



Studies on HIV-1 DNA Integration

Nick Vandegraaff

B.Sc. (Hons)

Infectious Diseases Laboratories
Institute of Medical and Veterinary Science
Adelaide
South Australia

Department of Molecular Biosciences
University of Adelaide
Adelaide
South Australia

A thesis submitted to the University of Adelaide in fulfilment of the
requirements for the degree of Doctor of Philosophy

February, 2002

Abstract	vii
Declaration of originality	ix
Acknowledgements	x
Abbreviations	xi
Publications and presentations arising	xiii

Introduction	1
1.1 Background	1
1.1.1 Historical Aspects of Acquired Immunodeficiency Syndrome (AIDS) and Human Immunodeficiency Virus (HIV)	1
1.1.2 Classification and Genetic Organisation of HIV	2
1.2 HIV Replication	3
1.2.1 Entry	4
1.2.2 Reverse Transcription of the Viral Genome	5
1.2.3 Transport of Viral DNA to the Nucleus	7
1.2.4 Integration and Circularisation of Newly Synthesised Viral DNA	8
1.2.5 Transcription of the Integrated Provirus	8
1.2.6 Translation, Assembly and Budding	10
1.2.7 The Alternative Route: Cell-to-Cell infection	11
1.3 HIV Integration	12
1.3.1 Background	12
1.3.2 The Integration Reaction	12
1.3.3 The Integrase Protein	15
1.3.4 Sequence Specificity	17
1.3.5 Host Proteins Involved in HIV DNA Integration	17
1.3.6 <i>In Vivo</i> Significance of Integrated HIV DNA: Persistent Viral Reservoirs ...	18
1.4 Extrachromosomal HIV DNA Forms	21
1.4.1 Transcriptional Activity of Extrachromosomal HIV DNA	21
1.4.2 <i>In Vivo</i> Significance of Extrachromosomal HIV DNA	21
1.5 The Kinetics of Viral Nucleic Acid Metabolism in Cell Culture	22

1.5.1 Total and Circular Viral DNA Accumulation.....	22
1.5.2 Integrated Viral DNA Accumulation.....	23
1.6 Inhibition of HIV Integration.....	24
1.6.1 Antiviral Treatment of HIV Infected Individuals.....	24
1.6.2 HIV Integration as a Target for Intervention.....	25
1.6.3 Current Screening and Characterisation of Inhibitors of Integration.....	26
1.6.4 Inhibitors of Integration.....	30
1.7 Whole Cell Assays for the Direct Detection of Integrated HIV DNA.....	32
1.7.1 Electrophoresis/Electro-elution.....	32
1.7.2 Inverse PCR.....	33
1.7.3 Nested- <i>Alu</i> PCR.....	34
1.8 Quantification using the Polymerase Chain Reaction (PCR).....	35
1.8.1 Background.....	35
1.8.2 PCR as a Quantitative Tool.....	36
1.9 Aims of this thesis.....	39
Materials and Methods.....	41
2.1 Materials.....	41
2.1.1 Cells and Cell Culture.....	41
2.1.2 Virus Stocks.....	41
2.1.3 Bacterial Culture.....	42
2.1.4 Plasmid Vectors.....	42
2.1.5 Oligonucleotide Sequences.....	42
2.1.6 Commonly used buffers and solutions.....	43
2.2 Methods.....	45
2.2.1 Cell-free HIV Infection.....	45
2.2.2 Detection of HIV P24 Antigen.....	45
2.2.3 HIRT DNA Extraction.....	46
2.2.4 DNA Standard Preparations.....	47
2.2.5 Standard PCR Procedures.....	49
2.2.6 Southern Transfer and Hybridisation Techniques.....	52
2.2.7 Preparation of Probes for Southern hybridisation.....	53
2.2.8 Sequence analysis.....	56
2.2.9 Phenol extraction and ethanol precipitation of DNA.....	57

2.2.10 Agarose gel electrophoresis and purification of DNA from agarose.....	57
2.2.11 Restriction enzyme digestions.....	57
2.2.12 Drugs and Cytotoxicity Assays.....	58
2.3 List of Suppliers	58

Development of an Assay for Integrated HIV DNA:

Degenerate Primer Polymerase Chain Reaction (DP-PCR).. 60

3.1 Introduction.....	60
3.1.1 Background	60
3.1.2 DP-PCR Principle	61
3.2 Demonstrating the DP-PCR Principle.....	64
3.2.1 Generating constructs mimicking integrated and linear HIV DNA.....	64
3.2.2 Designing a Degenerate Primer: RDM#2	65
3.2.3 Confirming the Basic Principle of DP-PCR	66
3.2.4 Optimising Selectivity: [MgCl ₂] and Annealing Temperature (T° _{anneal})	66
3.3 Background Chromosomal DNA and DP-PCR Optimisation	67
3.3.1 Background HuT-78 DNA.....	67
3.3.2 Optimising Degenerate Primer Composition and Length.....	68
3.4 Enhancing the Selectivity of DP-PCR: Incorporation of <i>Bsr</i> 1.....	71
3.4.1 The Principle of <i>Bsr</i> 1 Digestion.....	71
3.4.2 Benefit of <i>Bsr</i> 1 digestion in DP-PCR	72
3.5 Increasing the Sensitivity of DP-PCR.....	73
3.5.1 Optimising the Nested PCR	73
3.5.2 Sensitivity and Selectivity of the Nested DP-PCR Procedure	74
3.6 Characterisation of DP-PCR with Additional Control Constructs.....	75
3.6.1 Preparation of HA8 Integrated HIV DNA Standards	75
3.6.2 Preparation of Additional Constructs Mimicking Extrachromosomal HIV DNA	75
3.6.3 Equating New Control Standards.....	76
3.6.4 Characterising DP18 using New Controls	77
3.6.5 Sequence Analysis of the H3B, ACH-2 and 8E5 Cell Lines	78
3.7 Detection of Integrated DNA Following HIV Infection of HuT-78 T Cells	79
3.8 Discussion	80
3.8.1 Sensitivity of DP-PCR	80

