GENERATION OF SOLUBLE, CATALYTICALLY ACTIVE COVALENT HIV-1 SUBTYPE C INTEGRASE-DNA COMPLEXES TO IDENTIFY NOVEL STRAND TRANSFER INHIBITORS

Grant James Beyleveld

Dissertation submitted to the Faculty of Health Sciences, University of the Witwatersrand, Johannesburg, in fulfillment of the requirements for the degree of Master of Science in Medicine.

Johannesburg, 2011

Declaration

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

28th day of August, 2012

Abstract

The HIV-1 integrase (IN) enzyme is an integral part of the viral replication cycle and has no known human homologues, making it an ideal target for antiretroviral therapy. To date, only one inhibitor of IN strand transfer activity (Raltegravir, Isentress^{TM}) is available for human use. However, the inevitable emergence of antiretroviral drug resistance requires ongoing research into new/novel therapies. There are currently no assays to screen for IN inhibitors against HIV-1 subtype C in South Africa (and worldwide), therefore, the overall objective of this study was to generate and characterize locally relevant, soluble, functional recombinant HIV-1 subtype C IN proteins for use in strand transfer assays. Recombinant *integrase* genes, including a soluble HIV-1 subtype C mutant (05ZAFV6 with C56S, C65S, W131D, F185D and C280S) and HIV-1 subtype C Y143C mutant (05ZAFV6 soluble with Y143C) were designed, generated and cloned in frame into pET15b. Optimal bacterial expression conditions for the expression of these constructs as well as an HIV-1 subtype C wild type (05ZAFV6), subtype B wild type (NL4-3), and subtype B soluble (NL4-3 with F185K and C280S; as controls) IN, in *E.coli* BL21 cells were determined. All five recombinant IN were successfully purified using nickel affinity chromatography, and subsequently used to establish a strand transfer assay to assess their activity and their response to two well-known integrase inhibitors, L-Chicoric acid and Raltegravir. All five recombinant IN proteins were found to be biologically active, with IN^{Y143C} (116.67%) showing equivalent activity to IN^{Bwt} (117.37%), while IN^{Csol} (52.96%) was the lowest. The IC_{50} values of L-Chicoric acid were higher than the expected values for all five

recombinant IN, with the subtype B and C IN solubility mutations contributing to an increased resistance to inhibition by L-Chicoric acid.

The dose responses to Raltegravir for IN^{Cwt} and IN^{Bsol} were as expected, with IC_{50}'s in line with published data, and the IN^{Y143C} mutant (known mutation conferring resistance to Raltegravir) was resistant to inhibition of strand transfer activity at all Raltegravir concentrations tested except the highest (50 µM).

Finally, methods to complex the IN^{Y143C} mutant to thiolated-DNA were evaluated, however definitive data could not be obtained. Future work should focus on optimization of the purification and characterization of the IN-DNA complexes. Overall, this study has led to the establishment of functional strand transfer assays based on HIV-1 subtype C recombinant IN proteins, and established a framework for screening of novel HIV-1 subtype C IN inhibitors.

Acknowledgments

I hereby thank the following individuals and organizations for their contribution to this work:

Professor Maria Papathanasopoulos

Dr Raymond Hewer

Dr Salerwe Mosebi

Miss Telisha Traut

Funding from:

Mintek

HIV Pathogenesis Research Laboratory, Department of Molecular Medicine and Haematology, University of the Witwatersrand, Johannesburg

Poliomyelitus Research Foundation

Table of Contents

List of Figures

List of Tables

Nomenclature

°C	Degrees Celsius
Å	Angstroms
AIDS	Acquired Immune Deficiency Syndrome
AP	Alkaline phosphatase
ARV	Antiretroviral
BAF	Barrier to autointegration factor-1
BSA	Bovine serum albumin
bp	Base pair
CA/p24	Capsid Protein/ p24 Antigen
CCD	Catalytic Core Domain
CCR5	C-C chemokine receptor type 5
CD4	Cluster of Differentiation 4
CDC	Centres for Disease Control and Prevention, USA
cDNA	Complementary DNA
CTD	Carboxy Terminal Domain
CXCR4	C-X-C chemokine receptor type 4/ Fusin / Cluster of Differentiation 184 (CD184)
DKA	β-diketo acid
DNA	Deoxyribonucleic Acid
DIG	Digoxigenin
DTNB	5,5'-dithiobis-(2-nitrobenzoic acid)
DTT	Dithiothreitol
E.coli	*Escherichia coli*
FDA	Food and Drug Administration, USA
HAART	Highly Active Antiretroviral Therapy
HBV	Hepatitis B Virus
HIV-1	Human Immunodeficiency Virus Type 1
HIV-2	Human Immunodeficiency Virus Type 2
HMG-A1	High mobility group protein A1
HRP	Horseradish Peroxidase
HTLV	Human T-cell Leukemia Virus
IC_{50}	50% inhibition concentration
IN/p32	Integrase
INI1	Integrase interactor 1
IPTG	Isopropyl β-D-1-thiogalactopyranoside
Kb	Kilobases
kDa	Kilodaltons
LAV	Lymphadenopathy Associated Virus
LB	Luria Bertani (media)
LEDGF/p75	Lens Epithelial-Derived Growth Factor/p75
LTR	Long Terminal Repeat
mA	Milliamps

MA/p17	Matrix
MCC	Medicines Control Council, South Africa
ml	Milliliter
MMWR	Morbidity and Mortality Weekly Report
MPa	Mega Pascals
N-terminal	Amino terminal
NC/p7	Nucleocapsid Protein
NIH	National Institutes of Health, USA
nm	Nanometer
NRTI	Nucleoside Reverse Transcriptase Inhibitor
NNRTI	Non-Nucleoside Reverse Transcriptase Inhibitor
NTD	Amino Terminal Domain
OD	Optical density
p55	Gag
PBS	Phosphate buffered saline
PCP	*Pneumocystis carinii* pneumonia
PIC	Preintegration Complex
PI	Protease inhibitor
PFV	Prototype foamy virus
PR/p15	Protease
R5	CCR5 tropic (virus)
RNA	Ribonucleic acid
RNH/p66	RNase H
RNP	Ribonuclear Portion
RT/p51	Reverse Transcriptase
SDS-PAGE	Sodium Dodecyl Sulfate-Polyacrylamide Gel Electrophoresis
ST	Strand transfer
STI	Strand Transfer Inhibitor
SU/gp120	Envelope Surface Glycoprotein
T-TBS	Tween-Tris Buffered Saline
TAE Buffer	Tris-Acetate-EDTA Buffer
TB	Tuberculosis
TBS	Tris Buffered Saline
Vpr	Viral protein R
WHO	World Health Organization
X4	CXCR4 tropic (virus)
xg	Times gravity
µl	Microliter
µM	Micromolar
UNAIDS	Joint United Nations Programme on HIV/AIDS
UV	Ultraviolet

CHAPTER ONE

1 INTRODUCTION

1.1 Human Immunodeficiency Virus and the Global Epidemic

Acquired Immune Deficiency Syndrome (AIDS) was initially described following a Centres for Disease Control and Prevention, USA (CDC) Morbidity and Mortality Weekly Report (MMWR) outlining several cases of *Pneumocystis carinii* pneumonia (PCP) in young "previously healthy" homosexual male patients in 1981 (1-3). Infection with Human Immunodeficiency Virus type 1 (HIV-1) was later defined as the causative agent of AIDS (4). The virus is a primate lentivirus of the family Retroviridae, and was discovered simultaneously by two groups in 1983 (5,6). Gallo *et al.* (1983) originally designated HIV as part of the Human T-cell Leukemia Virus (HTLV) group (5), while Barré-Sinoussi *et al.* (1983) termed the virus a Lymphadenopathy Associated Virus (LAV) (6), but it was eventually termed HIV (7) in a 1986 letter to Nature. HIV quickly spread amongst the heterosexual population (initially via intravenous drug users and hemophiliacs receiving infected blood transfusions).

HIV targets human CD4+ T-cells (8,9), macrophages and dendritic cells (10), and results in a severe attenuation of the immune system due to destruction of CD4+ T cells, which ultimately leads to AIDS (11). A variety of opportunistic infections ensues and ultimately cause the death of the patient. HIV is often not detected early as initial symptoms are non-specific (12) so the move to test a patient relies heavily on clinical suspicion founded by a history of possible exposure, but more often only following the presentation of AIDS symptoms late in disease progression (particularly in a resource limited setting). More recently, there has been a move to perform HIV testing regardless of any clinical indicators, resulting

in the early detection of more infections. A variety of antibody, antigen and nucleic acid based HIV diagnostic tests are available. Infection is due to transmission by bodily fluids (blood, semen or breast milk) and infection is most likely to occur when transmission occurs at a mucosal surface (13), especially during sexual intercourse.

The disease burden on the world is significant - over 33.3 million people were living with HIV infection by the end of 2009 (14), while over 36 million people have died of AIDS since the beginning of the pandemic. While the pandemic is globally important, nowhere is it so pertinent as in Sub-Saharan Africa: according to the Joint United Nation Programme on HIV/AIDS (UNAIDS), Sub-Saharan Africa has the highest number of HIV-1 infections in the world (over 22.5 million people living with HIV infection, representing 67% of the worlds infections), while South Africa holds the ominous distinction of having the highest number of infections in a single country: 5.7 million people (14). During 2009, there were at least 1.8 million deaths worldwide attributed to AIDS, the vast majority of which (1.3 million) were in sub-Saharan Africa.

1.2 The Biology of HIV

1.2.1 HIV Genetic Diversity

There are two types of HIV in circulation, HIV type 1 (HIV-1) and HIV type 2 (HIV-2). HIV-1 originated following zoonotic transmission of the Simian Immunodeficiency Virus (SIV) from the common Chimpanzee (*Pan troglodytes*

troglodytes) (15). HIV-2 occurred similarly by zoonosis from the Sooty Mangabey monkey (*Cercoebus atys*) (16).

HIV-1 is of greater importance due to its higher virulence and infectivity, and reduced immune control (16). HIV-1 originated in West Central Africa, and is now the dominant HIV type globally. HIV-2 remains prevalent mostly in West Africa (16). HIV-1 can be further subdivided into groups M (Main), O (Outlier) (17), N (New, Non-M or Non-O) (18) and P (19). Group M is the collective term for the various subtypes A – F (but not E) (20,21), G – J (but not I) (22-25), and K (26). Subtype E was found to be a recombinant between other subtypes (21,27-29), as was subtype I (30,31). Further, there are currently 49 circulating recombinant forms (CRF's) (Figure 1.1) (32). The different subtypes may show up to 30% difference in their amino acid sequence (33).

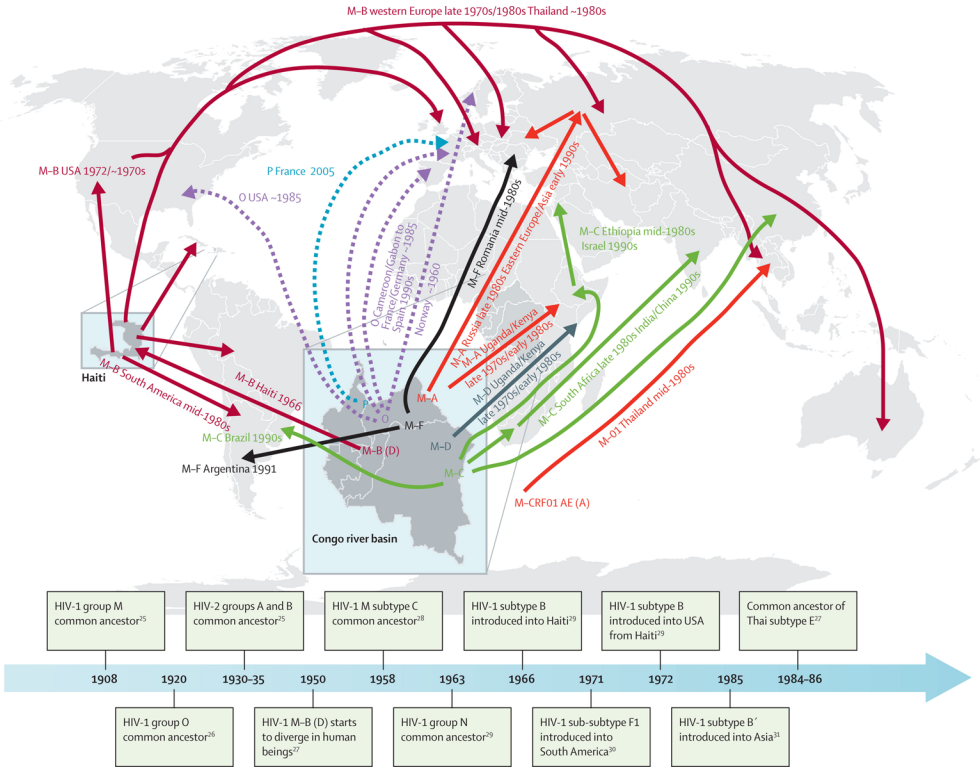

Figure 1.1: Chronology and geographical distribution and evolution of main HIV types, groups and subtypes. Main epicenters are enlarged. Copied from: Tebit and Arts (2011) (34).

1.2.2 The Structure of the HIV-1 Virus

The structure of HIV-1 consists of a cone shaped core of capsid (CA/p24) protein, which encloses two identical single stranded HIV-1 ribonucleic acid (RNA) molecules (plus strand) (35). Also contained within the p24 core is reverse transcriptase (RT/p51), integrase (IN/p32), protease (PR/p15) and RNase H (RNH/p66) (36). PR is involved in viral maturation (37), whilst IN, RT and RNH are all involved in the early steps of viral infection (38). Nucleocapsid (NC) protein p7 binds to viral RNA (constituting the ribonuclear portion, RNP) to prevent degradation by RNH whilst in the virion (36). The viral core is surrounded by the matrix (MA/p17), which forms a 7nm thick shell immediately inside of the lipid membrane (39). The lipid membrane is derived from the host cell, and also contains the exposed trimeric viral envelope surface glycoprotein (SU/gp120) and

membrane associated glycoprotein gp41, which are responsible for host receptor binding and fusion (36).

1.2.3 Genome Organisation of HIV-1

HIV-1 encodes 9 genes, flanked by identical long terminal repeat (LTR) regions (approximately 9.7 kilobase (kb) genome) - this remarkable efficiency in gene expression from a small genome is achieved through the use of multiple overlapping reading frames, resulting in the transcription of the genes *gag*, *pol*, *env*, *tat*, *rev*, *vif*, *vpr*, *vpu* and *nef* (32). The *gag* gene encodes p55 which is processed by viral proteolytic cleavage into p17 (MA), p24 (CA), p7 (NC) and p6 (40). The *pol* gene codes for the gag-pol precursor protein which, when cleaved by the viral protease, results in the proteins p15 (PR), p51 (RT), p66 (RNH) and p31 (IN) (32). The *env* gene codes for the precursor viral glycoprotein gp160, which is processed by host proteases into gp120 and gp41 (41) and which subsequently interact with each other and form a non-covalently bound gp120-gp41 trimer on the viral envelope surface (41). The *gag*, *pol* and *env* genes together encode the structural proteins of the HIV-1 genome. The *tat* and *rev* genes encode the regulatory factors Tat and Rev that control HIV-1 gene expression (42). Rev is encoded from the 3' end of the *tat* gene, although in an alternate reading frame. Viral Infectivity Factor (Vif), Viral protein R (Vpr), Viral Protein U (Vpu) and Negative Regulatory Factor (Nef) are the accessory proteins and are all encoded from individual genes, and alternate reading frames are employed here too.

1.2.4 The HIV-1 Life Cycle

HIV-1 executes its life cycle in a fairly systematic manner, and various steps in the process have been identified as antiretroviral (ARV) drug targets. These steps include fusion and entry, reverse transcription, integration, transcription and translation, assembly, and budding and maturation. Initially viral gp120 binds to the CD4 receptor on the surface of host cells, and gp41 mediates fusion of the viral membrane with the host cell membrane following the sequential binding of gp120 to the co-receptor CCR5 or CXCR4 (38). Strains of virus designated R5 are able to use the CCR5 co-receptor, which is present on dendritic cells at the sites of mucosal infection (12,13) (as well as T cells, macrophages and microglia cells (43)) and thus are more prevalent during the early acute stages of infection. R5 viruses are commonly termed “macrophage tropic” (12,44). X4 virus strains utilize the CXCR4 co-receptor (12,44), are considered “T cell tropic” and subsequently predominate during the later chronic stages of infection (44).

Upon entering the CD4+ T cell, the HIV-1 p24 core antigen disassembles, allowing reverse transcription of the viral RNA genome to complementary deoxyribonucleic acid (cDNA) by the viral RT enzyme (45). Immediately following this, the viral cDNA assembles together with a number of associated viral and cellular proteins into a preintegration complex (PIC) (45), and moves to the nucleus. There, the PIC facilitates the integration of the viral cDNA into the host chromosome, forming a functional provirus. This integration event is defined by two key processes: 3’ end processing and strand transfer (see Section 1.3.3). Viral proteins are transcribed from the integrated proviral cDNA, and assemble into new immature virions which

bud out of the host cell. During the budding process, these immature virions mature into infectious virions when the Gag proteins are cleaved by the viral PR (37).

1.3 HIV-1 IN Structure and Function

1.3.1 Structure and Domains on the HIV-1 IN Monomer

The HIV-1 IN enzyme, a 32 kilodalton (kDa) protein made up of 288 amino acids, performs two vital functions within its role in the viral life cycle; it processes the viral deoxyribonucleic acid (DNA) before integration and facilitates strand transfer, thereby establishing a provirus. IN is also capable of disintegration, wherein it facilitates the reversal of the strand transfer process, although the cellular relevance of this process is still unknown (46). Some authors have also found a requirement for IN in the RT nucleoprotein complex, suggesting an indirect role for IN in viral DNA replication (47,48). IN exists as a homodimer in solution, and it is proposed that larger functional tetramers are formed *in vivo* and *in vitro* (49,50), as well as other non-functional higher order complexes.

The enzyme can be divided into three domains, the amino terminal domain (NTD), the catalytic core domain (CCD) and the carboxy terminal domain (CTD). The NTD (residues 1-50) contains the HHCC motif (histidine 12, histidine 16, cysteine 40 and cysteine 43), and a zinc finger motif which chelates a single zinc atom (51,52). This zinc atom serves to stabilize the NTD, which in turn is required for IN-IN interactions (53-55). The CCD (residues 51-212) contains the universally conserved catalytic triad or DDE motif (aspartic acid 64, aspartic acid 116 and

glutamic acid 152), a negatively charged motif which coordinates the divalent magnesium ions at the active site by means of the carboxylates of these residues (56-58). These divalent metal ions have been shown to be required for IN enzyme activity (59). The CCD is responsible for binding the viral DNA ends in a sequence specific manner (60-64). The CTD is the least conserved of the three domains (49), and is responsible for nonspecific binding to the host DNA by means of a motif that shows homology to SH3 DNA binding motifs. Mutational studies of residues in this domain have shown that IN-IN interactions involving the CTD influence integration efficiency (65).

1.3.2 Models of the HIV-1 IN Tetramer

As yet, no crystal structures of the full length HIV-1 IN protein or any multimers have emerged. There are, however, several single domain (52,66-68) and two-domain structures (69,70) available, and a proposed tetramer has been assembled from this information (Figure 1.2). Complementation studies (wherein IN enzymes with different domain mutated to be defective were subsequently rendered active by mixing them together) showed that the functional IN enzyme was multimeric (49,58,71,72). Since the IN enzyme processes and integrates two viral ends, one can envision that two active sites are required to achieve this feat. Mu transposases, which are functionally and structurally homologous to HIV-1 IN (73), occur as functional tetramers wherein only two of the four active sites are reactive (74,75). IN dimers assemble and associate with the viral DNA ends (Figure 1.2 C), and these dimers are responsible for the 3' processing reaction (76-79). These dimers then associate to form a tetramer (Figure 1.2 A and B,

where the blue and pink subunits are reactive) which is considered the functional form and is capable of performing full strand transfer (65,80).

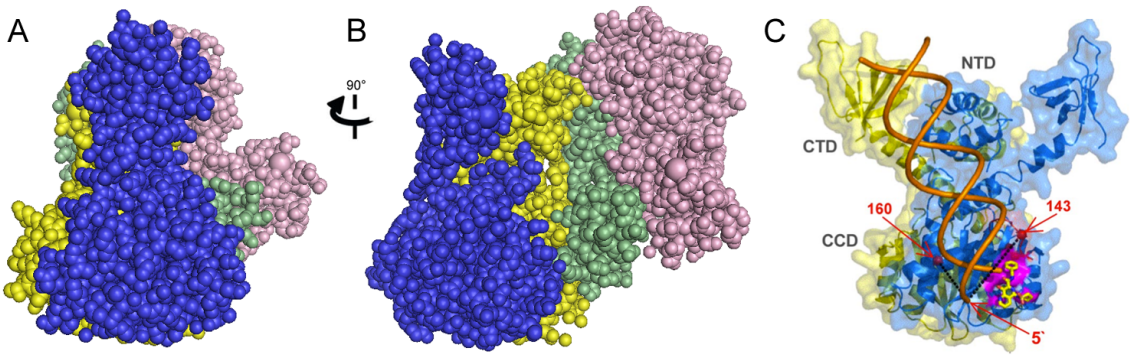

Figure 1.2: Comparison between IN models. A and B are adapted from Wang *et al.* (2001) (70) (PDB ID 1K6Y) and the colouring was changed for clarity. C is copied from Alian *et al.* (2009) (81). The blue and yellow subunits in A and B are analogous to the blue and yellow subunits in C, although for clarity the CTD regions have been omitted in A and B. The pink and green subunits in A and B represent a second dimer, making up a tetramer. DNA binding is not shown in A and B.

Most recently, the structure of the prototype foamy virus (PFV) tetramer in complex with its cognate DNA substrate was resolved by X-ray crystallography (82) (Figure 1.3), and this is thought to bear significance for the HIV-1 tetramer due to structural, functional and sequence based homology between these two IN enzymes. The model suggests a significantly different structure to that proposed by other groups: an asymmetric tetramer consisting of two dimers is predicted, and within each dimer the inner monomer (shown in green and blue in Figure 1.3) binds DNA, while the outer monomers (both yellow in Figure 1.3) stabilize the structure, contribute to activity and participate in host DNA binding (the NTD's and CTD's of the outer monomers did not resolve on the electron density map). The CCD-CTD linker is extended and occurs parallel to the NTD-CCD linker, and together these "truss" the complex together (indicated by arrows in Figure 1.3). While HIV-1 IN has shorter inter-domain linkers, it is thought that these can extend to allow a similar structure in HIV-1 IN. The inner subunits in this model appear to

control the catalytic function of the enzyme as well as tetramerization, whilst the outer subunits have a supporting structural role.

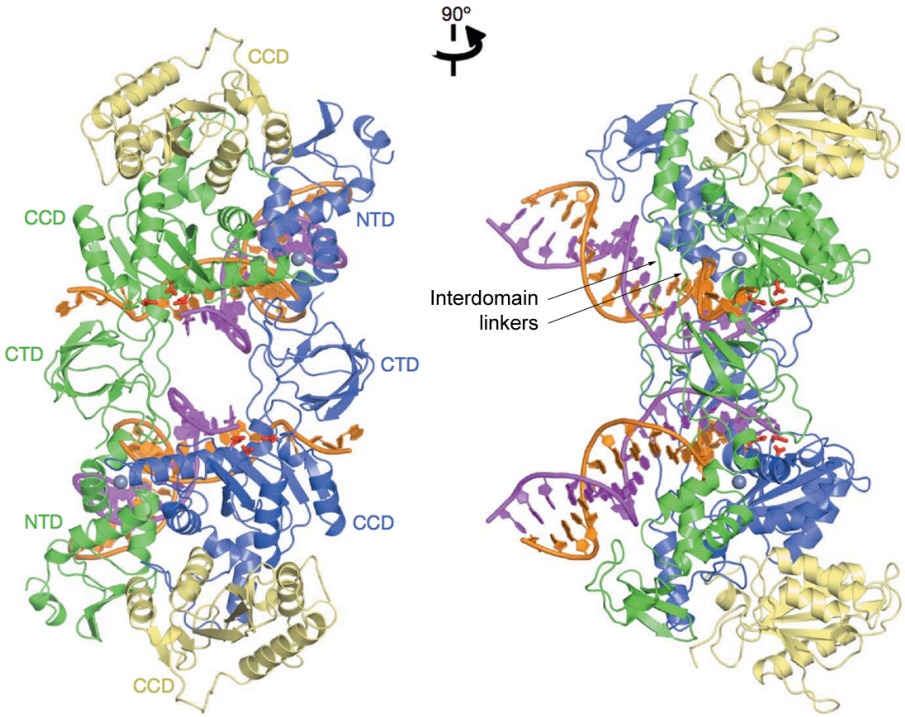

Figure 1.3: Model of the prototype foamy virus integrase tetramer with viral DNA ends. Inner monomers are coloured blue and green, DNA is coloured purple ad orange, outer monomers are coloured yellow (amino terminal and carboxy terminal domains did not resolve). Arrows indicate interdomain linkers. Source: Adapted from Hare *et al.* (2010) (82).

Several models for the IN-DNA interaction have been proposed (70,82-86), and in general they agree on a few key factors. First, IN dimers bind to each of the viral DNA ends (Figure 1.2 C), and then these come together to form an IN tetramer (Figure 1.2 A and B, DNA not shown, Figure 1.3). Second, models for this tetramer are constrained by the 5 base pair (bp) spacing between the integration events. Thus the active sites that bind the DNA (one from each dimer) must be in the order of 20 Å apart in the tetramer. Third, known interactions must be accounted for, such as the non-specific contacts between the CTD and the viral DNA (60,62) and interactions between the CCD and the host DNA. IN has a propensity to form higher order oligomers without bound DNA, but it has been shown that DNA

binding induces a dissociation of these DNA-free oligomeric forms before re-association into a tetrameric DNA complex (76). In fact, the tetrameric form cannot bind DNA but rather must form after dimers have associated with DNA already (78). Indeed, this characteristic has been employed in an attempt to inhibit IN function by shifting the equilibrium towards the tetrameric form and thus inhibiting DNA binding (87). This specific order of events has also been suggested to play a role in substrate specificity, because the functional tetrameric complex (and hence an integration event) will only form in the presence of two viral DNA ends which must first associate with IN dimers specifically (79). The model of PFV IN (82) shows elaborate IN-DNA interactions, with as much as 10 000 $Å^2$ of molecular surface being involved, with extensive interaction taking place in the terminal nucleotides of the viral DNA within the active site (where the DNA is shown to diverge considerably from the native helix form). The CTD appears to make contact with the DNA molecules associated with both dimers, which acts to stabilize the tetramer.

1.3.3 The HIV-1 Integration Reaction

The integration reaction itself is carried out in two discrete steps (Figure 1.4). As mentioned before first the IN enzyme dimer performs 3' processing wherein IN recognizes specific conserved LTR's in the viral DNA and cleaves a GT dinucleotide from these termini, leaving 2 bp CA dinucleotide overhangs with exposed 3' reactive hydroxyl (OH) groups (88,89). Subsequently, the IN tetramer coordinates the nucleophilic attack of the host DNA by the reactive 3' ends of the viral DNA and ligates the viral DNA to the host DNA. Insertion of both ends (double

end transfer) occurs at a 5 bp stagger on the host chromosome, and host DNA repair mechanisms are relied upon to join the viral and host strands and seal the nicks (90-94).

