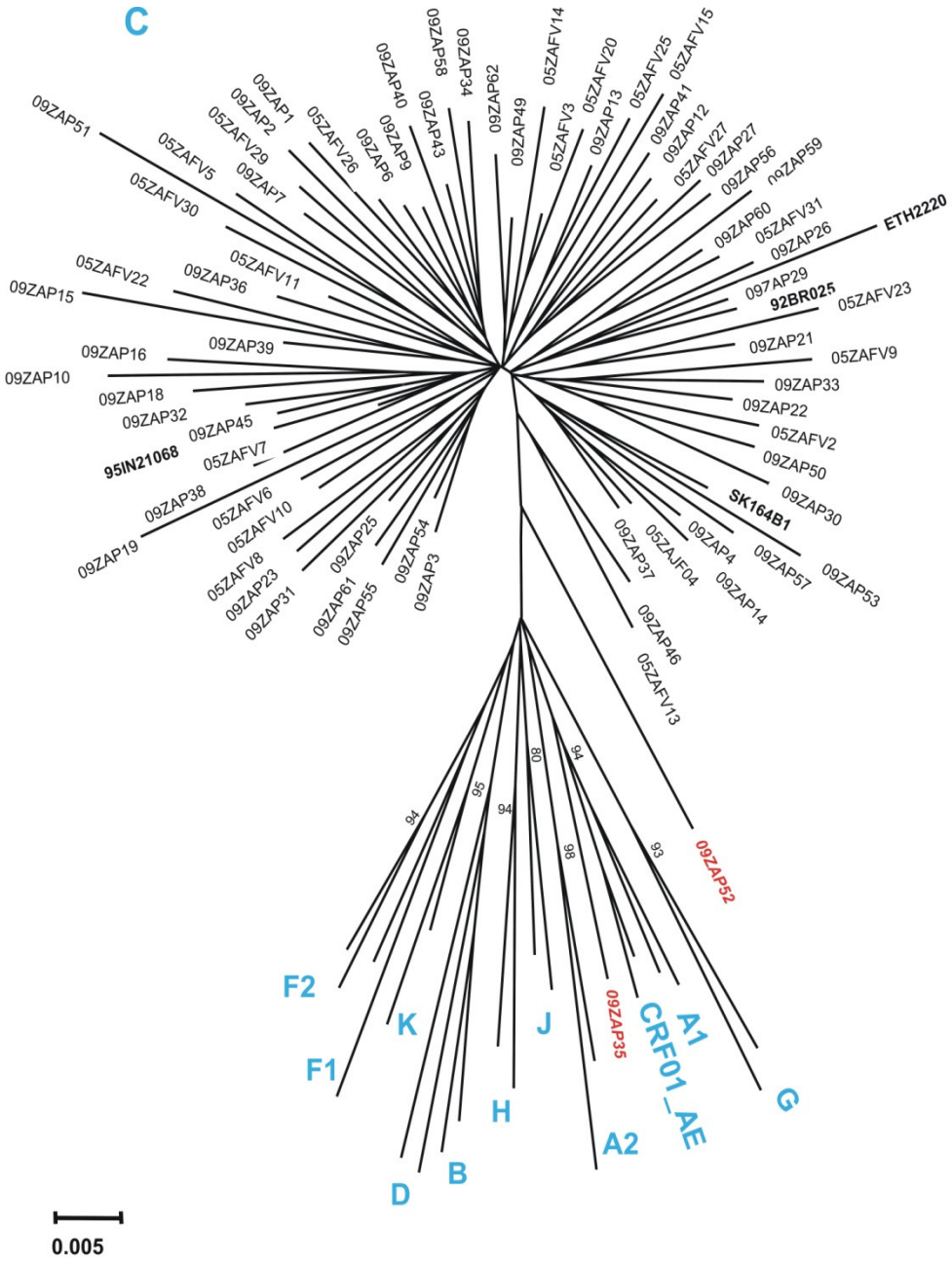


Figure 3.3: Phylogenetic relationships of the 73 newly characterized full-length integrase sequences with HIV-1 subtype reference sequences from the Los Alamos database (http://hiv-web.lanl.gov). The unique recombinant sequences are shown in red. Phylogenetic trees were constructed from nucleotide sequence alignments, using neighbor joining and the Kimura two-parameter distance matrix. The reliability of the branching order was estimated from 1000 bootstrap replicates, and only bootstrap values of 70% or higher are shown.

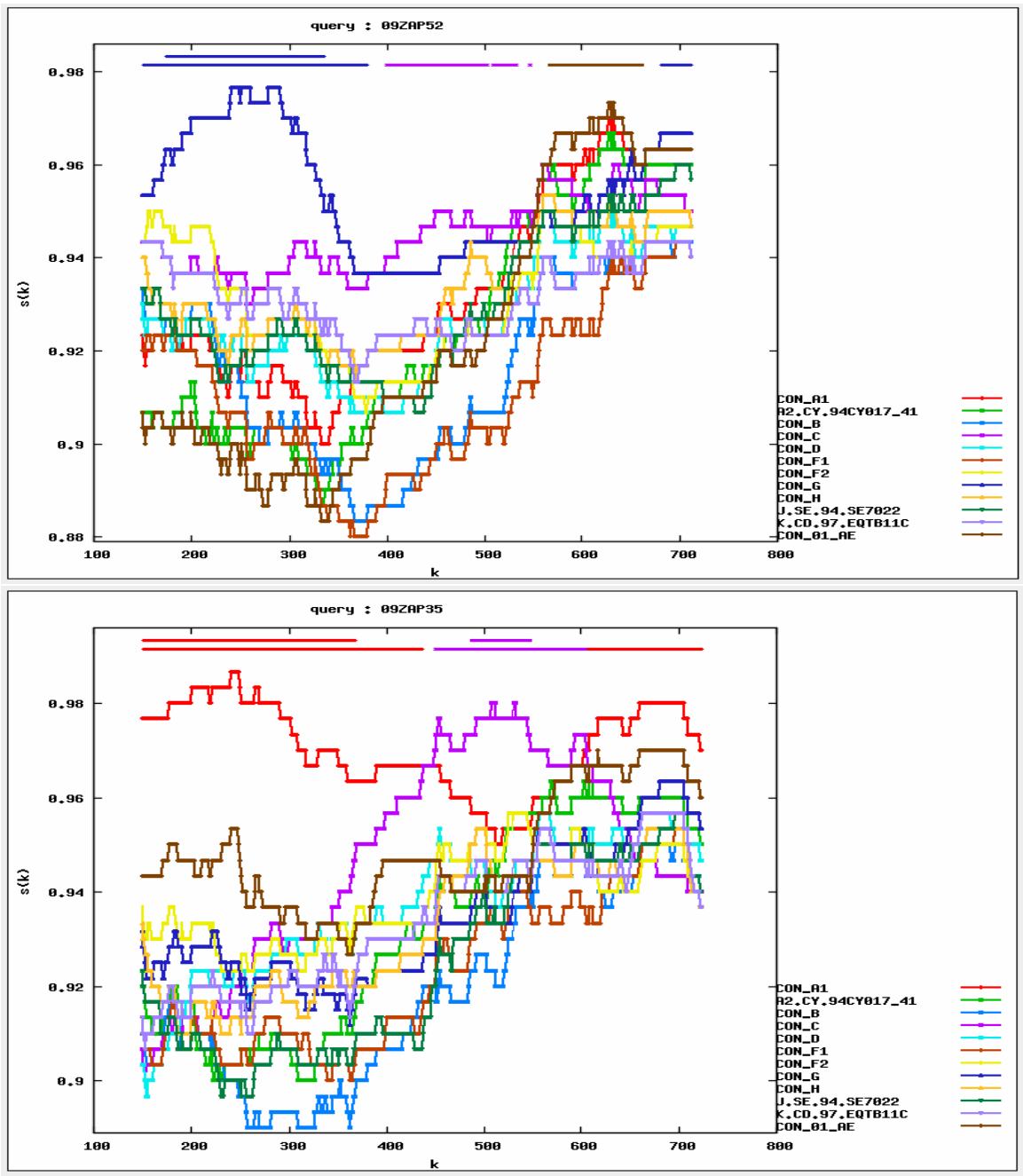


Figure 3.4: RIP 3.0 analysis of 09ZAP52 (top) and 09ZAP35 (bottom). The program compares the query sequence to a database of HIV-1 subtypes and selected recombinants and plots the similarity (s(K)) of the query sequence to background sequences (see key for different subtypes tested) at each nucleotide position (k). 09ZAP52 is shown to be a GC recombinant with part of its sequence not generally clear but possibly holding similarities to an AE recombinant around nucleotides 570 to 670. 09ZAP35 is shown to be an AC recombinant but also has high similarities to an AE recombinant between nucleotides 600 and 730.



Figure 3.5: Distribution of variants amongst the 73 newly characterized full-length integrase sequences. The consensus subtype C sequence is shown at the top of each 30 amino acid sequence section, below which is the number of annotated sequences containing an unambiguous amino acid at the indicated position. Beneath are the variants with the superscript numbers indicating the amount of times that they occurred. Domains and important residues are highlighted. The N-terminal domain is highlighted in yellow, the catalytic core domain in white and the C-terminal domain in blue. The catalytic triad residues are indicated in red and relevant integrase inhibitor resistance mutations are in green.

3.2. HIV-1 subtype C cell line infections, and detection of Preintegration Complexes

Attempts were made to infect all four cell lines (SupT1, CEM-SS, CEM-GFP and MT-2) with the HIV-1 subtype C FV3 viral isolate (CXCR4 utilizing). Despite repeated attempts, the CEM-SS and CEM-GFP cell lines were impervious to infection with FV3 (data not shown). By contrast, SupT1 and MT-2 cell lines were easily infected by FV3. Initial infection of SupT1 cells took approximately 21 days before the infection was detected by p24 antigen ELISA. FV3 virus adapted for growth on SupT1 cells was able to infect MT-2 cells within 5 days. Thereafter, viral infection of both MT-2 and SupT1 cells was allowed to continue up to day 13, as determined by p24 antigen ELISA (Figure 3.6) before use in all the SupT1 co-culture experiments.

PICs were isolated from SupT1/SuptT1 and MT-2/SuptT1 co-cultures, as described in sections 2.4.1 and 2.4.2 and analyzed by SDS-PAGE and Western blot analysis. However, despite repeated attempts, Western blot analysis of PIC isolation extracts, using the HIV-1 IN Monoclonal Antibody (8G4) antibody were negative for IN (results not shown). Further optimization of co-culturing conditions and different PIC isolation protocols were unsuccessful. Thus, purification of recombinant IN was undertaken.

3.3. HIV-1 subtype B recombinant Integrase Protein Isolation

Expression of the HIV-1 subtype B recombinant protein in transformed bacteria (*E. coli* BL21 pINSD.His) was detected as early as one hour post-induction with 1M IPTG, as determined by SDS-PAGE and Western blot analysis (Figure 3.7). Expression levels were sufficient after three hours post-induction with 1mM IPTG, and used to harvest bacterial cells in all subsequent experiments.

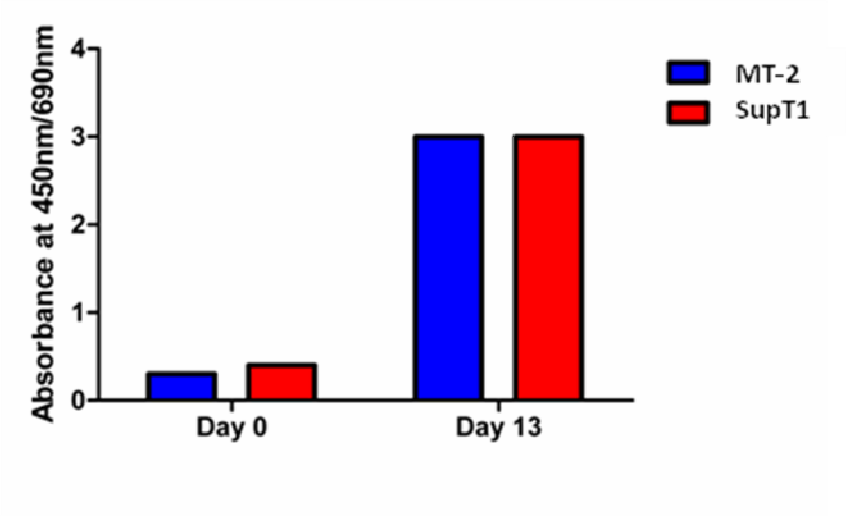


Figure 3.6: OD readings of the p24 antigen ELISA results showing infection of MT-2 cells and SupT1 with the FV3 viral isolate. OD readings of samples on day 0 and day 13 were taken. The key indicates cell lines infected. Levels of p24 were above the upper limit of OD reading of 3 on day 13.

The control HIV-1 HXB2 (subtype B) recombinant protein (IN_{wt}) was approximately 32kDa, as expected, whereas the expressed recombinant subtype B proteins was approximately 34.5kDa, due to the presence of extra amino acids from the His-tag and thrombin cleavage site.

The IN_{wt} was then tested for solubility by observing the bacterial cell fraction in which the protein appeared to be concentrated. SDS-PAGE and Western blot analysis of *E. coli* BL21 pINSD.His supernatant and pellet samples indicated that the protein was not secreted into the culture media (Figure 3.8.). Furthermore, the supernatant and pellet samples obtained after sonication of the cell lysate showed very low protein concentrations in the sonicated supernatant fraction, while most of the IN protein was concentrated in the sonicated pellet fraction (Figure 3.8). The presence of the protein, predominantly in the sonicated pellet fraction, indicates that the protein is insoluble. No attempts were made to purify IN from the insoluble fraction, and the soluble fraction became the focus of subsequent purification procedures for IN_{wt}.

The soluble fraction of the IN_{wt} was successfully purified by Nickel Chelating Column Chromatography, as determined by SDS-PAGE and Western Blot analysis (Figure 3.9). The protein purification yielded virtually uncontaminated soluble IN protein in the eluate. However, a significant amount of the expressed IN_{wt} was lost in the flow through or remained attached to the resin after elution (Figure 3.9).

In an attempt to optimize the purification protocol, the recombinant IN was eluted with different concentrations of imidazole. Figures 3.9 and 3.10 compare elution of IN at concentrations of 350mM and 500mM, respectively. When eluted with a higher concentration of imidazole there was no improvement of IN_{wt} detaching from the resin; however, an increased number of contaminating proteins was noted (Figure 3.10.). In addition, when the 350 and 500mM eluates were concentrated through the 10kDa size exclusion column they formed large aggregates. Because the purified recombinant protein yields were adequate for the purposes of this study, we did not attempt to further optimize the purification protocol.