3.8.2 Selectivity of DP-PCR	84
3.8.3 Conclusion	86

Development of a Quantitative HIV-1 Integrated Proviral Assay: Linker-Primer Polymerase Chain Reaction Method

(LP-PCR)	87
4.1 Introduction	87
4.1.1 Background	87
4.1.2 The Principle of the Linker-Primer PCR Assay: Amplification of Integrated HIV DNA	87
4.1.3 The Principle of the Linker-Primer PCR Assay: Selection Against Amplification of Extrachromosomal HIV DNA	89
4.1.4 Relative Restriction Frequencies of <i>Nla</i> III and <i>Bgl</i> II	89
4.2 Optimisation of Digestion Conditions	90
4.2.1 Optimisation of <i>Bgl</i> II Digestion	91
4.2.2 Optimisation of <i>Nla</i> III Digestion	92
4.2.3 Final Digestion Conditions	93
4.3 Proof of Principle	93
4.4 Optimisation of Linker Ligation	94
4.5 Detection of HIV DNA in the Presence of Background Human Lymphocyte DNA	96
4.5.1 Standard PCR Technique	96
4.5.2 Hot Start PCR Technique	97
4.6 Overcoming Background <i>Nla</i> III-fragment Amplification	98
4.6.1 Background <i>Nla</i> III-fragment Amplification	98
4.6.2 Incorporation of a Asymmetric or "Linear" PCR	99
4.6.3 Dynabead™ selection of the HIV LTR DNA fragments	102
4.6.4 Genuine Hot-start LP-PCR + Nested PCR	104
4.6.5 Amplification of Integrated HIV DNA Present in Cell Lines	106
4.7 Optimising the 1 st -Round PCR	107
4.8 Selection against Extrachromosomal forms	108
4.8.1 Addressing the Poor Selectivity	109
4.8.2 Incorporation of the Klenow-Mediated "Fill-In" Step	110
4.9 Finalised LP-PCR Protocol for the Detection of Integrated HIV-1 Forms	112

4.9.1 Linker Ligation Efficiency	113
4.9.2 Selectivity of LP-PCR for Integrated HIV-1 DNA.....	114
4.9.3 Sensitivity of LP-PCR.....	115
4.10 Comparison of LP-PCR with the nested- <i>Alu</i> PCR Protocol	116
4.10.1 The nested- <i>Alu</i> PCR protocol for the Detection of Integrated HIV DNA .	116
4.10.2 Comparison of LP-PCR and nested- <i>Alu</i> PCR: Clonal Cell Line DNA	117
4.11 Discussion	119

Kinetics of HIV Nucleic Acid Accumulation Following a

One-Step Infection of T Cells..... 121

5.1 Introduction	121
5.1.1 Background	121
5.1.2 Aims	122
5.2 Establishing a Synchronous Cell-Free Infection Model	122
5.2.1 Synchronous Entry	122
5.2.2 Optimising the Infection Efficiency: DEAE-Dextran Pre-treatment of HuT-78 Cells.....	123
5.2.3 Centrifugal Enhancement of Virus Infection	123
5.3 Kinetics of Integrated, Total and 2-LTR Viral DNA Accumulation Following Infection of T Cells	125
5.3.1 Kinetics of Viral DNA Accumulation: P24 Release.....	125
5.3.2 β -globin Analysis and Adjustment of HIRT Pellet (Chromosomal) DNA Fractions.....	126
5.3.3 Mitochondrial DNA Analysis of HIRT Supernatant (Extrachromosomal) DNA Fractions	127
5.3.4 Kinetics of Total Viral DNA Accumulation	128
5.3.5 Kinetics of Integrated Viral DNA Accumulation: LP-PCR.....	130
5.3.6 Kinetics of Integrated Viral DNA Accumulation: Nested <i>Alu</i> -PCR.....	132
5.3.7 Kinetics of 2-LTR Viral DNA Accumulation.....	132
5.4 Discussion	134

Specific Inhibition of HIV-1 Integration in Cell Culture:

Evaluating Putative Inhibitors of HIV-1 Integrase..... 137

6.1 Background	137
----------------------	-----

6.2 The Cell-free Infection Model	139
6.3 The Effect of L17 and AR177 on HIV Integration	139
6.3.1 Cytotoxicity.....	140
6.3.2 Infection in the Presence of L17 and AR177	140
6.3.3 Viral Nucleic Acid Accumulation.....	141
6.4 The Effects of Quercetin Dihydrate and Diketo Acid Inhibitors on HIV Integration	142
6.4.1 Cytotoxicity.....	143
6.4.2 Infection in the Presence of Quercetin Dihydrate and Diketo Acid Inhibitors	143
6.4.3 Viral Nucleic Acid Accumulation in the Presence of Quercetin Dihydrate and Diketo Acid Inhibitors.....	144
6.5 Discussion	145
General Discussion.....	148
7.1 Summary and Discussion.....	148
7.2 Future Work	153
Bibliography	156