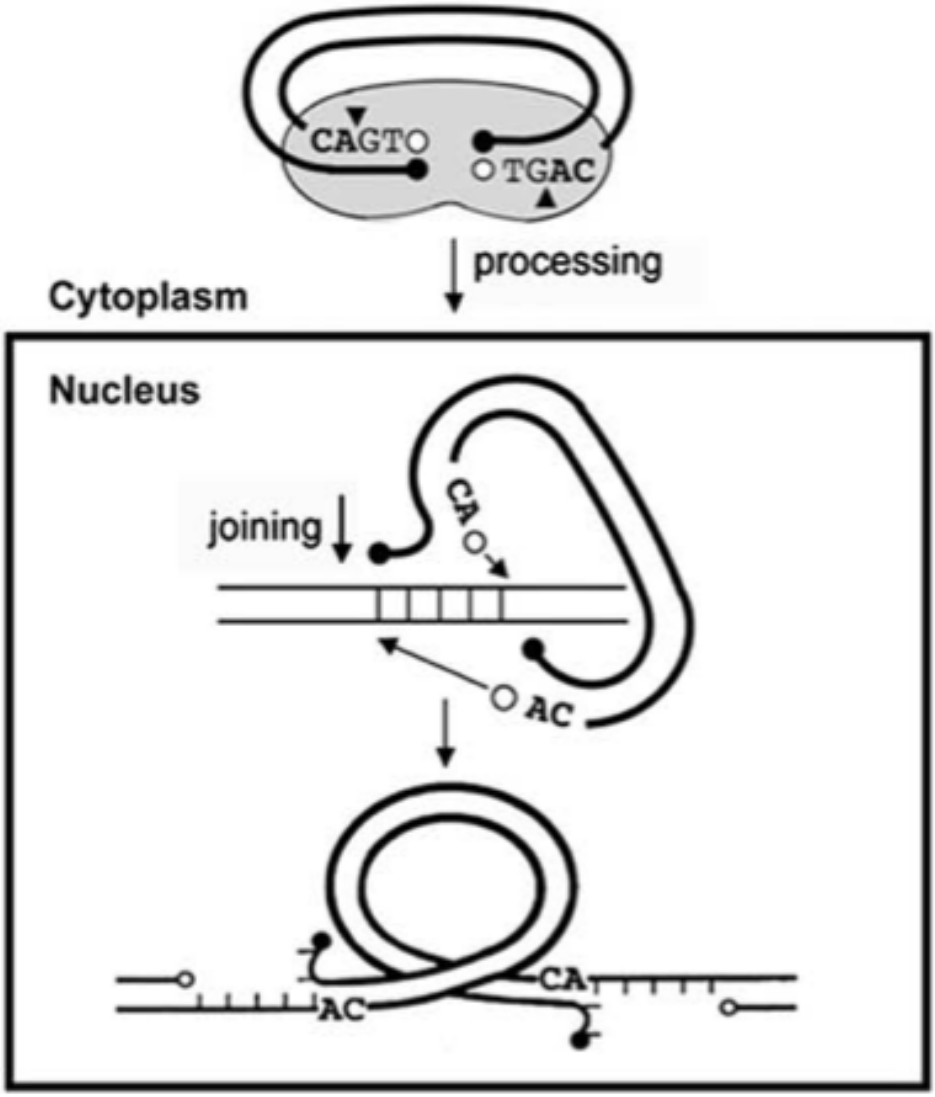

Figure 1.4: Schematic representation of the *in vivo* strand transfer reaction. Copied from: Merkel *et al.* (2009) (95). Top, the 3′ processing reaction is represented, showing the 5′ termini as filled circles. The position of the dinucleotide cleavage is indicated by arrows. The G and T nucleotides (grey) are removed leaving the highly conserved CA dinucleotide (in black). Bottom, the strand transfer reaction is shown: the 3′ ends of the viral DNA are inserted into the host chromosome, and the resulting 5bp gap is repaired by the host DNA repair mechanisms.

1.3.4 The Preintegration Complex

Following 3' processing, the IN dimer remains bound to the viral DNA and associates with the IN dimer on the other terminus of the viral DNA to form a tetramer. Various viral and host proteins associate with the tetramer to form the PIC, which serves as the functional unit for the strand transfer reaction described above (96). The complex consists of IN and viral DNA, as well as barrier to

autointegration factor-1 (BAF) (97), heat shock protein 60 (HSP60) (98), IN interactor 1 (INI1) (99), high mobility group protein A1 (HMG-A1) (100-102) and lens epithelial derived growth factor (LEDGF/p75) (103-105). Viral proteins present in the PIC are MA, Vpr, NC (p7) and RT (101). The PIC forms in the cytoplasm and by its association with LEDGF (104) and MA (106) it moves into the nucleus (Figure 1.4). LEDGF also plays a role in tethering the PIC to host DNA (103,107-109).

1.4 Antiretroviral Therapy

1.4.1 Disease Progression

Following infection with the HIV-1 virus, individuals undergo disease progression to AIDS over time (Figure 1.5). Phase I is know as the Acute Primary Infection phase, and lasts from a couple of weeks to six months. Patients sometimes present with flu-like symptoms, although this so called “seroconversion illness” is not often properly diagnosed. Patients in this phase typically have high viral loads, as the virus is allowed to replicate uncontrolled, and there is extensive depletion of the CD4+ T cells. Phase II is known as the Clinically Asymptomatic Infection phase, wherein the patient displays no outward symptoms of disease. The viral load lowers to a “set-point” and the CD4+ T cells recover to a more normal level. This phase may last up to 10 years, and there is a gradual decline in CD4+ T cell numbers during this time. Phase III is the Symptomatic HIV infection and AIDS phase, and is characterized by a variety of opportunistic infections and cancers associated with a decline in immune function and depletion of CD4+ T cells to

below 20 cell/µl. Without treatment, it is unlikely that an AIDS patient will survive for more than two years (110).

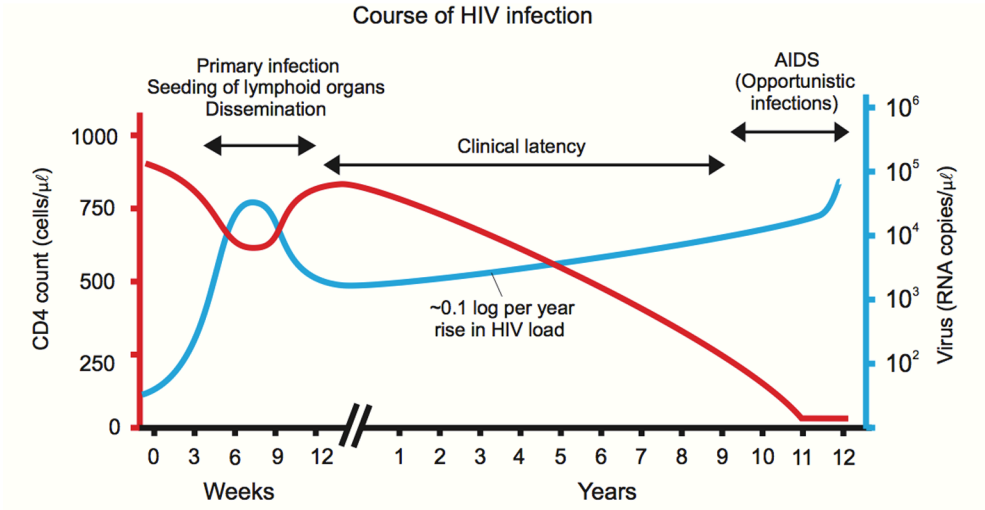

Figure 1.5: Schematic of HIV infection, Showing CD4+ T Cell Count and Viral Load as a function of time. Source: Stevens and Papathanasopoulos (2009) (110).

1.4.2 Current Treatment Options

ARV therapy or HAART (Highly active antiretroviral therapy) has proven successful in treating HIV-1 infected patients, and reducing the morbidity and mortality associated with AIDS. However, the development of viral drug resistance to ARV drugs has meant that many of the therapies are no longer effective for treatment experienced patients (111). Further, many of the ARV therapies have significant side effects and complex treatment strategies which affect patient compliance to the regimen (111). HIV replication is primarily driven by three enzymes: RT, protease and IN, and the common ARV therapies used today are mostly small molecule inhibitors of RT and protease. There are ten protease inhibitors (PI's) (eg. Saquinavir) (112), eight nucleoside RT inhibitors (NRTI's) (eg. AZT or zidovudine) (113), four NRTI combination drugs (eg. Combivir) (114,115), six non-nucleoside RT inhibitors (NNRTI's) (116, 120,121), two entry inhibitors;

one co-receptor antagonist (Maraviroc) (117) and one peptidic inhibitor of viral fusion (Enfuvirtide) (118), and one multiclass combination consisting of two NRTI's and one NNRTI (Atripla) (119). Because IN has no known human homologue, it was identified as an ideal drug target. However, to date there is only a single FDA approved drug available that inhibits IN function (Raltegravir, Isentress™). Thus, IN is an attractive target for ongoing ARV drug discovery.

1.4.3 ARV Therapy Regimens

The current World Health Organization (WHO) guidelines for HIV treatment are as follows: treatment commences when the CD4+ T cell count is below 350 cells/µL, unless the patient is diagnosed with a condition from the clinical stage 3 or 4 in the WHO Clinical Staging of Disease outline (see Appendix A), has a co-infection with Tuberculosis (TB) or Hepatitis B Virus (HBV), or is pregnant (122). The commonly used first line regimen consists of either zidovudine or tenofovir (both NRTI's) in combination with a second NRTI (either lamivudine or emtricitabine) and an NNRTI (efavirenz or nevirapine; efavirenz is not indicated for use in pregnant women during the first trimester). The second line involves a switch from tenofovir to zidovudine (or *vice versa* based on the original regimen), maintenance of lamivudine and emtricitabine and the addition of PI's (azatanavir or lopinavir, in combination with ritonavir).

In South Africa the approach is different (123): treatment generally commences when the CD4+ T cell count is below 200 cells/µL, although in the case of co-infection with tuberculosis (TB) or pregnancy this is adjusted to <350 cells/µL. Further, treatment begins irrespective of the CD4 count if the patient presents with

any condition from stage IV AIDS (see Appendix A). The current first line regimen consists of tenofovir (unless the patient is already on a stavudine regimen) or zidovudine (in the case of contraindication to tenofovir) in combination with a second NRTI (either lamivudine or emtricitabine) and a NNRTI (efavirenz or nevirapine). If the patient experiences virologic failure on a stavudine/zidovudine regimen, the second line regimen consists of tenofovir, lamivudine or emtricitabine, and a combination of PI's (lopinavir and ritonavir). If the patient failed treatment on a tenofovir regimen, the above second line therapy is used exchanging tenofovir for zidovudine (123). Only a handful of countries employ Raltegravir in their treatment strategies (for example Botswana and Mexico both use Raltegravir as a third line therapy) (124), and only the United States includes Raltegravir in their first and second line therapies (125). Raltegravir only obtained Medicines Control Council (MCC) approval for use in South Africa in May 2011 (F Venter, *Pers comm*), so it is expected that it will soon be incorporated in treatment regimens.

1.4.4 HIV-1 IN as a Target for Inhibition

IN and the strand transfer (ST) process have been deemed difficult targets for inhibition because only two events are required to take place in every cell in order for successful integration to occur (i.e. the integration of the two viral ends). This is not the case for other targets, such as reverse transcription where thousands of enzymatic reactions take place in order to carry out the viral life cycle (126). It is known that the strand transfer reaction is the primary target in order to halt the viral lifecycle, and not the 3' processing reaction (127). Pommier *et al.* (2005) describes four characteristics of a ST inhibitor (STI) in a 2005 review (94): firstly

the compound should be active at the appropriate point in the viral lifecycle, which gives a 4-16 hour window post infection. Secondly, following treatment there should be an accumulation of 2-long terminal repeat (2-LTR) circles, which arise when a build up of unintegrated viral DNA is circularized by host enzymes. Third, integration and the subsequent measurable amount of proviral DNA should decrease. Lastly, treatment should result in the selection of resistance mutations in the IN gene, which confer lower susceptibility when induced into recombinant IN and tested *in vitro*.

There are several published platforms to screen for integrase inhibitors (128-133). PICs isolated from infected cells have been shown to catalyze 3' processing and the integration of two viral DNA ends into target DNA, but they are notoriously difficult to isolate and work with. An alternative is the use of recombinant integrase. In solution, recombinant integrase exists as monomeric, dimeric, tetrameric and oligomeric structures (78,134). Only the tetramers are active for 3' processing and complete strand transfer *in vivo*, while dimers are capable of carrying out 3' processing and half site integration *in vitro*. Furthermore, integrase exhibits low solubility and so a combination of high salt buffers during the purification procedure and a variety of mutations to reduce surface hydrophobicity have been employed to enhance the yield. Unfortunately these high salt concentrations interfere with DNA binding, making it difficult to assess the activity of integrase when bound to DNA (81).

1.4.5 The Development of HIV-1 Strand Transfer Inhibitors

Early attempts at identifying STI's looked at ribozymes (135,136) and triple helix forming oligonucleotides (137) amongst others, although none of these concepts were tested beyond the basic science arena due to high levels of *in vitro* cellular toxicity. In 1994 a breakthrough was made at the MERCK laboratories when Hazuda *et al.* (1994) developed a sensitive, high-throughput, *in vitro* assay to assess strand transfer (132). This allowed for high throughput screening of potentially inhibitory compounds against recombinant IN using a shotgun approach. This led to a host of discoveries around IN: the contributions of specific domains to IN function, the identification of specific catalytic residues (such as the DDE motif) and most importantly for this research, the first small molecule inhibitors. Indeed, this assay is still in use today in the IN drug discovery pipeline.

1.4.6 Strand Transfer Assays

The first strand transfer assay to emerge was a simple electrophoretic assay (138-140), wherein a short radio labelled donor DNA (which mimicks the viral ends) is incubated with IN, target DNA and the required metal cofactors. The IN protein then joins the donor DNA to a target DNA and the resulting products are resolved on a polyacrylamide gel. If integration takes place, the products can be seen as higher molecular weight bands. The advantage of this method is the ability to directly visualize and quantify the substrates and products of the reaction. The main disadvantages are the time and labour intensive nature of this assay and its low sensitivity. Despite these disadvantages, this assay was the foundation for

much early work in HIV IN and became somewhat of a gold standard in IN activity studies, even today.

To monitor IN binding to DNA, surface plasmon resonance (SPR) (141) and fluorescent anisotropy (76,134) were used. Several methods to assess DNA processing were developed: monitoring the release of a radio labelled CA dinucleotide (142), anisotropy using a fluorescent CA dinucleotide (77,143,144) and the use of a fluorescent quencher on the 5' DNA end adjacent to the fluorescent labelled 3' CA dinucleotide (145). Monitoring the joining reaction has received the most attention, and the use of biotinylated DNA substrates has proven a versatile tool in this respect. A streptavidin coated microplate (146) and streptavidin coated magnetic beads (142) are employed to capture the reaction products after denaturation, but this still relies on electrophoresis for quantification. A scintillation proximity assay (SPA) was developed, wherein biotinylated anti-IN antibodies are used to capture IN onto special SPA beads, and then radio labelled drug compounds are screened; the SPA bead facilitates an increase in the scintillation count due to the proximity to the drug when it is bound to IN (127). This method is particularly amenable to high-throughput applications, although it relies on radio active detection methods. Finally, a method was devised to measure the amount of joining product formed without electrophoresis; using biotinylated donor DNA immobilized on a microplate, target DNA labelled with either a radioisotope, fluorescence or Digoxigenin (DIG) for detection using an enzyme linked antibody. This forms the basis for the assay used in this study (described below).

In the assay employed by (132): a 5’ biotinylated double stranded donor DNA representing the viral DNA end is bound to a streptavidin coated microwell plate. The biotinylated strand contains the CAGT motif at its 3’ end. Recombinant IN is added to the microwell and binds and processes the 3’ end of the donor DNA, resulting in the removal for the terminal GT dinucleotide, leaving behind the highly conserved CA dinucleotide (Figure 1.6 A). Next, a labelled target DNA is added (representing the random fragment of human DNA into which the virus can integrate its own DNA) and by action of the IN enzyme the strand transfer reaction takes place (Figure 1.6 B), integrating the bound donor DNA into the free target DNA. Using an enzyme or fluorophore conjugated antibody to detect the labelled target DNA, the amount of strand transfer can be quantified using various calorimetric or fluorescent detection methods. In order to assess how drug candidates may abrogate this process, compounds are added in a step before the target DNA is added. A drop in observed strand transfer activity indicates inhibition by that particular drug candidate.

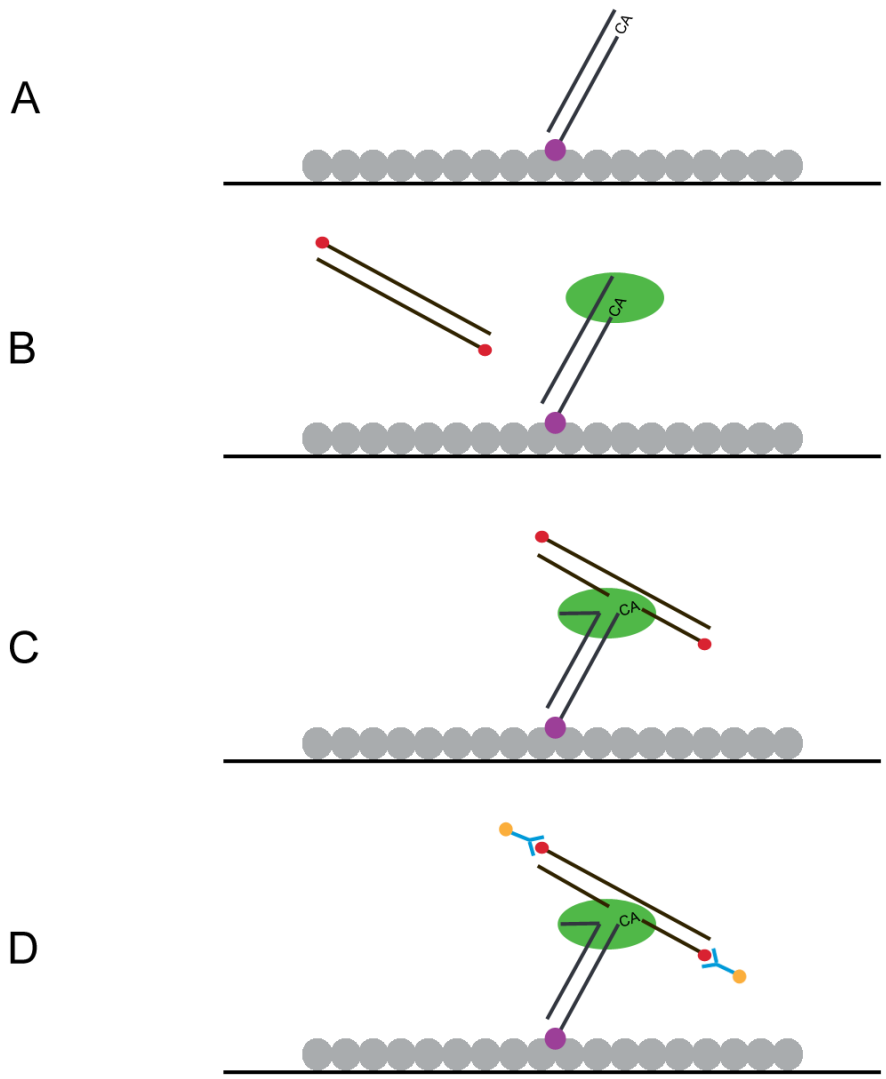

Figure 1.6: Schematic representation of the *in vitro* strand transfer reaction. Grey circles represent streptavidin coated surface of microwell. Purple circle represents biotin label on donor DNA, with the CA dinucleotide shown. (A) shows the first step, where the donor DNA is allowed to bind to the surface of the plate. In (B) the integrase enzyme (green) has bound to the donor DNA, and the target DNA has been added (molecular labels on target DNA shown as red circles). In (C) strand transfer has taken place. Enzyme conjugated antibodies (blue and orange, D) bind to the labelled ends of the target DNA and allow for quantification of strand transfer.

1.4.7 Early STI's and the Development of Raltegravir and Elvitegravir

Using the method described above, Robinson *et al.* (1996ab) studied a range of dicaffeoylquinic acids, which showed promising results against ST (147,148). Of these, L-Chicoric acid showed the greatest potency, with a 50% inhibitory concentration (IC_{50}) of approximately 0.6 µg/ml (0.126 µM) (147). Modelling work

with this inhibitor suggested binding took place at the active site in the catalytic core domain (148) and it was later found that the G140S mutation conferred resistance, further suggesting that binding occurs near the catalytic triad (149). Other early IN inhibitors contained the β-diketo acid (DKA) motif (126). Shionogi discovered 5-CITEP (150) and S-1360 (151) (licensed to GlaxoSmithKline as GSK-1349572) initially, the former being the first compound crystallized in complex with the IN CCD and the latter being the first compound to enter human clinical trials, although neither went beyond phase 2 testing. It was later suggested that this early model of 5-CITEP in complex with IN was not an accurate representation (152). At Merck, L-731,988 and L-870,812 were found, but again neither went to human clinical trials (153). L-870,810 (MERCK) was taken to phase 2 but hepatotoxicity observed in dogs ended further development (154). In 2007 MERCK attained FDA approval for the first IN inhibitor (an STI), Raltegravir (MK-0518, known commercially as Isentress™), for use in treatment experienced patients (155). In June 2009, they received FDA approval for Raltegravir as a first line therapy (126) based on early results from an ongoing phase 3 trial (156). Shortly after the FDA approval of Raltegravir Merck developed another lead, MK-2048, which holds promise as a second generation IN inhibitor (157), and shows good activity against at least four Raltegravir resistant IN mutants (158,159).

Japan Tobacco discovered another STI, JTK-303, which was subsequently licensed to Gilead Sciences (under the name GS-9137, also known as Elvitegravir) and has been in clinical development since 2005 (126). Phase 1 and phase 2 trials

showed clinical promise (160,161), and phase 3 trials are currently underway (126).

1.4.8 Modeling the IN drug interaction

In the PFV IN model by Hare *et al.* (2010) (82) (Figure 1.7), a proposed mechanism of action is put forward for Raltegravir and Elvitegravir since it has been shown that PFV is susceptible to these drugs (162). In this work it is evident that oxygen molecules in the drug structure interact specifically with the divalent metal ions within the active site and that the halobenzyl group common to both drugs displace the 3' terminal adenine of the viral DNA (Figure 1.7, panels B and C). The drugs have intimate interactions with the 3' CA dinucleotide of the viral DNA and with several highly conserved amino acid residues within the active site. Together, these effects disarm the active site of the IN enzyme and inhibit strand transfer.

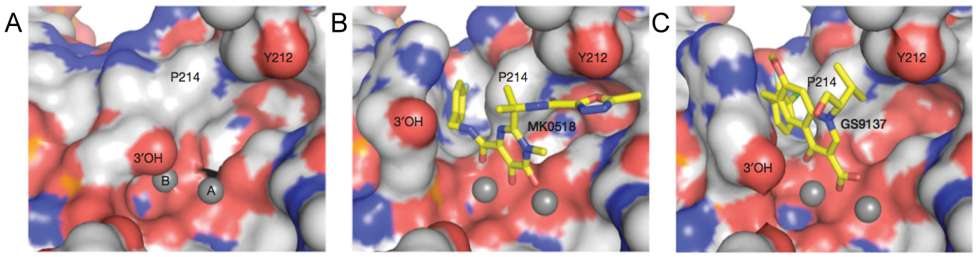

Figure 1.7: Surface models of the Prototype foamy virus active site, showing the divalent metal ions as grey spheres and drug molecules as stick models. (A) shows the active site without a drug present, (B) shows the active site with Raltegravir (MK0518) present and (C) shows the active site with Elvitegravir (GS9137) present. In (B) and (C) the oxygen molecules of the drug are shown in red. The halobenzyl group projects backwards into the active site. Copied from Hare *et al.* (2010) (82).

1.4.9 Non-Typical Strand Transfer Inhibitors

Other more unusual strand transfer inhibitors have also been developed. Typically, inhibitors are expected to be highly specific for their target, although several

inhibitors exist that target both IN and RT, such as V-165 (163-165), β-thujaplicinol (166-168) and Madurahydroxylactone and hydroxyisoquinalone derivatives (169,170). A range of peptide based inhibitors have been studied, including the hexapeptide HCKFWW (171-173), EBR28 (174), Integramides A and B (175,176), Indolicidin (177-179) and Pep1 (180). The mode of action and level of efficacy of these peptide based inhibitors varies, with the most effective in this list being the lysine linked tetramer of Indolicidin, which inhibits both 3' end processing and ST with an IC_{50} of 0.6 μM (178).

Developing peptidic inhibitors based on the IN sequence yielded several strong inhibitors, such as INH1 and INH5 which are based on the α1 and α5 helices of the IN monomer (IC_{50}'s of between 60 nM and 85 nM) (181) and K159 which was derived from the α4 helix of IN (182,183). Following this, an antibody against the K159 region of the IN monomer was developed with an IC_{50} of between 16 nM and 25 nM (184). Peptidic inhibitors based on other viral or host cellular proteins were studied: the S6 fragment of INI1 (185-187), the IN binding domain of LEDGF (188) and a small molecule (CHIBA-3003) that interferes with LEDGF-IN interaction (189). Screening of a library of peptides derived from the viral genome yielded several short peptides which have STI activity: 4268 and 4321 from RT (190), Rev13-23 and Rev53-67 (191) and Vpr61-75 (192), of which the Vpr derived peptides showed the greatest STI activity. Antibody based inhibitors were further explored, following the generation of a library of monoclonal antibodies against HIV-1 IN (193). The antibody mAb17 (194) and mAb33 (195,196) showed similar inhibition properties, while mAb32 was around 10 fold less effective (197). Several

oligonucleotide IN inhibitors have also been developed, with the most promising being those that mimic the viral U5 LTR and containing 6-oxocytosine substitutions (198,199). Unfortunately none of the compounds mentioned have progressed significantly through clinical trials, and most have been abandoned due to low efficacy, toxicity or the potential for side effects.

1.4.10 ARV Drug Resistance and STI's

Unfortunately as is the case for all previous ARVs, IN inhibitors are susceptible to the emergence of ARV drug resistance. HIV-1 uses an inherently low fidelity RT enzyme for viral replication, and thus the incidence of mutations within the viral genome is alarmingly high. For this reason, it is relatively easy for the virus to acquire mutations that confer resistance under the selective pressure of ARV treatment.

The mutations associated with resistance to IN inhibitors can be classified as primary or secondary. Primary resistance mutations occur within the CCD of the IN enzyme, and the three most commonly reported resistance pathways for Raltegravir are Q148H/R/K (with or without G140A/S) and N155H (with or without E92Q) and Y143C/R (without or without T97A) (200). All of these mutations are relatively near the active site of the protein. There is incomplete cross resistance with Elvitegravir, where for example Y143R/C/H mutations have no effect on Elvitegravir susceptibility but T66I shows high Elvitegravir resistance and has no effect on Raltegravir susceptibility (201). Secondary mutations include T66I/A/K, S153Y and R263K for Elvitegravir resistance, F121Y, E138A/K, G140A/S and S147G for both Elvitegravir and Raltegravir (although the latter has a greater

impact on Elvitegravir resistance), and Y143H for Raltegravir resistance (201). Secondary mutations do not always confer high levels of resistance on their own, although they contribute to resistance when they co-occur with the primary mutations.

The primary and secondary mutation combination of Q148H with G140S shows the greatest fold resistance (greater than 1000 fold against Raltegravir and Elvitegravir) (126). Due to the success of Raltegravir in the clinical setting and the inevitable emergence of drug resistant strains, there is intensified global interest in the identification of novel IN inhibitors, especially those that do not exhibit cross resistance with Raltegravir and Elvitegravir.

1.5 HIV-1 IN-DNA complexes

A fundamental problem facing all research into IN inhibitors is the ability to accurately model the native IN molecule within an *in vitro* system. As Alian *et al.* (2009) (81) proposed, the IN tetramer adopts a different conformation when bound to viral DNA. This conformational change may affect the properties of the drug interaction, and thus may negatively influence a drug's affinity for IN. Consequently, the screening assays used in this study may be identifying too many compounds that will only bind weakly in the native cellular system and thus fail out of cell based assays, and perhaps more importantly drug candidates that bind weakly to the non-native structure may be missed, but would be more effective in a cell based assay which they never reach.

As for the study of IN bound to its viral DNA, this has been greatly hampered by the proteins low inherent affinity for DNA. Furthermore, the low solubility of the protein requires that it be purified under very high salt conditions (1 M), which disrupts the IN-DNA interaction and makes it impossible to purify IN-DNA complexes. Recently work has emerged that explored the possibility of disulfide crosslinking of proteins with DNA, by incorporating cysteine mutations within the protein that interact with a thiolated-DNA. This technique was first explored by Huang *et al.* (202) when they generated disulfide crosslinked RT-DNA complexes, which were subsequently resolved by X-ray crystallography. Gao *et al.* further explored the technique (83) but only in 2009 was it used to properly crosslink IN with DNA for biochemical analysis (81). In this work, the authors generated an HIV-1 IN cysteine mutant capable of covalently crosslinking with thiolated-DNA.