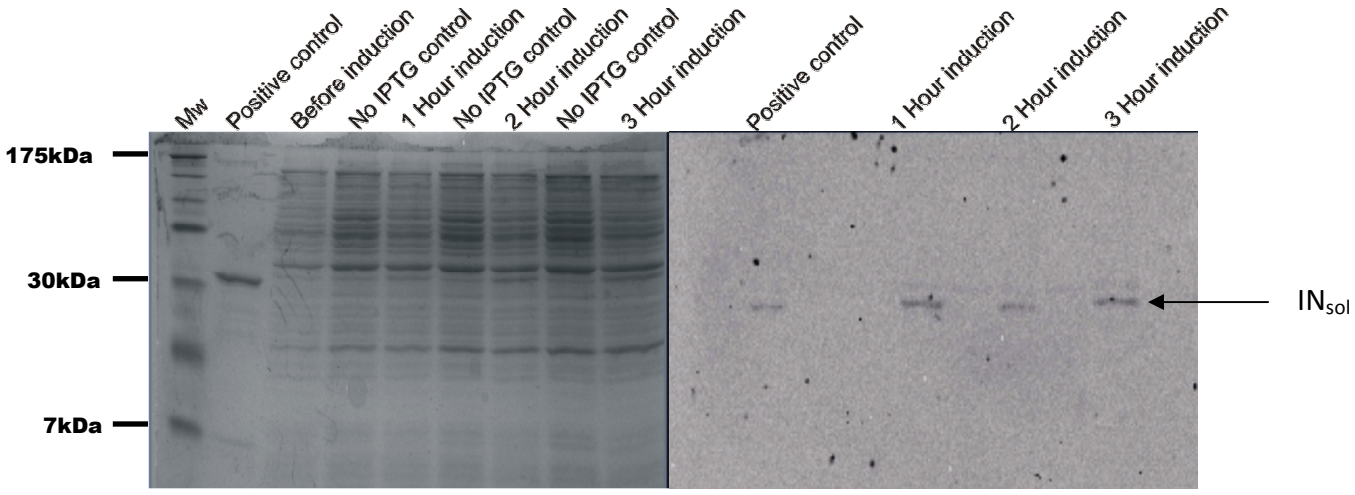


Figure 3.7: SDS-PAGE (left) and corresponding Western Blot (right) of induced BL21 pINSD.His cultures. Samples were taken at 1 hour intervals up to 3 hours after 1mM IPTG was added to the bacterial culture. Protein molecular weight markers (in kDa; kilodalton) are indicated on the left. The arrow indicates the expressed IN protein.

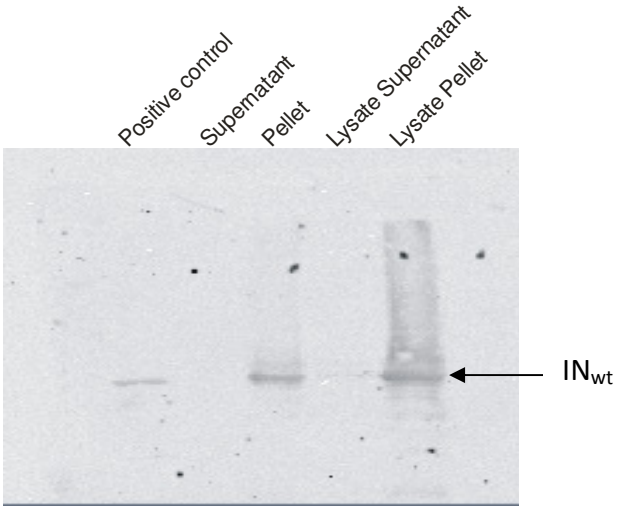


Figure 3.8: Western Blot analysis of induced *E. coli* BL21 pINSD.His bacterial samples. Lanes demarcated as Supernatant and Pellet were samples from *E. coli* BL21 pINSD.His that were not sonicated prior to centrifugation. Lanes demarcated as Lysate Supernatant and Lysate Pellet were samples taken after sonication and centrifugation. The Pellet sample indicates total protein (soluble and insoluble). The arrow indicates the expressed IN protein (approximately 34.5kDa).

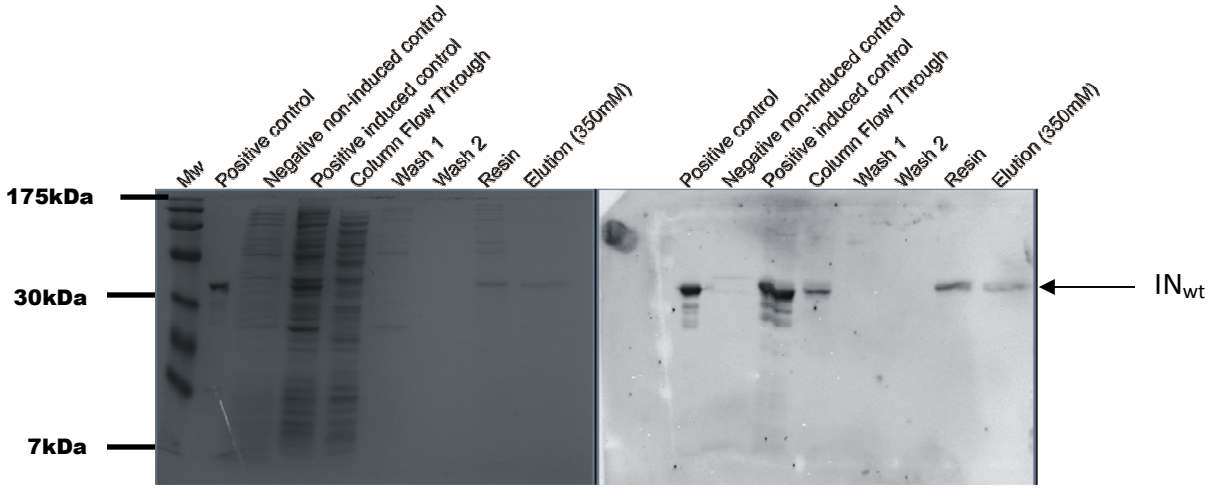


Figure 3.9: SDS-PAGE Gel (left) and corresponding Western Blot (right) of IN protein purification by nickel chelating column chromatography. This sample was eluted with 350mM imidazole. Protein molecular weight markers (in kDa; kilodalton) are indicated on the left. The arrow indicates the expressed and purified IN protein.

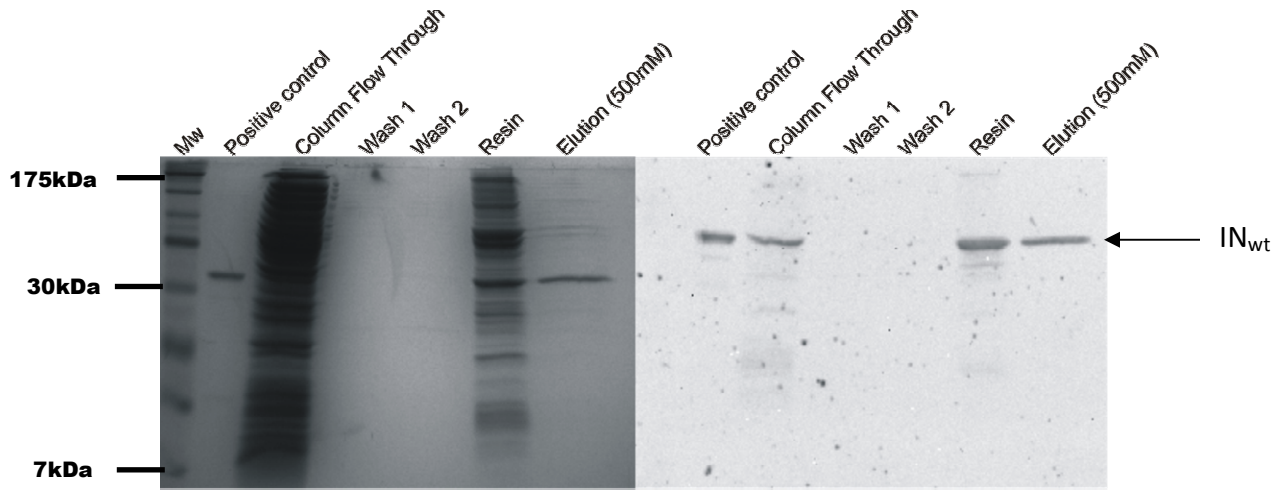


Figure 3.10: SDS-PAGE (left) and corresponding Western Blot (right) of IN protein purified by nickel chelating column chromatography. This sample was eluted with 500mM imidazole. Protein molecular weight markers (in kDa; kilodalton) are indicated on the left. The arrow indicates the expressed and purified IN protein.

The recombinant soluble IN (IN_{sol}) was successfully purified, as confirmed by SDS-PAGE (Figure 3.11). Because the recombinant IN_{sol} was purified on the AKTA system, analysis of different fractions (flow rate of 1.5ml/min) was possible. The bulk of protein eluted between fractions 23 to 60, with the majority detected in fractions 24 and 25. Some contaminating proteins were detected in the early eluted fractions (Figure 3.11). However, IN_{sol} was detected as late as fraction 60 with little or no contamination visible on the gel. For this reason, fractions 23 to 30 and 31 to 60 were pooled and analyzed separately.

Quantification of the purified IN_{wt} and IN_{sol} showed the samples were concentrated to 559µg/ml and 740µg/ml, respectively as per results of the BCA assay, and using the Beer-Lamberts law: $A=\varepsilon lc$. Where the extinction coefficient (ε) of IN is calculated at 50 460 M^{-1} cm^{-1} (Jenkins *et al.* 1996) at 280nm for IN, The Molar (M) concentrations were calculated for IN_{wt} and IN_{sol} to be 2µM and 6µM as per spectrophotometric results, respectively.

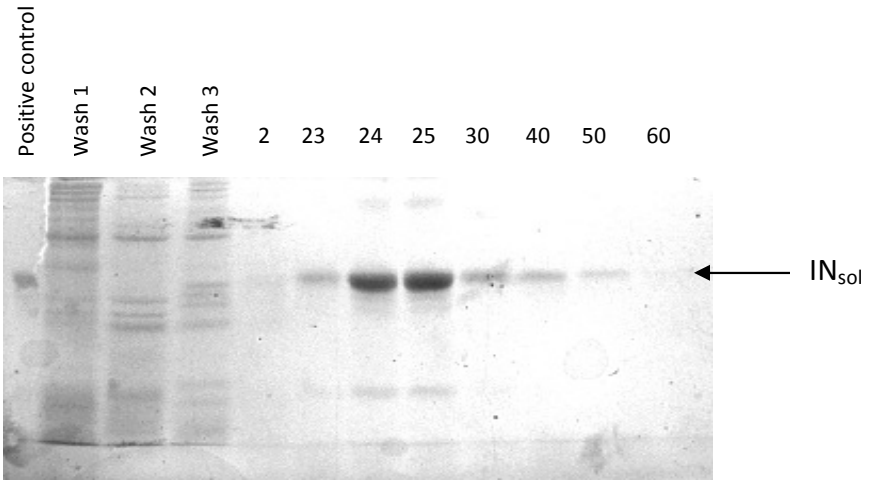


Figure 3.11: SDS-PAGE analysis of the purification of IN_{sol}. The enzyme is clearly noticed in the 24th and 25th fractions, however, contamination is noticed in these fractions. The protein eluted between fractions 23 and 60. The arrow indicates the purified IN protein.

3.4. Recombinant IN and IN_{sol} activity

3.4.1. High throughput microtitre plate HIV-1 integration assay

After several attempts and optimization of concentrations of donor DNA, Target DNA and IN, the high throughput microtitre plate HIV-1 integration assay was successfully set up.

The purified IN_{wt} and IN_{sol} were both tested for functional activity in the high throughput microtitre plate HIV-1 integration assay. Activities of the IN_{wt} and IN_{sol} with a single concentration of Target DNA (150nM), as measured by fluorescence, are shown in Figure 3.12A. Results confirmed that IN_{sol} showed enzymatic activity (Figure 3.12A; yellow bar) compared to the blank control ($p=0.0136$), whereas the IN_{wt} showed little to no enzymatic activity (Figure 3.12A; red bar) compared to the blank control ($p=0.1381$), as compared to the background (Figure 3.12A; blue bar). Overall, IN_{wt} showed little to

no enzymatic activity, and was excluded from further assays. IN_{sol} can initiate strand transfer and was used in all subsequent experiments.

A dose response curve of IN_{sol} to varying Target DNA concentrations is shown in Figure 3.12B. IN_{sol} showed an increase in activity proportional to increasing Target DNA concentrations, reaching a plateau at high concentrations of Target DNA. These results serve to further confirm that IN_{sol} is functional.

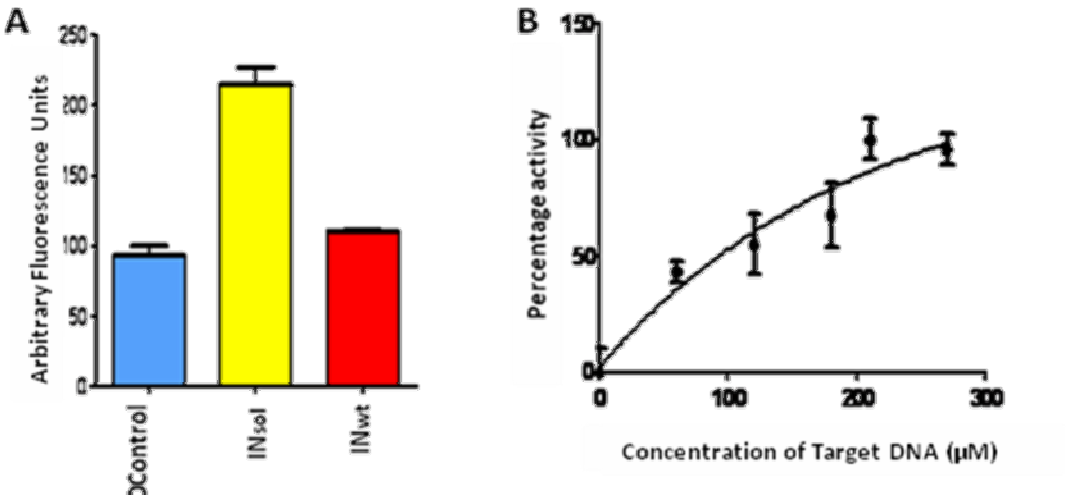


Figure 3.12 : HIV-1 recombinant integrase activity from the high throughput microtitre plate HIV-1 integration assay, as measured in arbitrary fluorescence units (excitation 485nm and emission 535nm). A) Purified recombinant IN activity results using only a single Target DNA concentration (150nM). The **blue**, **yellow** and **red** bars indicate the fluorescence readings for the negative control (no-IN), test containing 1 µM of IN_{sol}, and test containing 1µM of IN_{wt}, respectively, showing that IN_{sol} is active. All tests were done in duplicate. B) Dose response of IN_{sol} at different ranges of Target DNA concentrations. Results were normalized to a negative control containing no IN at maximum Target DNA concentration (270nM), and show that IN_{sol} activity increases as the Target DNA concentration increases. The curve has a R^2 value of 0.8629. All tests were done in duplicate.