Appendices

Abstract

There is limited information about the kinetics and efficiency of HIV-1 integration in cell culture primarily due to the lack of an assay that can clearly distinguish between the integrated and extrachromosomal HIV DNA forms. Here, the development of two novel PCR-based assays for the detection and quantification of integrated HIV-1 proviral forms are described. The degenerate-primer PCR (or DP-PCR; Chapter 3) protocol depended on the ability of semi-degenerate primers to anneal over the junction of the integrated 5' HIV U3 region and upstream cellular sequence, while not annealing to the corresponding 3' U3 region found in all HIV DNA forms. Although demonstrated to selectively amplify integrated HIV DNA present within a control construct and chromosomal extracts made from a mix of the ACH-2, 8E5 and H3B cell lines, the efficiency with which HIV DNA was independently amplified in each of these samples varied greatly. Furthermore, DP-PCR amplification of randomly integrated DNA (produced by infecting HuT-78 cells with HIV_{HXB2}) was shown to be highly inefficient. Therefore, while the DP-PCR procedure could direct the selective amplification of certain integrants, it was ultimately judged to have limited use as a quantitative tool for measuring randomly integrated HIV DNA following infection. In contrast, the linker-primer PCR (LP-PCR) assay (Chapter 4) routinely detected 10 copies of integrated HIV DNA in a background of 2×10^5 cell equivalents of human chromosomal DNA and gave only $\approx 7.5\%$ of this signal when detecting equimolar amounts of unintegrated HIV DNA. Using LP-PCR, it was shown that the specific anti-integration compound L-731,988 completely abolished the accumulation of integrated HIV DNA following acute HIV infection while allowing the synthesis of extrachromosomal viral DNA. Furthermore, evidence was presented that the proportion of random integration events detected using LP-PCR is significantly higher than that detected using the nested-*Alu* PCR technique. The LP-PCR assay was then used to study the kinetics and efficiency of HIV viral DNA integration into the host cell chromosome following infection with cell-free virus (Chapter 5). Using a near synchronous, one-step infection model in a T cell line, integrated HIV DNA was first detected one hour after the first appearance of full-length reverse transcribed product and three hours prior to the appearance of the 2-LTR circular form. Within the time-frame of a single round of HIV replication, $\approx 11\%$ of the total reverse transcribed product integrated into the host chromosome while only $\approx 0.7\%$

circularised to form the 2-LTR episomal DNA form. These results highlight the efficient nature of integrase-mediated HIV integration in infected T cells. The implications of these results are discussed.

To study the effect of potential HIV-1 integrase inhibitors during virus replication in cell culture, we used a modified nested *Alu*-PCR assay to quantify integrated HIV DNA and a standard PCR assay to measure extrachromosomal HIV DNA (Chapter 6). The two diketo acid integrase inhibitors (L-708,906 and L-731,988) blocked the accumulation of integrated HIV-1 DNA in T cells following infection but did not alter levels of newly-synthesised extrachromosomal HIV DNA. In contrast, we demonstrated that L17 (a member of the bisaroyl hydrazine family of integrase inhibitors) and AR177 (an oligonucleotide inhibitor) blocked the HIV replication cycle at, or prior to, reverse transcription, although both drugs had previously been shown to inhibit integrase activity in cell-free assays. Quercetin dihydrate (a flavone) was shown not to have any antiviral activity in our system despite reported anti-integration properties in cell-free assays. This refined *Alu*-PCR assay for HIV provirus is a useful tool for screening anti-integration compounds identified in biochemical assays for their ability to inhibit the accumulation of integrated HIV DNA in cell culture, and may be useful for studying the effects of these inhibitors in clinical trials.

Declaration

All work except for sequencing reported in this thesis was done by myself unless otherwise stated. All DNA sequencing was performed by Mr. Arthur Mangos in the Institute of Medical and Veterinary Science.

This thesis contains no material that has been accepted for the award of any other degree or diploma in any university or other tertiary institution.

To the best of my knowledge, this thesis contains no material previously published or written by any other person, except where due reference is made in the text of the thesis.

In accordance with the University of Adelaide regulations, I give my consent to this thesis being made available for photocopying and loan if accepted for the award of the degree.

Nicholas Andrew Vandegraaff

Acknowledgments

To my supervisors Dr. Li Peng, Dr. Jim McInnes and Prof. Chris Burrell for their expert guidance throughout the course of this PhD. In particular to my principal supervisor Dr. Li Peng, who has always been receptive to my ideas and has given me a great deal of confidence in my chosen discipline.....

To my family, for their counsel and utmost support for all my decisions during my studies.....

To the entire HIV lab, both present (Zilla, Whizzey and Timmmeee, Ra-uul, Sleazy, Melon, Mindamundy, Kel, Jilleeehhh and Feng-Feng) and past, for the incredibly enjoyable working environment and helpful discussions. In particular, to Raman Sharma, Alice Stephenson, Louis Mercure and Linda Mundy, and to my very dear friend Helen Hocking.....

To Prof. John Mills for his keen interest in this work and assistance with collaborations.....

To my great friends Heather McDonald, John Karlis, Dave and Juanita Ottaway and Grant "Porky" Buchanan.....

And finally to Mel, whose love and support over the last 3 years has provided me with the drive required to complete my studies.....

~ Thankyou ~

This work was in part supported by a University of Adelaide Scholarship, and part supported by a grant from the Australian Commonwealth AIDS Research Grant Programme awarded to myself.

Abbreviations

≈	approximately
<	less than
>	greater than
°C	degrees Celcius
AIDS	Acquired Immune Deficiency Syndrome
bp	base pair
Ci	curie
dATP	2'-deoxyadenosine 5'-triphosphate
dCTP	2'-deoxycytidine 5'-triphosphate
DDW	double distilled water
dGTP	2'-deoxyguanosine 5'-triphosphate
DNA	deoxyribonucleic acid
dNTP	2'-deoxynucleoside 5'-triphosphate
DTT	dithiothreitol
dTTP	2'-deoxythymidine 5'-triphosphate
EDTA	ethylene-diamine-tetra-acetic acid
g(units)	gravity force
HIV-1	Human Immunodeficiency Virus Type 1
h	hour
IN	integrase
kb	kilobase-pair
LB	Luria-Bertani medium
μg	microgram
μM	micromolar (micromoles per litre)
min	minute(s)
ml	millilitre
mM	millimolar (millimoles per litre)
ng	nanogram
OD	optical density
nt	nucleotide
³² P	radioactive phosphorous (mass number: 32)

PAGE	polyacrylamide gel electrophoresis
pBS	plasmid pBluescript (KS)
PCR	polymerase chain reaction
PIC	preintegration complex
RNA	ribonucleic acid
rpm	revolutions per minute
RT	reverse transcriptase
SDS	sodium dodecyl sulphate
s	second(s)
SSC	standard saline citrate
ssDNA	salmon sperm DNA
TCID ₅₀	50% tissue culture infectious dose
T4 PNK	T4 polynucleotide kinase

Publications and presentations arising

Publications:

Vandegraaff N, Kumar R, Hocking H, Burke TR Jr, Mills J, Rhodes D, Burrell C and Li P. (2001) Specific Inhibition of HIV-1 Integration in Cell Culture: Evaluating Putative Inhibitors of HIV-1 Integrase. *Antimicrob. Agents Chemother.* 2001 Sep;45(9):2510-6.