The HIV-1 IN construct generated in their work contained four mutations previously shown to reduce surface hydrophobicity (69): C56S, W131D, F185D and C280S. Interestingly, C65S was added to reduce potential reactivity with thiolated-DNA in an unwanted manner outside of the active site. The authors then identified several clusters of residues known to interact with or be in close proximity to the viral DNA end within and near the active site. Their work revealed Y143C (a primary resistance mutation) as the most promising candidate due to the high proportion of complex formation compared with other mutations, strong strand transfer activity and its ability to complex with blunt ended DNA as well as preprocessed DNA. The ability to purify this IN-DNA complex allowed for strand transfer analysis in the context of a functional IN-DNA interaction to take place, as well as other

biochemical analyses on IN-DNA complexes. This IN-DNA complex forms stable tetramers, catalyzes the strand transfer step, and can bind strand transfer inhibitors.

1.6 Aims of the Study

While there is considerable effort internationally to identify novel integrase inhibitors, those efforts have been primarily focused on HIV-1 subtype B. Local research conducted into HIV-1 subtype C IN has been limited: research into this field falls considerably behind international efforts and numerous assays that are routinely practiced internationally have not yet been attempted in South Africa. Thus the development of assays to allow for the identification of novel integrase inhibitors is urgently required in South Africa.

Previous work conducted in our laboratory showed that expression and purification of wild type HIV-1 subtype C IN resulted in a non-functional, predominantly insoluble recombinant protein. It is suspected that the limited solubility of the protein is impacting on its activity, and thus mutations which increase the solubility of the wild type IN would simplify and enhance the purification process as well as potentially increase the activity.

Thus, the overall aim of this study was to generate and characterize a locally relevant, soluble, functional recombinant HIV-1 subtype C IN cysteine mutant (IN^{Y143C}) which can form disulfide bonds with thiolated-DNA for use in strand

transfer assays to allow for the identification of novel IN inhibitors. This will be achieved by the following objectives:

1. To generate HIV-1 subtype C IN mutant constructs containing five solubility mutations as well as the described cysteine mutation Y143C.
2. To express and purify the HIV-1 subtype C recombinant IN.
3. To characterize and compare the functional activity of HIV-1 subtype C recombinant IN to HIV-1 subtype B recombinant IN.
4. To assemble the IN-thiolated DNA complexes, and test their ability to catalyze the strand transfer step of integration.

CHAPTER TWO

2 METHODS

2.1 Reagents Used in this Study

In our laboratory, pET-15b expression vectors (Novagen, Germany) encoding the following three constructs were available: HIV-1 subtype B wild type IN (Bwt, pINDS.His, National Institutes of Health (NIH) AIDS Research and Reference Reagents Program catalogue number 2957, contributor Dr. Robert Craigie), HIV-1 subtype B soluble mutant IN (Bsol, pINSD.His.Sol, NIH AIDS Research and Reference Reagents Program catalogue number 2958, contributor Dr. Robert Craigie) and HIV-1 subtype C wild type (GBCwt, previously cloned in our laboratory), which were used to transform *Escherichia coli (E.coli)* HB101, DH5α and BL21 cells. The Bwt IN sequence is from the HIV-1 NL4-3 isolate (203). The Bsol is the same sequence as the Bwt with F185K and C280S mutations (203). The GBCwt sequence is derived from the HIV-1 subtype C 05ZAFV6 IN (204).

Protein for positive controls in SDS-PAGE and Western blot experiments was either HIV-1_{NL4-3} IN protein (F185H/C280S) (NIH AIDS Research and Reference Reagents Program catalogue number 9420, contributor Dr. Robert Craigie), or IN^{Cwt} protein (available in our laboratory).

The antibodies used in the western blot experiments were mouse anti-HIV-1 IN Monoclonal Antibody (8G4) (NIH AIDS Research and Reference Reagents Program catalogue number 7375, contributor Dr Robert Craigie) and HisProbe-Horseradish Peroxidase (HRP, Thermo Scientific, USA).

All oligonucleotides used in this study are listed in Table 2.1.

Table 2.1: Oligonucleotides used in the Strand Transfer Assay and Complexation experiments

Name	Sequence	Experimental use
Donor(biotin)	5′-Biotin- ACC CTT TTA GTC AGT GTG GAA AAT CTC TAG CA-3′	Strand Transfer
	5′-ACT GCT AGA GAT TTT CCA CAC TGA CTA AAA G-3′	
Target(FITC)	5′-TGA CCA AGG GCT AAT TCA CT- FITC-3′	Strand Transfer
	5′-AGT GAA TTA GCC CTT GGT CA-FITC-3′	
Target(DIG)	5′-TGA CCA AGG GCT AAT TCA CT-DIG-3′	Strand Transfer
	5′-AGT GAA TTA GCC CTT GGT CA-DIG-3′	
Thiolated-DNA	5′-Thio-6C-linker-ACT GCT AGA GAT TTT CCA CAC TGA CTA AAA-3′	Complexation
	5′-Biotin-ACT GCT AGA GAT TTT CCA CAC TGA CTA AAA G-3′	
Short Thiolated-DNA	5′-Thio-6C-linker-ACT GCT AGA GAT TTT CCA CA-3′	Complexation
	5′-Biotin-TGT GGA AAA TCT CTA GCA-3′	

Donor(biotin) and Target(FITC) were synthesized by Inqaba Biotechnical Industries, South Africa. Thiolated-DNA and Short Thiolated-DNA were synthesized by IDT, USA.

All oligonucleotides were ordered as single stranded DNA oligonucleotides (as shown in Table 2.1). The lyophilized oligonucleotides were resuspended in Duplex Buffer (Appendix B) to a concentration of 100 µM, and combined in equimolar amounts for annealing. Annealing was performed using an Eppendorf Mastercycler gradient PCR thermocycler with the following protocol: 95°C for 2 minutes, followed by 12 steps of 10 minutes each starting at 68°C and dropping by 5°C each step, and finally cooling to 4°C before storing the annealed oligonucleotides at –20°C until required.

The inhibitors used in the strand transfer assay included: L-Chicoric Acid (Sigma Aldrich, Germany) and Raltegravir (NIH AIDS Research and Reference Reagents Program catalogue number 11680, contributor Merck & Company, Inc.).

2.2 Generation of Recombinant IN Constructs

2.2.1 Design of GBC Mutant Constructs

The mutations (C56S, C65S, W131D, F185D, C280S) described in the literature (81) were incorporated into the 05ZAFV6 subtype C wild type IN sequence using Sequencher 4.8 software (Gene Codes, USA) to design the GBCsol construct. Furthermore, a second construct with these same mutations as well as the Y143C mutation was created, designated GBCY143C. The constructs were provided in lyophilized form, cloned into pET-15b in frame with the N-terminal His-Tag with a start codon inserted at position one to facilitate translation. Synthesis was done by GeneArt (Germany) and the nucleotide sequences were codon optimized for bacterial expression.

2.2.2 IN Protein Modeling

The co-ordinates for a model of an IN monomer (INBwt) interacting with the viral DNA end were generously obtained from Miss T. Traut (Mintek, South Africa). The IN model was manipulated using the PyMOL Molecular Graphics System (Version 1.3, Schrödinger, LLC). The amino acid differences between HIV-1 subtype B and subtype C IN were incoporated into the model and highlighted, and the solubility mutations of the various mutant constructs were also incorporated and highlighted. The Y143C mutation was incoporated into the GBCsol construct, and the distance

between the terminal 5' adenine of the viral DNA and the C143 residue was measured.

2.2.3 Transformation of E.coli *BL21 Cells with HIV-1 IN constructs*

The Bsol mutant was available in *E.coli* HB101 cells, so a plasmid preparation (see 2.2.3.1) was performed to transform chemically competent *E.coli* BL21 cells (see 2.2.3.3) for expression. The lyophilized HIV-1 subtype C mutant constructs GBCsol and GBCY143C were resuspended and chemically competent *E.coli* BL21 cells were immediately transformed with this construct upon reconstitution (see 2.2.3.3). The Bwt and GBCwt constructs were both already available in *E.coli* BL21 cells.

2.2.3.1 Plasmid Preparation

The Bsol expression plasmid was obtained from the NIH AIDS Research and Reference Reagents Program as *E.coli* HB101 stocks. Recombinant plasmid was extracted from these cells in order to transform *E.coli* BL21 cells for purification purposes, as follows (205): overnight cultures were made by inoculating 10 µl of bacterial stock into 10 ml of Luria Bertani (LB) medium containing 1X Ampicillin (concentration 100 µg/ml; Appendix C) and incubating for 16 hours in a 37°C shaking incubator. Glycerol stocks of these cultures were made for long term storage as described in section 2.2.3.4. Recombinant plasmid was extracted from the remaining culture using a GenElute Plasmid Miniprep Kit (Sigma Aldrich, Germany) as per the manufacturer's instructions. Briefly, two 1 ml aliquots of each bacterial cell stock were centrifuged at 12 000 xg for 1 minute and the supernatant

discarded. The cells were lysed, and the alkaline pH of the lysis buffer ensured the plasmid remained in solution while a neutralization step precipitated the genomic DNA and other cellular debris (which was removed by centrifugation). The plasmid DNA in solution was purified using the supplied column and eluted with distilled water following washing. The corresponding samples were combined and quantified using a NanoDrop ND1000 spectrophotometer (Thermo Fisher Scientific, USA).

2.2.3.2 Generation of Chemically Competent BL21 Bacterial Cells

Chemically competent *E.coli* BL21 and *E.coli* DH5α cells were prepared as follows (206); an overnight culture of each was prepared by inoculating 10 µl of *E.coli* BL21 or *E.coli* DH5α glycerol stock (available in our laboratory) in 10 ml LB medium and incubating for 16 hours in a 37°C shaking incubator. A 1 in 40 dilution of this overnight culture was made in LB medium and incubated in a 37°C shaking incubator until the optical density (OD, measured at 600nm) was approximately 0.25 (measured on a Biowave S2100 Diode Array spectrophotometer, Biochrom, United Kingdom). Once they were at the correct OD, the cultures were centrifuged at 1500 x*g* for 10 minutes at 4°C. The supernatants were discarded and the cells were re-suspended in 10 ml of Transformation Buffer each (Appendix C) and incubated on ice for 30 minutes. The cells were then centrifuged at 1000 x*g* for 10 minutes at 4°C. The supernatants were discarded and the cells were re-suspended in 1 ml of Transformation Buffer, from which 100 µl aliquots of each were prepared and subsequently frozen at -80°C, until use.

2.2.3.3 Transformation of Chemically Competent E.coli BL21

Chemically competent *E.coli* DH5α and *E.coli* BL21 cells were transformed with each plasmid by the addition of 5 µl plasmid (using 5 µl distilled water as a negative control and 5 µl native pET-15b vector as a positive control) to a 100 µl aliquot of chemically competent *E.coli* cells of each strain. These were mixed gently and incubated on ice for 30 minutes. The cells were heat shocked at 42°C for 90 seconds and spread plated onto LB agar plates containing 1X Ampicillin (and additionally 1X Chloramphenicol for BL21 cells, Appendix C). The plates were incubated at 37°C for approximately 16 hours. A single isolated colony of transformed *E.coli* BL21 or *E.coli* DH5α cells was removed using a sterile inoculating loop and subsequently used to inoculate 50 ml of LB medium (containing both 1X Ampicillin and 1X Chloramphenicol for the former, and only 1X Ampicillin for the latter).

2.2.3.4 Generation of recombinant E.coli bacterial cell stocks

Stocks of all recombinant *E.coli* cells were made as follows: 800 µl of bacterial culture and 200 µl of sterile glycerol (Saarchem, South Africa) were mixed thoroughly and frozen slowly at –80°C, until used.

2.3 IN Protein Expression

2.3.1 Small Scale Induction and Optimization of HIV-1 IN Protein Expression

An overnight culture for each of the five *E.coli* BL21 cell stocks was made by inoculating 10 ml of 1X Ampicillin 1X Chloramphenicol LB medium with 1 ml of

each particular bacterial cell stock, and incubating in a 37°C shaking incubator for approximately 16 hours. Four 1 in 40 dilutions of each overnight culture were made, each in 20 ml of LB medium (containing 1X Ampicillin 1X Chloramphenicol), and incubated in a 37°C shaking incubator and the OD (measured at 600nm) was monitored (measured on a Biowave S2100 Diode Array Spectrophotometer, Biochrom, United Kingdom). Isopropyl β-D-1-thiogalactopyranoside (IPTG, Calbiochem, Germany) was added over a range of optical densities and the cultures were further incubated in a 37°C shaking incubator. Expression was optimized for various optical densities (0.4, 0.6 and 0.8), various IPTG concentrations (0.4 mM, 0.7 mM, 1.0 mM, 1.25 mM, 1.50 mM, 1.75 mM and 2.0 mM), various harvest times (3 hours, 6 hours and 24 hours post induction) and various temperatures (20°C and 37°C). Samples of each small scale induction were taken for Sodium dodecyl sulfate (SDS)-Polyacrylamide Gel Electrophoresis (PAGE) and western blot analysis. Uninduced cell cultures grown at the equivalent temperatures and harvested at the corresponding time points were used as a negative control.

2.3.2 SDS-PAGE

Samples were prepared for SDS-PAGE by combining a 1:1 equal volume of induced and uninduced bacterial culture and 2X Loading Buffer (Appendix C) and heating to 70°C for 15 minutes. Twenty microliters of prepared sample was loaded per well, with the exception of the molecular weight marker and the positive control, where 10 µl were loaded. The molecular weight marker used in all SDS-PAGE experiments was the Prestained Protein Marker, Broad Range (7-175 kDa)

(New England Biolabs, USA). The positive control was the HIV-1 NL4-3 IN Protein (F185H/C280S).

SDS-PAGE was carried out with a 1.5 mm 12.5% acrylamide resolving mini-gel and a 4% acrylamide stacking mini-gel (see Appendix C) according to the discontinuous gel method described by Laemmli (207). The Mini-PROTEAN 3 Cell system (Bio-Rad, USA) was used to cast and run the gels. Samples were resolved at a constant current of 10 milliamps (mA) per gel (using an EPS301 Electrophoresis Power Supply from Amersham Biosciences, United Kingdom) until the samples had entered the resolving gel, whereupon the current was adjusted to 20 mA per gel. Gels were resolved in duplicate, where one was stained with Coomassie Brilliant Blue staining solution (Appendix C) overnight followed by soaking in destaining solution 1(Appendix C) for 30 minutes and destaining solution 2 (Appendix C) until destained (208). The Coomassie images were captured using a GelDoc (Bio-Rad, USA). The duplicate gel was used for western blotting.

2.3.3 Western Blot Analysis

Proteins on the unstained SDS-PAGE gels were transferred using a Trans-Blot SD Semi-Dry Transfer Cell (Bio-Rad, USA) according to the manufacturer's instructions. Briefly, gels were equilibrated in Transfer Buffer (Appendix C), and then transferred onto Hybond-C Extra nitrocellulose membrane (Amersham Biosciences, United Kingdom) at 40 mA for 50 minutes. Membranes were then blocked with 10 mg/ml Bovine Serum Albumin (BSA) Fraction V (Roche Diagnostics, Germany) in Tween-TBS (Appendix C) overnight at 4°C.

A 1:10000 dilution of the HisProbe-HRP (Thermo Scientific, USA) was made in Tween-Tris Buffered Saline (T-TBS), and 50 µl/cm^2 applied to the membrane. The antibody was incubated statically on the membrane for 60 minutes at room temperature. This was followed by five 7-minute washes with approximately 150 ml of T-TBS each, shaking. SuperSignal West Pico Chemiluminescent Substrate (Pierce Protein Research Products, Thermo Scientific, USA) was used as per manufacturer's instructions: equal parts of the peroxide solution and luminol/enhancer solution were combined and 50 µl/cm^2 was added to the membrane and incubated statically for 15 minutes at room temperature. Western blot images were captured using the GelDoc (Bio-Rad, USA) imaging system, using the Live Acquire mode (60 exposures of 30 seconds each) in the QuantityOne (v4.6.1 build 055) software (Bio-Rad, USA).

2.3.4 Large Scale Expression

Large scale expression of each of the IN recombinant proteins was undertaken using the conditions optimized in section 2.3.1. An overnight culture for each of the five *E.coli* BL21 cell stocks was made by inoculating 10 ml of LB medium (containing 1X Ampicillin 1X Chloramphenicol) with 1 ml glycerol stock of each particular bacterial cell stock and incubating in a 37°C shaking incubator for approximately 16 hours. A 1 in 10 dilution of each culture was made to 1000 ml in the same medium and incubated in a shaking incubator at 37°C until the OD reached approximately 0.8 (measured on a Biowave S2100 Diode Array Spectrophotometer, Biochrom, United Kingdom). IPTG was added to each culture to a final optimized concentration of 1.0 mM and the cultures were incubated for a

further 3 hours. Samples of the culture were resolved on SDS-PAGE in order to confirm expression of the protein.

2.3.5 HIV-1 IN Protein Purification Protocol

Upon confirmation of IN expression from the large scale cultures, the cells were harvested by centrifugation at 2744 xg for 30 minutes (using a JA-10 rotor in a Beckman Optima centrifuge), and re-suspended in 10 ml Re-suspension Buffer (Appendix C). The cells were freeze-thawed twice, and homogenized using a Bandelin Sonopuls Ultrasonic Homogenizer HD3100 at 80% amplitude for 6 cycles of 1 minute (1.0 second pulse, 0.5 second pause) each. The homogenized suspensions were centrifuged at 40000 x*g* for 30 minutes at 4°C. The clarified supernatant was collected for purification of the expressed recombinant IN protein.

The proteins were purified using an ÄKTAprime Plus purification system (GE Healthcare Life Sciences, United Kingdom). Purification conditions were optimized individually for the constructs expressed in 1 l of culture (GBCwt, Bwt and Bsol) and those expressed in 9 l of culture (GBCsol and GBCY143C). Only the wash durations were altered for the different starting culture volumes. Briefly, the supernatant was loaded onto a 5 ml HisTrap HP column (GE Healthcare Life Sciences, United Kingdom) at a flow rate of 0.5 ml/min with a maximum pressure of 0.3 mega pascals (Mpa) using the ÄKTAprime Plus. The column was subsequently washed with either 6 or 30 column volumes (approximately 30 ml or 150 ml) of Binding Buffer (Appendix C) respectively. A step gradient of Binding Buffer (5 mM imidazole, see Appendix C) and Elution Buffer (600 mM imidazole, see Appendix C) was subsequently employed to wash the column with either 6 or

30 column volumes each at imidazole concentrations of 60 mM, 100 mM, and 125 mM. All washes were performed at a flow rate of 5 ml/min, with a maximum pressure of 0.3 Mpa. The protein was eluted using an imidazole gradient that ranged from 125 mM (the final wash) to 600 mM (Elution Buffer, Appendix C) over 45 ml, collecting 3 ml fractions (all at 3 ml/min, 0.3 MPa). UNICORN v5.20 Software (Build 500) (GE Healthcare Life Sciences, United Kingdom) was used to monitor the ultraviolet (UV) absorbance at 280 nm of the elution in order to identify which fractions contained the protein. The HisTrap column was washed as per manufacturer's instructions between each use.

2.3.6 Analysis of Eluted Protein Fractions

Fractions that contained the recombinant IN protein (as determined by UV absorbance monitoring during the purification process) were analyzed by SDS-PAGE and western blot analysis in order to assess the purity of the fractions. All fractions that did not contain aberrant bands were pooled and concentrated.

2.3.7 Recombinant HIV-1 IN Protein Concentration

The pooled fractions for each recombinant IN protein were concentrated using an Amicon Ultra-15 Centrifugal Device (Millipore, USA). The filter was prepared by adding 15 ml of Phosphate Buffered Saline (PBS, Sigma Aldrich, Germany) and centrifuging at 3000 x*g* for approximately 5 minutes (until the PBS had run through the filter). Fifteen ml of the pooled fractions was added to the filter and centrifuged at 3000 x*g* for approximately 10 minutes, whereupon the flow through was discarded and the process repeated until all the pooled fractions had been added.

In order to exchange the buffer, five 15 ml changes with Storage Buffer (Appendix C) were performed (also at 3000 x*g*, times varied).

The concentrated protein was quantified using a NanoDrop ND1000 Spectrophotometer (Thermo Fisher Scientific, USA). The molar extinction coefficients and molecular weights were calculated using the Peptide Calculator (v1.00) at the Center for Biotechnology, Northwestern University (www.basic.northwestern.edu/biotools/proteincalc.html).

The percentage purity of each recombinant IN was determined by densitometry analysis of SDS-PAGE experiments stained with Coomassie Brilliant Blue. IN samples were diluted at a range from 100% to 65%, and combined at a 1:1 ratio with Loading Buffer (Appendix C). Samples were resolved on Mini-PROTEAN TGX Precast gels (4-15%, 10 wells, 30 µL, Bio-Rad, USA) as in section 2.3.2. The gels were stained in Coomassie Brilliant Blue Staining Solution (Appendix C) overnight, and destaining was standardized across the experiment as follows: 30 minutes in Destaining Solution 1 (Appendix C) and four and a half hours in Destaining Solution 2 (Appendix C).

Images were captured using a MiniBis Pro Gel Imager (DNR Bio-Imaging Systems, Israel) and GelCapture Acquisition Software (Version 5.8, DNR Bio-Imaging Systems, Israel). Densitometry analysis was performed using GelQuant Pro (Version 4.1, DNR Bio-Imaging Systems, Israel) as follows: Lanes and bands were automatically detected using software presets, and the densitometry data for

each lane was totaled. The percentage purity of the IN sample was expressed as the densitometry value for the IN band as a percentage of the entire lane:

$$\% \text{ Purity} = (\text{IN density})/(\text{sum of all densities}) \text{ X } 100$$

The specific concentrations (in mg/ml and µM) were calculated for each protein using the following formula:

$$\text{Specific concentration} = \% \text{ Purity X Concentration}$$

2.4 Strand Transfer Assay

2.4.1 Functional Strand Transfer Assay

Two variations of the standard strand transfer assay were conducted using two different labeled target DNAs, and the differences are indicated where appropriate. The assay was performed as follows: Donor(biotin) DNA (table 2.1) was added to each well of a streptavidin coated microplate (R&D Systems, USA) at a concentration of 0.14 µM (100 µl per well) and incubated while shaking at 150 rpm at 22°C for 1 hour. The wells were aspirated and washed briefly three times with 350 µl PBS. Each recombinant IN protein to be tested was diluted to a concentration of 1 µM using Reaction Buffer (Appendix B) and 100 µl was added per well and incubated while shaking at 150 rpm at 22°C for 30 minutes (using reaction buffer only as a no-protein negative control). The protein was aspirated and the wells were washed twice with 200 µl Reaction Buffer, while shaking at 150 rpm at 22°C for 5 minutes each. Ninety microliters of Reaction Buffer was added to

each well, and Target(FITC) DNA or Target(DIG) DNA (10 µl at 2.5 µM diluted with Reaction Buffer, Table 2.1) was added to bring the final volume to 100 µl and incubated at 37°C for 1 hour. The wells were aspirated and washed three times with 300 µl 2X SSC buffer (Appendix B) incubated while shaking at 150 rpm at 22°C for 10 minutes each. The monoclonal anti-FITC Alkaline Phosphatase (AP) antibody produced in mouse (Sigma Aldrich, Germany) or the monoclonal anti-DIG (poly) fab fragment antibody (Roche Diagnostics, Germany) was diluted 1 in 10000 in Stabilzyme AP (Surmodics, USA) and 200 µl added to each well and incubated on the bench for 1 hour and then overnight at 4°C. Following incubation the antibody solution was aspirated and the wells were washed three times with 300 µl TBS (Appendix C) incubated while shaking at 150 rpm at 22°C for 10 minutes each. BluePhos Microwell Phosphatase substrate (KPL, USA) or 1-Step Ultra TMB-ELISA (Thermo Scientific, USA) substrate was prepared as per the manufacturer's instructions, and 100 µl was added per well and incubated while shaking at 150 rpm at 37°C for 1 hour. Absorption was read at 620 nm (for BluePhos) and 450 nm (for TMB) using an xMark microplate spectrophotometer (Bio-Rad, USA) and Microplate Manager 6.0 Software (Build 2.2.7, Bio-Rad 2008).

All experiments were carried out in triplicate in order to verify the assay and confirm functionality of the proteins. In each experiment, the average absorbance of the no-protein control was subtracted from the experimental data to normalize against the background signal. Each experimental value was then calculated as a percentage of the positive control (IN^{Bsol} available in our laboratory, confirmed to be active) using the following formula:

$$\% \text{ Activity} = (\text{Experiment}/\text{IN}^{\text{Bsol}}) \times 100$$

2.4.2 Inhibition Strand Transfer Assay

The strand transfer assay was performed as described in section 2.4.1 except following the removal of IN and washing with Reaction Buffer, the drug compounds were added in 90 µl Reaction Buffer (final drug concentration was calculated using the 100 µl final volume) and incubated while shaking at 150 rpm at 37°C for 30 minutes. Following this incubation, target DNA was added as before and the assay was continued as normal.

Single dose responses to 10 µM L-chicoric acid diluted with Dimethyl sulfoxide (Sigma Aldrich, Germany) (in reaction buffer, Appendix B) were carried out in biological duplicate and experimental triplicate. After normalizing the results against the background signal as before, the average activity for each protein was compared with the average activity for the drug-free control to calculate a percentage inhibition as per the following formula:

$$\% \text{ Inhibition} = [1-(\text{drug}/\text{control})] \times 100$$

The IC_{50}, defined as the concentration at which activity is inhibited to 50% of the no drug control, was determined using a drug concentration range from 0.1 µM to 25 µM for L-Chicoric acid and from 0.1 nM to 10 µM (and 50 µM for IN^{Y143C}) for Raltegravir, comparing each experimental value to the drug free control as before and averaging the biological duplicate. Dose ranges were selected to include a range of concentrations on either side of the expected IC_{50}. A dose response

curve was constructed (as described in section 2.6) from the average of the experimental duplicate data, and an IC_{50} value was extrapolated from the curve.

2.5 HIV-1 IN-DNA Complexes

The thiolated-DNA to be used in complexation studies was annealed as described in section 2.1. The reactivity of the thiol groups on the terminal adenine of the DNA was increased by modification using reducing agents. The disulfide form was present following synthesis, and was reduced to the more reactive sulfhydryl form by treatment with Dithiothreitol (DTT). The thiolated-DNA (in Duplex Buffer, Appendix B) was desalted using an Illustra ProbeQuant G-50 Micro Column (GE Healthcare, United Kingdom), as per the manufacturer's instructions. This column contains G-50 Sephadex DNA grade F gel, which allows for the purification of large biomolecules (in the case of this product, larger than 20 bases) and buffer exchange into the buffer with which the column is equilibrated. Briefly, the column was equilibrated by passing three exchanges of 250 µl each of Activation Buffer 1 through the column using a Hettich Mikro 220 Centrifuge (Hettich Industries, USA) at 590 xg for 1 minute each. Next, 50 µl of thiolated-DNA was added to the column, followed by a centrifugation at 590 xg for 2 minutes to elute the thiolated-DNA in Activation Buffer 1 (Appendix B). This solution was then incubated at 37°C for 16 hours.

The sulfhydryl form of the thiol group on the DNA was further reduced by 5,5'-dithiobis-(2-nitrobenzoic acid) (DTNB, Sigma Aldrich, Germany) to yield the most reactive DTNB-modified form. The thiolated-DNA and Activation Buffer 1

(Appendix B) was desalted using the same procedure as before, substituting Activation Buffer 2. This solution was incubated at room temperature for 1 hour. Finally the thiolated-DNA was desalted again, using Complexation Buffer (Appendix B) to equilibrate the column. The eluted DTNB-modified thiolated-DNA was used immediately.