The modifications to the high throughput microtitre plate HIV-1 integration assay protocol using IN_{sol} allowed for a more absolute quantification of substrate reactivity (Figure 3.13). Results obtained from the modified protocol were also more reproducible and showed a greater significant difference between the blank control and enzymatic test ($p=\leq 0.0001$) compared to the unmodified assay.

‘

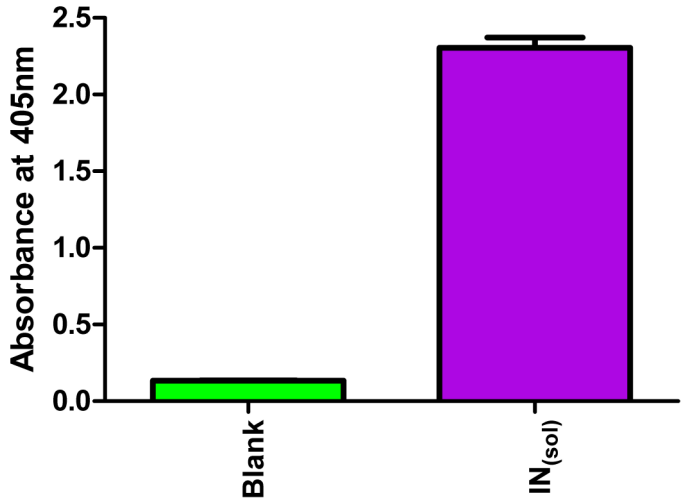


Figure 3.13: HIV-1 recombinant integrase activity from the modified high throughput microtitre plate HIV-1 integration assay, as measured by absorbance (405nm). Results show a definitive difference between IN_{sol} and the blank control. Heights of the columns indicate an average mean of a triplicate.

The modified protocol was used to test the chicoric acid control inhibitor (works similarly to Raltegravir in that it binds to the IN active site). Results confirmed that the assay can successfully be used to test IN inhibitors, since they showed a dose response curve: as the chicoric acid concentration increases, the recombinant IN activity decreases (Figure 3.14). The IC_{50} was calculated at 101.5nM of inhibitor.

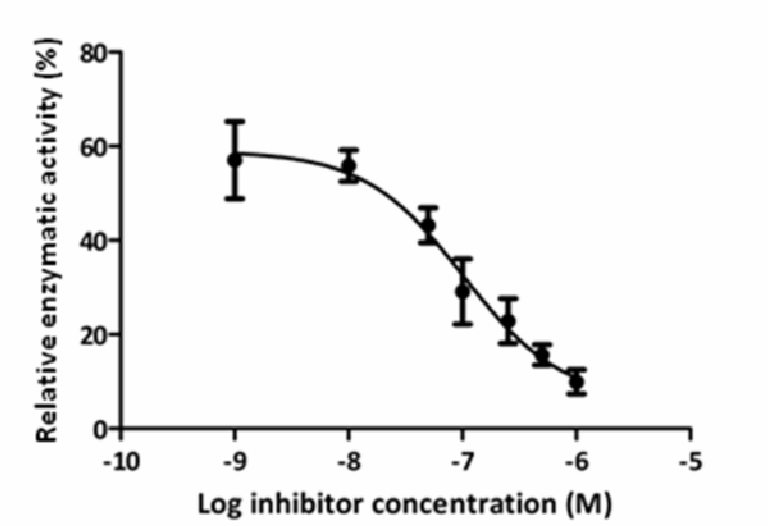


Figure 3.14: HIV-1 recombinant integrase (IN) activity from the modified high throughput microtitre plate HIV-1 integration assay, as measured by absorbance (405nm) and normalized to a positive control (%). As the concentration of the chicoric acid inhibitor increases, the recombinant IN activity decreases. Points indicate average mean of a triplicate. IC_{50} is calculated as 101.5nM on GraphPad Prism 5.1.

3.4.2. HIV-1 Integrase Urea-PAGE Based Functional Analysis

Despite repeated attempts, the Urea-PAGE strand transfer assay was not successfully established.

Figure 3.15 is a representative autoradiograph of the Urea-PAGE strand transfer assay. The Target DNA was 5'-radiolabelled on both the top and bottom strands. Radiolabelling of DNA worked, since Target DNA, as well as the control donor DNA was detected for all experiments. Because strand transfer can excise the Target DNA at any point to integrate the donor DNA, and a single Target DNA can accept more than one integration product from different donor stands, successful strand transfer detection on the autoradiograph would show ladder formation/pattern above and below the labelled Target DNA substrate. This was not evident on any developed autoradiographs, although in the experiment depicted in Figure 3.15, the donor DNA negative control (Figure 3.15; lane 4) gave a false positive result, and had the type of ladder pattern that is expected from ST activity.

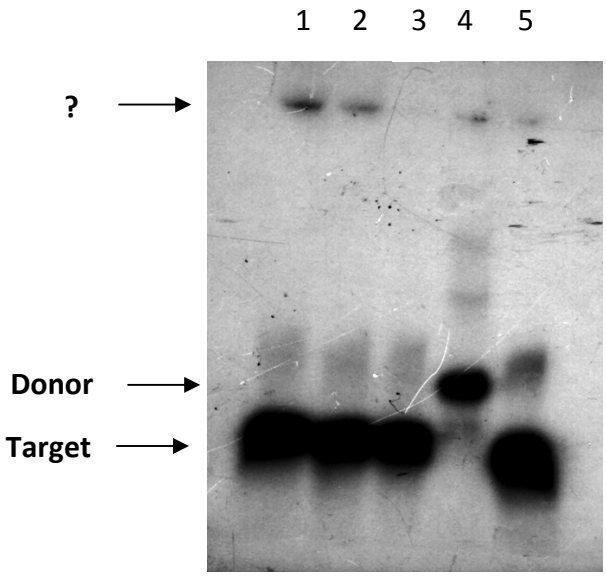


Figure 3.15: Representative autoradiography of the Urea-Gel based assay using Assay Buffer 2. Labelled Target DNA and Donor (32 mer) DNA were mixed with IN_{sol} from purification A (lane 1), IN_{sol} from purification B (lane 2), or a no IN enzyme control (lane 3). Labelled Donor DNA only and labelled Target DNA only controls are shown in lanes 4 and 5, respectively.

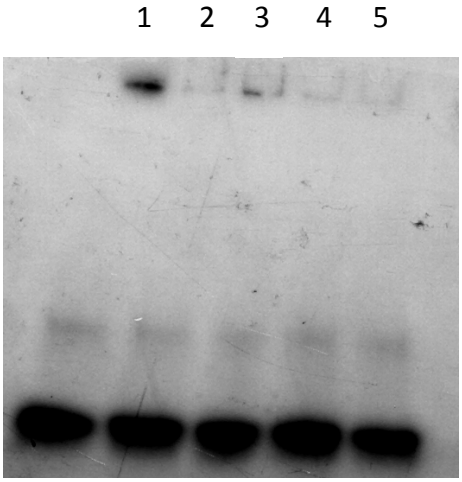


Figure 3.16: Representative autoradiography of the Urea-Gel based assay depicting use of Assay Buffer 1. Lanes are indicated. Labelled Target DNA and labelled Donor 1 (Table 2.2) DNA were mixed with IN_{sol} from purification A (lane 1), IN_{sol} from purification B (lane 2), or a no IN enzyme control (lane 3). Labelled Donor DNA only and labelled Target DNA only controls are shown in lanes 4 and 5, respectively. Donor DP (Table 2.2) concentration was 60nM and target concentration was 100nM.

The uppermost top bands detected in Figure 3.15 (labelled "?") may be Target DNA-IN aggregates that have not entered the gel during electrophoresis. Similar results were experienced when using shorter donor substrates (Table 2.2, DP; Figure 3.16).

3.4.3. SPA bead Strand Transfer Assay

The SPA bead strand transfer assay proved more difficult to set up and optimize, as compared to the high throughput microtitre plate HIV-1 integration assay. The IN_{sol} was tested for functional activity in the SPA bead assay, and measured in counts per million (CPM; Figure 3.17). Preliminary results show CPM values obtained for the SPA beads complexed with IN_{sol} indicated a higher scintillation count as compared to the negative control containing no-IN enzyme. These results were, however, not reproducible. The control inhibitor, chicoric acid, was used in the SPA bead strand transfer assay. As the inhibitor concentrations increased the enzymatic activity decreased, indicated by a sigmoidal curve (Figure 3.18) fitted to readings taken in duplicate. The IC_{50} was calculated to be 248.5nM, but these results were not repeated in an independent experiment.

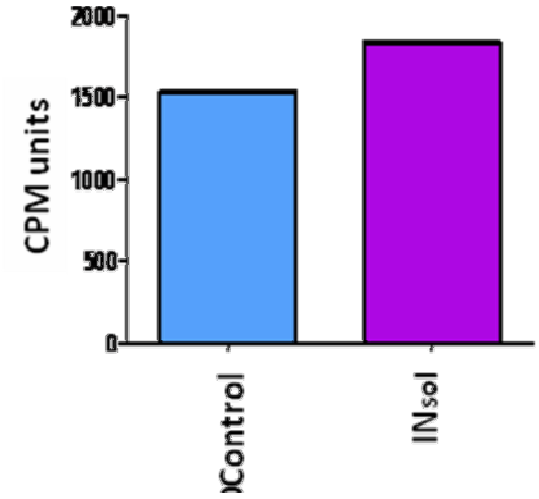


Figure 3.17: SPA bead assay activity results. The bead assay results indicate that there is ST activity when comparing the IN_{sol} to the Blank control.

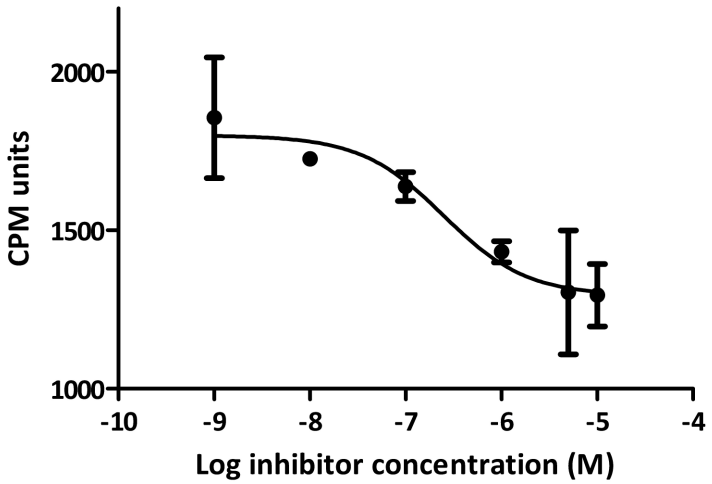


Figure 3.18: SPA bead assay with increasing inhibitor concentration. As the concentration of the chicoric acid inhibitor increases, the recombinant IN_{sol} activity decreases. Points indicate the average mean of a duplicate. IC_{50} is calculated as 248.5nM on GraphPad Prism 5.1.

CHAPTER 4

Discussion

4. Discussion

This study describes the successful establishment of an HIV-1 subtype C specific *integrase* genotyping assay and description of baseline mutations in *integrase* in HIV-1 subtype C (and unique recombinant) infected integrase inhibitor-naive individuals in South Africa. In addition, recombinant IN was expressed and purified and tested in three strand transfer assays, of which the high throughput microtitre plate HIV-1 integration assay was optimized.

RT-PCR and PCR amplification was successful for 91.3% of all samples, and included 51 of the 58 patient plasma samples (not dependent on viral load), and all 22 proviral DNA samples, respectively. All 73 of the amplified samples were successfully sequenced. The *integrase* amplification and sequencing success rates are similar to those reported previously (Van Laethem *et al.* 2008; Eshleman *et al.* 2009; Van Baelen *et al.* 2009). Nucleotide sequences of the 73 newly characterized full-length *integrase* genes ranged from 867 to 876 base pairs, with intact open reading frames confirming the presence of functional *integrase* genes in all patients and primary virus isolates. The *integrase* intrasubtype nucleotide diversity for all 73 newly characterized samples varied between 2.2% and 7.9%, similar to what has been reported previously (Rhee *et al.* 2008).

As expected, the predominant circulating subtype within the South African population was HIV-1 subtype C (Figure 3.3), while the two samples (09ZAP35 and 09ZAP52) were unique AC and CG recombinants in *integrase* (Figure 3.4). Since subtype A predominates in East, Central and West Africa, and subtype G predominates in Central and West Africa it is possible these patients originated from these countries and a recombination event occurred in South Africa.

The distribution of IN variants analyzed in Figure 3.5, showed that the newly characterized IN proteins varied from 288 to 292 amino acids in length, primarily due to the presence of an unusual tetrapeptide insertion (QNME) in the C terminus of 23

samples. This unusual insertion has been noted previously in two HIV-1 subtype C samples obtained from South Africa (Hunt *et al.* 2003; Tzitzivacos *et al.* 2009). The C-terminus has been shown to be important for IN dimerization and DNA binding. This is believed to be crucial for proper function and positioning of the catalytic domain close to its target site (Michel *et al.* 2009). A CTD IN deletion mutant (terminal 15 amino acids) has been shown to be 80% catalytically impaired (Dar *et al.* 2009), therefore this extension may play an interesting role in IN catalytic function. This finding warrants further investigation to determine the impact of the insertion on the function of the integrase enzyme. It would be useful to obtain a recombinant IN protein with this extension and conduct functional assessments determining any differences in catalytic activity. This could be achieved by cloning a representative subtype C IN enzymes with this extension into a suitable vector, expressing, purifying and testing the functionality using the modified high throughput microtitre plate HIV-1 integration assay developed in this study.

The sequences were analyzed for the presence of integrase mutations associated with reduced *in vitro* susceptibility to RAL (E92Q, F121Y, E138AK, G140AS, Y143RCH, S147G, Q148HRK, N155HS) and ELV (T66IAK, E92Q, F121Y, E138AK, G140AS, S147G, Q148HRK, S153Y, N155HS, R263K) (Stanford Database, updated 15 September, 2009). None of the samples contained the mutations associated with primary resistance (Y143RCH, Q148HRK and N155HS) (Figure 3.5).

The E157Q mutation was observed among 4.1% (n=3) of patient samples, including two subtype C patient samples (09ZAP34 and 09ZAP53) and one recombinant (09ZAP52) patient sample (Figure 3.5). Until recently the E157Q mutation was listed among the mutations associated with reduced *in vitro* susceptibility to raltegravir and elvitegravir (http://hivdb.stanford.edu), but is now described as a minimally polymorphic mutation that is selected *in vitro* by elvitegravir and rarely *in vivo* by raltegravir, having minimal, if any effect on integrase inhibitor susceptibility. Interestingly, amongst several other

studies characterizing a total of 501 subtype C IN sequences, (Rhee *et al.* 2008; Van Laethem *et al.* 2008; Eshleman *et al.* 2009; Passaes *et al.* 2009) the E157Q mutation was only reported in one sample (1.4%) (Eshleman *et al.* 2009).