Vandegraaff N, Kumar R, Burrell C and Li P. Kinetics of Human Immunodeficiency Virus Type 1 (HIV) DNA Integration in Acutely Infected Cells as Determined Using a Novel Assay for Detection of Integrated HIV DNA. *J. Virol.* 2001 Nov;75(22):11253-11260.

Vandegraaff N, Kumar R (Co-first author), Burrell C and Li P. Evaluation of PCR-Based Methods for the Quantification of Integrated HIV-1 DNA. *Submitted.*

Conference Presentations:

Vandegraaff N, R Kumar, H Hocking, TR Burke Jr, J Mills, D Rodes, CJ Burrell and P Li (2001) Specific inhibition of HIV-1 integration in cell culture: evaluating putative inhibitors of HIV-1 integrase. Poster presented at the 13th Australasian Society for HIV Medicine Conference, Melbourne, Australia, 2001.

Vandegraaff N, Kumar R, Burrell C and Li P. Kinetics of HIV-1 Integration in Acutely Infected Cells Using a Novel Assay for the Detection of Integrated HIV-1 DNA. Poster presented at the 1st Australian Virology Group Meeting, Fraser Island, Australia, 2001.

Vandegraaff N, Kumar R, Mundy L, Mills J, Burrell C and Li P. HIV Integration Kinetics and Integrase Inhibitor Analysis in Cell Culture. Speaker, 12th Annual Conference of the Australasian Society of HIV Medicine, Melbourne, Australia, 2000.

Kumar R, Cheney K, Mundy L, **Vandegraaff N**, Shaw D, Burrell CJ and Li P. Double-Negative T Cells as Substantial Reservoirs of HIV in Patients Receiving Highly Active Antiretroviral Therapy (HAART). Talk given by Dr. Kumar at the 12th Annual Conference of the Australasian Society of HIV Medicine, Melbourne, Australia, 2000.

Vandegraaff N, Burrell CJ and Li P. A Novel Assay For the Detection of Integrated HIV DNA: Degenerative Primer PCR (DP-PCR). Poster presented at the Keystone Symposia on HIV Pathogenesis and Vaccine Development, Keystone, Colorado, 1999.

Chapter 1

Introduction

1.1 Background

1.1.1 Historical Aspects of Acquired Immunodeficiency Syndrome (AIDS) and Human Immunodeficiency Virus (HIV)

Acquired immunodeficiency syndrome (AIDS) was formally defined in 1981 following the description of five opportunistic cases of *Pneumocystis carinii* in severely immunocompromised men (1). The epidemiology and patient demography associated with subsequent reports of this immune disorder suggested that a transmissible agent spread through genital secretions and blood was responsible for the disease. However, the causative viral agent (termed the human immunodeficiency virus or HIV) was not identified until 1983 (9, 96, 201). Since then, HIV has become the most intensively studied human pathogen. HIV infection in humans is characterised by the loss of T-lymphocyte function that ultimately results in the progression to a severe state of immuno-suppression. Recent evidence has suggested that the immediate viral progenitor of HIV initially evolved in the species *Pan troglodytes troglodytes* (or chimpanzee) found in the sub-Saharan jungles of Africa (97). The viral strain responsible for the species jump was determined to have been endemic in populations of chimpanzee for centuries without causing significant disease. Since these animals are a food source to humans living in this region of Africa, HIV most likely arose as a zoonotic infection through the exposure of open wounds on humans to contaminated chimpanzee blood.

In a survey conducted in 1998, 33 million people worldwide were infected with HIV and approximately 16000 new infections were occurring every day (6 million per year). More than 95% of these new infections occurred in developing countries (particularly sub-Saharan Africa) where access to health care continues to be limited. In addition to the psychological costs of HIV infection (both to the individual and family), the

economic burden (globally \$14 billion every year) associated with the prevention and treatment of AIDS is ever increasing (189).

1.1.2 Classification and Genetic Organisation of HIV

HIV is a member of the family *Retroviridae*, a large family of viruses that predominantly infect vertebrate animals. Viruses within this family have been associated with both rapid and long-term disease states comprising neoplastic wasting, neurological and immunodeficiency diseases.

All retroviruses contain a two identical copies of a '+' sense RNA genome that is between 7 and 12kb in length. The genomic RNA molecules exist in close association with a number of viral proteins, including the nucleocapsid protein, the viral replicase (reverse transcriptase), the integrase and the protease proteins (Fig. 1.1 and Fig. 1.2). Viral capsid proteins encapsulate this nucleoprotein complex and together form a highly condensed core structure that can be easily distinguished by electron microscopy. Another protein layer (composed of the matrix protein) serves to tether the core structure to the host-cell derived lipid envelope, which is embedded with viral surface glyco-proteins essential for the attachment and entry of the virus. Retroviral virions can generally be observed by electron microscopy as particles of between 80-100 nm in diameter.

Unlike members of the family *Picornaviridae* (for example, poliovirus), the '+' sense retroviral RNA genome does not serve as mRNA upon entry into the host cell. It is instead reverse transcribed by the viral replicase into double stranded (ds) viral DNA and integrated into the host cellular chromosome. It is this ability of retroviruses to generate a dsDNA molecule from the genomic RNA molecule that has given this family of viruses their name. The integrated viral DNA (or provirus) is subsequently used as a transcriptional template by the host cells' RNA Polymerase II to generate mRNA and ultimately viral proteins. Both the processes of reverse transcription (see section 1.2.2) and integration (see section 1.3) are unique to the retroviral life cycle.

Figure 1.1 The HIV-1 Virion

Artistic impression of the HIV virion showing the key structural components.