The complexation reaction was carried out as follows: approximately 5 µg IN^{Y143C} (in 10 µl Storage Buffer, Appendix C) was added with 4 µl freshly prepared DTNB-modified thiolated-DNA (50 µM stock) and 23 µl Complexation Buffer (Appendix B), and 4 µl 100mM $MnCl_2$ stock (added last). The reaction was incubated for 1 hour at room temperature. The entire reaction was combined with an equal volume of Loading Buffer (containing no β-mercaptoethanol, Appendix C) and heated to 70°C for 15 minutes before analysis by SDS-PAGE and Western Blotting.

2.6 Data Handling and Statistical Analyses

All strand transfer assay results were captured using Microsoft Excel for Mac 2011 (Version 14.0.0, 100825). Average values and standard deviations for all biological replicates were determined, and percentage activity or percentage inhibition for each protein was calculated and compared between experiments. A two-tailed two-sample unequal variance (heteroscedastic) students t-test was carried out between all possible pairs of proteins to determine if there were significantly different results for the drug free, single dose and dose response assays for all drugs tested. A significance level of 95% was used, indicating that a *p* value of

less than 0.05 was significant. IC_{50} values were calculated using Prism for Macintosh (GraphPad Software, Inc., version 4.0b).

CHAPTER THREE

3 RESULTS

3.1 Design of HIV-1 Subtype C IN Mutant Constructs

An amino acid sequence alignment showing the introduced mutations in the HIV-1 subtype C mutant constructs (as well as the HIV-1 subtype B constructs from the NIH and GBCwt) is shown in Figure 3.1. The in frame His-Tag is not shown in this alignment. The initial methionine residue was added to the constructs to allow expression because the IN protein is natively expressed as a polyprotein, and is not shown in the figure. The Bwt and Bsol constructs differ from each other only in positions 185 and 280 (the solubility mutations), and differ from the GBCwt sequence in 20 other positions (~6.9% amino acid difference). Of these differences, more than half are significant changes in amino acid properties (such as a change of charge or hydrophobicity) and only one represents a significant change in the size of the amino acid side chain (G284R).

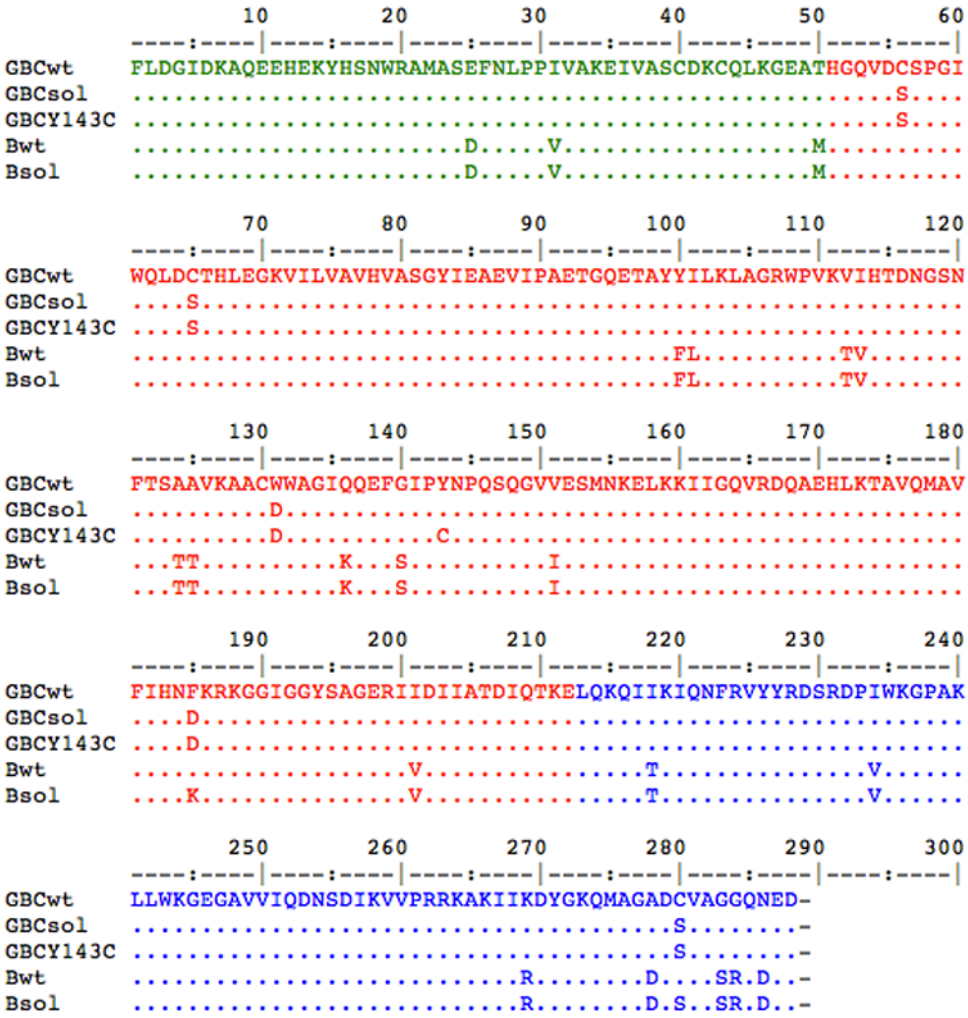

Figure 3.1: Alignment of integrase amino acid sequences used in this study. The amino terminal domain is represented in green, the carboxy terminal domain is represented in blue and the catalytic core domain is represented in red. Dots indicate the same amino acid as the GBCwt sequence, variations are shown.

Modeling of IN^{Bsol} with DNA bound shows the overall structure similar to that of other models (81,209) (Figure 3.2). Residue 185, the site of the solubility mutations F185K/D, is visible on the surface of the CCD. The CTD is attenuated by 18 amino acids as these failed to resolve in the original crystallography experiments and thus are not included in the model. Consequently, the C280S mutation is not visible in this model.

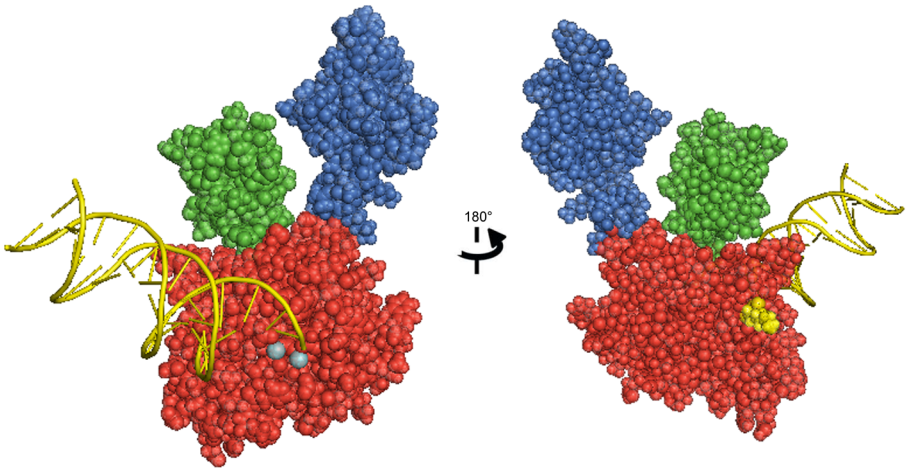

Figure 3.2: Front (left) and back (right) view of DNA binding by IN^{Bsol} monomer (Miss T. Traut, personal communication). The carboxy terminal domain is represented in green, the amino terminal domain is represented in blue and the catalytic core domain is represented in red. Grey spheres in front view (left) are magnesium ions co-ordinated by the DDE motif in the active site of the catalytic core domain. The viral DNA ends (3′ processed) are shown in yellow. The yellow motif visible in the back view (right) is residue 185.

The natural variations between the IN^{Cwt} and IN^{Bwt} amino acid sequences are highlighted in yellow in Figure 3.3. The variations present in the terminal region of the CTD are not shown as this region is not present on the model. Several of the mutations are buried within the tertiary structure of the protein and so are not visible in this figure.

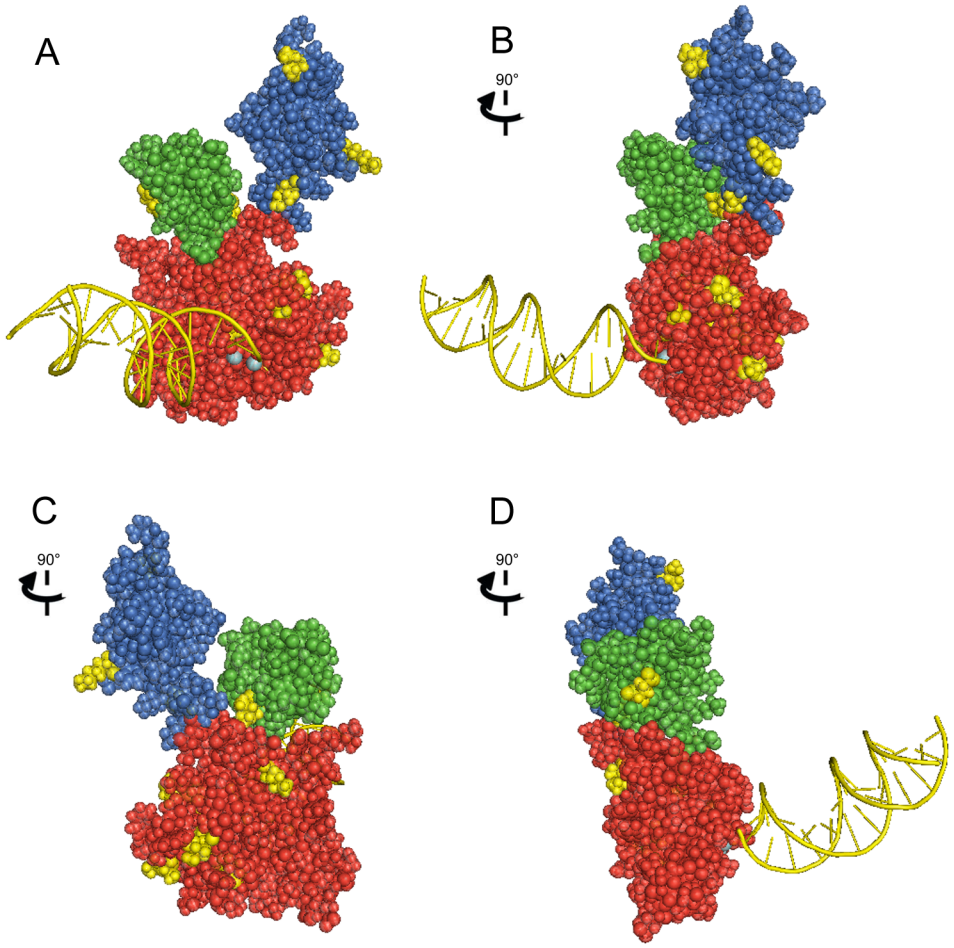

Figure 3.3: DNA binding by IN^{Cwt} shown from four different angles (90 degree left rotation). Domains and features are coloured as in Figure 3.2. Amino acid differences between IN^{Bwt} and IN^{Cwt} are shown in yellow.

The active site of IN^{Y143C} is shown in Figure 3.4. The residues shown in white are the solubility mutations specific to IN^{Csol} and IN^{Y143C}. A distance of approximately 7.1 Å between the terminal 5' adenine of the viral DNA (with attached thiol group) and the C143 residue was calculated. The C65S mutation (white) is near the active site of the protein, in close proximity to the 5' end of the viral DNA. The solubility mutations C56S and W131D are visible in the figure, while F185D is on the other side of the protein (not visible in the figure), while residue 280 did not resolve in this structure.

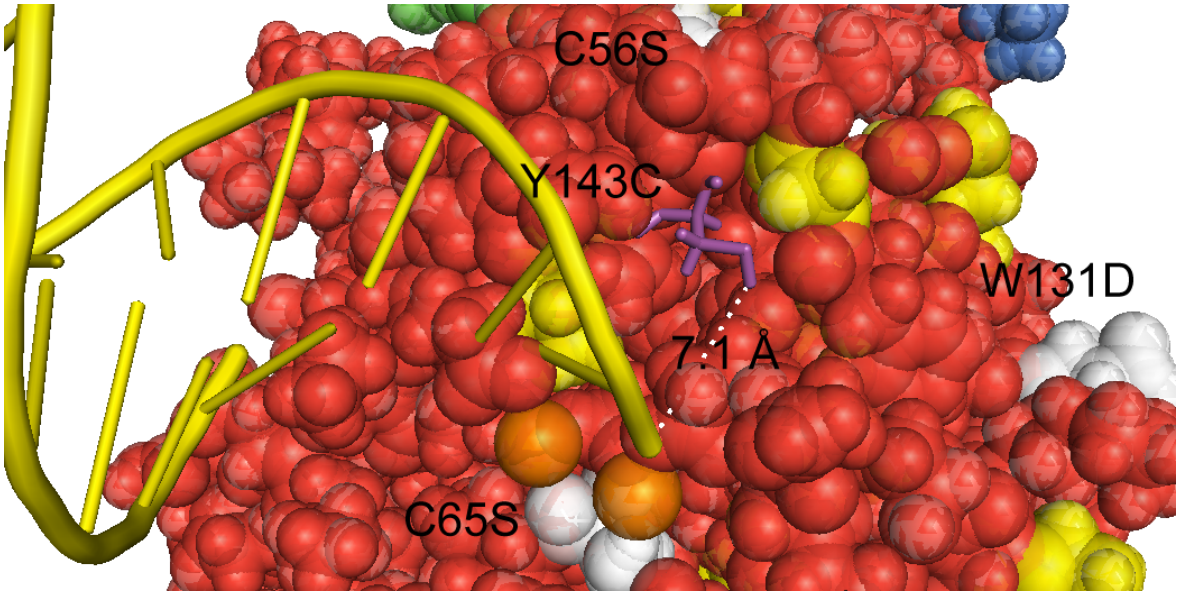

Figure 3.4: The active site of HIV-1 subtype C IN^{Y143C}. Domains and features are coloured as in Figure 3.2, except the magnesium II ions are represented as orange spheres. Solubility mutations specific to the subtype C IN proteins are shown in white. The Cysteine at position 143 is shown as a purple stick model. The thiol group (not depicted) is attached to the 5′ adenine present in the active site, and the distance to C143 is indicated by a dashed line (~7.1Å). Co-ordinates obtained from Miss T. Traut.

A plasmid preparation of the Bsol construct was obtained, and successfully transformed into *E.coli* BL21 cells for expression studies. The resuspended GBCsol and GBCY143C constructs were also transformed into *E.coli* BL21 cells (results not shown). Both Bwt and GBCwt were already available in *E.coli* BL21 cells.

3.2 Recombinant IN Expression

3.2.1 Small Scale Recombinant IN Expression

Optimal expression conditions were determined for all five recombinant IN proteins in *E.coli* BL21 cells. Both IN^{Bwt} and IN^{Cwt} expressed well at 1 mM IPTG induced at an OD of 0.8. SDS-PAGE and western blot experiments of IN^{Cwt} (Figure 3.5) showed a band at 0.7 mM IPTG at the correct molecular weight (~34 kDa), but detection was not possible by Western blotting. At 1.0 mM IPTG, a band at the correct molecular weight was observed by SDS-PAGE and Western blotting, confirming the expression of the desired protein. The differences in molecular weights seen between the positive control IN and the recombinant IN proteins expressed in this study are attributed to the in-frame expression with the His-Tag (2.18 kDa).

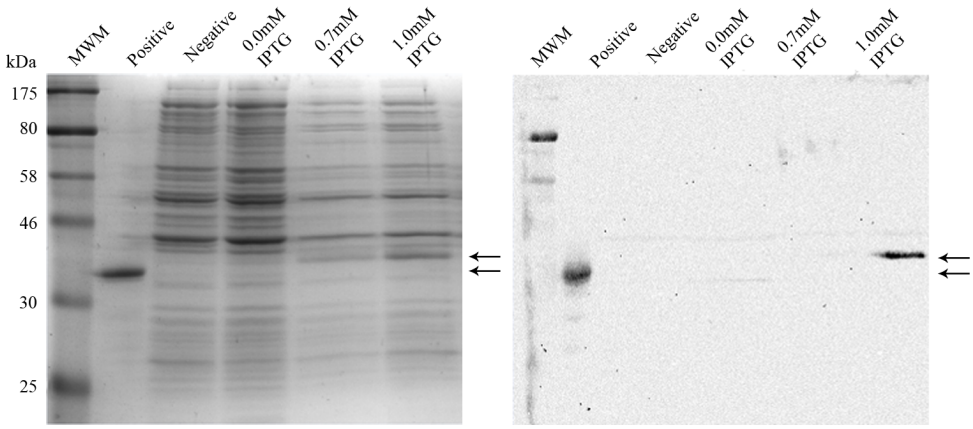

Figure 3.5: Coomassie stained SDS-PAGE gel (left) and corresponding Western blot (right) analysis of small scale expression of IN^{Cwt}. IPTG concentrations used are listed at the top of the lanes. Molecular weight marker sizes are indicated on the left. Arrows indicate the positive control ($HIV\text{-}1_{NL4\text{-}3}$ IN protein, bottom arrow) and His-tagged IN^{Cwt} (top arrow). The negative control lane contains *E.coli* BL21 not transformed with the expression vector. The mouse anti-HIV-1 IN Monoclonal Antibody (8G4) was used in the Western blot experiment.

Representative SDS-PAGE of IN^{Bwt} and IN^{Bsol} expression are shown in Figures 3.6 and 3.7 respectively. IN^{Bwt} and IN^{Bsol} were expressed well at 1.0 mM IPTG concentrations and were of the expected size (~34 kDa).

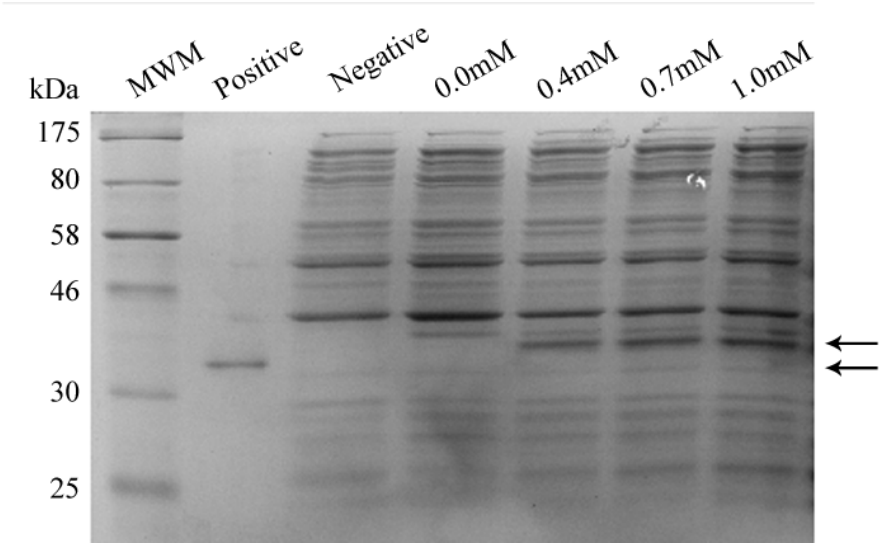

Figure 3.6: Coomassie stained SDS-PAGE analysis of a representative small scale induction of IN^{Bwt}. The IPTG concentrations used are listed at the top of each lane. Molecular weight markers are indicated on the left. Arrows indicate the positive control IN protein (HIV-1_{NL4-3} IN protein, bottom arrow) and the His-tagged IN^{Bwt} (top arrow). Negative control lane contains *E.coli* BL21 not transformed with an expression vector. IN bands were observed at the correct molecular weight (~34kDa) at 0.4mM, 0.7mM and 1.0mM IPTG concentrations.

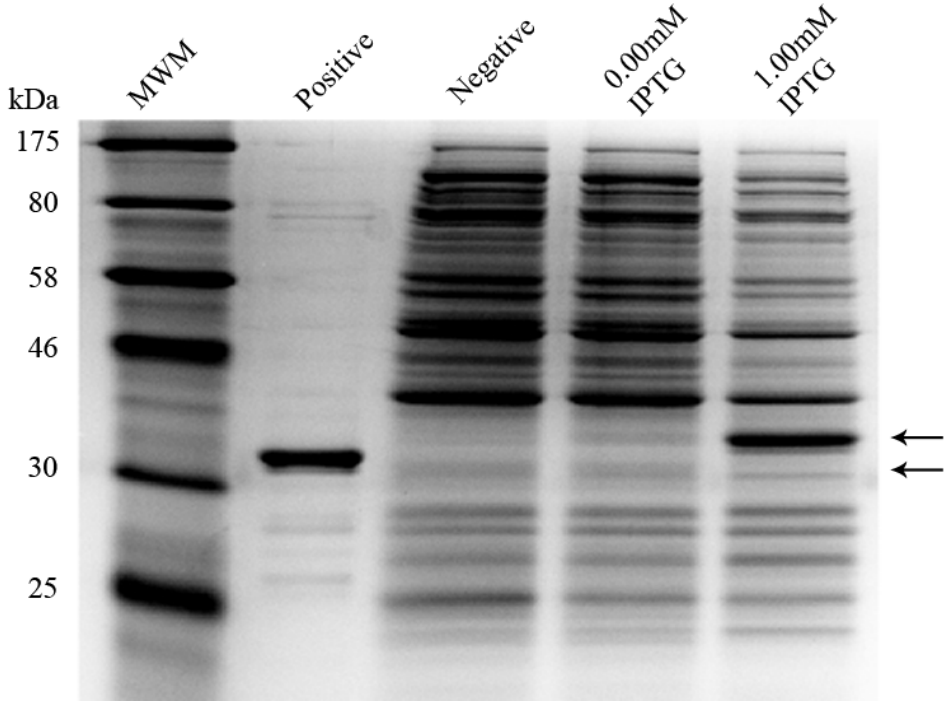

Figure 3.7: Coomassie stained SDS-PAGE gel of a representative small scale induction of IN^{Bsol}. The IPTG concentrations used are listed at the top of each lane. Molecular weight markers are indicated on the left. Arrows indicate the positive control IN protein (HIV-1_{NL4-3} IN protein, bottom arrow) and the His-tagged IN^{Bsol} (top arrow). The negative control lane contains *E.coli* BL21 not transformed with an expression vector. At 1.0mM IPTG an IN band was observed at the correct molecular weight (~34kDa).

IN bands were not observed on Coomassie stained SDS-PAGE in either raw culture, concentrated cells or supernatant (Figure 3.8) for both IN^{Csol} and IN^{Y143C}, however IN expression was confirmed by western blot (Figure 3.9). After extensive optimization efforts, no appreciable difference in expression levels at the various IPTG concentrations (Figure 3.9) or temperatures (results not shown) tested were noted for these two constructs, and thus expression parameters of 1.0 mM IPTG induced at an OD of 0.8 and incubated at 37°C were used, with a corresponding scale-up in culture volumes to account for low expression levels of IN^{Csol} and IN^{Y143C}.

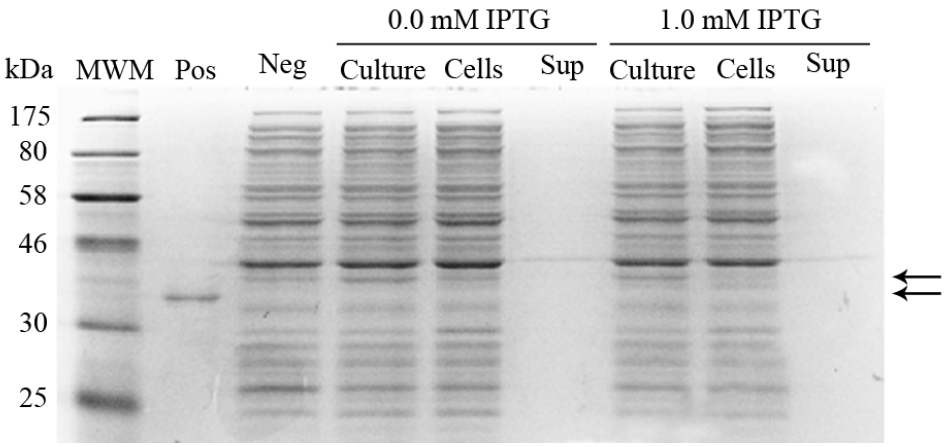

Figure 3.8: Coomassie stained SDS-PAGE analysis showing culture, concentrated cells and supernatant from a representative small scale induction of IN^{Csol}. The IPTG concentrations used are shown at the top of each lane. Molecular weight markers are indicated on the left. Arrows indicate the positive control IN protein (HIV-1_{NL4-3} IN protein, bottom arrow) and expected position (~34kDa) of His-tagged IN^{Csol} (top arrow). The negative control lane contains *E.coli* BL21 not transformed with an expression vector. No IN bands were visible by Coomassie staining in any fraction. Similar results were achieved for IN^{Y143C}.

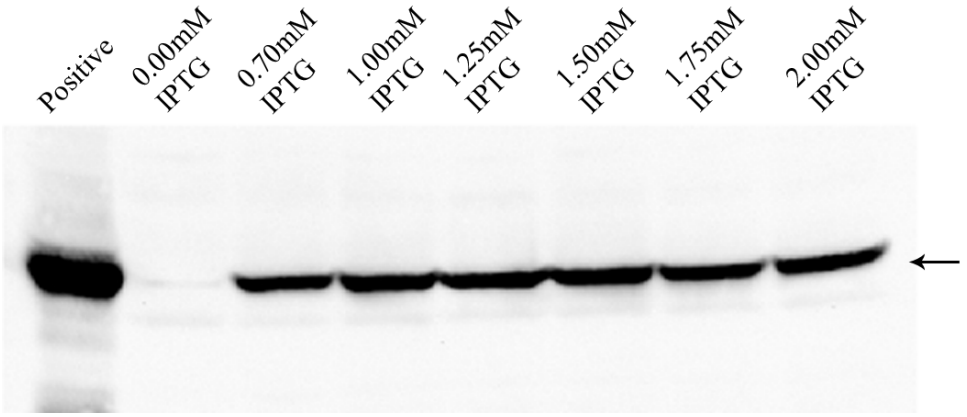

Figure 3.9: Western blot analysis of a representative small scale induction of IN^{Csol} over a range of IPTG concentrations (indicated at the top of each lane). The positive control is the previously expressed IN^{Cwt} protein. The arrow indicates the position of the IN bands. The antibody used for detection was the HisProbe-HRP antibody, allowing for the detection of the IN^{Cwt} positive control. Similar results were achieved for IN^{Y143C}.

Overall, optimal recombinant IN protein expression conditions of all 5 constructs in *E.coli* BL21 cells included induction of protein expression with 1 mM IPTG when the cell growth had reached an OD of 0.8, followed by three hours incubation at 37°C with shaking. These conditions were subsequently used for all large scale recombinant IN protein expression.

3.2.2 Large Scale Recombinant IN Expression

Expression of the IN recombinant proteins was upscaled, and the IN^{Cwt}, IN^{Bwt} and IN^{Bsol} recombinant proteins were routinely expressed in 1 l cultures, while IN^{Csol} and IN^{Y143C} recombinant proteins were expressed in 9 l cultures to obtain enough bacterial cells for the subsequent purification protocols. Large scale recombinant protein expression was repeated as required.

3.2.3 Recombinant IN Purification

For IN^{Bwt}, IN^{Bsol} and IN^{Cwt}, one litre of culture was processed as described in section 2.3.5, whereas for IN^{Csol} and IN^{Y143C} 9 l of culture was processed in the same manner to establish a purification protocol. After several attempts, an optimal purification protocol for the soluble IN fraction using a Nickel II charged column was established (refer to section 2.3.5 for details). Wash volumes were adjusted on a purification to purification basis, by monitoring the UV absorbance levels of the eluent coming off the column. From previous studies in our laboratory it was known that the IN protein elutes at imidazole concentrations of around 200 mM and higher, so incremental washes were carried out at lower imidazole concentrations. During each wash, an early absorbance peak was generally observed corresponding to non-specifically bound proteins. Once the absorbance reading leveled off, the imidazole concentration was increased using a step gradient. This process was repeated 4 or 5 times at concentrations between 60 mM and 150 mM imidazole. The volume required to effectively wash the column of unwanted protein varied between IN^{Bwt}, IN^{Bsol} and IN^{Cwt} which were purified from 1 l of culture and IN^{Csol} and IN^{Y143C} which were purified from 9 l of culture, with the latter requiring larger volumes. Recombinant IN was then eluted using a linear imidazole gradient from 150 mM to 600 mM over a volume of approximately 50 ml, collecting 4 ml fractions. The recombinant IN began eluting off the column at approximately 160 mM imidazole and peaked at approximately 320 mM imidazole. A representative elution profile of IN^{Bwt} is shown in Figure 3.10. A symmetrical bell shaped UV absorbance elution profile was obtained with corresponding SDS-

PAGE (Figure 3.11) and Western blotting (Figure 3.12) analysis confirming purification of the recombinant IN.