The isoleucine at position 72 (I72) was identified at a high frequency amongst the subtype C samples and the 05ZAP52 recombinant (67.1%; n=49), whereas V72 and T72 was found in 28.7% (n=21, including the 05ZAP35 recombinant) and 1.4% (n=1) of the patients, respectively, confirming that it is a naturally occurring subtype C polymorphism. This confirms the findings by Rhee *et al.* (2008) who characterized naturally occurring variation in previously published HIV-1 group M integrase sequences, and found that 390 of 432 (90.3%) integrase-inhibitor-naive subtype C samples analyzed contained isoleucine at position 72, and only 9.4% and 0.2% contained a valine and threonine, respectively. The polymorphic mutation, T97A, associated with integrase inhibitors was observed in one primary virus sample (05ZAFV27), while K156N, another accessory integrase inhibitor resistance mutation was found in patient 09ZAP34. The uncommon polymorphism K156N was not reported in subtype C, but occurs in 9% of subtype B and 5% of subtype D sequences (Rhee *et al.* 2008). By contrast, T97A was observed in subtypes A, B, C, D and CRF01_AE and CRF02_AG, albeit at low frequencies (Rhee *et al.* 2008).

V201I has also been described as a resistance mutation that affects other earlier IN inhibitors (Rhee *et al.* 2008), however this mutation was not seen in any of the newly characterized patient samples. However, I201 is present in all 73 sequences, which indicates that HIV-1 subtype C is possibly predisposed to resistance to the earlier IN inhibitors such as S-1360. Similarly, the V165I, which was present in three samples (05ZAFV27, 05ZAP3 and 05ZAP54) is a resistance mutation possibly conferring resistance to the earlier IN inhibitors.

Because the antiretroviral drug history of the patient samples analyzed in 2009 was unknown, we could not address the influence of prior ARV therapy on integrase polymorphisms. Ceccherini-Silberstein (Ceccherini-Silberstein *et al.* 2007) reported several integrase polymorphisms, for example, M154I, V165I and M185L, that were positively associated with specific reverse transcriptase mutations (F227L and T215Y) in treated patients. In our study, the V165I polymorphism was noted in 3 samples, 09ZAP3, 09ZAP54 (unknown drug history) and 05ZAFV29 (antiretroviral drug naive). This highlights the need for ongoing surveillance of integrase mutations, as well as HIV-1 subtype C specific phenotypic studies to elucidate the relevance of the abovementioned polymorphisms (as well as naturally occurring polymorphisms) in the absence of the signature integrase mutations on susceptibility/resistance to integrase inhibitors.

Overall, none of the three previously reported major mutations (Y143R/C/H, Q148H and N155H) associated with resistance to integrase inhibitors were observed in our cohort, indicating that integrase inhibitors may be used as a treatment option in South Africa. This genotyping assay is available and can be implemented if/when raltegravir becomes available in South Africa.

Extensive IN sequence analysis from the 22 viral isolates led to the identification of suitable HIV-1 subtype C candidate viruses for PIC isolation and activity assays. The initial methodologies for HIV-1 infection and PIC isolations were identical to those described by (Brooun *et al.* 2001), with the exception of using a different viral isolate. Because isolate 05ZAFV3 was not capable of infecting CEM-SS cells, other cell lines were used in co-culture experiments. Isolation of PICs from SupT1/SupT1 and SupT1/MT-2 co-culture experiments were perceived to be a failure by SDS-PAGE and Western blot analysis (results not shown). This may be attributed to lower concentrations of isolated PICs from the SupT1/SupT1 and SupT1/MT-2 co-culture experiments, or the suitability of Western blots as a detection method. In addition, the monoclonal antibody used was specific for the HXB2 strain of HIV-1 subtype B (Nilsen *et al.* 1996). Therefore, the

antibody may not have bound optimally to the subtype C IN enzyme present in the purified PICs.

The published PIC isolation methods are all similar/identical, therefore future optimization studies should focus on using the same viral isolate (HXB2) used in the publication by (Brooun *et al.* 2001) as control. In addition, a more sensitive detection method such as the real time PCR assay and Southern blot analysis that detects integrated oligonucleotide products can be established (Engelman *et al.* 2008).

Because our attempt to isolate HIV-1 subtype C PIC's was unsuccessful, the study focus shifted to the isolation of recombinant IN enzymes, IN_{wt} and IN_{sol} for the establishment and use in strand transfer assays. Both proteins were successfully expressed and purified (Figures 3.9 to 3.11). Because several groups have shown that the soluble (Bushman *et al.* 1993; Hazuda *et al.* 1994; John *et al.* 2005) and insoluble (Bushman *et al.* 1993) fractions of IN_{wt} are active, and sufficient yields of IN_{wt} were obtained from the soluble fraction in this study, the insoluble aggregates were not refolded. The Bushman *et al.* (1993) and John *et al.* (2005) groups used the same expression plasmid (pINSD.His) as the one used in this study, and showed that the purified soluble IN fraction had enzymatic activity in gel-based assays and microtitre plate assays. In addition, the purification of the soluble fraction of the IN_{wt} enzyme employed in this study was similar to the methods described by Bushman *et al.* (1993) and John *et al.* (2005), with the exception of centrifugation speeds and thrombin cleavage of the His-tag. The IN_{sol} (double mutant) was purified under native conditions as described previously (with the exception of the His-tag cleavage), where it was shown to have catalytic activity similar to that of IN_{wt} enzyme (Jenkins *et al.* 1995; Jenkins *et al.* 1996). The F185K and the C280S mutations of IN_{sol} increase the enzyme's solubility when expressed in bacterial cells, and allow for easier handling when purifying and concentrating the recombinant enzyme.

The recombinant IN_{wt} and IN_{sol} proteins in this study were expressed with an N-terminal His-tag, and there was no thrombin cleavage step to remove the His-tag. Bushman *et al.* (1993) removed the His-tag from their purified IN_{wt} protein, whereas John *et al.* (2005) did not. This implies that removal of the His-tag does not impact on the IN_{wt} or IN_{sol} enzymatic function. Furthermore, His-tag removal from IN_{sol} and a subtype C IN (cloned from pINDIE-C1) was shown to be unnecessary and does not impact on IN activity (Bar-Magen *et al.* 2009).

There are two major published assays (with numerous modifications) that utilize recombinant IN enzymes to monitor ST activity (Chow 1997; Andrake *et al.* 2009). These being the gel-based assays and high throughput solid support assays. Three assays which cover the range of the available assays were selected to be set-up and optimized in this study. These included the high throughput microtitre plate HIV-1 integration assay (Hazuda *et al.* 1994; Vink *et al.* 1994), HIV-1 Integrase Urea-PAGE Based Functional Analysis (Bushman and Craigie 1991; Jenkins *et al.* 1996) and the SPA bead ST assay (Grobler *et al.* 2009).

The rationale behind the high throughput microtitre plate HIV-1 integration assay utilizes oligonucleotides that mimic the U5 LTR region of HIV-1 (Bushman and Craigie 1991) and is known as the Donor DNA, which is biotinylated at the 5'-end (Table 2.2). The Donor DNA is designed to be preprocessed (Bar-Magen *et al.* 2009), therefore, the assay does not test 3'-P and only specifically tests ST activity. During the assay, the IN dimer assembles and is stabilized on the donor substrate (Ellison and Brown 1994). The Target DNA is then added, and the assembled IN/Donor then nicks the target at a non-specified location, but prefers GC rich areas (Hazuda *et al.* 1994). IN also has nonspecific nuclease activity (Engelman and Craigie 1995) and therefore the Target DNA has a FITC label on both 3'-ends of the double stranded oligonucleotide so as not to lose the signal. After the Target DNA is ligated into the Donor DNA, the IN is washed off and only the strand transfer products remain in the well. Fluorescence can thereafter be measured to

detect ST activity, or as in the case of the modified protocol, an HRP-conjugated antibody against FITC could be used with subsequent detection via absorbance of the HRP substrate.

The high throughput microtitre plate HIV-1 integration assay and a modification thereof, was set up according to published methods (Hazuda *et al.* 1994; Chow 1997), with several exceptions. The experimental differences of the present study with published work included different assay conditions and the oligonucleotide sequences used. The assay conditions used in this study included two different buffers tested, i.e. Assay Buffer 1 and Assay Buffer 2 (see section 2.6.). Assay Buffer 1 was shown to allow for better IN activity in the high throughput microtitre plate HIV-1 integration assay, as compared to Assay Buffer 2 (data not shown). These buffers were chosen based on magnesium efficiency (Engelman and Craigie 1995) and previously published protocols (Hazuda *et al.* 1994; Jenkins *et al.* 1996; Bar-Magen *et al.* 2009). Comparison of the conditions for activity assays of published work showed that the maintenance of a pH between 6.5 and 7.5 is important (John *et al.* 2005) for IN enzymatic activity *in vitro*, where Good's buffers (Good *et al.* 1966) with pKa values close to these pH ranges are ideal. The ionic strength plays a role in activity depending on the metal ions present, and should not exceed 100mM NaCl. Addition of a mild reducing agent such as DTT or β-mercaptoethanol is essential to maintain the solubility of the IN enzyme. The critical components required for strand transfer are the metal ions such as $MgCl_2$ or $MnCl_2$ which interact with the active site of the IN enzyme. Glycerol and $ZnCl_2$ were also shown to enhance activity (Engelman and Craigie 1995; Zheng *et al.* 1996) and these were also included in the buffers used in this study (See section 2.6 and Table S6.1). In addition, DMSO, PEG and Nonidet-P40 (now known as Igepal) are also commonly used reagents in buffers (Jenkins *et al.* 1996), and save for Nonidet-P40 the prior two components were used in Assay Buffer 2. Although the compositions of buffers seem to vary between published methods, the abovementioned principles were adhered to in the present study. Besides the buffer conditions, the oligonucleotide donor substrates used in this

study were longer (32 mer) than commonly used (21 mer) substrates. However, the longer substrates are still a mimic of the U5 LTR sequence that IN is specific for, and oligonucleotide donor substrates of up to 35 mer have been used in activity assays before and shown to improve concerted integration (Hazuda *et al.* 1994; Engelman and Craigie 1995; Chow 1997; Andrake *et al.* 2009; Bar-Magen *et al.* 2009).

Results from the high throughput microtitre plate HIV-1 integration assay using IN_{wt} and IN_{sol} purified during the course of this study, showed differing results for the IN_{wt}, but similar results for the IN_{sol} (Hazuda *et al.* 1994; John *et al.* 2005; Bar-Magen *et al.* 2009). The enzymatic activity of the soluble fraction of IN_{wt} could not be reproduced in this study (Figure 3.12). Because the expressed/purified recombinant IN_{wt} protein was the correct size (Figures 3.9 and 3.10), the lack of enzymatic activity could most likely be attributed to incorrect folding of the protein. The addition of reducing agents such as DTT or β-mercaptoethanol in the assay buffers (see sections 2.6.1 - 2.6.2) helps to limit aggregation. However, it is possible that the concentration of the purified recombinant IN_{wt} protein was far too high, causing it to aggregate. These aggregates were noticed during concentration with the size exclusion columns (section 2.5.). The subsequent dialysis with a buffer containing DTT or β-mercaptoethanol may not have been sufficient to maintain the solubility of the protein. Future purification procedures may require a reducing agent to be present throughout the purification.

By contrast, the purified IN_{sol} showed about a two-fold activity over a negative blank (Figure 3.12A) and about a ten-fold activity over the negative in the modified protocol (Figure 3.13). In addition, activity of the enzyme showed a dose dependency curve with increasing concentrations of Target DNA, where these results indicate that the IN enzyme begins to get saturated with target and only a limited amount of substrate is integrated (Figure 3.12A). These results are similar to those reported previously (Hazuda *et al.* 1994; John *et al.* 2005; Wang *et al.* 2005). Where Hazuda *et al.* (1994) showed a hyperbolic curve of activity with increasing concentration of Target DNA. Wang *et al.*

(2005) showed a two-fold increase in activity after one hour over a control using a similar method to the high throughput microtitre plate HIV-1 integration assay used in this study. John *et al.* (2005) showed, under similar conditions to the modification of the high throughput microtitre plate HIV-1 integration assay used in this study, close to a 10 fold increase in activity. Chicoric acid is an inhibitor of ST that binds to the active site of IN (Reinke *et al.* 2002). Increasing concentrations of chicoric acid in the high throughput microtitre plate HIV-1 integration assay confirmed that the assay can be used for the rapid screening of potential HIV-1 IN inhibitors (see Figure 3.14). The IC_{50} of 101.5nM obtained here is similar to previously published data ranging from 100nM – 500nM (Charvat *et al.* 2006).

Because the high throughput microtitre plate HIV-1 integration assay is an indirect measure of IN activity, the purified recombinant IN_{sol} was then tested in a urea gel assay which allows for the visualization of ST products, and is thus a more direct indication of IN_{sol} enzymatic activity. Despite repeated attempts (see section 3.5 and supplementary data), it was not possible to successfully set up the assay. The assay conditions were similar to published assays that have shown good activity with the same purified IN_{sol} protein used in this study (Jenkins *et al.* 1996; Bar-Magen *et al.* 2009). The current assay conditions included long Donor substrates (Figure 3.15, 32 mer) and short Donor substrates (Figure 3.16, 21 mer). Further parameters that were changed in an attempt to optimize the assay included using a range of buffers, and different concentrations of substrates and enzyme as indicated in the supplementary section (Figures S1-S5 and SupTable 5.1). Of interest, most of the analyzed radiographs contained strong signals in the wells of each gel. This obstruction may be attributed to oligomers alone, or a complex of oligomers and IN that could not enter the gel.

Most published methods use oligomers that are PAGE gel purified (Bar-Magen *et al.* 2009), however, this was not carried out in this study. Future attempts should include the use of PAGE gel purified oligomers. Unincorporated ^{32}P-ATP is usually removed from

the substrates before use (Jenkins *et al.* 1996; Bar-Magen *et al.* 2009). This was carried out, but the removal of unincorporated ^{32}P-ATP greatly reduced the radioactive signal. After removal of unincorporated ^{32}P-ATPs, intensifying screens 200 (Carl Roth) were used to increase the sensitivity (Figures S2, S3 and S5). Intensifying screens have also been mentioned to increase sensitivity of detection during exposure of the X-ray films (Walker *et al.* 2008). However, the radioactive signal only increased marginally. Most published papers use a phosphorimager for detection of ST products (Jenkins *et al.* 1996; Bar-Magen *et al.* 2009). The phosphorimager is more sensitive due to having a more linear detection and a higher cut-off compared to X-ray film and may also be used for quantification (http://www.downstate.edu/biochemistry/phosphor_lab.html). In addition, pre-flashing the X-ray film with a short burst of low energy light also increases the linear range of the X-ray film and may increase the sensitivity. Future recommendations would include the use of pre-flashed X-ray film or a phosphorimager to enhance the sensitivity of detection in the assay.