Produced by Russell Knightly Media (rkm.com.au)

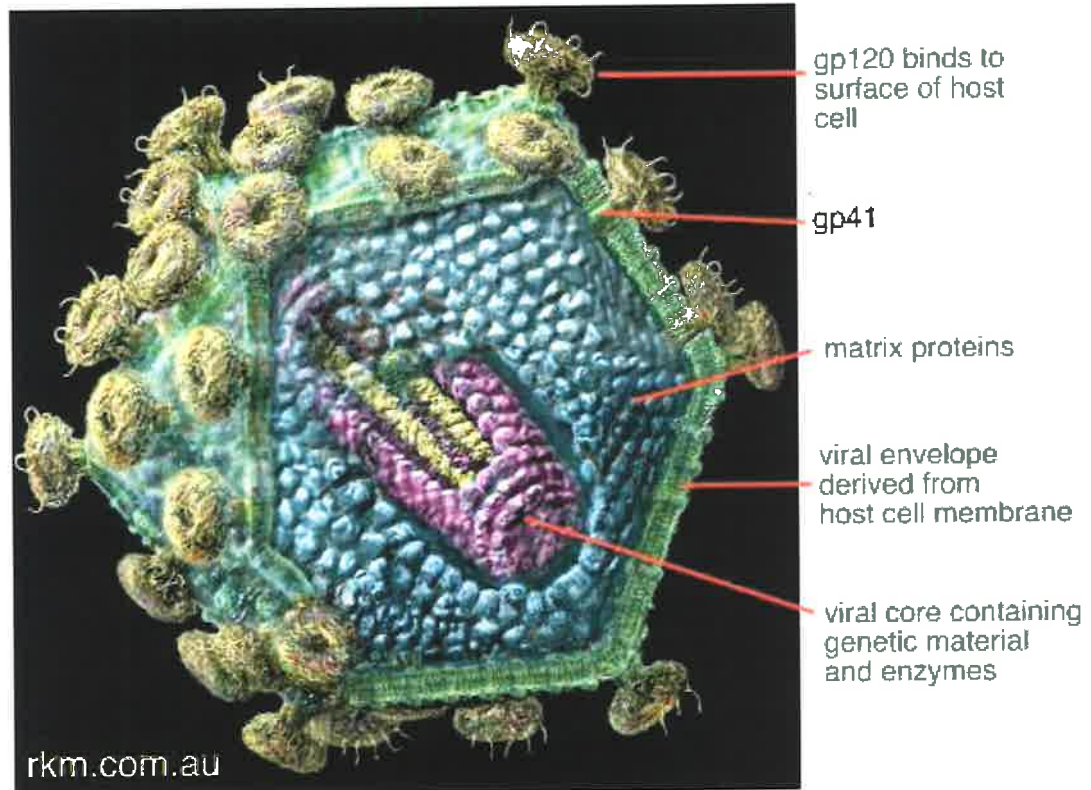
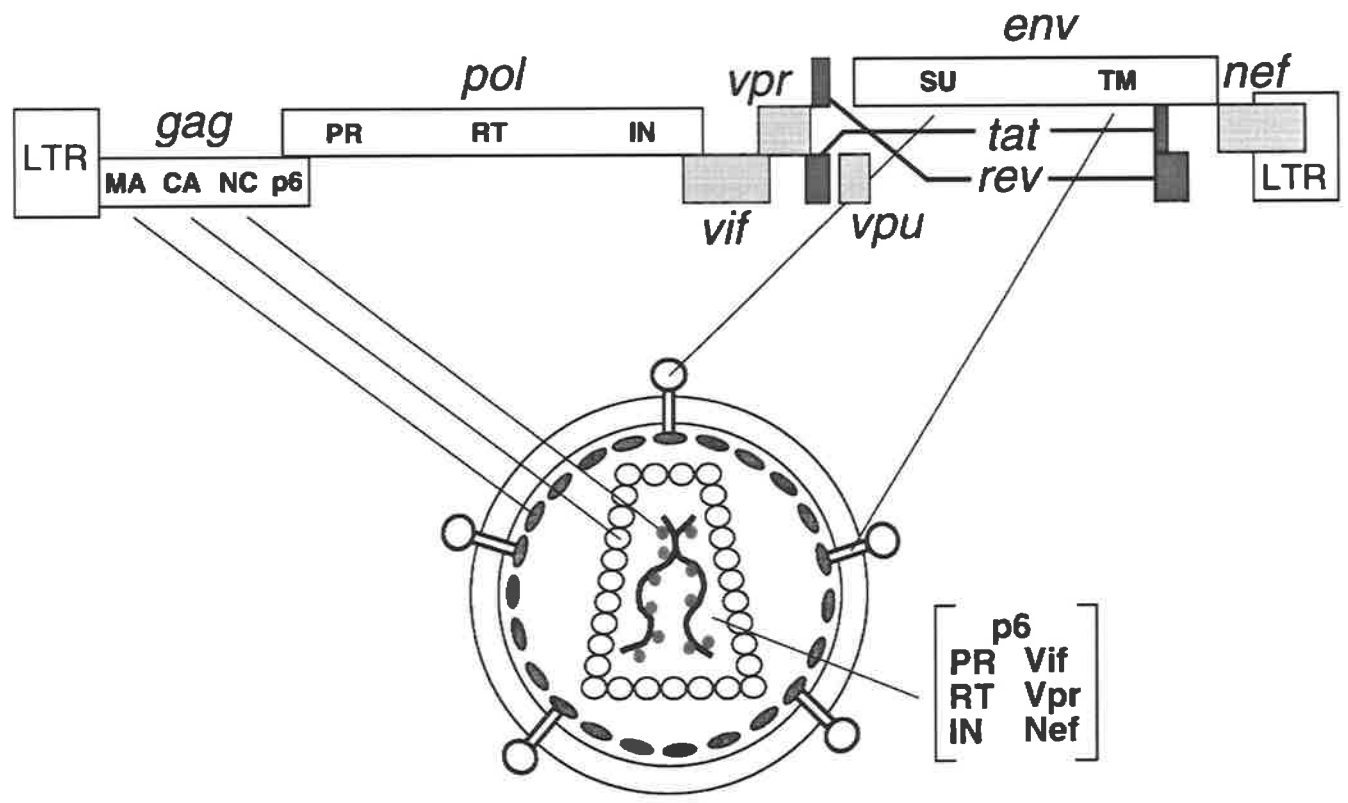


Figure 1.2 HIV-1 Genome Organisation and Virion Structure

Relative size and position of HIV genes encoded by the 9718 bp HIV genome. The relative positions of the proteins encoded by each gene within the virion is indicated.