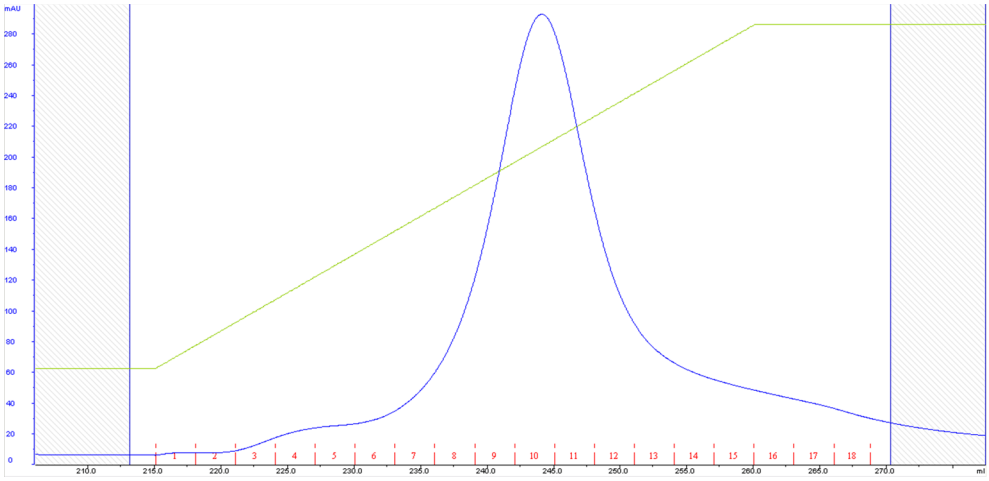

Figure 3.10: Representative elution profile from large scale purification of IN^{Bwt}. The green line represents the imidazole concentrations, starting at 20mM on the left and moving in a gradient over 45ml to 600mM imidazole on the right. The blue line represents UV absorbance.

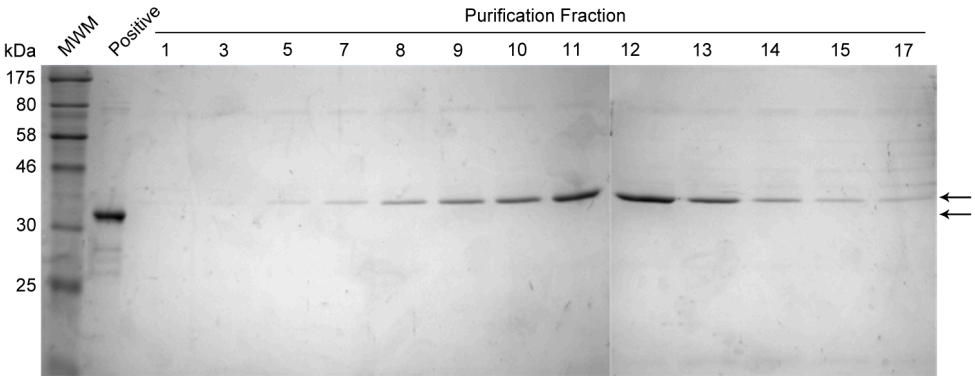

Figure 3.11: Coomassie stained SDS-PAGE analysis showing purification fractions from a representative IN^{Bwt} large scale purification experiment. Arrows indicate the positive control IN protein ($HIV\text{-}1_{NL4\text{-}3}$ IN protein, bottom arrow) and the expected position (~34kDa) of His-tagged recombinant IN^{Bwt} (top arrow). Fraction numbers are indicated on the top of each lane and molecular weight markers are indicated on the left.

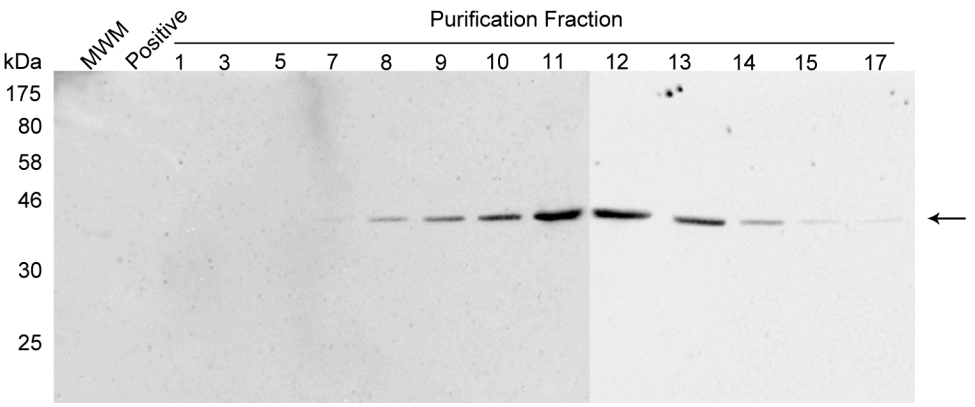

Figure 3.12: Western blot showing purification fractions from a representative large scale purification experiment of IN^{Bwt}. The arrow indicates the expected position (~34kDa) of the His-tagged IN^{Bwt}. Fraction numbers are indicated on the top of each lane. Molecular weight markers are indicated on the left. The antibody used for detection was the HisProbe-HRP antibody, thus the positive control protein HIV-1_{NL4-3} IN could not be detected.

The highest concentration of IN was found in fractions 10 and 11, while IN was present (as determined by SDS-PAGE, Figure 3.11) in fractions 5-19, although fractions 15-19 appear to contain several contaminating bands. Western blot analysis (Figure 3.12) showed that IN was present in fractions 7-17 but based on SDS-PAGE results only fractions 7-13 were pooled and concentrated. Similar results were obtained for the other four recombinant IN proteins.

3.2.4 Recombinant IN Concentration

Protein concentration and buffer exchange using the Amicon Ultra-15 Centrifugal Device yielded between 0.5 ml and 1.5 ml of concentrated IN. Densitometry analysis of SDS-PAGE experiments was carried out to determine the percentage purity of the IN proteins. The results are summarized in Table 3.1.

Table 3.1: Recombinant IN protein properties and concentrations

	Concentration (mg/ml)	Concentration (μM)	Molar Extinction Coefficient (ε, m^2/mol)	Molecular Weight (kDa)	Percentage purity	Specific concentration (mg/ml I μM)
IN^{Cwt}	1.68	49.191	52070	34.282	72.22%	1.21 I 35.39
IN^{Csol}	0.81	23.73	46020	34.130	15.52%	0.13 I 3.68
IN^{Y143C}	0.46	13.50	44860	34.070	18.82%	0.087 I 2.54
IN^{Bwt}	0.75	21.73	50790	34.522	50.81%	0.38 I 11.03
IN^{Bsol}	5.86	170	50670	34.487	83.4%	4.89 I 141.78

A representative SDS-PAGE of the percentage purity experiments is shown in Figure 3.13.

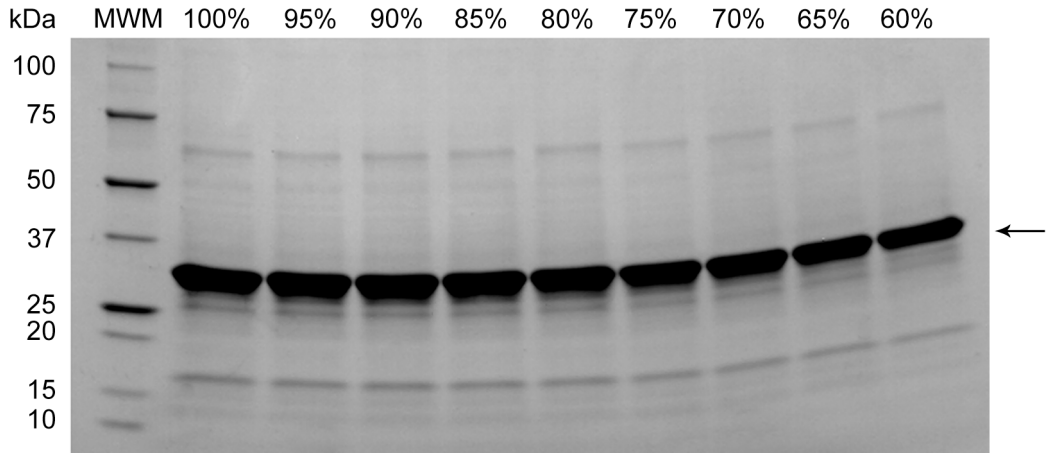

Figure 3.13: Coomassie stained SDS-PAGE of concentrated IN^{Cwt} at various concentrations. The arrow indicates the expected position (~34kDa) of the His-tagged IN^{Cwt} .The percentage concentration of IN is indicated at the top of each lane. The molecular weight marker used was the Precision Plus Protein Unstained Standard (Bio-Rad, USA), and the weights are indicated on the left.

All five concentrated recombinant IN proteins were subsequently used in the strand transfer assays, and the IN^{Y143C} mutant was used in complexation experiments.

3.3 Strand Transfer Analysis

3.3.1 Activity Studies

The strand transfer assay was set up using FITC or DIG labelled target DNA. After several attempts, it was established that the Target(FITC) and anti-FITC antibody combination showed a higher sensitivity and yielded more reproducible results as compared to the Target(DIG) and anti-DIG antibody, so this version of the strand transfer assay was used in all subsequent experiments.

Strand transfer analysis was carried out on all five recombinant IN proteins using the IN^{Bsol} as the positive control reference protein (therefore arbitrarily designated at 100% activity). The percentage activity of the remaining four recombinant IN proteins relative to the IN^{Bsol} is shown in Figure 3.14. Activity ranged from ~53% to ~117% for IN^{Csol} and IN^{Bwt} respectively. IN^{Y143C} and IN^{Bwt} had higher activity than the IN^{Bsol}, while IN^{Cwt} and IN^{Csol} had lower activity. Similarly, the IN^{Csol} had lower activity than the IN^{Cwt}.

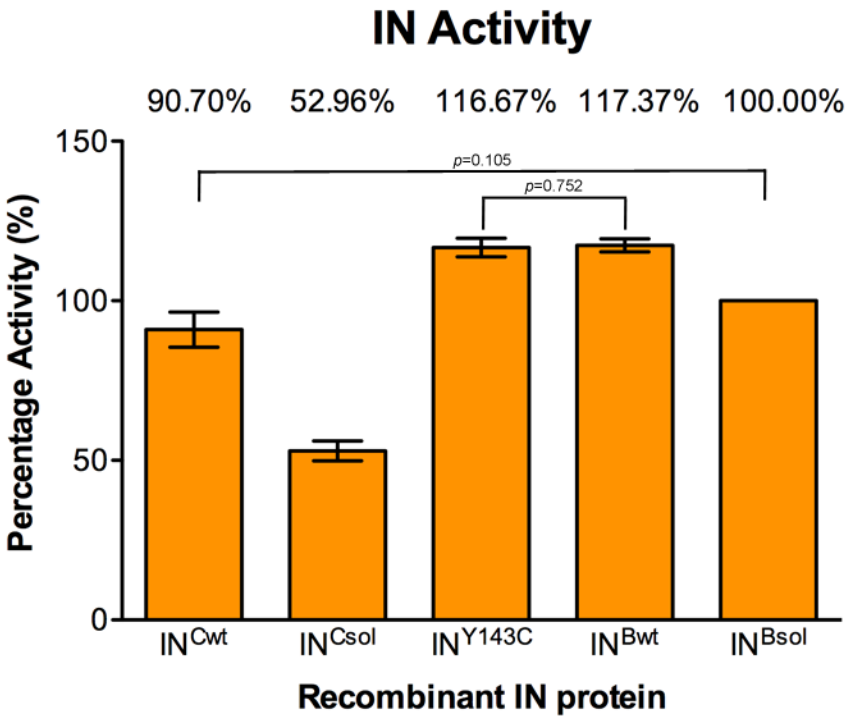

Figure 3.14: Recombinant IN activity as measured by strand transfer assay. Activity is expressed as a percentage, relative to IN^{Bsol} (indicated above each column on the graph). The error bars indicate the standard deviation over 3 separate experiments (n=9). Statistically similar results are indicated by brackets, with the *p* values given above.

Statistical analysis (by means of a student's T-test) of the data from three experiments provided evidence that the activities of both IN^{Cwt} and IN^{Bsol} were not statistically different (p=0.105, 95% confidence) and the activities of IN^{Y143C} and IN^{Bwt} were not statistically different (p=0.752, 95% confidence), while all other combinations were determined to be significantly different (p<0.05).

Overall, all five recombinant IN proteins were shown to be functional and capable of facilitating strand transfer.

3.3.2 Single Dose Studies

The established strand transfer assay was subsequently used to determine the effect of selected IN inhibitors on the activity of all five recombinant IN proteins.

The response to single dose administration of 10 µM L-Chicoric acid was measured for all five recombinant IN proteins and the percentage inhibition is summarized in Figure 3.15. L-Chicoric acid inhibited the strand transfer activity of all five recombinant IN proteins, with inhibition values ranging from between 87% (IN^{Csol}) and 96% (IN^{Bsol}).

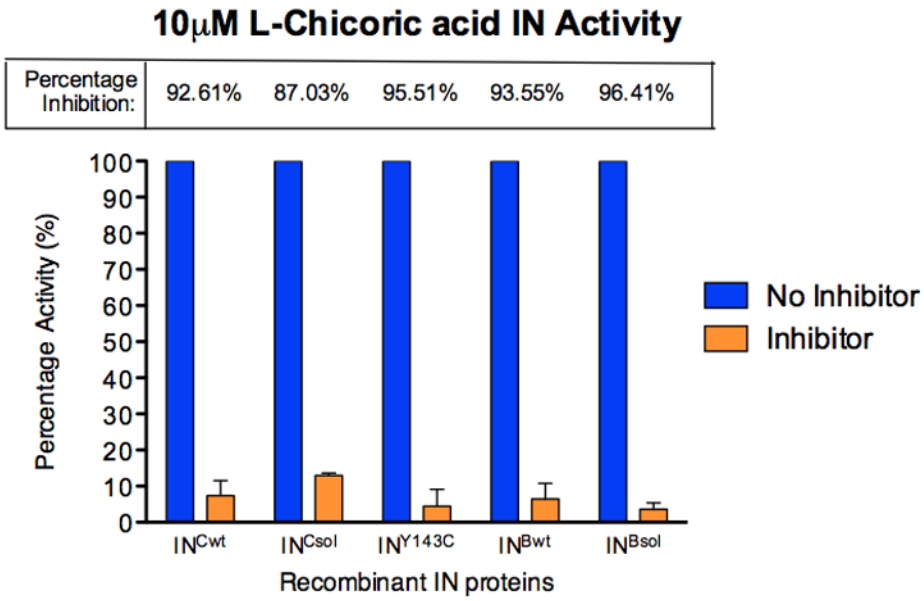

Figure 3.15: Recombinant IN activity with 10µM L-Chicoric acid, normalized against the no inhibitor control. The percentage inhibition is indicated above each column pair on the graph. The error bars indicate the standard deviation over 3 separate experiments.

Statistical analysis (by means of a students T-test) showed that inhibition of IN^{Csol} and IN^{Bsol} by 10 µM L-Chicoric acid were significantly different (p=0.006, 95% confidence) and there was no evidence of a significant difference between the level of inhibition of any other recombinant IN proteins ($p>0.05$).

3.3.3 Dose Response Studies

The dose responses to Raltegravir and L-Chicoric acid were determined for all five recombinant IN proteins, and an IC_{50} value for each was calculated.

The combined dose curves for all recombinant IN proteins in response to L-Chicoric acid is shown in Figure 3.16. The IC_{50} values are summarized in Table 3.2.

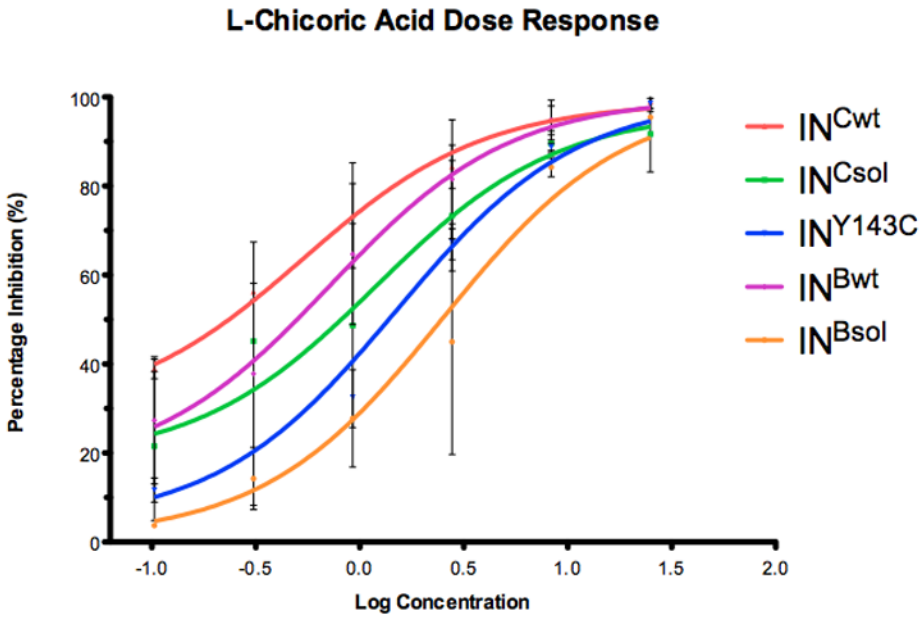

Figure 3.16: Combined dose response curves for all five recombinant integrase proteins in response to L-Chicoric acid. The error bars indicate the standard deviation over two separate experiments.

Table 3.2: IC_{50} values for all five recombinant IN proteins in response to L-Chicoric acid

	IN^{Cwt}	IN^{Csol}	IN^{Y143C}	IN^{Bwt}	IN^{Bsol}
IC_{50}	0.5378 µM	1.198 µM	1.495 µM	0.7202 µM	2.519 µM

Typical dose response curves to Raltegravir were obtained for IN^{Cwt} and IN^{Bsol} (Figure 3.17), allowing for the estimation of IC_{50} values (Table 3.3). Because the percentage inhibition of strand transfer activity for IN^{Bwt} was not higher than 50% at the highest Raltegravir concentration used, accurate IC_{50} values could not be extrapolated. The IN^{Y143C} mutant was generally resistant to inhibition of strand

transfer activity by Raltegravir, with the exception of the highest concentration used.

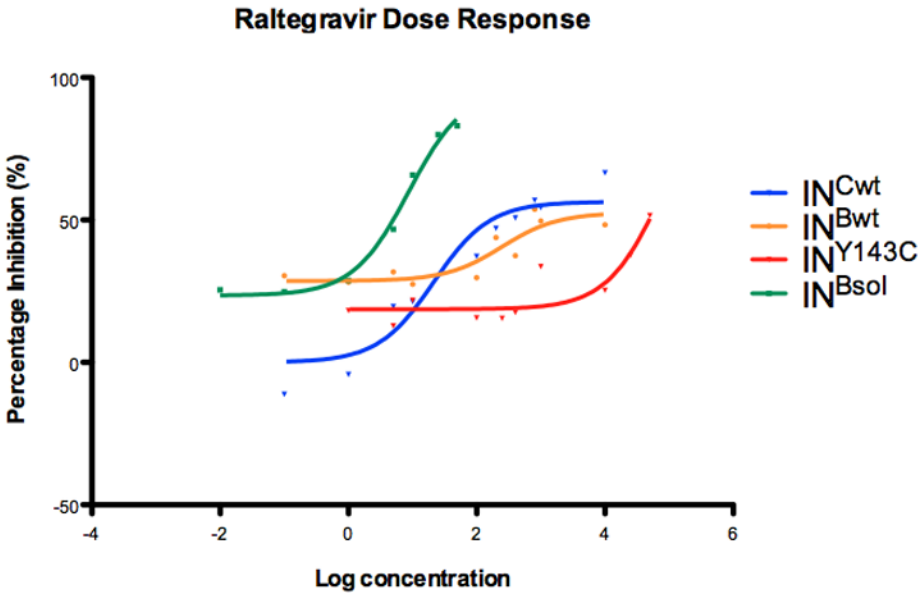

Figure 3.17: Combined dose response curves for IN^{Cwt}, IN^{Bwt} and IN^{Y143C} in response to Raltegravir.

The IC_{50} values for these proteins in response to Raltegravir are summarized in Table 3.3.

Table 3.3: IC_{50} values for IN^{Cwt}, IN^{Bwt} and IN^{Y143C} in response to Raltegravir

	IN^{Cwt}	IN^{Y143C}	IN^{Bwt}	IN^{Bsol}
IC_{50}	21.35 nM	≥50 nM	247.5 nM	8.72 nM

3.4 Complexation

The complexation experiment was performed several times with IN^{Y143C}, using IN^{Csol} as the control. However upon destaining of the SDS-PAGE gel, several contaminating bands were noted in both the test (IN^{Y143C}) and control (IN^{Csol})

lanes, and this made confirmation of IN-DNA complexes impossible. Thus, the results were inconclusive (data not shown).

CHAPTER FOUR

4 DISCUSSION

HIV-1 IN performs the task of integrating the viral DNA into the host chromosome, an obligate step in the lifecycle of the virus. For this reason, IN was identified as a potential drug target and the search for inhibitory compounds began. Using an assay to assess the strand transfer activity of IN, many inhibitors have been identified but to date only one has progressed past FDA approval and onto the shelf. In this study, five HIV-1 IN proteins were expressed and purified for use in one such assay: they are three subtype C proteins (which are unique to this study) and two subtype B proteins (both from the NIH AIDS Research and Reference Reagents Program). The HIV-1 subtype C wild type protein (IN^{Cwt}) was cloned in our laboratory from the 05ZAFV6 sequence, and the two subtype C mutants (IN^{Csol} and IN^{Y143C}) were generated using the wild type 05ZAFV6 sequence and inserting specific mutations from the literature (see section 1.6 and below). Strand transfer analysis was carried out on these proteins to assess their functionality, and their response to two well-known IN inhibitors was measured. It was found that all of the IN proteins in this study were functional, although IN^{Csol} exhibited lower than expected functionality while IN^{Y143C} showed higher than expected functionality. The single dose and dose response data with L-Chicoric acid and Raltegravir were in-line with previously published data for all proteins in this study. The Y143C mutation present in IN^{Y143C} is intended to allow stable complex formation with thiolated-DNA, such that IN bound natively with it's DNA substrate can be purified for use in strand transfer experiments.

Alian *et al.* (2009) (81) list four solubility mutations in their paper: C56S, W131D, F185D and C280S (81), which serve to reduce surface hydrophobicity of the IN

protein and thus increase solubility of the protein (see Figures 3.1 and 3.4). Indeed, modeling of the IN protein revealed that all of these mutations are present on the surface of the protein. The original solubility mutations established for HIV-1 subtype B IN were F185K and C280S (Figures 3.1 and 3.2) (203), and are widely accepted as the standard solubility mutations for IN. Later C56S and W131D were added (69). While F185K is one of the original mutations found in the Jenkins Bsol mutant (203), Alian *et al.* (2009) (81) used F185D instead and do not supply a reason for this change. The original F185K mutation replaces the hydrophobic phenylalanine residue with a positively charged lysine residue (shown in yellow on Figure 3.2), whilst Alian *et al.* (2009) (81) (and this study) use a negatively charged aspartic acid residue. A 1999 paper suggests that the solubility of these two different mutations (F185K and F185D) are similar (210). Additionally, the C65S mutation serves to preclude potential unwanted reactivity with thiolated DNA and protein modeling reveals that this residue is present near the active site in close proximity to the viral DNA end (Figure 3.4).

These mutations were introduced into the GBCwt sequence and designated GBCsol. A second construct, additionally containing the Y143C mutation was generated and termed GBCY143C. This Cysteine mutation occurs near the active site of the functional IN protein (Figure 3.4), and allows for the formation of complexes with thiolated DNA. The covalent disulphide bond between the thiol group on the DNA and the sulfur on the Cysteine has been shown to withstand high salt conditions, allowing for the purification of IN-DNA complexes (81). It should be noted that this Y143C mutation is an established Raltegravir resistance

mutation, as discussed in section 1.4.10, and that it impacts negatively on IN activity in the absence of any strand transfer inhibitors (discussed below).

Modeling of the protein structure reveals the locations of the various amino acids within the context of the tertiary structure of the protein, and allows for several conclusions to be drawn. Of the 20 amino acid differences between IN^{Cwt} and IN^{Bwt}, 16 could be visualized using the available model, while the remaining four occur within a region of the CTD not contained in the model. Subsequently, it was determined that all of the mutations present in the NTD are on the surface of the protein, although only M50T is a significant change in amino acid properties (a hydrophobic amino acid is replaced with a polar uncharged residue). Amino acid changes that alter the overall surface hydrophobicity or charge of the protein are expected to hold consequences for protein solubility or dimerization (if present on the dimer-dimer interface), however the NTD is not catalytic, rather serving dimerization functions and thus changes here are not expected to influence protein functionality to any great degree (53-55)

There were 10 mutations present in the CCD, and four of these were on the outer surface of the protein. I151V is of particular interest because it is within the active site (Figure 4.1 B), and occurs immediately adjacent the Glutamic acid residue that makes up part of the DDE motif. This motif is responsible for co-ordinating the magnesium II ions (shown in light blue in Figure 4.1), and is highly conserved in HIV-1 IN. The Isoleucine to Valine change is not expected to incur serious consequences as there is no significant change in amino acid side chain properties or size.

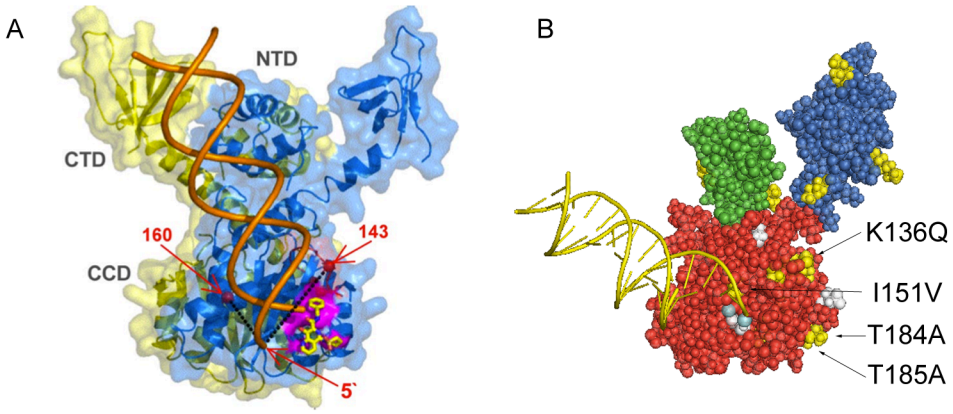

Figure 4.1: Comparison between the model presented by Alian *et al.* (2009) (81) (A) and the IN^{Csol} model used in this study (B). In A, IN monomers are shown in blue and yellow. The domains and features in B are coloured as in Figure 3.2, except the magnesium II ions are shown as light blue spheres. The mutations K136Q, I151V, T184A and T185A are indicated in B with arrows. The monomer in B is analogous to the blue monomer in A.

The T184A and T185A mutations are also on the surface of the protein (Figure 4.1 B), and in this instance IN^{Cwt} loses two polar side chains for hydrophobic side chains. Based on the IN dimer model presented by Alian *et al.* (2009) (81) (Figure 4.1 A), the position of these mutations does not appear to be on the dimer-dimer interaction surface so will not likely interfere with this interaction, although they would be expected to affect protein solubility.

Neither K136Q (Figure 4.1 B) or V201I (on the reverse side of the protein in Figure 4.1 B, not visible) are expected to have any serious consequences for the protein, as neither are near the active site or present significant amino acid side chain property changes. The remaining five mutations in the CCD are buried within the protein so they are not solvent accessible, do not participate in dimer-dimer interactions, and are also not near the active site.