A further modification to the microplate assay is the SPA bead assay which was set up using the purified recombinant IN_{sol} according to specifications from the literature (Grobler *et al.* 2009). This assay measures the direct binding of Target DNA to the immobilized donor DNA. The rationale behind this assay utilizes the scintillation properties of the SPA beads and activation of this scintillation when they are in close contact with a radioactive source. The radiolabel on the Target DNA substrate when it is in solution would not induce scintillation as strongly compared to if it were attached to the beads. Therefore, kinetic binding properties of the Target DNA to Donor DNA could be further elucidated using this assay. Results show levels of IN_{sol} activity and inhibition comparable to those showed by Grobler *et al.* (2009) with the IC_{50} value of 248.5nM for chicoric acid inhibition within the published range of 100nM – 500nM mentioned above (Charvat *et al.* 2006). These results are preliminary, however, and require further verification before its usage can be made routine. Another modification could be done using magnetic beads as a platform to immobilize donor DNA (Albanese *et al.* 2008).

These modifications all aid in the efficiency of the assay, and show the versatility of the application.

Overall, this study describes the establishment of a high throughput microtitre plate HIV-1 integration assay that can be used to screen for novel IN inhibitors developed in South Africa. In addition, a genotyping assay to screen for the emergence of IN mutations as a result of IN inhibitor use was successfully set up, and can be implemented by the genotyping laboratory when IN inhibitors become available in the public sector. While the high throughput microtitre plate HIV-1 integration assay was successfully set up and used to screen for chicoric acid inhibitor activity, the SPA bead strand transfer assay, once validated, can replace the microtitre plate HIV-1 integration assay since it has higher throughput capabilities required to screen chemical libraries.

Future work should focus on the optimization of the Urea-PAGE strand transfer assay. This assay needs to be optimized in order to confirm that each batch of purified recombinant IN has ST activity prior to its inclusion in screening assays such as the high throughput microtitre plate HIV-1 integration assay or SPA bead strand transfer assay. Lastly, future efforts in isolation of an HIV-1 subtype C recombinant IN for use in these assays may allow for the screening and identification of a locally relevant IN inhibitor.

CHAPTER 5

References

5. References

Albanese, A., D. Arosio, M. Terreni and A. Cereseto (2008) HIV-1 Pre-Integration Complexes Selectively Target Decondensed Chromatin in the Nuclear Periphery. *PLoS ONE*. **3**, e2413.

Anderson, J. L., E. M. Campbell, X. Wu, N. Vandegraaff, A. Engelman and T. J. Hope (2006) Proteasome Inhibition Reveals that a Functional Preintegration Complex Intermediate Can Be Generated during Restriction by Diverse TRIM5 Proteins. *J VIROL*. **80**, 9754-9760.

Andrake, M., J. Ramcharan, G. Merkel, X. Zhao, T. Burke and A. Skalka (2009) Comparison of metal-dependent catalysis by HIV-1 and ASV integrase proteins using a new and rapid, moderate throughput assay for joining activity in solution. *AIDS RES THERAPY*. **6**, 14.

Bally, F., R. Martinez, S. Peters, P. Sudre and A. Telenti (2004) Polymorphism of HIV Type 1 Gag p7/p1 and p1/p6 Cleavage Sites: Clinical Significance and Implications for Resistance to Protease Inhibitors. *AIDS RES HUM RETROV.* **16**, 1209-1213.

Bar-Magen, T., R. Sloan, V. Faltenbacher, D. Donahue, B. Kuhl, M. Oliveira, H. Xu and M. Wainberg (2009) Comparative biochemical analysis of HIV-1 subtype B and C integrase enzymes. *RETROVIROLOGY*. **6**, 103.

Baranova, S., F. V. Tuzikov, O. D. Zakharova, N. A. Tuzikova, C. Calmels, S. Litvak, L. Tarrago-Litvak, V. Parissi and G. A. Nevinsky (2007) Small-angle X-ray characterization of the nucleoprotein complexes resulting from DNA-induced oligomerization of HIV-1 integrase. *NUCL ACIDS RES*. **35**, 975-987.

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

Brooun, A., D. D. Richman and R. S. Kornbluth (2001) HIV-1 Preintegration Complexes Preferentially Integrate into Longer Target DNA Molecules in Solution as Detected by a Sensitive, Polymerase Chain Reaction-based Integration Assay. *J BIOL CHEM.* **276**, 46946-46952.

Bukrinsky, M., N. Sharova, M. Dempsey, T. Stanwick, A. Bukrinskaya, S. Haggerty and M. Stevenson (1992) Active nuclear import of human immunodeficiency virus type 1 preintegration complexes. *P NATL ACAD SCI USA*. **89**, 6580 - 6584.

Bukrinsky, M., N. Sharova, T. McDonald, T. Pushkarskaya, W. Tarpley and M. Stevenson (1993) Association of integrase, matrix, and reverse transcriptase antigens of human immunodeficiency virus type 1 with viral nucleic acids following acute infection. *P NATL ACAD SCI USA*. **90**, 6125 - 6129.

Bushman, F. D. and R. Craigie (1991) Activities of human immunodeficiency virus (HIV) integration protein in vitro: specific cleavage and integration of HIV DNA. *P NATL ACAD SCI USA.* **88**, 1339-1343.

Bushman, F. D., A. Engelman, I. Palmer, P. Wingfield and R. Craigie (1993) Domains of the integrase protein of human immunodeficiency virus type 1 responsible for polynucleotidyl transfer and zinc binding. *P NATL ACAD SCI USA.* **90**, 3428-3432.

Cai, M., R. Zheng, M. Caffrey, R. Craigie, G. M. Clore and A. M. Gronenborn (1997) Solution structure of the N-terminal zinc binding domain of HIV-1 integrase. *NAT STRUCT MOL BIOL*. **4**, 567-577.

Carteau, S., C. Hoffmann and F. Bushman (1998) Chromosome Structure and Human Immunodeficiency Virus Type 1 cDNA Integration: Centromeric Alphoid Repeats Are a Disfavored Target. *J VIROL*. **72**, 4005-4014.

Ceccherini-Silberstein, F., I. Malet, L. Fabeni, V. Svicher, C. Gori, S. Dimonte and *e. al*. (2007) Specific mutations related to resistance to HIV-1 integrase inhibitors are associated with reverse transcriptase mutations in HAART-treated patients. *ANTIVIR THER*. **12**, S6.

Charvat, T. T., D. J. Lee, W. E. Robinson and A. R. Chamberlin (2006) Design, synthesis, and biological evaluation of chicoric acid analogs as inhibitors of HIV-1 integrase. *BIOORGAN MED CHEM*. **14**, 4552-4567.

Chen, J. C.-H., J. Krucinski, L. J. W. Miercke, J. S. Finer-Moore, A. H. Tang, A. D. Leavitt and R. M. Stroud (2000) Crystal structure of the HIV-1 integrase catalytic core and C-terminal domains: A model for viral DNA binding. *P NATL ACAD SCI USA.* **97**, 8233-8238.

Cherepanov, P., D. Surratt, J. Toelen, W. Pluymers, J. Griffith, E. De Clercq and Z. Debyser (1999) Activity of recombinant HIV-1 integrase on mini-HIV DNA. *NUCL ACIDS RES*. **27**, 2202-2210.

Cherepanov, P., A. L. B. Ambrosio, S. Rahman, T. Ellenberger and A. Engelman (2005) Structural basis for the recognition between HIV-1 integrase and transcriptional coactivator p75. *P NATL ACAD SCI USA.* **102**, 17308-17313.

Chow, S. A. (1997) In Vitro Assays for Activities of Retroviral Integrase. *METHODS*. **12**, 306-317.

Ciuffi, A., M. Llano, E. Poeschla, C. Hoffmann, J. Leipzig, P. Shinn, J. R. Ecker and F. Bushman (2005) A role for LEDGF/p75 in targeting HIV DNA integration. *NAT MED*. **11**, 1287-1289.

Ciuffi, A. and F. D. Bushman (2006) Retroviral DNA integration: HIV and the role of LEDGF/p75. *TRENDS GENET.* **22**, 388-395.

Connell, B. J., K. Michler, A. Capovilla, W. D. Venter, W. S. Stevens and M. A. Papathanasopoulos (2008) Emergence of X4 usage among HIV-1 subtype C: evidence for an evolving epidemic in South Africa. *AIDS*. **22**, 896-899

Dar, M., B. Monel, L. Krishnan, M.-C. Shun, F. Di Nunzio, D. Helland and A. Engelman (2009) Biochemical and virological analysis of the 18-residue C-terminal tail of HIV-1 integrase. *RETROVIROLOGY*. **6**, 94.

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

Devroe, E., A. Engelman and P. A. Silver (2003) Intracellular transport of human immunodeficiency virus type 1 integrase. *J CELL SCI*. **116**, 4401-4408.

Dvorin, J. D., P. Bell, G. G. Maul, M. Yamashita, M. Emerman and M. H. Malim (2002) Reassessment of the Roles of Integrase and the Central DNA Flap in Human Immunodeficiency Virus Type 1 Nuclear Import. *J VIROL*. **76**, 12087-12096.

Ellison, V. and P. O. Brown (1994) A stable complex between integrase and viral DNA ends mediates human immunodeficiency virus integration in vitro. *P NATL ACAD SCI USA.* **91**, 7316-7320.

Engelman, A. and R. Craigie (1995) Efficient magnesium-dependent human immunodeficiency virus type 1 integrase activity. *J VIROL*. **69**, 5908-5911.

Engelman, A. and P. Cherepanov (2008) The Lentiviral Integrase Binding Protein LEDGF/p75 and HIV-1 Replication. *PLoS Pathog*. **4**, e1000046.

Eshleman, S. H., S. E. Hudelson, P. Smith, J. Hackett, V. Holzmayer, P. Swanson, S. G. Devare and N. Marlowe (2009) Analysis of pol integrase sequences in diverse HIV type 1 strains using a prototype genotyping assay. *AIDS RES HUM RETROV.* **25**, 343-5.

Farnet, C. M. and W. A. Haseltine (1990) Integration of human immunodeficiency virus type 1 DNA in vitro. *P NATL ACAD SCI USA.* **87**, 4164-4168.

Farnet, C. M., B. Wang, J. R. Lipford and F. D. Bushman (1996) Differential inhibition of HIV-1 preintegration complexes and purified integrase protein by small molecules. *P NATL ACAD SCI USA.* **93**, 9742-9747.

Faust, E., A. Garg, L. Small, A. Acel, R. Wald and B. Udashkin (1996) Enzymatic capability of HIS-tagged HIV-1 integrase using oligonucleotide disintegration substrates. *J BIOMED SCI.* **3**, 254-265.

Gallay, P., T. Hope, D. Chin and D. Trono (1997) HIV-1 infection of nondividing cells through the recognition of integrase by the importin/karyopherin pathway. *P NATL ACAD SCI USA*. **94**, 9825-9830.

Gallo, R., S. Salahuddin, M. Popovic, G. Shearer, M. Kaplan, B. Haynes, T. Palker, R. Redfield, J. Oleske, B. Safai and a. et (1984) Frequent detection and isolation of cytopathic retroviruses (HTLV-III) from patients with AIDS and at risk for AIDS. *SCIENCE*. **224**, 500-503.

Goldgur, Y., F. Dyda, A. B. Hickman, T. M. Jenkins, R. Craigie and D. R. Davies (1998) Three new structures of the core domain of HIV-1 integrase: An active site that binds magnesium. *P NATL ACAD SCI USA.* **95**, 9150-9154.

Good, N., D. Winget, W. Winter, T. Connolly, S. Izawa and R. Singh (1966) Hydrogen Ion Buffers for Biological Research. *BIOCHEMISTRY*. **5**, 467-477.

Goto, T., M. Nakai and K. Ikuta (1997) The life-cycle of human immunodeficiency virus type 1. *MICRON*. **29**, 123-138.

Grobler, J. A., K. A. Stillmock and D. J. Hazuda (2009) Scintillation proximity assays for mechanistic and pharmacological analyses of HIV-1 integration. *METHODS*. **47**, 249-253.

Gulizia, J., M. P. Dempsey, N. Sharova, M. I. Bukrinsky, L. Spitz, D. Goldfarb and M. Stevenson (1994) Reduced nuclear import of human immunodeficiency virus type 1 preintegration complexes in the presence of a prototypic nuclear targeting signal. *J VIROL*. **68**, 2021-2025.

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

Haseltine, W. (1991) Molecular biology of the human immunodeficiency virus type 1. *FASEB J*. **5**, 2349-2360.

Hazuda, D., J. Hastings, A. Wolfe and E. Emini (1994) A novel assay for the DNA strand-transfer reaction of HIV-1 integrase. *NUCL ACIDS RES*. **22**, 1121 - 1122.

Hobaika, Z., L. Zargarian, R. Maroun, O. Mauffret, T. Burke and S. Fermandjian (2009) HIV-1 Integrase and Virus and Cell DNAs: Complex Formation and Perturbation by Inhibitors of Integration. *NEUROCHEM RES.* **35**, 888-893.

Hunt, G., M. Papathanasopoulos, G. Gray and C. Tiemessen (2003) Characterisation of Near-full Length Genome Sequences of Three South African Human Immunodeficiency Virus Type 1 Subtype C Isolates. *VIRUS GENES*. **26**, 49-56.

Iordanskiy, S., R. Berro, M. Altieri, F. Kashanchi and M. Bukrinsky (2006) Intracytoplasmic maturation of the human immunodeficiency virus type 1 reverse transcription complexes determines their capacity to integrate into chromatin. *RETROVIROLOGY*. **3**, 4.

Jenkins, T., A. Hickman, F. Dyda, R. Ghirlando, D. Davies and R. Craigie (1995) Catalytic domain of human immunodeficiency virus type 1 integrase: Identification of a soluble mutant by systematic replacement of hydrophobic residues. *P NATL ACAD SCI USA.* **92**, 6057 - 6061.

Jenkins, T., A. Engelman, R. Ghirlando and R. Craigie (1996) A soluble active mutant of HIV-1 integrase: involvement of both the core and the C-terminal domains in multimerization. *J BIOL CHEM*. **271**, 7712 - 7718.

John, S., T. M. Fletcher, III and C. B. Jonsson (2005) Development and Application of a High-Throughput Screening Assay for HIV-1 Integrase Enzyme Activities. *J BIOMOL SCREEN*. **10**, 606-614.

Kalpana, G., S. Marmon, W. Wang, G. Crabtree and S. Goff (1994) Binding and stimulation of HIV-1 integrase by a human homolog of yeast transcription factor SNF5. *SCIENCE*. **266**, 2002-2006.

Kolegraff, K., P. Bostik and A. A. Ansari (2006) Characterization and Role of Lentivirus-Associated Host Proteins. *EXP BIOL MED*. **231**, 252-263.