Abbreviations are as follows:

LTR	Long Terminal Repeat
MA	matrix protein (p17)
CA	capsid protein (p24)
NC	nucleocapsid protein (p9)
PR	protease protein (p11)
RT	reverse transcriptase protein (p66/p51)
IN	integrase protein (p32)
SU	surface envelope protein (gp120)
TM	transmembrane protein (gp41)



Retroviruses have been further divided into distinct genera based primarily on genomic relationships (see Table 1.1 (“Retroviruses” Table 1)). Although not closely related to each other at the genetic level, five genera (groups 1-5 in Table 1.1) represent retroviruses with oncogenic potential and until recently were collectively referred to as the oncoviruses. The genetic composition and arrangement of all retroviruses is highly conserved, with all non-defective viruses containing three major genetic regions:

- 1) GAG (**group specific antigen**), encoding structural proteins
- 2) POL (**polymerase**), encoding enzymic proteins
- 3) ENV (**envelope**), encoding surface glyco-proteins.

Long terminal repeat (LTR) sequences invariably flank these three genetic regions (see Fig. 1.2). With the exception of the HTLV-BLV genera, oncogenic retroviruses usually carry only this basic information and are therefore termed simple retroviruses. In contrast, members of the HTLV-BLV genera, Lentiviruses and Spumaviruses all encode additional regulatory proteins that act in *trans* to both coordinate and enhance the infection process *in vivo* (see section 1.2.5). Viruses (such as HIV) classified within these genera are termed complex retroviruses.

HIV-1 is classified within the subfamily Lentivirinae, a group of complex retroviruses that cause slow degenerative diseases and encode a number of regulatory and accessory proteins (in addition to Gag, Pol and Env) within the genome. The genetic map of HIV-1, and the positions of the structural gene products within the virion, is schematically illustrated in Figure 1.2. Isolates of HIV-1 can be further divided into at least 8 genotypic sub-types (clades A-H) based on sequence similarities between the gag and env genes. The predominant viral strain isolated from HIV-infected patients in the developed countries of Australasia, USA and Europe is classified within clade B (Fig. 1.3).

1.2 HIV Replication

The two major cell types infected by HIV *in vivo* are macrophages and CD4-bearing T lymphocytes. Virus isolates frequently exhibit either a macrophage-tropic or T cell-

Table 1.1 Classification of Retroviruses

Grouping of Retroviruses into genus based on genome similarities. Previously, groups 1 to 5 were collectively known as the Oncoviruses.

^aDistinctive features seen in transmission electron micrographs

Adapted from Coffin *et al.*, 1997 (55)

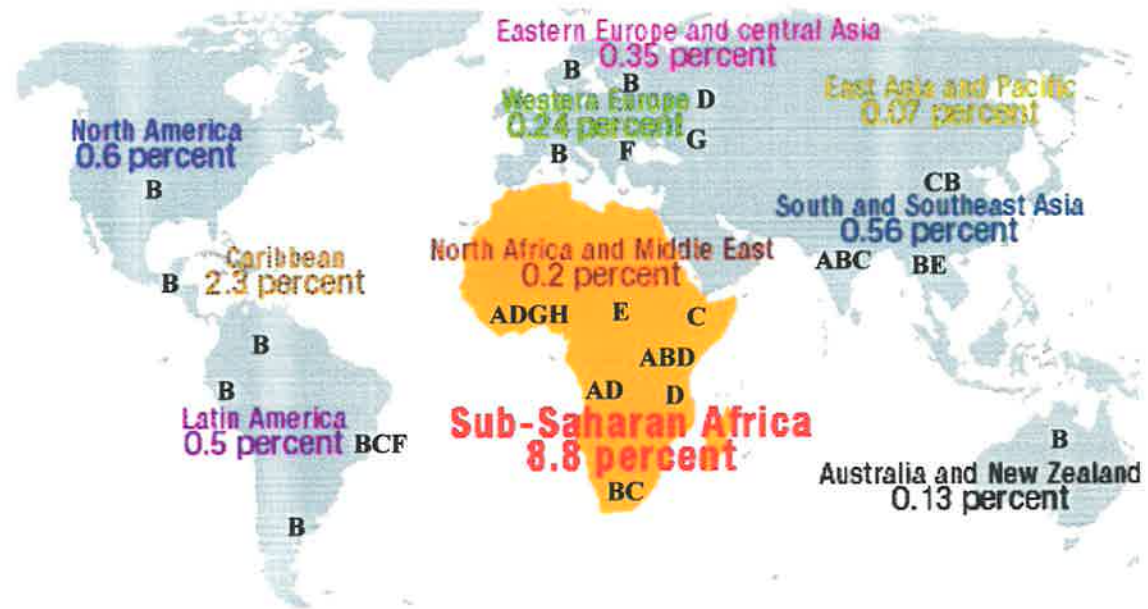
Genus	Example	Virion morphology ^a	Genome
1. Avian sarcoma and leukosis viral group	Rous sarcoma virus	central, spherical core	simple
2. Mammalian B-type viral group	mouse mammary tumor virus	eccentric, spherical core	simple
3. Murine leukemia-related viral group	Moloney murine leukemia virus	central, spherical core	simple
4. Human T-cell leukemia- bovine leukemia viral group	human T-cell leukemia virus	central, spherical core	complex
5. D-type viral group	Mason-Pfizer monkey virus	cylindrical core	simple
6. Lentiviruses	human immunodeficiency virus	cone-shaped core	complex
7. Spumaviruses	human foamy virus	central, spherical core	complex

Figure 1.3 Global Adult Prevalence and Clade Distribution of HIV-1

The adult prevalence rate is the proportion (expressed as a percentage) of adults (aged between 15 and 49) living with HIV/AIDS in 2000, using 2000 population numbers.

The major HIV-1 clades (A-H) responsible for infection in various regions are indicated.