There are three mutations within the part of the CTD contained in this model and all are present on the surface of the protein, however only T218I is important as it

is a change from a polar side chain to a hydrophobic side chain. The amino acid is not present on the dimer-dimer interface predicted by Alian *et al.* (2009) (81), but it may have implications for protein solubility. The remaining four mutations in the CTD (namely R270K, D278A, S283G and R283G) are part of a region of the CTD not present in the model, so no speculation can be made on the location and effects of these mutations. It is worth noting, however, that all of these mutations incur a change in amino acid side chain properties.

The six carbon thiol linker present on the terminal 5' adenine of the Thiolated-DNA is approximately 9 Å in length (81), which is sufficient to make the crosslink according to our model which predicts the distance between the terminal 5' adenine and the C143 residue to be 7.1 Å (Figure 3.4). Further, residue 143 occurs on a flexible loop (residues 139-150) within the CCD (66,81), which suggests that this binding length need not be exact in order for a crosslink to occur. The model also shows the proximity of the C56S mutation to the viral DNA, supporting the inclusion of this mutation to reduce potential unwanted interactions with the thiolated DNA.

For the expression of all recombinant IN constructs, a bacterial expression system was selected due to the low cost, ease of operation, simple scale-up amenability, rapid culture time and high yield attained with this method. The primary disadvantages with this type of system include decreased protein solubility, a complete lack of post-translational modification and the possibility of impaired folding (211). While these disadvantages are overcome with the use of mammalian expression systems, the benefits are negated by the slow culture

times, high cost, reduced yield and complicated technique (88). Indeed, most publications involving IN expression, beginning with the first reports in 1990 and continuing to present day, describe the use of bacterial expression systems and report on the functionality of the protein produced.

IN is known to have a low solubility *in vitro* and so Sherman *et al.* (1990) (88) developed the method that would become the standard for all future work: the bacterial cells are harvested by centrifugation and lysed under high salt conditions in order to solubilize the IN protein (88,203,212). Indeed, this technique has been used successfully before in our laboratory.

Several publications describe the use of the pET-15b vector system for the expression of recombinant HIV-1 IN (49,61,203,212,213), and this vector has also been successfully used to express functional HIV-1 IN in our laboratory. The IN nucleotide sequence was inserted downstream of the T7 promoter and the lac operator in the pET-15b vector for all five IN constructs. The T7 promoter is only recognized by the T7 RNA polymerase, thus facilitating expression using the bacterial *E.coli* BL21 (DE3) expression system. These cells have been transduced by the λ bacteriophage T7 and thus contain the T7 RNA polymerase necessary to transcribe the messenger RNA from a gene under control of the T7 promoter. This particular RNA polymerase has high promoter specificity, RNase activity (necessary for high fidelity transcription), is not dependent on other protein factors to initiate transcription and has a higher rate of transcription than other *E.coli* RNA polymerases (214) and so this promoter has become popular for expressing recombinant proteins (215). The lac operator is a regulatory element which

enables temporal control expression; it is bound by the lac repressor (LacI) which interferes with binding on the RNA polymerase and subsequently prevents transcription (216). LacI is displaced by lactose, or in this case the lactose analogue IPTG (which cannot be metabolized by *E.coli* and thus remains at a constant concentration), and transcription and subsequent translation are allowed to take place. *E.coli* BL21 (DE3) contains the *lacI* gene, and so is capable of repressing expression of genes under control of the lac operator in the absence of IPTG.

The IN constructs expressed in this study are located in frame with a His-Tag, which results in expression of a fusion protein with a His-Tag at the amino terminal (N-terminal) of the protein. As a result, purification by immobilized metal affinity chromatography (217) was possible. The Histidine amino acids that comprise the N-terminal His-Tag form a strong interaction with the immobilized metal ions (Nickel II in this study) on the matrix, and thus recombinant IN is bound to the matrix (218). The His-Tag contributes approximately 2.18 kDa to the molecular weight of the recombinant IN with which it is fused (http://www.basic.northwestern.edu/biotools/proteincalc.html).

Small scale inductions were carried out to ascertain whether recombinant IN was being expressed by the expression vectors. Optimization efforts demonstrated that an IPTG concentration of 1.0 mM IPTG induced at an OD of 0.8 and incubated for 3 hours resulted in high levels of expression for constructs Bsol, Bwt and GBCwt. Unfortunately the GBCsol and GBCY143C constructs did not express to the same degree, although these same induction conditions were used (with a scale up in

culture volume to account for reduced expression). Induction for 24 hours appeared to make little impact on expression levels and yielded low molecular weight IN bands on western blot indicative of protein degradation (results not shown). There does not appear to be any leaky expression, as confirmed by the absence of bands in all negative control lanes (Figures 3.5, 3.6 and 3.7).

GBCsol and GBCY143C did not show this same high level of expression under these conditions (see Figure 3.8). Initially, IN bands were not distinguishable on Coomassie stained SDS-PAGE, so the induction conditions were optimized in an effort to maximize protein expression for these constructs. It was found that expression was best when induced at an OD of 0.8 (results not shown), although IPTG concentration appeared to have little or no impact on expression levels as shown by western blot analysis (Figure 3.9).

Previous experiments with GBCwt had resulted in good yields when expression was scaled up to 1 l volumes, and as such this was done for Bwt and Bsol. Due to the lower levels of expression for GBCsol and GBCY143C, expression was scaled up to 9 l for these constructs.

Following large scale expression, the recombinant IN was purified from the soluble cellular fraction. The washing steps contained low concentrations of imidazole to displace weakly bound Histidine containing proteins. Imidazole has a similar chemical structure to that of Histidine, and thus competes with Histidine-containing proteins and His-Tagged proteins for binding to the Nickel II charged resin in the column. An excess concentration of imidazole was finally used to elute the His-

Tagged IN protein from the column. The characteristics of this affinity column mean that other bacterial proteins that might contain several Histidine residues in close proximity may also bind to the column and co-elute with the recombinant IN, and for this reason an imidazole gradient was employed and small elution fractions were captured in order to minimize the degree of contamination from proteins of this nature.

Wash volumes were adjusted on a purification to purification basis, by monitoring the UV absorbance levels of the eluant coming off the column. From previous studies in our laboratory it was known that the IN protein would elute at imidazole concentrations of around 200 mM and higher, so incremental washes were carried out at lower imidazole concentrations. During each wash, an early absorbance peak was generally observed corresponding to non-specifically bound proteins. Once the absorbance reading leveled off, the imidazole concentration was increased using a step gradient. This process was repeated 4 or 5 times at concentrations between 60 mM and 150 mM imidazole. The volume required to effectively wash the column of unwanted protein varied between IN^{Cwt}, IN^{Bwt} and IN^{Bsol} which were purified from 1 l of culture and IN^{Csol} and IN^{Y143C} which were purified from 9 l of culture, with the latter requiring larger volumes. Recombinant IN was then eluted using a linear imidazole gradient from 150 mM to 600 mM over a volume of approximately 50 ml, collecting 4 ml fractions. The UV absorbance profile clearly indicated when the IN protein was being displaced from the column (Figure 3.10). SDS-PAGE and western blot analysis confirmed the presence of IN in these fractions (Figures 3.11 and 3.9).

The His-Tag used for affinity chromatography in this study contains the Thrombin cleavage site, which allows for the removal of the His-Tag using Thrombin. There is evidence in the literature that the removal of the His-Tag has no influence on IN activity, so the His-Tag was not removed in this study (219).

The percentage purity was determined by densitometry analysis, and the results are summarized in Table 3.1. A representative SDS-PAGE of the percentage purity experiment is shown in Figure 3.13. In this figure it is clear there are several contaminating bands of both higher and lower molecular weight. This finding was consistent for all the protein purified in this study, although the degree of contamination varied between proteins. IN^{Csol} and IN^{Y143C} show the lowest purity, and these two proteins also exhibited low expression levels. Because of the low expression levels, the culture volumes for these two proteins were scaled up, and this presumably caused the increase in contaminating protein during the purification process. While the wash steps were scaled up, the amount of protein that elutes at or near the imidazole concentration at which IN elutes was increased proportionately, resulting in this significant drop in purity.

As described in the section 2.4, two distinct sets of reagents were evaluated for the strand transfer assay. It was found that the reaction using FITC-labelled target DNA and subsequent detection using monoclonal anti-FITC AP antibody yielded more sensitive and reproducible results. AP has a lower rate constant than HRP and so substrate turnover occurs more slowly. Also, there is no substrate inhibition of AP so the reaction can continue for long periods of time, and observes a nearly linear development profile (220). Consequently the absorbance values of the

assay were read after 1 hour and 2 hours at 37°C, and then again after spending the remaining 24 hours at 4°C. It was found that the results after 1 hour at 37°C were satisfactory (results not shown), so all subsequent experiments were incubated for this amount of time.

In order to assess the functionality of the constructs purified in this study, drug free strand transfer assays were carried out. All constructs were compared to an IN^{Bsol} protein already established to be functional in our laboratory. This IN^{Bsol} protein is widely accepted as the control IN protein, and is used extensively in the literature (48,53,60,76,221-236) in *in vitro* studies on HIV-1 IN (203).

The assay was performed in biological and experimental triplicate, and the data from these experiments confirm the activity levels of the proteins. The standard deviations of the raw absorbance readings did not exceed 5% of the mean for any experiment and no outliers were detected or excluded, indicating there was no significant variation within each experiment. With IN^{Bsol} arbitrarily designated at 100% activity, the relative activity of the other constructs was determined individually for each of the three experiments (Figure 3.14). The standard deviation for the mean activity of each construct over the three experiments was less than 5% in all cases.

A two-tailed two-sample unequal variance (heteroscedastic) students t-test was carried out to determine if the mean activities between each possible pair of IN proteins in this study were statistically similar. It was found that only IN^{Cwt} and IN^{Bsol} ($p=0.105$, 95% confidence) and IN^{Y143C} and IN^{Bwt} ($p=0.752$, 95%

confidence) were statistically similar, whereas all other proteins showed significantly different levels of activity.

These results show clearly that IN^{Y143C} and IN^{Bwt} have significantly higher activity than that of IN^{Bsol} (while not being significantly different from each other). In the case of the HIV-1 subtype B IN proteins, one expects the wild type protein to have a higher activity than that of the soluble mutant due to the absence of any artificial mutations. HIV-1 is a highly mutable virus due to the low fidelity replication afforded it by RT, and as such the virus undergoes evolution at an extremely rapid rate. Because of this, one expects the virus to have the most efficient version of the IN protein already available to itself. The solubility mutations introduced into Bsol by Jenkins *et al.* (1996) (203) were designed *in silico* so as to reduce the surface hydrophobicity of the protein in order to increase yields and facilitate crystallization, and were selected on the basis that they incurred the lowest reduction in IN activity. These results, however, do not align well with the data published in 1996; Jenkins *et al.* (1996) (203) found no significant difference between the activity of the soluble mutant and the wild type protein.

The hypothesis that the soluble mutant should have lower activity than the wild type protein is supported by subtype C proteins in this study. We find the activity of IN^{Csol} to be 38% lower than that of the IN^{Cwt}, and no previous data exists comparing these two proteins. The IN^{Csol} mutant contains a different set of solubility mutations than those of IN^{Bsol} (see Figure 3.1) and the strand transfer activity of this set of solubility mutations was not originally assessed (69). When

Alian *et al.* (2009) (81) added C65S and changed F185K to F185D in their work, they found the activity to be approximately 80% of the wild type protein (81). Thus, in light of the evidence supplied by Alian *et al.* (2009) (81), we expect the activity of the IN^{Csol} mutant to be lower than that of the wild type.

It is interesting to note that there is no significant difference in activity between IN^{Cwt} and the IN^{Bsol} mutant protein. No previously published work on HIV-1 subtype C wild type IN is available, and no studies have specifically explored the difference in activity between subtype B and subtype C proteins. The data presented here shows clearly that HIV-1 subtype C IN functions at a reduced efficiency when compared to HIV-1 subtype B IN, particularly for the wild type proteins but as discussed this result is complicated by the different mutations present in the soluble mutants. It is apparent that the additional mutations imposed on the IN^{Csol} construct (namely C56S, W131D, F185D and C65S) have a detrimental effect on IN activity.

A particular curiosity in this data is IN^{Y143C}. In light of the fact that the mutations in IN^{Csol} appear to abrogate functionality to some degree, it remains to be explained why the Y143C mutation should restore the functionality of the protein back to that of the subtype B wild type level ($p > 0.05$). The literature states that infectivity of the Y143C virus is only 10% of that of the wild type and that there is a 10 fold reduction in the amount of integrated viral DNA (measured in cell culture based experiments) (237). *In vitro* data from this same study revealed that the Y143C mutant IN showed only 21% of the wild type strand transfer activity (measured

using a gel based strand transfer assay). Since the Y143C mutation exists within the context of the other solubility mutations used in this study, the most accurate representation of the expected result would be that provided by Alian *et al.* (2009) (81), where they mention that the activity of the various cysteine mutations was between 64% and 90% of the wild type enzyme, although this was carried out in HIV-1 subtype B IN. It is worth mentioning that both Delelis *et al.* (2010) (237) and Alian *et al.* (2009) (81) used a gel based assay to measure strand transfer activity while the data presented in this study is based on an electrophoresis free assay (132).

A low percentage purity implies that there is a considerable amount of contaminating protein present in the IN solution. It is interesting to note that the percentage purity of IN^{Csol} and IN^{Y143C} are comparatively low (~15% and ~19% respectively, Table 3.1), although this does not appear to influence the functionality of the proteins: IN^{Csol} does show the lowest activity (see Figure 3.14) which could be attributed to its lower purity but IN^{Y143C} has particularly high activity despite also having a low purity. IN^{Y143C} has a statistically similar level of activity to IN^{Bwt} which has a much higher purity (~50%) comparatively. Similarly, the susceptibility of these proteins to L-Chicoric acid is not compromised by the purity as seen in the dose response curves Figure 3.16. Taken together these results suggest that the amount of impurities in the IN protein solution does not affect the activity. Similarly, the reduced concentration of the IN protein does not seem to have any bearing on the activity of the protein: the percentage purity of the IN proteins affects the specific purity of each protein, with the net result that the

actual concentration of IN present in the strand transfer assay is not what is expected. This is due to the fact that contaminating proteins will also absorb UV light at a wavelength of 280 nm, and this will subsequently result in over-reporting of the concentration by the spectrophotometer. This is the case for all the proteins used in this study, where the specific concentration of IN used in the assay is between 15% and 83%, however as mentioned before there seems to be no correlation between the percentage purity and the activity of the protein.

According to the literature, a percentage inhibition in the region of 90% is expected for L-Chicoric acid at 5 µg/ml (147). In our studies, a single dose of 10 µM (4.74 µg/ml) L-Chicoric acid inhibited IN^{Bwt} and IN^{Bsol} by ~94% and ~96% respectively.

Following analysis by a two-tailed two-sample unequal variance (heteroscedastic) students t-test on all possible pairs of IN proteins, it was found that only IN^{Csol} and IN^{Bsol} are significantly different ($p=0.006$, 95% confidence) with respect to their susceptibility to L-Chicoric acid. According to this statistical data, at a 95% confidence level all the proteins responded in the same way to L-Chicoric acid with the exception of the IN^{Bsol} and IN^{Csol}. As discussed before, the solubility mutations in these two constructs are not identical and so direct conclusions about the differences between subtype B and subtype C responses to L-Chicoric acid cannot be made. It is worth noting that at 87.03%, IN^{Csol} showed the lowest sensitivity to L-Chicoric acid whilst IN^{Bsol} showed the greatest sensitivity at almost 97% inhibition, giving a range of only ~10% across all the recombinant IN variants tested in this study.

The typical IC_{50} value for L-Chicoric acid is reported to be between 0.104 µM and 0.210 µM (147), although the data generated in this study appears to be much higher; ranging from a 2 fold increase (IN^{Cwt}, 0.538 µM) all the way to a 10 fold increase (IN^{Bsol}, 2.519 µM). Robinson *et al.* (1996) (147) use the same HIV-1 subtype B wild type protein as used in this study, although the IC_{50} value obtained in this study was 3-4 fold higher. This is believed to be due to variations between batches of purified protein, specifically in terms of concentration and purity. Also, Robinson *et al.* (1996) (147) performed an electrophoresis based assay for IN activity, and the inherent differences between that assay and the one used in this study may account for these differences in IC_{50} values.

The maximum percentage inhibition achieved by the highest concentration of L-Chicoric acid is marginally higher for IN^{Cwt} and IN^{Bwt} than for the other recombinant IN proteins tested in this study (Figure 3.16), and the IC_{50} values are similarly lower than the other proteins. Thus, the solubility mutations (namely F185K and C280S for the subtype B protein, and F185D, W131D, C56S and C280S for the subtype C proteins) appear to contribute mildly to resistance to L-Chicoric acid. Interestingly, the Y143C mutation present in IN^{Y143C} does not contribute to L-Chicoric acid resistance to the degree expected. Furthermore, the HIV-1 subtype B proteins both contain the G140S mutation (see Figure 3.1) which has been shown to confer resistance to L-Chicoric acid by 600 fold in tissue culture based experiments (149), although this same level of resistance is not seen in biochemical assays. These findings warrant further investigation, and open the door to future work to determine the impact of residue 140 on L-Chicoric acid

resistance *in vitro* using HIV-1 subtype B and subtype C recombinant proteins, and how these translate to actual resistance in a clinical setting. The findings presented here indicate that the solubility mutations appear to have the greatest effect on resistance to this compound, and indeed this has been noticed for other IN inhibitors; a paper by Marchand *et al.* (2003) (238) describes the resistance profiles of F185K and C280S double-mutant IN proteins.

The IC_{50} values for IN^{Cwt}, IN^{Y143C}, IN^{Bwt} and IN^{Bsol} in response to Raltegravir are summarized in Table 3.3. According to published data, Raltegravir has an IC_{50} for strand transfer of approximately 15nM (155). The HIV-1 subtype C recombinant IN appears to have a slightly higher IC_{50} in response to Raltegravir, which is as expected (239). Bar-Magen *et al.* (2010) (239) found that HIV-1 subtype B and subtype C exhibited differential resistance profiles in response to IN inhibitors.

By contrast, the IN^{Bwt} has a much higher IC_{50} than expected (~250 nM), while IN^{Bsol} is slightly below the published values with an IC_{50} of 8.72 nM (Table 3.3). The G140S mutation present in the subtype B IN proteins is expected to result in an increase in the IC_{50} to approximately 30 nM (240), whereas the result in this study is approximately 10 fold higher for IN^{Bwt}.

There appeared to be a low level of inhibition of IN^{Y143C} strand transfer activity (~20%) across the range of Raltegravir concentrations, with the exception of the highest concentration (50 µM, Figure 3.17). Subsequently, no IC_{50} value could be obtained for IN^{Y143C}. However, this is to be expected and the data shows clearly

that the Y143C mutation is capable of conferring a high level of resistance to Raltegravir. The implications for drug screening using this protein are significant - if candidate molecules are found that inhibit this protein it implies these molecules do not share cross resistance with Raltegravir and Elvitegravir, and could possibly replace these drugs upon failure of therapy due to resistance.

Despite optimization attempts, SDS-PAGE analysis of complexation experiments was inconclusive, and consequently these results were not included. Future work should focus on optimizing the experiments in order to fully characterize the IN-DNA complexes. Larger volumes of protein and DNA substrate may also contribute to the success of the experiment. Increasing the percentage purity of the IN^{Y143C} will help to reduce the presence of contaminating bands and allow for easier interpretation of the results and western blot analysis should be performed to verify the results; using an enzyme linked streptavidin molecule or anti-Biotin antibody to detect the DNA substrates will confirm the presence of complexes if the bands coincide with the approximate location of the IN protein. Further, this method will allow for detection of uncomplexed DNA as well. Of particular interest in these experiments will be whether higher molecular weight IN dimers and tetramers can be detected by non-reducing SDS-PAGE, especially in complex with thiolated-DNA.

Following confirmation of complexation, purification of the IN-DNA complexes will be necessary to allow for further characterization. Alian *at al.* (2009) (81) purified their complexes by size exclusion chromatography although a simpler method using Amicon Ultra-15 Centrifugal Devices to buffer exchange may remove excess

thiolated-DNA (which would be allowed through the filter due to its smaller size), and this would be sufficient for strand transfer analysis. When analyzing the strand transfer abilities of the IN-DNA complex, the first step is the addition of the IN-DNA complex as the thiolated-DNA has biotin on its distal terminus for attachment of the whole complex onto the streptavidin coated plate. The removal of the excess thiolated-DNA is critical, as it would bind to the plate. Excess (uncomplexed) IN protein is not likely to be problematic in a strand transfer experiment as it would not bind and would be removed during the washing step.

Overall, this study is the first to report on the expression and purification of functional recombinant IN from South African HIV-1 subtype C, and a comparison with established HIV-1 subtype B IN proteins. These can all serve as reagents for further research into the discovery of novel IN inhibitors. The basis for complexation experiments involving a thiolated-DNA and IN^{Y143C} has been established, and this area of research holds promise for the future study of HIV-1 subtype C IN in complex with viral DNA.

CHAPTER FIVE

5 REFERENCES

1. Gottlieb MS, Schroff R, Schanker HM, Weisman JD, Fan PT, Wolf RA, et al. *Pneumocystis carinii* pneumonia and mucosal candidiasis in previously healthy homosexual men. New England Journal of Medicine. 1981 305(24):1425–1431.

2. Gottlieb MS. Pneumocystis Pneumonia--Los Angeles. American Journal of Public Health. 2006 96(6):980–981.

3. Brennan RO, Durack DT. Gay compromise syndrome. The Lancet. 1981 2(8259):1338.

4. Pasteur Institut. Human Immunodeficiency Virus (HIV) associated with Acquired Immune Deficiency Syndrome (AIDS), A Diagnostic Method for AIDS and Pre-AIDS, and a Kit Therefore. 1987 (5135864).

5. Gallo RC, Sarin PS, Gelmann EP, Robert-Guroff M, Richardson E, Kalyanaraman VS, et al. Isolation of human T-cell leukemia virus in acquired immune deficiency syndrome (AIDS). Science. 1983 220(4599):865.

6. Barré-Sinoussi F, Chermann JC, Rey F, Nugeyre MT, Chamaret S, Gruest J, et al. Isolation of a T-lymphotropic retrovirus from a patient at risk for acquired immune deficiency syndrome (AIDS). Science. 1983 220(4599):868.

7. Coffin J, Haase A, Levy JA, Montagnier L, Oroszlan S, Teich N, et al. What to call the AIDS virus? Nature. 1986 321(6065):10.

8. Dalgleish AG, Beverley PC, Clapham PR, Crawford DH, Greaves MF, Weiss RA. The CD4 (T4) antigen is an essential component of the receptor for the AIDS retrovirus. Nature. 1984 312(5996):763.

9. Klatzmann D, Champagne E, Chamaret S, Gruest J, Guetard D, Hercend T, et al. T-lymphocyte T4 molecule behaves as the receptor for human retrovirus LAV. Nature. 1984 (312):767–768

10. Spira AI, Marx PA, Patterson BK, Mahoney J, Koup RA, Wolinsky SM, et al. Cellular targets of infection and route of viral dissemination after an intravaginal inoculation of simian immunodeficiency virus into rhesus macaques. The Journal of Experimental Medicine. 1996 183(1):215.

11. Stevenson M. HIV-1 pathogenesis. Nature Medicine. 2003 9(7):853–860.

12. Kahn JO, Walker BD. Acute human immunodeficiency virus type 1 infection. New England Journal of Medicine. 1998 339(1):33.

13. Royce RA, Sena A, Cates W, Cohen MS. Sexual transmission of HIV. New England Journal of Medicine. 1997 336(15):1072.

14. Update AE. December 2009. Joint United Nations Programme on HIV/AIDS and World Health Organization. 2009.

15. Gao F, Bailes E, Robertson DL, Chen Y, Rodenburg CM, Michael SF, et al. Origin of HIV-1 in *Pan troglodytes troglodytes*. Nature. 1999 397(6718):436–441.

16. Reeves JD, Doms RW. Human immunodeficiency virus type 2. Journal of General Virology. 2002 83(6):1253.

17. Gurtler LG, Hauser PH, Eberle J, Brunn Von A, Knapp S, Zekeng L, et al. A new subtype of human immunodeficiency virus type 1 (MVP-5180) from Cameroon. The Journal of Virology. 1994 68(3):1581.

18. Simon F, Mauclère P, Roques P, Loussert-Ajaka I, Müller-Trutwin MC, Saragosti S, et al. Identification of a new human immunodeficiency virus type 1 distinct from group M and group O. Nature Medicine. 1998 4(9):1032–1037.

19. Plantier JC, Leoz M, Dickerson JE, De Oliveira F, Cordonnier F, Lemée V, et al. A new human immunodeficiency virus derived from gorillas. Nature Medicine. 2009 15(8):871–872.

20. Myers G, MacInnes K, Korber B. The emergence of simian/human immunodeficiency viruses. AIDS Research and Human Retroviruses. 1992 8(3):373–386.

21. Louwagie J, McCutchan FE, Peeters M, Brennan TP, Sanders-Buell E, Eddy GA, et al. Phylogenetic analysis of gag genes from 70 international HIV-1 isolates provides evidence for multiple genotypes. AIDS. 1993 7(6):769.

22. Janssens W, Heyndrickx L, Fransen K, Motte J, Peeters M, Nkengasong JN, et al. Genetic and phylogenetic analysis of env subtypes G and H in Central Africa. AIDS Research and Human Retroviruses. 1994 10(7):877–879.

23. Kostrikis LG, Bagdades E, Cao Y, Zhang L, Dimitriou D, Ho DD. Genetic analysis of human immunodeficiency virus type 1 strains from patients in Cyprus: identification of a new subtype designated subtype I. The Journal of Virology. 1995 69(10):6122.

24. Leitner T, Alaeus A, Marquina S, Lilja E, Lidman K, Albert JAN. Yet another subtype of HIV type 1? AIDS Research and Human Retroviruses. 1995 11(8):995–997.

25. Louwagie J, Janssens W, Mascola J, Heyndrickx L, Hegerich P, van der Groen G, et al. Genetic diversity of the envelope glycoprotein from human immunodeficiency virus type 1 isolates of African origin. The Journal of Virology. 1995 69(1):263.

26. Triques K, Bourgeois A, Vidal N, Mpoudi-Ngole E, Mulanga-Kabeya C, Nzilambi N, et al. Near-full-length genome sequencing of divergent African HIV type 1 subtype F viruses leads to the identification of a new HIV type 1 subtype designated K. AIDS Research and Human Retroviruses. 2000 16(2):139–151.

27. McCutchan FE, Hegerich PA, Brennan TP, Phanuphak P, Singharaj P, Jugsudee A, et al. Genetic Variants of HIV-1 in Thailand. AIDS Res. Hum. Retroviruses AIDS Research and Human Retroviruses. 1992 8(11):1887–1895.

28. Carr JK, Salminen MO, Koch C, Gotte D, Artenstein AW, Hegerich PA, et al. Full-length sequence and mosaic structure of a human immunodeficiency virus type 1 isolate from Thailand. The Journal of Virology. 1996 70(9):5935.

29. Gao F, Robertson DL, Morrison SG, Hui H, Craig S, Decker J, et al. The heterosexual human immunodeficiency virus type 1 epidemic in Thailand is caused by an intersubtype (A/E) recombinant of African origin. The Journal of Virology. 1996 70(10):7013.

30. Gao F, Robertson DL, Carruthers CD, Li Y, Bailes E, Kostrikis LG, et al. An isolate of human immunodeficiency virus type 1 originally classified as subtype I represents a complex mosaic comprising three different group M subtypes (A, G, and I). The Journal of Virology. 1998 72(12):10234.

31. Nasioulas G, Paraskevis D, Magiorkinis E, Theodoridou M, Hatzakis A. Molecular analysis of the full-length genome of HIV type 1 subtype I: evidence of A/G/I recombination. AIDS Research and Human Retroviruses. 1999 15(8):745–758.