Marchand, C., A. A. Johnson, E. Semenova and Y. Pommier (2006) Mechanisms and inhibition of HIV integration. *DRUG DISCOV TODAY*. **3**, 253-260.

Marshall, H. M., K. Ronen, C. Berry, M. Llano, H. Sutherland, D. Saenz, W. Bickmore, E. Poeschla and F. D. Bushman (2007) Role of PSIP1/LEDGF/p75 in Lentiviral Infectivity and Integration Targeting. *PLoS ONE*. **2**, e1340.

Mazumder, A., A. Engelman, R. Craigie, M. Fesen and Y. Pommier (1994) Intermolecular disintegration and intramolecular strand transfer activities of wild-type and mutant HIV-1 integrase. *NUCL ACIDS RES*. **22**, 1037-1043.

McColl, D. J. and X. Chen (2009) Strand transfer inhibitors of HIV-1 integrase: Bringing IN a new era of antiretroviral therapy. *ANTIVIR RES.* **85**, 101-118.

Michel, F., C. Crucifix, F. Granger, S. Eiler, J.-F. Mouscadet, S. Korolev, J. Agapkina, R. Ziganshin, M. Gottikh, A. Nazabal, S. Emiliani, R. Benarous, D. Moras, P. Schultz and M. Ruff (2009) Structural basis for HIV-1 DNA integration in the human genome, role of the LEDGF/P75 cofactor. *EMBO J*. **28**, 980-991.

Miller, R. J., J. S. Cairns, S. Bridges and N. Sarver (2000) Human Immunodeficiency Virus and AIDS: Insights from Animal Lentiviruses. *J VIROL*. **74**, 7187-7195.

Mizuuchi, K. (1992) Polynucleotidyl transfer reactions in transpositional DNA recombination. *J BIOL CHEM.* **267**, 21273-21276.

Nilsen, B., I. Haugan, K. Berg, L. Olsen, P. Brown and D. Helland (1996) Monoclonal antibodies against human immunodeficiency virus type 1 integrase: epitope mapping and differential effects on integrase activities in vitro. *J VIROL*. **70**, 1580-1587.

Passaes, C. B., M. L. Guimaraes, S. L. Fernandez, S. Lorete Rdos, S. L. Teixeira, J. C. Fernandez and M. G. Morgado (2009) Lack of primary mutations associated with integrase inhibitors among HIV-1 subtypes B, C, and F circulating in Brazil. *J ACQUIR IMMUNE DEFIC SYNDR*. **51**, 7-12.

Perelson, A. S., A. U. Neumann, M. Markowitz, J. M. Leonard and D. D. Ho (1996) HIV-1 Dynamics in Vivo: Virion Clearance Rate, Infected Cell Life-Span, and Viral Generation Time. *SCIENCE*. **271**, 1582-1586.

Petit, C., O. Schwartz and F. Mammano (2000) The Karyophilic Properties of Human Immunodeficiency Virus Type 1 Integrase Are Not Required for Nuclear Import of Proviral DNA. *J VIROL*. **74**, 7119-7126.

Reinke, R. A., P. J. King, J. G. Victoria, B. R. McDougall, G. Ma, Y. Mao, M. G. Reinecke and W. E. Robinson (2002) Dicaffeoyltartaric Acid Analogues Inhibit Human Immunodeficiency Virus Type 1 (HIV-1) Integrase and HIV-1 Replication at Nontoxic Concentrations. *J MED CHEM* **45**, 3669-3683.

Rhee, S., T. Liu, M. Kiuchi, R. Zioni, R. Gifford, S. Holmes and R. Shafer (2008) Natural variation of HIV-1 group M integrase: implications for a new class of antiretroviral inhibitors. *RETROVIROLOGY*. **5**, 74.

Rice, P. A. and T. A. Baker (2001) Comparative architecture of transposase and integrase complexes. *NAT STRUCT MOL BIOL*. **8**, 302-307.

Sato, M., H. Kawakami, T. Motomura, H. Aramaki, T. Matsuda, M. Yamashita, Y. Ito, Y. Matsuzaki, K. Yamataka, S. Ikeda and H. Shinkai (2009) Quinolone Carboxylic Acids as a Novel Monoketo Acid Class of Human Immunodeficiency Virus Type 1 Integrase Inhibitors. *J MED CHEM.* **52**, 4869-4882.

Schröder, A. R. W., P. Shinn, H. Chen, C. Berry, J. R. Ecker and F. Bushman (2002) HIV-1 Integration in the Human Genome Favors Active Genes and Local Hotspots. *CELL*. **110**, 521-529.

Sherman, P. and J. Fyfe (1990) Human immunodeficiency virus integration protein expressed in Escherichia coli possesses selective DNA cleaving activity. *P NATL ACAD SCI USA.* **87**, 5119 - 5123.

Sherman, M. P., C. M. C. de Noronha, M. I. Heusch, S. Greene and W. C. Greene (2001) Nucleocytoplasmic Shuttling by Human Immunodeficiency Virus Type 1 Vpr. *J VIROL*. **75**, 1522-1532.

Sherman, M. P. and W. C. Greene (2002) Slipping through the door: HIV entry into the nucleus. *MICROBES INFECT*. **4**, 67-73.

Sinha, S. and D. P. Grandgenett (2005) Recombinant Human Immunodeficiency Virus Type 1 Integrase Exhibits a Capacity for Full-Site Integration In Vitro That Is Comparable to That of Purified Preintegration Complexes from Virus-Infected Cells. *J VIROL*. **79**, 8208-8216.

Tzitzivacos, D. B., C. T. Tiemessen, W. S. Stevens and M. A. Papathanasopoulos (2009) Viral Genetic Determinants of Nonprogressive HIV Type 1 Subtype C Infection in Antiretroviral Drug-Naive Children. *AIDS RES HUM RETROV.* **25**, 1141-1148.

Van Baelen, K., E. Rondelez, V. Van Eygen, K. Ariën, M. Clynhens, P. Van den Zegel, B. Winters and L. J. Stuyver (2009) A combined genotypic and phenotypic human immunodeficiency virus type 1 recombinant virus assay for the reverse transcriptase and integrase genes. *J VIROL METHODS.* **161**, 231-239.

Van Laethem, K., Y. Schrooten, K. Covens, N. Dekeersmaeker, P. De Munter, E. Van Wijngaerden, M. Van Ranst and A. M. Vandamme (2008) A genotypic assay for the amplification and sequencing of integrase from diverse HIV-1 group M subtypes. *J VIROL METHODS*. **153**, 176-81.

Van Maele, B., K. Busschots, L. Vandekerckhove, F. Christ and Z. Debyser (2006) Cellular co-factors of HIV-1 integration. *TRENDS BIOCHEM SCI*. **31**, 98-105.

Vink, C., M. Banks, R. Bethell and R. H. A. PIasterk (1994) A high-throughput, non-radioactive microtitre plate assay for activity of the human immunodeficiency virus integrase protein. *NUCL ACIDS RES*. **22**, 2176-2177.

Walker, J. M., R. Rapley and B. M. Harvey (2008) Autoradiography and Fluorography. *Molecular Biomethods Handbook*, pp 396-410, Humana Press

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

Wang, Y., H. Klock, H. Yin, K. Wolff, K. Bieza, K. Niswonger, J. Matzen, D. Gunderson, J. Hale, S. Lesley, K. Kuhen, J. Caldwell and A. Brinker (2005) Homogeneous High-Throughput Screening Assays for HIV-1 Integrase 3'-Processing and Strand Transfer Activities. *J BIOMOL SCREEN.* **10**, 456-462.

Yung, E., M. Sorin, A. Pal, E. Craig, A. Morozov, O. Delattre, J. Kappes, D. Ott and G. V. Kalpana (2001) Inhibition of HIV-1 virion production by a transdominant mutant of integrase interactor 1. *NAT MED*. **7**, 920-926.

Zhang, J.-q., J.-j. Wang, W.-j. Li, L. Huang, L. Tian, J.-l. Xue, J.-z. Chen and W. Jia (2009) Cellular protein TTRAP interacts with HIV-1 integrase to facilitate viral integration. *BIOCHEM BIOPH RES CO.* **387**, 256-260.

Zheng, R., T. M. Jenkins and R. Craigie (1996) Zinc folds the N-terminal domain of HIV-1 integrase, promotes multimerization, and enhances catalytic activity. *P NATL ACAD SCI USA.* **93**, 13659-13664.

CHAPTER 6

Supplementary

6. Supplementary

Table S6.1: Buffer recipes used in Urea-PAGE based assays

Name	Recipe
Assay Buffer 1	20mM HEPES pH 7.3, 75mM NaCl, 10 mM $MgCl_2$, 2 µM $ZnCl_2$, 1 % glycerol
Assay Buffer 2	25mM HEPES pH 7.5, 10mM $MgCl_2$, 5mM DTT, 10% PEG 8000, 10% DMSO, 20µM $ZnCl_2$, 75mM NaCl
Assay Buffer 3	20 mM HEPES pH 7.5, 30 mM NaCl, 1mM DTT, 125 µM $ZnCl_2$, 7.5 mM $MnCl_2$
Assay Buffer 4	50 mM MOPS pH 7.2, 10 mM $MgCl_2$, 14 mM B-ME

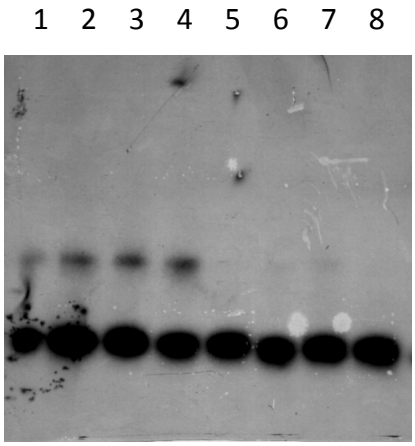


Figure S1: Autoradiograph of Urea PAGE based strand transfer activity assay using Buffer 3. Lanes 1-4 are no enzyme controls. Lanes 5-8 are enzymatic tests. All wells contained Target where only the target (T1 and T2) were radiolabeled. Order of donor additions are 1: DU/B1, 2: DP/B1, 3:DU/B2, 4:DP/B2 (same order for 5-8). Buffer 3 was used for the reactions. Donor (21 mer) concentration was 200 nM target concentration was 400 nM and enzyme concentration was 1 µM. Only target was radiolabelled

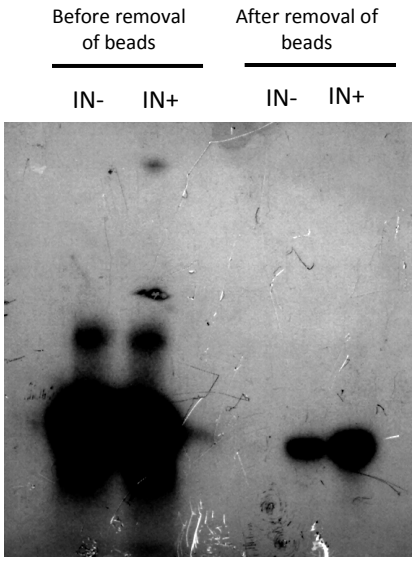


Figure S2: Autoradiograph of Gel based strand transfer activity assay depicting use of SPA Beads as substrate immobilizer. Lanes are indicated. All wells contained Target where only the target (T1 and T2) were radiolabeled. Donor (32 mer) concentration was 360 nM target concentration was 2 µM and enzyme concentration was 0.5 µM. Only target was radiolabelled.

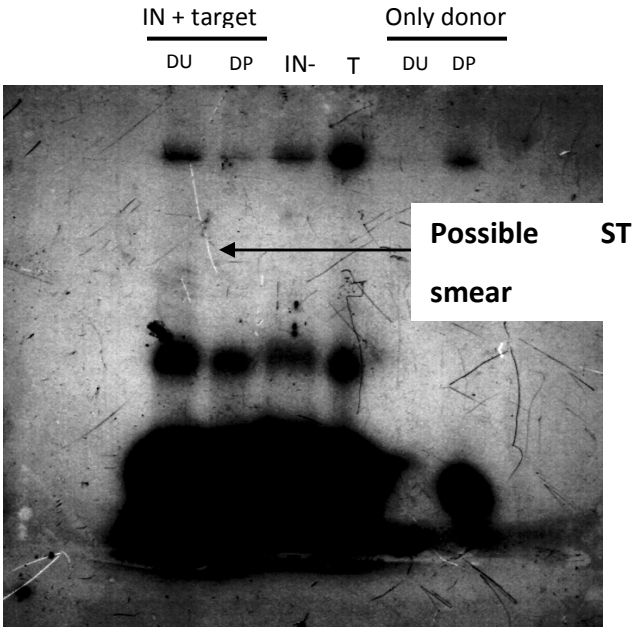


Figure S3: Autoradiograph of Gel based strand transfer activity assay utilizing buffer 4. Lanes are indicated. Donor concentration was 125 nM target concentration was 750 µM and enzyme concentration was 1 µM. Target and donor was radiolabelled. DU denotes unprocessed donor, DP denotes processed donor and T denotes target.

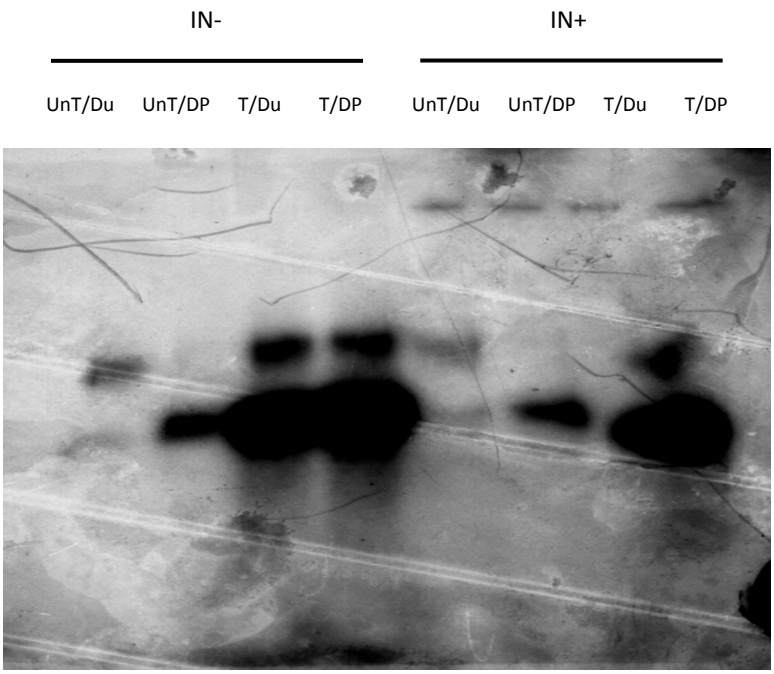


Figure S4: Results depicting use of buffer 1. Donor concentration was 60 nM and target concentration was 100 nM and enzyme concentration was 1 µM. UnT denotes non radiolabeled Target DNA. T denotes labelled target.