Adapted from Coffin *et.al.*, 1997 (55)



tropic phenotype based primarily on co-receptor usage (see section 1.2.1). Although the majority of circulating T cells exist in a resting (or quiescent) state, cellular activation appears to be required for a productive infection to occur in these cells (233, 238, 272, 273). Acute HIV infection in humans can result in greater than 1% of the total circulating CD4⁺ lymphocyte pool being infected (52). The predominant route by which HIV is thought to infect target cells *in vivo* is initiated by the attachment of free virus to the surface of the target cell (cell-free infection). A second route, involving the direct contact of an infected cell with an adjacent uninfected susceptible cell (cell-to-cell HIV infection), has been demonstrated *in vitro*, however the extent to which this infection mode contributes to viral dissemination *in vivo* is unknown (see section 1.2.2 below).

1.2.1 Entry

Cell-free HIV infection (Fig. 1.4) is initiated through the high affinity interaction of a trimeric gp120 protein complex embedded within the virion envelope with the CD4 receptor present on the target cells surface (175). Although the majority of infection events *in vivo* are believed to require this interaction, CD4 independent entry has been described. Recently, viruses isolated from CD8⁺ cells obtained from two AIDS patients were shown to infect CD4⁻ cells in a CD8-dependent fashion (216). Although the role and significance these strains play in viral pathogenesis *in vivo* remains unclear, the ability of the virus to enter cells through a variety of routes indicates a substantial level of functional plasticity associated with the gp120/gp41 protein complex.

Although in most cases essential, the interaction of the gp120 protein with the CD4 protein alone has been shown to be insufficient for viral entry (50, 51, 122). Fusion of the viral and cellular membranes has been shown to require an additional interaction of the gp120 protein with a member of the chemokine receptor family of proteins present on the susceptible cell surface (51). Supporting this, physiological ligands to a variety of co-receptors have been shown to be able to inhibit HIV entry into certain cell types (13, 42, 187, 230). Although CXCR4 (for T-tropic viruses) and CCR5 (for macrophage-tropic viruses) appear to be the two major co-receptors utilised by HIV, other members of the chemokine receptor family have been shown to be able to mediate HIV entry into cells *in vitro* (Table 1.2). The interaction of the variable V3 loop of gp120 with the

Figure 1.4 Life-Cycle of HIV-1

Cell-free HIV infection of a target cell. Although extrachromosomal viral DNA molecules (linear unintegrated, two-LTR circular and one-LTR circular) can be produced upon infection, integration of the linear unintegrated viral DNA into the host cell chromosome is required for a productive infection to occur.

Adapted from Furtado *et al.*, 1999 (94)

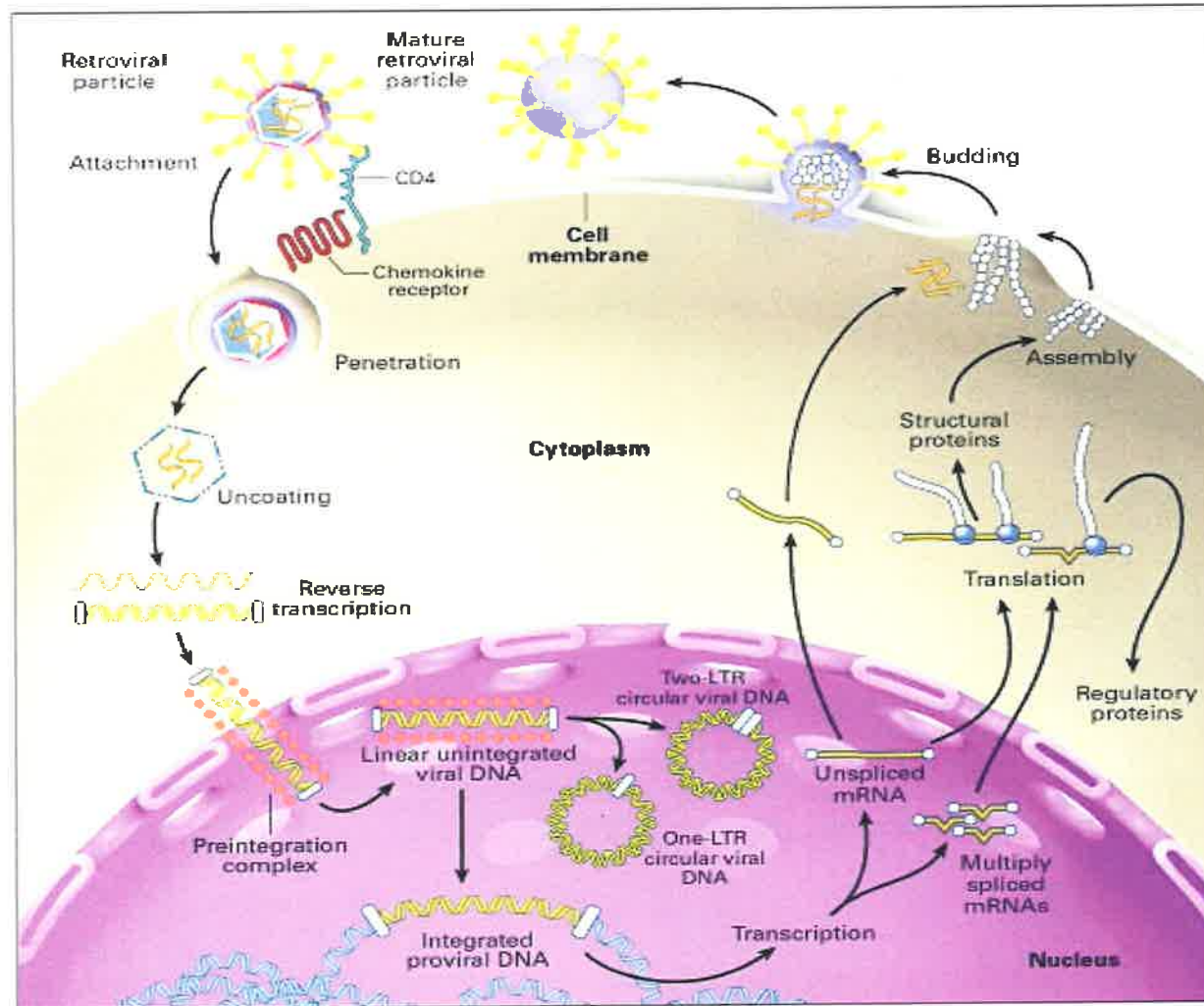


Table 1.2 HIV-1 Co-receptor Usage

Co-receptor utilisation by primate Lentiviruses.