32. Kuiken C, Foley B, Leitner T, Apetrei C. HIV sequence compendium 2010. Los Alamos National Laboratory. 2010.

33. Thomson MM, Pérez-Álvarez L, Nájera R. Molecular epidemiology of HIV-1 genetic forms and its significance for vaccine development and therapy. The Lancet Infectious Diseases. 2002 2(8):461–471.

34. Tebit DM, Arts EJ. Tracking a century of global expansion and evolution of HIV to drive understanding and to combat disease. The Lancet Infectious Diseases. 2011 11(1):45–56.

35. Takasaki T, Kurane I, Aihara H, Ohkawa N, Yamaguchi J. Electron microscopic study of human immunodeficiency virus type 1 (HIV-1) core structure: two RNA strands in the core of mature and budding particles. Archives of Virology. 1997 142(2):375–382.

36. Gelderblom HR. Assembly and morphology of HIV: potential effect of structure on viral function. AIDS. 1991 5(6):617.

37. Oroszlan S, Luftig RB. Retroviral proteinases. Current Topics in Microbiology and Immunology. 1990 157:153–185.

38. Cann AJ, Karn J. Molecular biology of HIV: new insights into the virus life-cycle. AIDS. 1989 3(1):S19.

39. Gelderblom HR, Özel M, Hausmann EHS, Winkel T, Pauli G, Koch MA. Fine structure of human immunodeficiency virus (HIV), immunolocalization of structural proteins and virus-cell relation. Micron and Microscopica Acta. 1988 19(1):41–60.

40. Göttlinger H. HIV-1 Gag: a molecular machine driving viral particle assembly and release. HIV Sequence Compendium. 2001.

41. Kwong PD, Wyatt R, Robinson J, Sweet RW, Sodroski J, Hendrickson WA. Structure of an HIV gp 120 envelope glycoprotein in complex with the CD4 receptor and a neutralizing human antibody. Nature. 1998 393(6686):648-659.

42. Karn J. Tat, a novel regulator of HIV transcription and latency. HIV Sequence Compendium. Los Alamos National Laboratory. 2000.

43. Samson M, Labbe O, Mollereau C, Vassart G, Parmentier M. Molecular Cloning and Functional Expression of a New Human CC-Chemokine Receptor Gene. Biochemistry. 1996 35(11):3362–3367.

44. Broder CC, Collman RG. Chemokine receptors and HIV. Journal of Leukocyte Biology. 1997 62(1):20.

45. Marchand C, Johnson AA, Semenova E, Pommier Y. Mechanisms and inhibition of HIV integration. Drug Discovery Today Disease Mechanisms. 2006 3(2):253–260.

46. Chow SA, Vincent KA, Ellison V, Brown PO. Reversal of integration and DNA splicing mediated by integrase of human immunodeficiency virus. Science. 1992 255(5045):723–726.

47. Wu X, Liu H, Xiao H, Conway JA, Hehl E, Kalpana GV, et al. Human immunodeficiency virus type 1 integrase protein promotes reverse transcription through specific interactions with the nucleoprotein reverse transcription complex. The Journal of Virology. 1999 73(3):2126–2135.

48. Zhu K, Dobard C, Chow SA. Requirement for integrase during reverse transcription of human immunodeficiency virus type 1 and the effect of cysteine mutations of integrase on its interactions with reverse transcriptase. The Journal of Virology. 2004 78(10):5045–5055.

49. Engelman A, Bushman FD, Craigie R. Identification of discrete functional domains of HIV-1 integrase and their organization within an active multimeric complex. The EMBO Journal. 1993 12(8):3269–3275.

50. Ellison V, Gerton J, Vincent KA, Brown PO. An essential interaction between distinct domains of HIV-1 integrase mediates assembly of the active multimer. Journal of Biological Chemistry. 1995 270(7):3320.

51. Burke CJ, Sanyal G, Bruner MW, Ryan JA, LaFemina RL, Robbins HL, et al. Structural implications of spectroscopic characterization of a putative zinc finger peptide from HIV-1 integrase. Journal of Biological Chemistry. 1992 267(14):9639–9644.

52. Cai M, Zheng R, Caffrey M, Craigie R, Clore GM, Gronenborn AM. Solution structure of the N-terminal zinc binding domain of HIV-1 integrase. Nature Structural and Molecular Biology. 1997 4(7):567–577.

53. Zheng R, Jenkins TM, Craigie R. Zinc folds the N-terminal domain of HIV-1 integrase, promotes multimerization, and enhances catalytic activity. Proceedings of the National Academy of Sciences. 1996 93(24):13659–13664.

54. Lee SP, Xiao J, Knutson JR, Lewis MS, Han MK. Zn2+ promotes the self-association of human immunodeficiency virus type-1 integrase in vitro. Biochemistry. 1997 36(1):173–180.

55. Lee SP, Han MK. Zinc stimulates Mg2+-dependent 3'-processing activity of human immunodeficiency virus type 1 integrase in vitro. Biochemistry. 1996 35(12):3837–3844.

56. Kulkosky J, Jones KS, Katz RA, Mack JP, Skalka AM. Residues critical for retroviral integrative recombination in a region that is highly conserved among retroviral/retrotransposon integrases and bacterial insertion sequence transposases. Molecular and Cellular Biology. 1992 12(5):2331.

57. Drelich M, Wilhelm R, Mous J. Identification of amino acid residues critical for endonuclease and integration activities of HIV-1 IN protein in vitro. Virology. 1992 188(2):459–468.

58. van Gent DC, Vink C, Groeneger AA, Plasterk RH. Complementation between HIV integrase proteins mutated in different domains. The EMBO Journal. 1993 12(8):3261–3267.

59. Engelman A, Craigie R. Identification of conserved amino acid residues critical for human immunodeficiency virus type 1 integrase function in vitro. The Journal of Virology. 1992 66(11):6361–6369.

60. Esposito D, Craigie R. Sequence specificity of viral end DNA binding by HIV-1 integrase reveals critical regions for protein-DNA interaction. The EMBO Journal. 1998 17(19):5832–5843.

61. Jenkins TM, Esposito D, Engelman A, Craigie R. Critical contacts between HIV-1 integrase and viral DNA identified by structure-based analysis and photo-crosslinking. The EMBO Journal. 1997 16(22):6849–6859.

62. Heuer TS, Brown PO. Mapping features of HIV-1 integrase near selected sites on viral and target DNA molecules in an active enzyme-DNA complex by photo-cross-linking. Biochemistry. 1997 36(35):10655–10665.

63. Drake RR, Neamati N, Hong H, Pilon AA, Sunthankar P, Hume SD, et al. Identification of a nucleotide binding site in HIV-1 integrase. Proceedings of the National Academy of Sciences. 1998 95(8):4170–4175.

64. Johnson AA, Santos W, Pais GCG, Marchand C, Amin R, Burke TR, et al. Integration requires a specific interaction of the donor DNA terminal 5'-cytosine with glutamine 148 of the HIV-1 integrase flexible loop. Journal of Biological Chemistry. 2006 281(1):461–467.

65. Li M, Craigie R. Processing of viral DNA ends channels the HIV-1 integration reaction to concerted integration. Journal of Biological Chemistry. 2005 280(32):29334–29339.

66. Maignan S, Guilloteau JP, Zhou-Liu Q, Clément-Mella C, Mikol V. Crystal structures of the catalytic domain of HIV-1 integrase free and complexed with its metal cofactor: high level of similarity of the active site with other viral integrases1. Journal of Molecular Biology. 1998 282(2):359–368.

67. Goldgur Y, Dyda F, Hickman AB, Jenkins TM, Craigie R, Davies DR. Three new structures of the core domain of HIV-1 integrase: an active site that binds magnesium. Proceedings of the National Academy of Sciences. 1998 95(16):9150.

68. Lodi PJ, Ernst JA, Kuszewski J, Hickman AB, Engelman A, Craigie R, et al. Solution structure of the DNA binding domain of HIV-1 integrase. Biochemistry. 1995 34(31):9826–9833.

69. Chen JCH, Krucinski J, Miercke LJW, Finer-Moore JS, Tang AH, Leavitt AD, et al. Crystal structure of the HIV-1 integrase catalytic core and C-terminal domains: a model for viral DNA binding. Proceedings of the National Academy of Sciences. 2000 97(15):8233.

70. Wang JY, Ling H, Yang W, Craigie R. Structure of a two-domain fragment of HIV-1 integrase: implications for domain organization in the intact protein. The EMBO Journal. 2001 20(24):7333–7343.

71. van den Ent FM, Vos A, Plasterk RH. Dissecting the role of the N-terminal domain of human immunodeficiency virus integrase by trans-complementation analysis. The Journal of Virology. 1999 73(4):3176–3183.

72. Fletcher TM, Soares MA, McPhearson S, Hui H, Wiskerchen M, Muesing MA, et al. Complementation of integrase function in HIV-1 virions. The EMBO Journal. 1997 16(16):5123–5138.

73. Rice P, Mizuuchi K. Structure of the bacteriophage Mu transposase core: a common structural motif for DNA transposition and retroviral integration. Cell. 1995 82(2):209–220.

74. Aldaz H, Schuster E, Baker TA. The interwoven architecture of the Mu transposase couples DNA synapsis to catalysis. Cell. 1996 85(2):257–269.

75. Williams TL, Jackson EL, Carritte A, Baker TA. Organization and dynamics of the Mu transpososome: recombination by communication between two active sites. Genes and Development. 1999 13(20):2725.

76. Deprez E, Tauc P, Leh H, Mouscadet JF, Auclair C, Hawkins ME, et al. DNA binding induces dissociation of the multimeric form of HIV-1 integrase: a time-resolved fluorescence anisotropy study. Proceedings of the National Academy of Sciences. 2001 98(18):10090–10095.

77. Guiot E, Carayon K, Delelis O, Simon F, Tauc P, Zubin E, et al. Relationship between the oligomeric status of HIV-1 integrase on DNA and enzymatic activity. Journal Biological Chemistry. 2006 281(32):22707–22719.

78. Faure A, Calmels C, Desjobert C, Castroviejo M, Caumont-Sarcos A, Tarrago-Litvak L, et al. HIV-1 integrase crosslinked oligomers are active in vitro. Nucleic Acids Research. 2005 33(3):977–986.

79. Baranova S, Tuzikov FV, Zakharova OD, Tuzikova NA, Calmels C, Litvak S, et al. Small-angle X-ray characterization of the nucleoprotein complexes resulting from DNA-induced oligomerization of HIV-1 integrase. Nucleic Acids Research. 2007 35(3):975–987.

80. Li M, Mizuuchi M, Burke TR, Craigie R. Retroviral DNA integration: reaction pathway and critical intermediates. The EMBO Journal. 2006 25(6):1295–1304.

81. Alian A, Griner SL, Chiang V, Tsiang M, Jones G, Birkus G, et al. Catalytically-active complex of HIV-1 integrase with a viral DNA substrate binds anti-integrase drugs. Proceedings of the National Academy of Sciences. 2009 106(20):8192.

82. Hare S, Gupta SS, Valkov E, Engelman A, Cherepanov P. Retroviral intasome assembly and inhibition of DNA strand transfer. Nature. 2010 464(7286):232-236

83. Gao K, Butler SL, Bushman F. Human immunodeficiency virus type 1 integrase: arrangement of protein domains in active cDNA complexes. The EMBO Journal. 2001 20(13):3565–3576.

84. Podtelezhnikov AA, Gao K, Bushman FD, McCammon JA. Modeling HIV-1 integrase complexes based on their hydrodynamic properties. Biopolymers. 2003 68(1):110–120.

85. Wielens J, Crosby IT, Chalmers DK. A three-dimensional model of the human immunodeficiency virus type 1 integration complex. Journal of Computer-Aided Molecular Design. 2005 19(5):301–317.

86. Ren G, Gao K, Bushman FD, Yeager M. Single-particle image reconstruction of a tetramer of HIV integrase bound to DNA. Journal of Molecular Biology. 2007 366(1):286–294.

87. Hayouka Z, Rosenbluh J, Levin A, Loya S, Lebendiker M, Veprintsev D, et al. Inhibiting HIV-1 integrase by shifting its oligomerization equilibrium. Proceedings of the National Academy of Sciences. 2007 104(20):8316–8321.

88. Sherman PA, Fyfe JA. Human immunodeficiency virus integration protein expressed in Escherichia coli possesses selective DNA cleaving activity. Proceedings of the National Academy of Sciences. 1990 87(13):5119–5123.

89. LaFemina RL, Callahan PL, Cordingley MG. Substrate specificity of recombinant human immunodeficiency virus integrase protein. The Journal of Virology. 1991 65(10):5624–5630.

90. Brown PO. Integration of retroviral DNA. Current Topics in Microbiology and Immunology. 1990 157:19–48.

91. Ellison V, Abrams H, Roe T, Lifson J, Brown P. Human immunodeficiency virus integration in a cell-free system. The Journal of Virology. 1990 64(6):2711.

92. Bushman FD, Fujiwara T, Craigie R. Retroviral DNA integration directed by HIV integration protein in vitro. Science. 1990 249(4976):1555–1558.

93. Bushman FD, Craigie R. Activities of human immunodeficiency virus (HIV) integration protein in vitro: specific cleavage and integration of HIV DNA. Proceedings of the National Academy of Sciences. 1991 88(4):1339–1343.

94. Pommier Y, Johnson AA, Marchand C. Integrase inhibitors to treat HIV/AIDS. Nature Reviews Drug Discovery. 2005 4(3):236–248.

95. Merkel G, Andrake MD, Ramcharan J, Skalka AM. Oligonucleotide-based assays for integrase activity. Methods. 2009 47(4):243–248.

96. Miller MD, Farnet CM, Bushman FD. Human Immunodeficiency Virus type 1 preintegration complexes: studies of organization and composition. The Journal of Virology. 1997 71(7):5382.

97. Lin C-W, Engelman A. The barrier-to-autointegration factor is a component of functional human immunodeficiency virus type 1 preintegration complexes. The Journal of Virology. 2003 77(8):5030–5036.

98. Bushman FD. Host Proteins in Retroviral DNA Integration. Advances in Virus Research. 1999 52:301–317.

99. Kalpana GV, Marmon S, Wang W, Crabtree GR, Goff SP. Binding and stimulation of HIV-1 integrase by a human homolog of yeast transcription factor SNF5. Science. 1994 266(5193):2002–2006.

100. Farnet CM, Bushman FD. HIV-1 cDNA integration: requirement of HMG I (Y) protein for function of preintegration complexes in vitro. Cell. 1997 88(4):483–492.

101. Gao K, Gorelick RJ, Johnson DG, Bushman F. Cofactors for human immunodeficiency virus type 1 cDNA integration in vitro. The Journal of Virology. 2003 77(2):1598.

102. Li L, Yoder K, Hansen MST, Olvera J, Miller MD, Bushman FD. Retroviral cDNA integration: stimulation by HMG I family proteins. The Journal of Virology. 2000 74(23):10965.

103. Llano M, Saenz DT, Meehan A, Wongthida P, Peretz M, Walker WH, et al. An essential role for LEDGF/p75 in HIV integration. Science. 2006 314(5798):461.

104. Maertens G, Cherepanov P, Pluymers W, Busschots K, De Clercq E, Debyser Z, et al. LEDGF/p75 is essential for nuclear and chromosomal targeting of HIV-1 integrase in human cells. Journal of Biological Chemistry. 2003 278(35):33528.

105. Van Maele B, Busschots K, Vandekerckhove L, Christ F, Debyser Z. Cellular co-factors of HIV-1 integration. Trends in Biochemical Sciences. 2006 31(2):98–105.

106. Gallay P, Swingler S, Song J, Bushman F, Trono D. HIV nuclear import is governed by the phosphotyrosine-mediated binding of matrix to the core domain of integrase. Cell. 1995 83(4):569–576.

107. Ciuffi A, Llano M, Poeschla E, Hoffmann C, Leipzig J, Shinn P, et al. A role for LEDGF/p75 in targeting HIV DNA integration. Nature Medicine. 2005 11(12):1287–1289.

108. Ciuffi A, Bushman FD. Retroviral DNA integration: HIV and the role of LEDGF/p75. Trends in Genetics. 2006 22(7):388–395.

109. Hombrouck A, De Rijck J, Hendrix J, Vandekerckhove L, Voet A, De Maeyer M, et al. Virus evolution reveals an exclusive role for LEDGF/p75 in chromosomal tethering of HIV. PLoS Pathogens. 2007 3(3):e47.

110. Stevens W, Papathanasopoulos M. Human Immunodeficiency Virus. In: Mendelow B, Ramsay M, Chetty N, Stevens W, editors. Molecular Medicine for Clinicians. Johannnesburg: Wits University Press; 2009. p. 311–324.

111. Hamers RL, Derdelinckx I, Van Vugt M, Stevens W, Dewit, R, Tobias F, et al. The status of HIV-1 resistance to antiretroviral drugs in sub-Saharan Africa. Antiviral Therapy. 2008 13(5):625–639.

112. Wensing AMJ, van Maarseveen NM, Nijhuis M. Fifteen years of HIV protease inhibitors: raising the barrier to resistance. Antiviral Research. 2010 85(1):59–74.

113. Mitsuya H, Weinhold KJ, Furman PA, St Clair MH, Lehrman SN, Gallo RC, et al. 3“-Azido-3-”deoxythymidine (BW A509U): an antiviral agent that inhibits the infectivity and cytopathic effect of human T-lymphotropic virus type III/lymphadenopathy-associated virus in vitro. Proceedings of the National Academy of Sciences. 1985 82(20):7096.

114. First combination drug for patients with HIV. American Family Physicians. 1997 56(9):2356.

115. What you need to know about Combivir. AIDS Alert. 1998 13(1)

116. Cihlar T, Ray AS. Nucleoside and nucleotide HIV reverse transcriptase inhibitors: 25 years after zidovudine. Antiviral Research. 2010 85(1):39–58.

117. Tilton JC, Doms RW. Entry inhibitors in the treatment of HIV-1 infection. Antiviral Research. 2010 85(1):91–100.

118. Lalezari JP, Henry K, O'Hearn M, Montaner JSG, Piliero PJ, Trottier B, et al. Enfuvirtide, an HIV-1 fusion inhibitor, for drug-resistant HIV infection in North and South America. New England Journal of Medicine. 2003 348(22):2175.

119. A once-daily combination tablet (Atripla) for HIV. The Medical Letter on Drugs and Therapeutics. 2006 48(1244):78–79.

120. Janssen PAJ, Lewi PJ, Arnold E, Daeyaert F, de Jonge M, Heeres J, et al. In search of a novel anti-HIV drug: multidisciplinary coordination in the discovery of 4-[[4-[[4-[(1E)-2-cyanoethenyl]-2,6-dimethylphenyl]amino]-2-pyrimidinyl]amino]benzonitrile (R278474, rilpivirine). Journal of Medical Chemistry. 2005 48(6):1901–1909.

121. Press Announcements - FDA approves new HIV treatment [Internet]. FDA. 2011 May 20 [cited 2011 Jul. 29]; Available from: http://www.fda.gov/NewsEvents/Newsroom/PressAnnouncements/ucm256087.htm

122. World Health Organisation World Health. Antiretroviral Therapy For HIV Infection in Adults and Adolescents. Austria: WHO; 2011.

123. National Department of Health, South Africa National Department of. The South African Antiretroviral Treatment Guidelines, 2010. 2011.

124. Summary Table of HIV Treatment Regimens [Internet]. [cited 2011 Jun. 20]; Available from: http://www.aidstar-one.com/focus_areas/treatment/resources/regimes_summary

125. Thompson MA, Aberg JA, Cahn P, Montaner JSG, Rizzardini G, Telenti A, et al. Antiretroviral treatment of adult HIV infection: 2010 recommendations of the International AIDS Society-USA panel. The Journal of the American Medical Association. 2010 304(3):321–333.

126. McColl DJ, Chen X. Strand transfer inhibitors of HIV-1 integrase: bringing IN a new era of antiretroviral therapy. Antiviral Research. 2010 85(1):101–118.

127. Espeseth AS, Felock P, Wolfe A, Witmer M, Grobler J, Anthony N, et al. HIV-1 integrase inhibitors that compete with the target DNA substrate define a unique strand transfer conformation for integrase. Proceedings of the National Academy of Sciences. 2000 97(21):11244–11249.

128. Brooun A, Richman DD, Kornbluth RS. HIV-1 preintegration complexes preferentially integrate into longer target DNA molecules in solution as detected by

a sensitive, polymerase chain reaction-based integration assay. Journal of Biological Chemistry. 2001 276(50):46946–46952.

129. Farnet CM, Haseltine WA. Integration of human immunodeficiency virus type 1 DNA in vitro. Proceedings of the National Academy of Sciences. 1990 87(11):4164–4168.

130. Shimura K, Kodama E, Sakagami Y. Broad antiretroviral activity and resistance profile of the novel human immunodeficiency virus integrase inhibitor elvitegravir (JTK-303/GS-9137). Journal of Virology. 2008 82(2):764-774.

131. Agosto L, Yu J, Dai J, Kaletsky R, Monie D. HIV-1 integrates into resting CD4+ T cells even at low inoculums as demonstrated with an improved assay for HIV-1 integration. Virology. 2007 368(1):60-72.

132. Hazuda DJ, Hastings JC, Wolfe AL, Emini EA. A novel assay for the DNA strand-transfer reaction of HIV-1 integrase. Nucleic Acids Research. 1994 22(6):1121.

133. Sinha S, Grandgenett DP. 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. The Journal of Virology. 2005 79(13):8208–8216.

134. Deprez E, Tauc P, Leh H, Mouscadet JF, Auclair C, Brochon JC. Oligomeric states of the HIV-1 integrase as measured by time-resolved fluorescence anisotropy. Biochemistry. 2000 39(31):9275–9284.

135. Sioud M, Drlica K. Prevention of human immunodeficiency virus type 1 integrase expression in Escherichia coli by a ribozyme. Proceedings of the National Academy of Sciences. 1991 88(16):7303–7307.

136. Raillard SA, Joyce GF. Targeting sites within HIV-1 cDNA with a DNA-cleaving ribozyme. Biochemistry. 1996 35(36):11693–11701.

137. Bouziane M, Cherny DI, Mouscadet JF, Auclair C. Alternate strand DNA triple helix-mediated inhibition of HIV-1 U5 long terminal repeat integration in vitro. Journal of Biological Chemistry. 1996 271(17):10359.

138. Katzman M, Katz RA, Skalka AM, Leis J. The avian retroviral integration protein cleaves the terminal sequences of linear viral DNA at the in vivo sites of integration. The Journal of Virology. 1989 63(12):5319.

139. Katz RA, Merkel G, Kulkosky J, Leis J, Skalka AM. The avian retroviral IN protein is both necessary and sufficient for integrative recombination in vitro. Cell. 1990 63(1):87–95.

140. Craigie R, Fujiwara T, Bushman F. The IN protein of Moloney murine leukemia virus processes the viral DNA ends and accomplishes their integration in vitro. Cell. 1990 62(4):829–837.

141. Asante-Appiah E, Skalka A. Divalent cations stimulate preferential recognition of a viral DNA end by HIV-1 integrase. Biochemistry. 1999 38(26):8458-68.

142. Müller B, Jones K, Merkel G. Rapid solution assays for retroviral integration reactions and their use in kinetic analyses of wild-type and mutant Rous sarcoma virus integrases. Proceedings of the National Academy of Sciences. 1993 90(24):11633-7.

143. Mazumder A, Pommier Y, Balis F. Incorporation of a fluorescent guanosine analog into olignoucleotides and its application to a real time assay for the HIV-1 integrase 3′-Processing reaction. Nucleic Acids Research. 1995 23(15):2872-80.

144. Smolov M, Gottikh M, Tashlitskii V, Korolev S. Kinetic study of the HIV– 1 DNA 3′-end processing. FEBS Journal. 2006 273(6):1137-51.

145. He H, Ma X, Liu B, Chen W. A novel high-throughput format assay for HIV-1 integrase strand transfer reaction using magnetic beads. Acta Pharmacologica. 2008 29(3):397-404.

146. Craigie R. A rapid in vitro assay for HIV DNA integration. Nucleic Acids Research. 1991 19(10):2729-34.

147. Robinson WE, Cordeiro M, Abdel-Malek S, Jia Q, Chow SA, Reinecke MG, et al. Dicaffeoylquinic acid inhibitors of human immunodeficiency virus integrase: inhibition of the core catalytic domain of human immunodeficiency virus integrase. Molecular Pharmacology. 1996 50(4):846.

148. Robinson WE, Reinecke MG, Abdel-Malek S, Jia Q, Chow SA. Inhibitors of HIV-1 replication that inhibit HIV integrase. Proceedings of the National Academy of Sciences. 1996 93(13):6326.

149. King PJ, Robinson WE Jr. Resistance to the anti-human immunodeficiency virus type 1 compound L-chicoric acid results from a single mutation at amino acid 140 of integrase. The Journal of Virology. 1998 72(10):8420.

150. Goldgur Y, Craigie R, Cohen GH, Fujiwara T, Yoshinaga T, Fujishita T, et al. Structure of the HIV-1 integrase catalytic domain complexed with an inhibitor: a platform for antiviral drug design. Proceedings of the National Academy of Sciences. 1999 96(23):13040.

151. Vandekerckhove L. GSK-1349572, a novel integrase inhibitor for the treatment of HIV infection. Current Opinion in Investigational Drugs. 2010 11(2):203.

152. Jaskolski M, Alexandratos J, Bujacz G. Piecing together the structure of retroviral integrase, an important target in AIDS therapy. FEBS Journal. 2009 276(11):2926-46.

153. Hazuda DJ, Felock P, Witmer M, Wolfe A, Stillmock K, Grobler JA, et al. Inhibitors of strand transfer that prevent integration and inhibit HIV-1 replication in cells. Science. 2000 287(5453):646–650.

154. Little S, Saah A, Drusano G, Schooley R, Hass D, Kumar P, et al. Antiviral effect of L-000870810, a novel HIV-1 integrase inhibitor [Internet]. In: CROI. Boston: 2005. Available from: http://retroconference.org/2005/cd/Abstracts/24615.htm

155. Summa V, Petrocchi A, Bonelli F, Crescenzi B, Donghi M, Ferrara M, et al. Discovery of raltegravir, a potent, selective orally bioavailable HIV-integrase inhibitor for the treatment of HIV-AIDS infection. Journal of Medical Chemistry. 2008 51(18):5843–5855.

156. Lennox JL, DeJesus E, Lazzarin A, Pollard RB, Madruga JVR, Berger DS, et al. Safety and efficacy of raltegravir-based versus efavirenz-based combination

therapy in treatment-naive patients with HIV-1 infection: a multicentre, double-blind randomised controlled trial. The Lancet. 2009 374(9692):796–806.

157. Serrao E, Odde S, Ramkumar K, Neamati N. Raltegravir, elvitegravir, and metoogravir: the birth of “me-too” HIV-1 integrase inhibitors. Retrovirology. 2009 6:25.

158. Vacca J, Wai J, Fisher T, Embrey M, Hazuda D. Discovery of MK-2048: subtle changes confer unique resistance properties to a series of tricyclic hydroxypyrrole integrase strand transfer inhibitors. WEPEA088. 4th IAS Conf.; 2007.

159. Wai J, Fisher T, Embrey M, Egbertson M, Vacca J, Hazuda D, Miller M, Witmer M, Gabryelski L, Lyle T. Next generation of inhibitors of HIV-1 integrase strand transfer inhibitor: structural diversity and resistance profiles. In *14th Conference on Retroviral and Opportunistic Infections, Los Angeles CA, USA*. 2007.

160. DeJesus E, Berger D, Markowitz M, Cohen C, Hawkins T, Ruane P, et al. Antiviral activity, pharmacokinetics, and dose response of the HIV-1 integrase inhibitor GS-9137 (JTK-303) in treatment-naive and treatment-experienced patients. Journal of Acquired Immune Deficiency Syndromes. 2006 43(1):1.

161. Zolopa AR, Lampiris H, Blick G, Walworth C, Zhong L, Chuck SL, et al. The HIV integrase inhibitor elvitegravir (EVG/r) has potent and durable activity in treatment-experienced patients with active optimized background therapy (OBT). In: *The 47th Annual Interscience Conference on Antimcrobial Agents and Chemotherapy*. 2007.

162. Valkov E, Gupta SS, Hare S, Helander A, Roversi P, McClure M, et al. Functional and structural characterization of the integrase from the prototype foamy virus. Nucleic Acids Research. 2009 37(1):243–255.