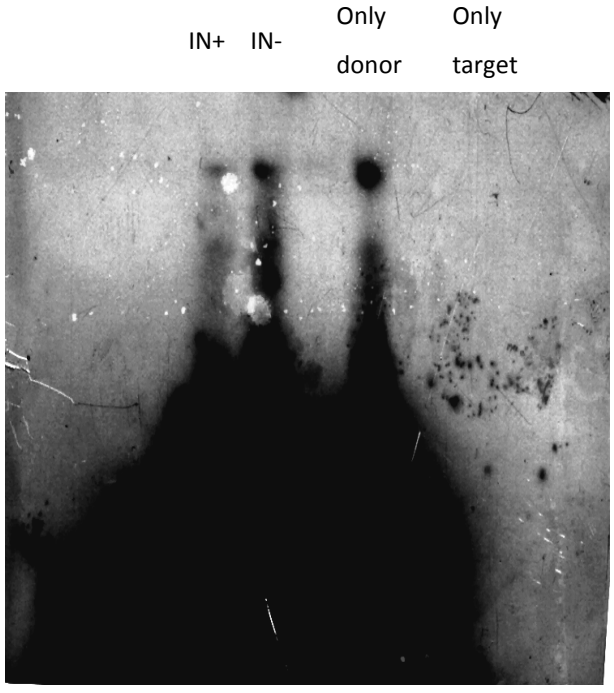


Figure S5: Results depicting use of buffer 3. Donor DNA was 500 nM and Target DNA was 2 µM and enzyme was 2 µM.

University
of the Witwatersrand,
Johannesburg

Human Research Ethics Committee (Medical)
(formerly Committee for Research on Human Subjects (Medical)

Secretariat: Research Office, Room SH10005, 10th floor, Senate House • Telephone: +27 11 717-1234 • Fax: +27 11 339-5708
Private Bag 3, Wits 2050, South Africa

Ref: **W-CJ-080827/2**

27/08/2008

TO WHOM IT MAY CONCERN:

Waiver: This certifies that the following research does not require clearance from the Human Research Ethics Committee (Medical).

Investigator: **Mr M Q Fish Student No 0413586R**

Project title: **Development of an Integrase enzyme strand transfer assay using HIV-1 subtype C preintegration complexes.**

Reason: This research is wholly laboratory using cell lines from a repository; there are no human participants.

UNIVERSITY OF THE WITWATERSRAND
PROF PE CLEATON-JONES
HREC (MEDICAL)
2008 -08- 27
JOHANNESBURG

Professor Peter Cleaton-Jones
Chair: Human Research Ethics Committee (Medical)

copy: Anisa Keshav, Research Office, Senate House, Wits

AIDS RESEARCH AND HUMAN RETROVIRUSES
Volume 26, Number 4, 2010
© Mary Ann Liebert, Inc.
DOI: 10.1089/aid.2009.0249

Natural Polymorphisms of *integrase* Among HIV Type 1-Infected South African Patients

Muhammad Q. Fish,[1] Raymond Hewer,[2] Carole L. Wallis,[1] Willem D.F. Venter,[3] Wendy S. Stevens,[1] and Maria A. Papathanasopoulos[1]

Abstract

An HIV-1 subtype C specific assay was established for *integrase* genotyping from 51 integrase inhibitor-naive patient plasma samples and 22 antiretroviral drug-naive primary viral isolates from South Africa. Seventy-one of the 73 samples were classified as HIV-1 subtype C and two samples were unique AC and CG recombinants in *integrase*. Amino acid sequence analysis revealed there were no primary mutations (Y143R/C/H, Q148H/R/K, and N155H/S) associated with reduced susceptibility to the integrase inhibitors raltegravir and elvitegravir. However, one sample had the T97A mutation, three samples had the E157Q and V165I mutations, and the majority of samples contained the polymorphic mutation V72I. The expected finding of no major integrase mutations conferring resistance to integrase inhibitors suggests that this new antiretroviral drug class will be effective in our region where HIV-1 subtype C predominates. However, the impact of E157Q and other naturally occurring polymorphisms warrants further phenotypic investigation.

The HIV-1 integrase enzyme (32 kDa, 288 amino acids) is responsible for the irreversible integration of the newly synthesized viral cDNA into the host cell chromosomal DNA, an essential step in the HIV-1 replication cycle.[1] Integrase is composed of three functional domains: the N-terminal domain (amino acids 1–50), which coordinates zinc binding, the catalytic core domain (amino acids 51–212), which contains three catalytic residues forming a DDE motif, and the C-terminal domain (amino acids 213–288), which is involved in host DNA binding and integrase oligomerization.[2]

Integrase functions in cells within the context of high-molecular-weight preintegration complexes that are formed in the cytoplasm after reverse transcription of the viral RNA. HIV-1 integration is subdivided into two steps with distinct kinetics and requirements. In the cytoplasm, a dinucleotide adjacent to a phylogenetically conserved CA is hydrolytically cleaved off the 3′ end of each strand of the linear viral cDNA by integrase in the preintegration complex (3′-processing).[3–7] After nuclear transport of the preintegration complex, two phosphodiester bonds on each strand of the chromosomal DNA are targeted by the 3′-hydroxyls at the ends of the viral DNA, resulting in its insertion within the host genome (strand transfer).[3–7] Removal of the two unpaired nucleotides on the 5′ viral DNA ends and gap filling of the interfaces between the viral and host genomic DNA are then done using host DNA repair machinery via mechanisms that are not yet fully understood. Completion of integration results in a fully functional provirus, which can then be used to initiate viral DNA transcription. There is no known human homologue to integrase, thus making it an ideal antiretroviral drug target.

Raltegravir (Isentress, Merck & Co., Inc., Whitehouse Station, NJ), the first integrase inhibitor to be approved by the U.S. Food and Drug Administration (October 2007), has successfully been used in combination with other antiretroviral drugs for treatment-experienced HIV-1 patients with multidrug-resistant strains. Raltegravir blocks viral replication by inhibiting DNA strand transfer.[8] However, drug resistance to integrase inhibitors has been shown to occur both *in vivo* and *in vitro*, and is the consequence of mutations that are selected in the viral *integrase* gene targeted by the strand transfer antiretroviral drugs such as raltegravir and elvitegravir (Gilead Sciences, Foster City, CA). To date, there are at least 30 substitutions that have been specifically associated with the development of resistance to raltegravir and/or elvitegravir (http://hivdb.stanford.edu). However, the main mutational pathways associated with raltegravir and elvitegravir resistance include signature mutations at integrase positions 143, 148, and 155.

Access to antiretroviral drugs has escalated dramatically in the sub-Saharan African region in the past decade. Regimens

[1]HIV Pathogenesis Research Laboratory, Department of Molecular Medicine and Haematology, University of the Witwatersrand Medical School, Parktown, Johannesburg, South Africa.
[2]AuTEK Biomed, Mintek, Private Bag X3015, Randburg, Johannesburg, South Africa.
[3]Reproductive Health and HIV Research Unit, University of the Witwatersrand Medical School, Parktown, Johannesburg, South Africa.

in this region almost universally place a thymidine analogue with lamivudine and a nonnucleoside reverse transcriptase inhibitor as first line therapy, with a protease inhibitor-based regimen for those failing first line therapies. However, these regimens carry significant toxicities, requiring drug substitutions. In addition, an increasing number of patients require alternative regimens as they fail first and second line regimens due to nonadherence, toxicity, or the development of drug resistance. Thus, the effectiveness of integrase inhibitors such as raltegravir needs to be investigated as alternative antiretroviral drug treatment options are required. Several groups have published genotypic assays for the amplification and sequencing of *integrase* from diverse HIV-1 group M subtypes,[9–11] however, we opted to establish a subtype C-specific *integrase* genotyping assay. This study describes the baseline mutations in *integrase* in HIV-1 subtype C (and unique recombinant) infected integrase inhibitor-naive individuals in South Africa.

A total of 80 samples from patients attending the Charlotte Maxeke Johannesburg Hospital were available for purposes of this study. These included 58 randomly selected patient plasma samples sent for routine viral load monitoring during 2009, and 22 well-characterized primary HIV-1 isolates from antiretroviral drug-naive AIDS patients isolated during 2005.[12] The patients viral loads ranged from 1500 to >3 million RNA copies/ml, and included antiretroviral drug-naive and treatment-experienced patients. Ethical clearance for the study was obtained for Research on Human Subjects (Medical) at the University of the Witwatersrand (clearance number M090575). Viral RNA was extracted from the 58 patient plasma samples using the automated Roche MagNa Pure LC analyzer and the MagNa Pure LC Total Nucleic Acid Isolation kit (Roche, Germany), according to the manufacturer's instructions. Proviral DNA from the 22 viral isolates was extracted from p24 antigen-positive cultures with a High Pure PCR Template Preparation kit (Roche, GmbH, Germany) according to the manufacturer's instructions. The full-length *integrase* gene was reverse transcriptase polymerase chain reaction (RT-PCR) amplified into cDNA from the extracted viral RNA samples using the reverse primer SQ9RC(3) (5′-TTGGATGTCTGCTTTCATAATGAT-3′)[13] and the Expand Reverse Transcriptase (RT) kit (Roche Diagnostics, Germany), as per the manufacturer's instructions. Amplification from synthesized cDNA or extracted proviral DNA was performed using primers SQ7F(2) (5′-AGGAGCAGAAACTTTCT ATGTAGATGG-3′)[13] and SQ9RC(3) with the Expand High Fidelity^{PLUS}kit (Roche, Germany), as per the manufacturer's instructions. PCR amplicons (approximately 1.3 kb in length) were purified using the High Pure PCR Product Purification kit (Roche), as per the manufacturer's instructions.

Primers spanning the full-length *integrase* (~880 bp) were used to sequence the PCR products in both directions (population-based sequencing). These included QFIN1F (5′-ATGGGTACCAGCACACAAAG-3′), QFIN2F (5′-TAGC AGGAAGATGGCCAGTC-3′), QFIN1R (5′-CTCTCCTTG AAATATACATATGGTGC-3′), and QFIN2R (5′-CTACTA CTCCTTGACTTTGGGG-3′). Purified amplicons were sequenced with the ABI Big Dye terminator system (v3.1) according to the manufacturer's instructions on the ABI Prism 3730 and 3100-*Avant* Genetic Analyzer (Applied Biosystems, Foster City, CA). Sequence data were edited using the Sequencing Analysis V3.3 program (Applied Biosystems), and the complete *integrase* sequences were assembled and manually edited using Sequencher V4.7 (Genecodes, Ann Arbor, MI). For subtyping, a multiple alignment of the full length *integrase* region with references from HIV-1 subtypes A to K, CRF01_AE, and CRF02_AG (http://hiv-web.lanl.gov) was generated in Clustal X. Aligned sequences were converted to MEGA V3.0 format and used in phylogenetic and molecular evolutionary analyses. A phylogeny reconstruction of each *integrase* gene was performed by neighbor-joining using the Kimura two-parameter distance matrix. The stability of the nodes was assessed by bootstrap analysis (1000 replicates) and bootstrap values >70% were considered significant. Nucleotide sequences were converted to amino acid sequences and compared to the 2004 consensus C sequence available from the Los Alamos website (http://hiv-web.lanl.gov). The amino acid sequence alignment was extensively analyzed for the presence of primary mutations, nonpolymorphic and polymorphic mutations associated with resistance to known integrase inhibitors (http://hivdb.stanford.edu).

RT-PCR and PCR amplification was successful for 91.3% of all samples, and included 51 of the 58 patient plasma samples (not dependent on viral load) and all 22 proviral DNA samples, respectively. All 73 of the amplified samples were successfully sequenced. The *integrase* amplification and sequencing success rates are similar to those reported previously.[9–11] Nucleotide sequences of the 73 newly characterized full-length *integrase* genes ranged from 867 to 876 base pairs, with intact open reading frames confirming the presence of functional *integrase* genes in all patients and primary virus isolates. The *integrase* intrasubtype nucleotide diversity for all 73 newly characterized samples varied between 2.2% and 7.9%, similar to what has been reported previously.[14] Phylogenetic analysis of the full-length *integrase* nucleotide sequences, together with viruses from the major subtypes, revealed that 71 sequences clustered within HIV-1 subtype C with a bootstrap value of 100%, whereas two samples (09ZAP35 and 09ZAP52) showed evidence of recombination (Fig. 1). These two sequences were further analyzed using the RIP 3.0 Program (http://hiv-web.lanl.gov), which showed that samples 09ZAP35 and 09ZAP52 were unique AC and CG recombinants in *integrase* (results not shown). Since subtype A predominates in East, Central, and West Africa, and subtype G predominates in Central and West Africa, it is possible that these patients originated from these countries and a recombination event occurred in South Africa.

An alignment of the 73 integrase amino acid sequences was extensively analyzed and compared, and a distribution of the variants is shown in Fig. 2. The proteins varied from 288 to 292 amino acids in length, primarily due to the presence of an unusual tetrapeptide insertion (QNME) in the C terminus of the protein in 23 patient and primary virus isolate samples. This unusual insertion has been noted previously.[15,16] The C terminus has been shown to be important for integrase dimerization and DNA binding. This is crucial for proper function and positioning of the catalytic domain close to its target site. This finding warrants further investigation to determine the impact of the insertion on the function of the integrase enzyme.