163. Bonnenfant S, Thomas CM, Vita C, Subra F, Deprez E, Zouhiri F, et al. Styrylquinolines, integrase inhibitors acting prior to integration: a new mechanism of action for anti-integrase agents. The Journal of Virology. 2004 78(11):5728–5736.

164. Pannecouque C, Pluymers W, Van Maele B, Tetz V, Cherepanov P, De Clercq E, et al. New class of HIV integrase inhibitors that block viral replication in cell culture. Current Biology. 2002 12(14):1169–1177.

165. Shaw-Reid CA, Munshi V, Graham P, Wolfe A, Witmer M, Danzeisen R, et al. Inhibition of HIV-1 ribonuclease H by a novel diketo acid, 4-[5-(benzoylamino)thien-2-yl]-2,4-dioxobutanoic acid. Journal of Biological Chemistry. 2003 278(5):2777–2780.

166. Budihas SR, Gorshkova I, Gaidamakov S, Wamiru A, Bona MK, Parniak MA, et al. Selective inhibition of HIV-1 reverse transcriptase-associated ribonuclease H activity by hydroxylated tropolones. Nucleic Acids Research. 2005 33(4):1249–1256.

167. Didierjean J, Isel C, Querré F, Mouscadet J-F, Aubertin A-M, Valnot J-Y, et al. Inhibition of human immunodeficiency virus type 1 reverse transcriptase, RNase H, and integrase activities by hydroxytropolones. Antimicrobial Agents and Chemotherapy. 2005 49(12):4884–4894.

168. Semenova EA, Johnson AA, Marchand C, Davis DA, Yarchoan R, Pommier Y. Preferential inhibition of the magnesium-dependent strand transfer reaction of HIV-1 integrase by alpha-hydroxytropolones. Molecular Pharmacology. 2006 69(4):1454–1460.

169. Marchand C, Beutler JA, Wamiru A, Budihas S, Möllmann U, Heinisch L, et al. Madurahydroxylactone derivatives as dual inhibitors of human immunodeficiency virus type 1 integrase and RNase H. Antimicrobial Agents and Chemotherapy. 2008 52(1):361–364.

170. Billamboz M, Bailly F, Barreca ML, De Luca L, Mouscadet J-F, Calmels C, et al. Design, synthesis, and biological evaluation of a series of 2-hydroxyisoquinoline-1,3(2H,4H)-diones as dual inhibitors of human immunodeficiency virus type 1 integrase and the reverse transcriptase RNase H domain. Journal of Medical Chemistry. 2008 51(24):7717–7730.

171. Puras Lutzke RA, Eppens NA, Weber PA, Houghten RA, Plasterk RH. Identification of a hexapeptide inhibitor of the human immunodeficiency virus

integrase protein by using a combinatorial chemical library. Proceedings of the National Academy of Sciences. 1995 92(25):11456–11460.

172. Long Y, Lung F, Pommier Y, Neamati N. Tryptophan-rich integrase inhibitory peptides and peptidomimetics. In: Martinez JaF, J-A, editors. Peptides 2000. Paris: EDK; 2001. pp. 725–726.

173. Krajewski K, Long Y-Q, Marchand C, Pommier Y, Roller PP. Design and synthesis of dimeric HIV-1 integrase inhibitory peptides. Bioorganic and Medicinal Chemistry Letters. 2003 13(19):3203–3205.

174. de Soultrait VR, Caumont A, Parissi V, Morellet N, Ventura M, Lenoir C, et al. A novel short peptide is a specific inhibitor of the human immunodeficiency virus type 1 integrase. Journal of Molecular Biology. 2002 318(1):45–58.

175. Singh SB, Herath K, Guan Z, Zink DL, Dombrowski AW, Polishook JD, et al. Integramides A and B, two novel non-ribosomal linear peptides containing nine C(alpha)-methyl amino acids produced by fungal fermentations that are inhibitors of HIV-1 integrase. Organic Letters. 2002 4(9):1431–1434.

176. De Zotti M, Formaggio F, Kaptein B, Broxterman QB, Felock PJ, Hazuda DJ, et al. Complete absolute configuration of integramide A, a natural, 16-mer peptide inhibitor of HIV-1 integrase, elucidated by total synthesis. ChemBioChem. 2009 10(1):87–90.

177. Robinson WE, McDougall B, Tran D, Selsted ME. Anti-HIV-1 activity of indolicidin, an antimicrobial peptide from neutrophils. Journal of Leukocyte Biology. 1998 63(1):94–100.

178. Krajewski K, Marchand C, Long Y-Q, Pommier Y, Roller PP. Synthesis and HIV-1 integrase inhibitory activity of dimeric and tetrameric analogs of indolicidin. Bioorganic and Medicinal Chemistry Letters. 2004 14(22):5595–5598.

179. Marchand C, Krajewski K, Lee H-F, Antony S, Johnson AA, Amin R, et al. Covalent binding of the natural antimicrobial peptide indolicidin to DNA abasic sites. Nucleic Acids Researcg. 2006 34(18):5157–5165.

180. Desjobert C, de Soultrait VR, Faure A, Parissi V, Litvak S, Tarrago-Litvak L, et al. Identification by phage display selection of a short peptide able to inhibit only

the strand transfer reaction catalyzed by human immunodeficiency virus type 1 integrase. Biochemistry. 2004 43(41):13097–13105.

181. Maroun RG, Gayet S, Benleulmi MS, Porumb H, Zargarian L, Merad H, et al. Peptide inhibitors of HIV-1 integrase dissociate the enzyme oligomers. Biochemistry. 2001 40(46):13840–13848.

182. Sourgen F, Maroun RG, Frère V, Bouziane M, Auclair C, Troalen F, et al. A synthetic peptide from the human immunodeficiency virus type-1 integrase exhibits coiled-coil properties and interferes with the in vitro integration activity of the enzyme. Correlated biochemical and spectroscopic results. European Journal of Biochemistry. 1996 240(3):765–773.

183. Krebs D, Maroun RG, Sourgen F, Troalen F, Davoust D, Fermandjian S. Helical and coiled-coil-forming properties of peptides derived from and inhibiting human immunodeficiency virus type 1 integrase assessed by 1H-NMR--use of NH temperature coefficients to probe coiled-coil structures. European Journal of Biochemistry. 1998 253(1):236–244.

184. Maroun RG, Krebs D, Roshani M, Porumb H, Auclair C, Troalen F, et al. Conformational aspects of HIV-1 integrase inhibition by a peptide derived from the enzyme central domain and by antibodies raised against this peptide. European Journal of Biochemistry. 1999 260(1):145–155.

185. Yung E, Sorin M, Pal A, Craig E, Morozov A, Delattre O, et al. Inhibition of HIV-1 virion production by a transdominant mutant of integrase interactor 1. Nature Medicine. 2001 (8):920–926.

186. Sorin M, Yung E, Wu X, Kalpana GV. HIV-1 replication in cell lines harboring INI1/hSNF5 mutations. Retrovirology. 2006 3:56.

187. Maroun M, Delelis O, Coadou G, Bader T, Ségéral E, Mbemba G, et al. Inhibition of early steps of HIV-1 replication by SNF5/Ini1. Journal of Biological Chemistry. 2006 281(32):22736–22743.

188. De Rijck J, Vandekerckhove L, Gijsbers R, Hombrouck A, Hendrix J, Vercammen J, et al. Overexpression of the lens epithelium-derived growth

factor/p75 integrase binding domain inhibits human immunodeficiency virus replication. The Journal of Virology. 2006 80(23):11498–11509.

189. De Luca L, Barreca ML, Ferro S, Christ F, Iraci N, Gitto R, et al. Pharmacophore-based discovery of small-molecule inhibitors of protein-protein interactions between HIV-1 integrase and cellular cofactor LEDGF/p75. ChemMedChem. 2009 (8):1311–1316.

190. Oz Gleenberg I, Avidan O, Goldgur Y, Herschhorn A, Hizi A. Peptides derived from the reverse transcriptase of human immunodeficiency virus type 1 as novel inhibitors of the viral integrase. Journal of Biological Chemistry. 2005 280(23):21987–21996.

191. Rosenbluh J, Hayouka Z, Loya S, Levin A, Armon-Omer A, Britan E, et al. Interaction between HIV-1 Rev and integrase proteins: a basis for the development of anti-HIV peptides. Journal of Biological Chemistry. 2007 282(21):15743–15753.

192. Gleenberg IO, Herschhorn A, Hizi A. Inhibition of the activities of reverse transcriptase and integrase of human immunodeficiency virus type-1 by peptides derived from the homologous viral protein R (Vpr). Journal of Molecular Biology. 2007 369(5):1230–1243.

193. Bizub-Bender D, Kulkosky J, Skalka AM. Monoclonal antibodies against HIV type 1 integrase: clues to molecular structure. AIDS Research and Human Retroviruses. 1994 10(9):1105–1115.

194. Yi J, Arthur JW, Dunbrack RL, Skalka AM. An inhibitory monoclonal antibody binds at the turn of the helix-turn-helix motif in the N-terminal domain of HIV-1 integrase. Journal of Biological Chemistry. 2000 275(49):38739–38748.

195. Yi J, Cheng H, Andrake MD, Dunbrack RL, Roder H, Skalka AM. Mapping the epitope of an inhibitory monoclonal antibody to the C-terminal DNA-binding domain of HIV-1 integrase. Journal of Biological Chemistry. 2002 277(14):12164–12174.

196. Ramcharan J, Colleluori DM, Merkel G, Andrake MD, Skalka AM. Mode of inhibition of HIV-1 Integrase by a C-terminal domain-specific monoclonal antibody. Retrovirology. 2006 3:34.

197. Marchand C, Maddali K, Métifiot M, Pommier Y. HIV-1 IN inhibitors: 2010 update and perspectives. Current Topics in Medical Chemistry. 2009 9(11):1016–1037.

198. Brodin P, Pinskaya M, Parsch U, Bischerour J, Leh H, Romanova E, et al. 6-oxocytidine containing oligonucleotides inhibit the HIV-1 integrase in vitro. Nucleosides, Nucleotides and Nucleic Acids. 2001 20(4-7):481–486.

199. Brodin P, Pinskaya M, Buckle M, Parsch U, Romanova E, Engels J, et al. Disruption of HIV-1 integrase-DNA complexes by short 6-oxocytosine-containing oligonucleotides. Biochemistry. 2002 41(5):1529–1538.

200. Blanco J-L, Varghese V, Rhee S-Y, Gatell JM, Shafer RW. HIV-1 integrase inhibitor resistance and its clinical implications. Journal of Infectious Diseases. 2011 203(9):1204–1214.

201. Integrase Inhibitor Resistance Notes [Internet]. HIV Drug Resistance Database – Stanford University. [cited 2011 Jun. 22] Available from: http://hivdb.stanford.edu//cgi-bin/INIResiNote.cgi

202. Huang H, Chopra R, Verdine GL, Harrison SC. Structure of a covalently trapped catalytic complex of HIV-1 reverse transcriptase: implications for drug resistance. Science. 1998 282(5394):1669.

203. Jenkins TM, Engelman A, Ghirlando R, Craigie R. A soluble active mutant of HIV-1 integrase: involvement of both the core and carboxyl-terminal domains in multimerization. Journal of Biological Chemistry. 1996 271(13):7712–7718.

204. Fish MQ, Hewer R, Wallis CL, Venter WDF, Stevens WS, Papathanasopoulos MA. Natural polymorphisms of integrase among HIV type 1-infected South African patients. AIDS Research and Human Retroviruses. 2010 26(4):489–493.

205. Birnboim HC, Doly J. A rapid alkaline extraction procedure for screening recombinant plasmid DNA. Nucleic Acids Res. 1979 Nov. 24;7(6):1513–1523.

206. Mandel M, Higa A. Calcium-dependent bacteriophage DNA infection. Journal of Molecular Biology. 1970 53(1):159–162.

207. Laemmli UK. Most commonly used discontinuous buffer system for SDS electrophoresis. Nature. 1970 227:680–685.

208. Wilson CM. Staining of proteins on gels: comparisons of dyes and procedures. Methods in Enzymology. 1983 91:236.

209. Zhao Z, McKee CJ, Kessl JJ, Santos WL, Daigle JE, Engelman A, et al. Subunit-specific protein footprinting reveals significant structural rearrangements and a role for N-terminal Lys-14 of HIV-1 Integrase during viral DNA binding. Journal of Biological Chemistry. 2008 283(9):5632–5641.

210. Maxwell KL, Mittermaier AK, Forman-Kay JD, Davidson AR. A simple in vivo assay for increased protein solubility. Protein Science. 1999 8(9):1908–1911.

211. Yin J, Li G, Ren X, Herrler G. Select what you need: a comparative evaluation of the advantages and limitations of frequently used expression systems for foreign genes. Journal of Biotechnology. 2007 127(3):335–347.

212. Bushman FD, Engelman A, Palmer I, Wingfield P, Craigie R. Domains of the integrase protein of human immunodeficiency virus type 1 responsible for polynucleotidyl transfer and zinc binding. Proceedings of the National Academy of Sciences. 1993 90(8):3428–3432.

213. Hickman AB, Palmer I, Engelman A, Craigie R, Wingfield P. Biophysical and enzymatic properties of the catalytic domain of HIV-1 integrase. Journal of Biological Chemistry. 1999 269(46):29279–29287.

214. Sastry SS, Ross BM. Nuclease activity of T7 RNA polymerase and the heterogeneity of transcription elongation complexes. Journal of Biological Chemistry. 1997 272(13):8644.

215. Studier FW, Moffatt BA. Use of bacteriophage T7 RNA polymerase to direct selective high-level expression of cloned genes. Journal of Molecular Biology. 1986 189(1):113–130.

216. Gilbert W, Müller-Hill B. Isolation of the lac repressor. Proceedings of the National Academy of Sciences. 1966 56(6):1891.

217. Porath J, Carlsson JAN, Olsson I, Belfrage G. Metal chelate affinity chromatography, a new approach to protein fractionation. Nature. 1975 258:598–599.

218. Terpe K. Overview of tag protein fusions: from molecular and biochemical fundamentals to commercial systems. Applied Microbiology and Biotechnology. 2003 60(5):523–533.

219. Bar-Magen T, Sloan RD, Faltenbacher VH, Donahue DA, Kuhl BD, Oliveira M, et al. Comparative biochemical analysis of HIV-1 subtype B and C integrase enzymes. Retrovirology. 2009 6:103.

220. ELISA Technical Guide [Internet]. KPL, Inc.; Available from: http://www.kpl.com/docs/techdocs/KPL%20ELISA%20Technical%20Guide.pdf

221. Mohammed KD, Topper MB, Muesing MA. Sequential deletion of the integrase (Gag-Pol) carboxyl terminus reveals distinct phenotypic classes of defective HIV-1. The Journal of Virology. 2011 85(10):4654–4666.

222. Gupta K, Diamond T, Hwang Y, Bushman F, Van Duyne GD. Structural properties of HIV integrase. Lens epithelium-derived growth factor oligomers. Journal of Biological Chemistry. 2010 285(26):20303–20315.

223. Tsiang M, Jones GS, Hung M, Mukund S, Han B, Liu X, et al. Affinities between the binding partners of the HIV-1 integrase dimer-lens epithelium-derived growth factor (IN dimer-LEDGF) complex. Journal of Biological Chemistry. 2009 284(48):33580–33599.

224. Nishitsuji H, Hayashi T, Takahashi T, Miyano M, Kannagi M, Masuda T. Augmentation of reverse transcription by integrase through an interaction with host factor, SIP1/Gemin2 Is critical for HIV-1 infection. PLoS ONE. 2009 4(11):7825.

225. Ebina H, Chatterjee AG, Judson RL, Levin HL. The GP(Y/F) domain of TF1 integrase multimerizes when present in a fragment, and substitutions in this domain reduce enzymatic activity of the full-length protein. Journal of Biological Chemistry. 2008 283(23):15965–15974.

226. Berthoux L, Sebastian S, Muesing MA, Luban J. The role of lysine 186 in HIV-1 integrase multimerization. Virology. 2007 364(1):227–236.

227. Guarné A, Brendler T, Zhao Q, Ghirlando R, Austin S, Yang W. Crystal structure of a SeqA-N filament: implications for DNA replication and chromosome organization. The EMBO Journal. 2005 24(8):1502–1511.

228. Phan AT, Kuryavyi V, Ma J-B, Faure A, Andréola M-L, Patel DJ. An interlocked dimeric parallel-stranded DNA quadruplex: a potent inhibitor of HIV-1 integrase. Proceedings of the National Academy of Sciences. 2005 102(3):634–639.

229. Lu R, Limón A, Devroe E, Silver PA, Cherepanov P, Engelman A. Class II integrase mutants with changes in putative nuclear localization signals are primarily blocked at a postnuclear entry step of human immunodeficiency virus type 1 replication. The Journal of Virology. 2004 78(23):12735–12746.

230. Basbous J, Bazarbachi A, Granier C, Devaux C, Mesnard J-M. The central region of human T-cell leukemia virus type 1 Tax protein contains distinct domains involved in subunit dimerization. The Journal of Virology. 2003 77(24):13028–13035.

231. Limón A, Devroe E, Lu R, Ghory HZ, Silver PA, Engelman A. Nuclear localization of human immunodeficiency virus type 1 preintegration complexes (PICs): V165A and R166A are pleiotropic integrase mutants primarily defective for integration, not PIC nuclear import. The Journal of Virology. 2002 76(21):10598–10607.

232. Petit C, Schwartz O, Mammano F. The karyophilic properties of human immunodeficiency virus type 1 integrase are not required for nuclear import of proviral DNA. The Journal of Virology. 2000 74(15):7119–7126.

233. Carteau S, Gorelick RJ, Bushman FD. Coupled integration of human immunodeficiency virus type 1 cDNA ends by purified integrase in vitro: stimulation by the viral nucleocapsid protein. The Journal of Virology. 1999 73(8):6670–6679.

234. Gerton JL, Ohgi S, Olsen M, DeRisi J, Brown PO. Effects of mutations in residues near the active site of human immunodeficiency virus type 1 integrase on specific enzyme-substrate interactions. The Journal of Virology. 1998 72(6):5046–5055.

235. Engelman A, Liu Y, Chen H, Farzan M, Dyda F. Structure-based mutagenesis of the catalytic domain of human immunodeficiency virus type 1 integrase. The Journal of Virology. 1997 71(5):3507–3514.

236. Taddeo B, Carlini F, Verani P, Engelman A. Reversion of a human immunodeficiency virus type 1 integrase mutant at a second site restores enzyme function and virus infectivity. The Journal of Virology. 1996 70(12):8277–8284.

237. Delelis O, Thierry S, Subra F, Simon F, Malet I, Alloui C, et al. Impact of Y143 HIV-1 integrase mutations on resistance to raltegravir in vitro and in vivo. Antimicrobial Agents and Chemotherapy. 2010 54(1):491–501.

238. Marchand C, Johnson AA, Karki RG, Pais GCG, Zhang X, Cowansage K, et al. Metal-dependent inhibition of HIV-1 integrase by beta-diketo acids and resistance of the soluble double-mutant (F185K/C280S). Molecular Pharmacology. 2003 64(3):600–609.

239. Bar-Magen T, Donahue DA, McDonough EI, Kuhl BD, Faltenbacher VH, Xu H, et al. HIV-1 subtype B and C integrase enzymes exhibit differential patterns of resistance to integrase inhibitors in biochemical assays. AIDS. 2010 24(14):2171–2179.

240. Delelis O, Malet I, Na L, Tchertanov L, Calvez V, Marcelin AG, et al. The G140S mutation in HIV integrases from raltegravir-resistant patients rescues catalytic defect due to the resistance Q148H mutation. Nucleic Acids Research. 2008 37(4):1193–1201.

CHAPTER SIX

6 APPENDICES

Appendix A: WHO Clinical staging of HIV disease

Clinical stage 1

Asymptomatic Persistent generalized lymphadenopathy

Clinical stage 2

Moderate unexplained weight loss

Recurrent respiratory tract infections (sinusitis, tonsillitis, otitis media, pharyngitis)

Herpes zoster

Angular cheilitis

Recurrent oral ulcerations

Papular pruritic eruptions

Seborrhoeic dermatitis

Fungal nail infections

Clinical stage 3

Unexplained severe weight loss (over 10% of presumed or measured body weight)

Unexplained chronic diarrhoea for longer than 1 month

Unexplained persistent fever (intermittent or constant for longer than 1 month)

Persistent oral candidiasis

Oral hairy leukoplakia

Pulmonary tuberculosis

Severe bacterial infections (e.g. pneumonia, empyema, meningitis, pyomyositis, bone or joint infection, bacteraemia, severe pelvic inflammatory disease)

Acute necrotizing ulcerative stomatitis, gingivitis or periodontitis

Unexplained anaemia, neutropenia and/or chronic thrombocytopenia

Clinical stage 4

HIV wasting syndrome

Pneumocystis jiroveci pneumonia

Recurrent severe bacterial pneumonia

Chronic herpes simplex infection (orolabial, genital or anorectal of more than 1 month's duration or visceral at any site)

Oesophageal candidiasis (or candidiasis of trachea, bronchi or lungs)

Extrapulmonary tuberculosis

Kaposi sarcoma

Cytomegalovirus disease (retinitis or infection of other organs, excluding liver, spleen and lymph nodes)

Central nervous system toxoplasmosis HIV encephalopathy

Extrapulmonary cryptococcosis including meningitis

Disseminated nontuberculous mycobacteria infection

Progressive multifocal leukoencephalopathy

Chronic cryptosporidiosis

Chronic isosporiasis

Disseminated mycosis (histoplasmosis, coccidiomycosis)

Recurrent septicaemia (including nontyphoidal *Salmonella*)

Lymphoma (cerebral or B cell non-Hodgkin)

Invasive cervical carcinoma

Atypical disseminated leishmaniasis

Symptomatic HIV-associated nephropathy or HIV-associated cardiomyopathy

Source: Antiretroviral therapy for HIV infection in Adults And Adolescents: Recommendations for a Public Health Approach (122).

Appendix B: Strand Transfer Assay and Complexation Reagents

20X Saline Sodium Citrate (SSC) Buffer

300mM NaCl (Sigma Aldrich Germany)

300mM Sodium Citrate (BDH Chemicals, England)

pH 7.0

Reaction Buffer

10mM $MnCl_2$ (Sigma Aldrich, Germany)

10mM $MgCl_2$ (Merck, Germany)

20mM Hepes (Sigma Aldrich, Germany)

75mM NaCl (Sigma Aldrich, Germany)

2μM ZnCl (Associated Chemical Enterprises, South africa)

1% v/v Glycerol (Saarchem, South Africa)

pH7.3

Filter sterilized

Annealing Buffer

10mM Tris (Merck, Germany)

100mM NaCl (Sigma Aldrich, Germany)

pH 8.0

Filter sterilized

200mM Phosphate Buffer (pH 8.0)

0.1884% w/v Monosodium phosphate, monohydrate

4.994% w/v Disodium phosphate, heptahydrate

Activation Buffer 1

170mM Phosphate Buffer (above)

40mM DTT (Sigma Aldrich, Germany)

1% Glycerol (Saarchem, South Africa)

Filter sterilized

Activation Buffer 2

100mM Phosphate Buffer (above)

1mM 5,5'-dithiobis-(2-nitrobenzoic acid) (DTNB) (Sigma Aldrich, Germany)

Complexation Buffer

50mM Tris (Merck, Germany)

500mM NaCl (Sigma Aldrich, Germany)

1mM $MnCl_2$ (only added at the final step from 100mM stock) (Sigma Aldrich, Germany)

pH 8.0

Appendix C: Protein Expression and Purification Reagents

Transformation Buffer

100mM $CaCl_2$ (Merck, Germany)

10mM PIPES (Boehringer, Germany)

15% v/v Glycerol (Saarchem, South Africa)

pH 7.0

1000X Ampicillin

10%w/v Ampicillin Sodium Salt (Melford Laboratories, England)

50%v/v Ethanol (Merck, Germany)

1000X Chloramphenicol

10% w/v Chloramphenicol (Calbiochem, Germany)

100% Ethanol (Merck, Germany)

Luria Broth Agar

1%w/v Pancreatic Digest of Casein (Tryptone powder) (Merck, Germany)

1% w/v Yeast Extract (Merck, Germany)

0.5%w/v NaCl (Sigma Aldrich, Germany)

1.5% w/v Agar (Sigma Aldrich, Germany) - For solid media only

autoclaved before use

4X Stacking Buffer

0.5M Tris(hydroxymethyl)-aminoethane (Merck, Germany)

pH 6.8

4X Running Buffer

1.5M Tris(hydroxymethyl)-aminoethane (Merck, Germany)

pH 8.8

4% Stacking Gel

1.3ml Monomer Solution (30% Acrylamide, 50% Bis-Acrylamide) (Sigma Aldrich, Germany)

2.5ml 4X Stacking buffer

100µL 10% w/v SDS (Merck, Germany)

5.9ml dH_2O

100µL 10% Ammonium persulfate (Sigma Aldrich, Germany)

10µL N,N,N,N'-Tetramethylethylene diamine (TEMED) (Sigma Aldrich, Germany)

12.5% Running Gel

6.25ml Monomer Solution (30% Acrylamide, 50% Bis-Acrylamide) (Sigma Aldrich, Germany)

3.75ml 4X Running buffer

150µL 10% w/v Sodium dodecyl sulfate (Merck, Germany)

4.7ml dH_2O

150µL 10% Ammonium persulfate (Sigma Aldrich, Germany)

10µL N,N,N,N'-Tetramethylethylene diamine (TEMED) (Sigma Aldrich, Germany)

2X Loading Buffer

25% v/v 4X Stacking Buffer

40% v/v 10% Sodium dodecyl sulfate (Merck, Germany)

20%v/v Glycerol (Merck, Germany)

1% v/v β-Mercaptoethanol (Sigma Aldrich, Germany)

0.001% w/v Bromothymol Blue (Merck, Germany)

5X Tank Buffer

125mM Tris(hydroxymethyl)-aminoethane (Merck, Germany)

967mM Glycine (Merck, Germany)

17mM Sodium dodecyl sulfate (Merck, Germany)

Destaining Solution 1

40% v/v Methanol (Merck, Germany)

7% v/v Acetic Acid (Merck, Germany)

Destaining Solution 2

7% v/v Acetic Acid (Merck, Germany)

5% v/v Methanol (Merck, Germany)

Coomassie Brilliant Blue Staining Solution

0.025% w/v Comassie Brilliant blue R250 (Merck, Germany)

40% v/v Methanol (Merck, Germany)

7% v/v Acetic Acid (Merck, Germany)

Transfer Buffer

20% v/v Methanol (Merck, Germany)

20% v/v 5X Tank Buffer

TBS

1% w/v NaCl (Sigma Aldrich Life Science, St Louis, USA)

3% w/v Tris(hydroxymethyl)-aminoethane (Merck, Germany)

0.2% w/v Potassium chloride (Calbiochem, USA)

pH 7.4

autoclaved before use

For T-TBS, 1%v/v Tween 20 (Saarchem, South Africa)

Re-suspension Buffer

1M NaCl (Sigma Aldrich, Germany)

20mM Tris(hydroxymethyl)-aminoethane (Merck, Germany)

5mM Imidazole (Sigma Aldrich, Germany)

10mM $MgCl_2$ (Merck Chemicals, South Africa)

0.25% v/v Triton-X (Merck, Germany)

1mM PMSF (Sigma Aldrich, Germany)

1000U Benzonase Nuclease (Novagen, Germany)

pH 7.9

Washing Buffer

20mM Hepes (Sigma Aldrich, Germany)

1M NaCl (Sigma Aldrich, Germany)

20mM Imidazole (Sigma Aldrich, Germany)

pH 7.5

Elution Buffer

20mM Hepes (Sigma Aldrich, Germany)

1M NaCl (Sigma Aldrich, Germany)

600mM Imidazole (Sigma Aldrich, Germany)

pH 7.5

Storage Buffer

1M Nacl (Sigma Aldrich, Germany)

20mM Hepes (Sigma Aldrich, Germany)

0.1mM EDTA (Boehringer, Germany)

1mM Dithiothreitol (DTT) (Sigma Aldrich, Germany)

1% Glycerol (Saarchem, South Africa)

pH 7.5