The sequences were analyzed for the presence of integrase mutations associated with reduced *in vitro* susceptibility to raltegravir (E92Q, F121Y, E138AK, G140AS, Y143RCH, S147G, Q148HRK, N155HS) and elvitegravir (T66IAK, E92Q, F121Y, E138AK, G140AS, S147G, Q148HRK, S153Y, N155HS,

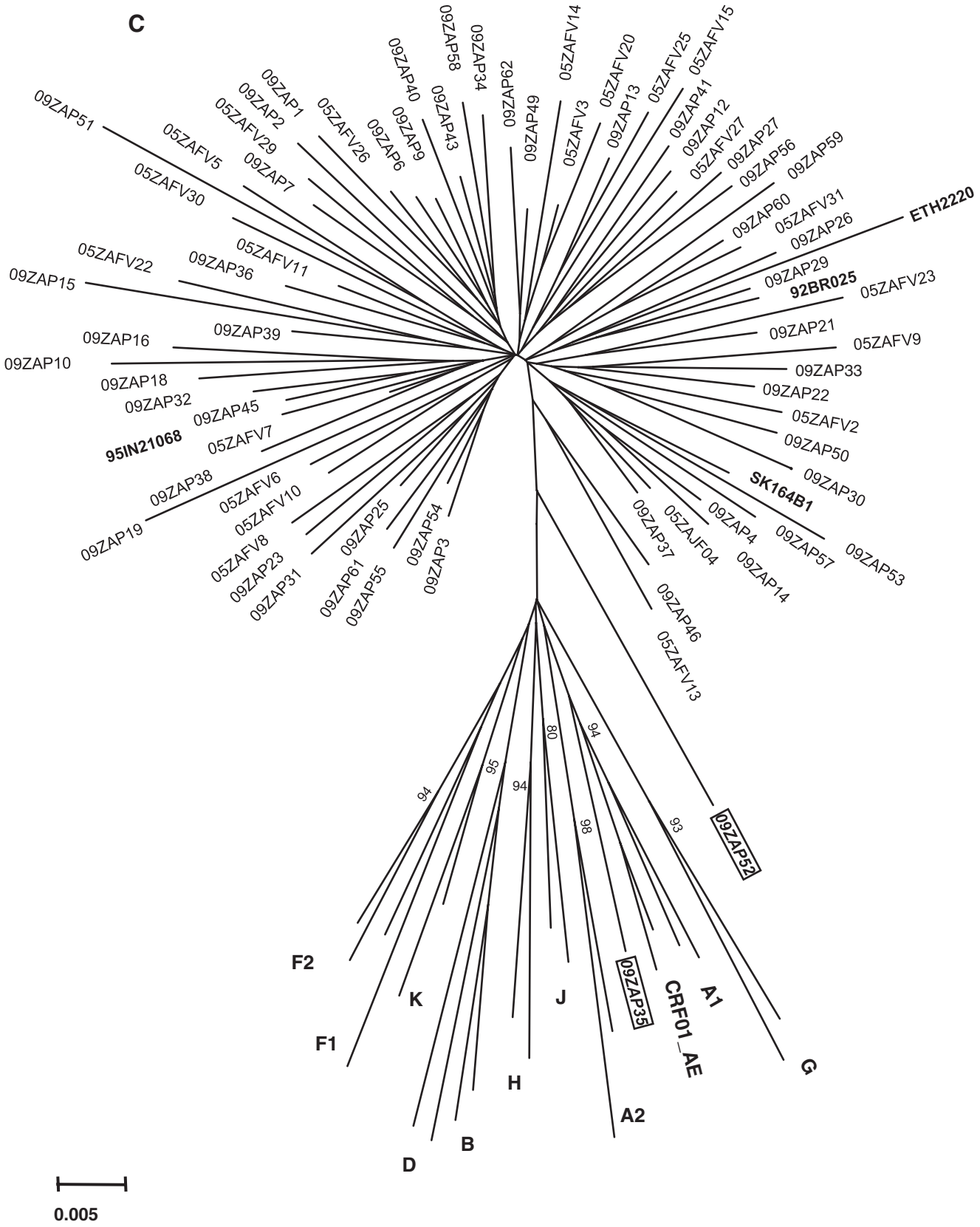


FIG. 1. Phylogenetic relationships of the 73 newly characterized full-length *integrase* sequences with HIV-1 subtype reference sequences from the Los Alamos database (http://hiv-web.lanl.gov). The unique recombinant sequences are boxed. Phylogenetic trees were constructed from nucleotide sequence alignments, using neighbor joining and the Kimura two-parameter distance matrix. The reliability of the branching order was estimated from 1000 bootstrap replicates, and only bootstrap values of 70% or higher are shown.

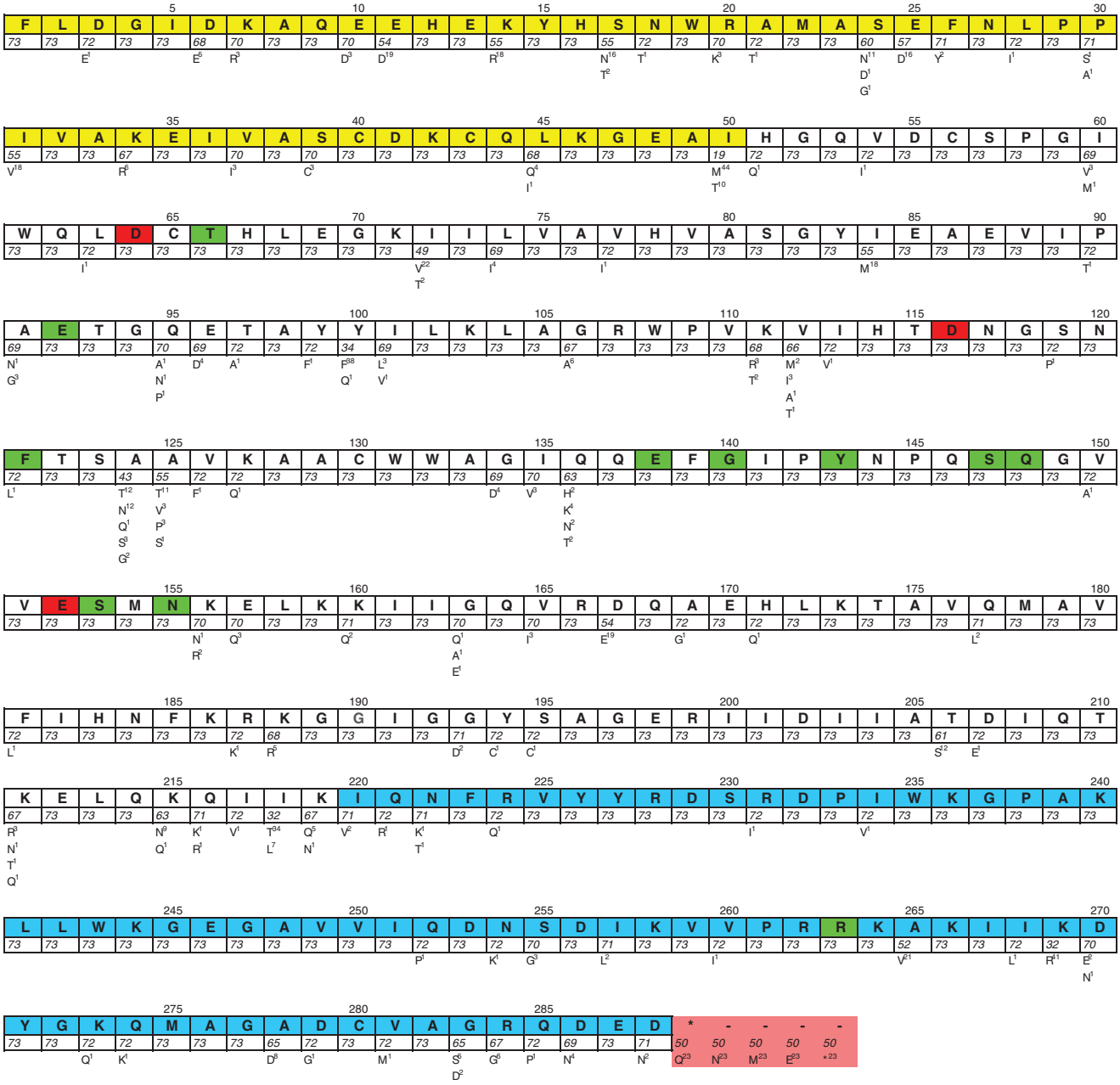


FIG. 2. Distribution of variants among the 73 newly characterized full-length integrase sequences. The 2002 consensus subtype C sequence (http://hiv-web.lanl.gov) is shown at the top of each 30 amino acid sequence section. Numbers below the consensus sequence indicate the number of annotated sequences (out of 73) containing an unambiguous amino acid at that position. Beneath that are the amino acid variants with superscript numbers indicating the amount of times they occur. Integrase functional domains and important residues are highlighted. The N terminal domain is highlighted in yellow, the catalytic core domain in white, and the C terminal domain in blue; the catalytic triad (DDE) residues are in red and relevant integrase inhibitor resistance mutations are in green.

R263K) (Stanford Database, updated 15 September, 2009). None of the samples contained the mutations associated with primary resistance (Y143RCH, Q148HRK, and N155HS).

The E157Q mutation was observed among 4.1% ($n = 3$) of patient samples, including two subtype C patient samples (09ZAP34 and 09ZAP53) and one recombinant (09ZAP52) patient sample (Fig. 2). Until recently the E157Q mutation was listed among the mutations associated with reduced *in vitro* susceptibility to raltegravir and elvitegravir (http://hivdb.stanford.edu), but is now described as a minimally polymorphic mutation that is selected *in vitro* by elvitegravir and rarely *in vivo* by raltegravir, having a minimal, if any effect on integrase inhibitor susceptibility. Interestingly, among several other studies characterizing a total of 501 subtype C integrase sequences,[9,11,14,17] the E157Q mutation was reported in only one sample (1.4%).[11]

The isoleucine at position 72 (I72) was identified at a high frequency among the subtype C samples and the 05ZAP52

recombinant (67.1%; $n = 49$), whereas V72 and T72 were found in 28.7% ($n = 21$, including the 05ZAP35 recombinant) and 1.4% ($n = 1$) of the patients, respectively, confirming their status as a naturally occurring subtype C polymorphism. This confirms the findings by Rhee *et al.*[14] who characterized naturally occurring variation in previously published HIV-1 group M integrase sequences, and found that 390 of 432 (90.3%) integrase-inhibitor-naive subtype C samples analyzed contained isoleucine at position 72, and only 9.4% and 0.2% contained a valine and threonine, respectively. The polymorphic mutation, T97A, associated with integrase inhibitors was observed in one primary virus sample (05ZAFV27), whereas K156N, another accessory integrase inhibitor resistance mutation, was found in patient 09ZAP34. The uncommon polymorphism K156N was not reported in subtype C, but occurs in 9% of subtype B and 5% of subtype D sequences.[14] By contrast, T97A was observed in subtypes A, B, C, D, and CRF01_AE and CRF02_AG, albeit at low frequencies.[14]

Because the antiretroviral drug history of the patient samples analyzed in 2009 was unknown, we could not address the influence of prior antiretroviral therapy on integrase polymorphisms. Ceccherini-Silberstein *et al.*[18] reported several integrase polymorphisms, for example, M154I, V165I, and M185L, that were positively associated with specific reverse transcriptase mutations (F227L and T215Y) in treated patients. In our study, the V165I polymorphism was noted in three samples, 09ZAP3, 09ZAP54 (unknown drug history), and 05ZAFV29 (antiretroviral drug naive). This highlights the need for ongoing surveillance of integrase mutations as well as HIV-1 subtype C-specific phenotypic studies to elucidate the relevance of the above-mentioned polymorphisms (as well as naturally occurring polymorphisms) in the absence of the signature integrase mutations on susceptibility/resistance to integrase inhibitors.

In conclusion, none of the three previously reported major mutations (Y143R/C/H, Q148H, and N155H) associated with resistance to integrase inhibitors was observed in our cohort, indicating that integrase inhibitors may be used as a treatment option in South Africa.

Sequence Data

The 73 *integrase* sequences were submitted to GenBank using Sequin V5.35 (http://www.ncbi.nlm.nih.gov), and are available under the accession numbers GQ872465 to GQ872538.

Acknowledgments

We are grateful for research funding from AuTEK Biomed (Mintek and Harmony Gold Mining Company) and the Department of Molecular Medicine and Haematology.

Author Disclosure Statement

No competing financial interests exist.

References

1. Engelman A, Mizuuchi K, and Craigie R: HIV-1 DNA integration: Mechanisms of viral DNA cleavage and DNA strand transfer. Cell 1991;67:1211–1221.
2. Van Maele B, Busschots K, Vandekerckhove L, Christ F, and Debyser Z: Cellular co-factors of HIV-1 integration. Trends Biochem Sci 2006;31:98–105.
3. LaFemina RL, Schneider CL, Robbins HL, *et al.*: Requirement of active human immunodeficiency virus type 1 integrase enzyme for productive infection of human T-lymphoid cells. J Virol 1992;66:7414–7419.
4. Sakai H, Kawamura M, Sakuragi J, *et al.*: Integration is essential for efficient gene expression of human immunodeficiency virus type 1. J Virol 1993;67:1169–1174.
5. Miller MD, Farnet CM, and Bushman FD: Human immunodeficiency virus type 1 preintegration complexes: Studies of organization and composition. J Virol 1997;71:5382–5390.
6. Van Maele B and Debyser Z: HIV-1 integration: An interplay between HIV-1 integrase, cellular and viral proteins. AIDS Rev 2005;7:26–43.
7. Yoder KE and Bushman FD: Repair of gaps in retroviral DNA integration intermediates. J Virol 2000;74:11191–11200.
8. Pommier Y, Johnson AA, and Marchand C: Integrase inhibitors to treat HIV/AIDS. Nat Rev Drug Discov 2005;4:236–248.
9. Van Laethem K, Schrooten Y, Covens K, *et al.*: A genotypic assay for the amplification and sequencing of integrase from diverse HIV-1 group M subtypes. J Virol Methods 2008;153:176–181.
10. Van Baelen K, Rondelez E, van Eygen V, *et al.*: A combined genotypic and phenotypic human immunodeficiency virus type 1 recombinant virus assay for the reverse transcriptase and integrase genes. J Virol Methods 2009;161:231–239.
11. Eshleman SH, Hudelson SE, Smith P, *et al.*: Analysis of pol integrase sequences in diverse HIV type 1 strains using a prototype genotyping assay. AIDS Res Hum Retroviruses 2009;25:343–345.
12. Connell BJ, Michler K, Capovilla A, *et al.*: Emergence of X4 usage amongst HIV-1 subtype C: Evidence for an evolving epidemic in South Africa. AIDS 2008;22:896–899.
13. Rousseau CM, Birditt BA, McKaya AR, *et al.*: Large-scale amplification, cloning and sequencing of near full-length HIV-1 subtype C genomes. J Virol Methods 2006;136:118–125.
14. Rhee SY, Liu TF. Kiuchi M, *et al.*: Natural variation of HIV-1 group M integrase: Implications for a new class of antiretroviral inhibitors. Retrovirology 2008;5:74.
15. Hunt GM, Papathanasopolous MA, Gray GE, and Tiemessen CT: Characterisation of near-full length genome sequences of three South African human immunodeficiency virus type 1 subtype C isolates. Virus Genes 2003;26:49–56.
16. Tzitzivacos DB, Tiemessen CT, Stevens WS, and Papathanasopoulos MA: Viral genetic determinants of nonprogressive HIV-1 subtype C infection in antiretroviral drug naïve children. AIDS Res Hum Retroviruses 2009;25:1141–1148.
17. Passaes CB, Guimaraes ML, Fernandez SLC, *et al.*: Lack of primary mutations associated with integrase inhibitors among HIV-1 subtypes B, C, and F circulating in Brazil. J Acquir Immune Defic Syndr 2009;51:7–12.
18. Ceccherini-Silberstein F, Malet I, Fabeni L, *et al.*: Specific mutations related to resistance to HIV-1 integrase inhibitors are associated with reverse transcriptase mutations in HAART-treated patients. Antivir Ther 2007;12:S6.

Address correspondence to:
M.A. Papathanasopoulos
HIV Pathogenesis Research Laboratory
Department of Molecular Medicine and Haematology
University of the Witwatersrand Medical School
7 York Road
Parktown, 2193, Johannesburg, South Africa

E-mail: Maria.Papathanasopoulos@wits.ac.za