HIV-1	Human Immunodeficiency Virus Type 1
HIV-2	Human Immunodeficiency Virus Type 2
SIV	Simian Immunodeficiency Virus

CORECEPTOR UTILIZATION BY PRIMATE LENTIVIRUSES

Receptor	Ligands	Viral usage	Expression pattern
Major			
Coreceptors			
CCR5	MIP-1 α , MIP-1 β , RANTES	HIV-1, HIV-2, SIV	Lymphocytes, macrophages, brain
CXCR4	SDF-1	HIV-1	Monocytes, T cells
Minor			
Coreceptors			
CCR2b	MCP-1, MCP-3	HIV-1, HIV-2	Monocytes, T cells
CCR3	Eotaxin, MCP-3, MCP-4	HIV-1	Eosinophils, microglia, Th2 T cells
CCR8	I-309	HIV-1, SIV	Monocytes, thymocytes
BOB/ GPR15	?	SIV, HIV-2	T cells, colon
Bonzo/ STRL33	?	SIV, HIV-2	T cells, monocytes, placenta
GPR1	?	SIV	Macrophages
APJ	?	HIV-1, SIV	Brain
CX ₃ CR1/ V28	Fractalkine/ neurotactin	HIV-1, HIV-2	Lymphocytes, brain
BLTR	Leukotriene B4	HIV-1	Thymus, spleen, monocytes

chemokine receptor protein drives conformational changes within the gp120 protein and exposes a hydrophobic fusion peptide in the gp41 protein required for membrane fusion to occur (50). Currently, peptide inhibitors designed to specifically associate with the fusion peptide and block its' insertion into the cellular membrane are being developed (61, 67, 152). Following fusion of the viral and cellular membranes, the viral core particle is internalised and reverse transcription of the genomic RNA to dsDNA commences.

1.2.2 Reverse Transcription of the Viral Genome

Reverse transcription of the viral RNA genome to dsDNA occurs within a structure derived from the core particle that is termed the replication or preintegration complex (15). In addition to viral RNA and reverse transcribed viral DNA, the replication complex has been shown by immunoprecipitation studies to comprise the viral reverse transcriptase (RT), integrase (IN), nucleocapsid (NC), matrix (MA) and vpr proteins (23, 24, 130, 178). Furthermore, cellular proteins such as histones, the barrier to autointegration factor (BAF) and a high-mobility group protein (HMG I(Y)) protein have also been determined to be present within the replication/preintegration complex (39, 80, 159).

The viral reverse transcriptase protein (RT) has both RNA-dependent and DNA-dependent DNA polymerase activities. In addition, an RNase H activity present in the C-terminal domain of RT is required to digest the RNA components of the viral RNA/DNA duplexes generated in the replication process. HIV reverse transcriptase does not have the potential for proofreading (3'-5' exonuclease⁻) and therefore the error rate associated with the reverse transcription process is high (one error every 10kb) (169). Since the size of the HIV genome is $\approx 10\text{kb}$, this error rate equates to an average of a single nucleotide mutation for each time reverse transcription of the genome occurs.

A model of the reverse transcription process resulting in synthesis of viral DNA is illustrated in Figure 1.5. Briefly, reverse transcription is initiated by the extension of a host-derived tRNA₃^{lys} that binds to the viral PBS (**p**rim**e**r **b**inding **s**ite) region present immediately downstream of the viral 5' LTR (see Fig. 1.5, panel A). This first extension

phase produces what is known as minus-strand strong-stop DNA, and may occur in the virion (64, 276, 277). Due to the RNase H activity associated with the reverse transcriptase enzyme, digestion of the template RNA molecule accompanies extension of viral DNA. The resulting single-stranded strong-stop DNA is then transferred to the other end of the viral genome in the first of two strand-transfer reactions catalysed by the RT enzyme (Fig. 1.5, panel B). The strong-stop DNA sequence pairs to the R sequence present at the 3'-end of the genomic RNA and is extended the length of the genome to generate a minus-strand viral DNA molecule (Fig. 1.5, panel B and C). Although the majority of RNA within the heteroduplex is degraded via the RNase H activity of the RT enzyme, two short regions (known as the polypurine tracts or PPTs) positioned immediately upstream to the 3'-U3 region are resistant to digestion and remain associated with the newly synthesised DNA. These short RNA oligonucleotides serve as primers to initiate plus-strand DNA synthesis that extends over the U3 region and attached tRNA₃^{lys} molecule of the newly synthesised minus-strand to regenerate the PBS region (Fig. 1.5, panel C and D). This plus-strand strong-stop DNA is then transferred to the 3'-end of the newly synthesised minus-strand where complementary PBS regions anneal, culminating in the extension of both minus-strand and plus-strand DNA to generate the dsDNA molecule (complete with duplicated U3-R-U5 (Long Terminal Repeat; LTR) regions (Fig. 1.5, panel E and F).

The process of retroviral reverse transcription is directed by the virally-encoded reverse transcriptase enzyme and is therefore an ideal target for antiviral drugs. Consequently, numerous drugs have been developed that interfere with this process, primarily by targeting the active site of the RT enzyme. Nucleoside analogs, such as azidothymidine (AZT) and dideoxyinositol (ddI), are converted to triphosphate forms *in vivo* and then compete with dNTPs by directly binding to the RT active site. In contrast, non-nucleoside inhibitors of RT (*eg.* nevirapine) have also been identified that bind to regions adjacent to the active site. Binding to these regions serves to rearrange the polymerase active site conformation and lock it in an inactive form (145, 210).

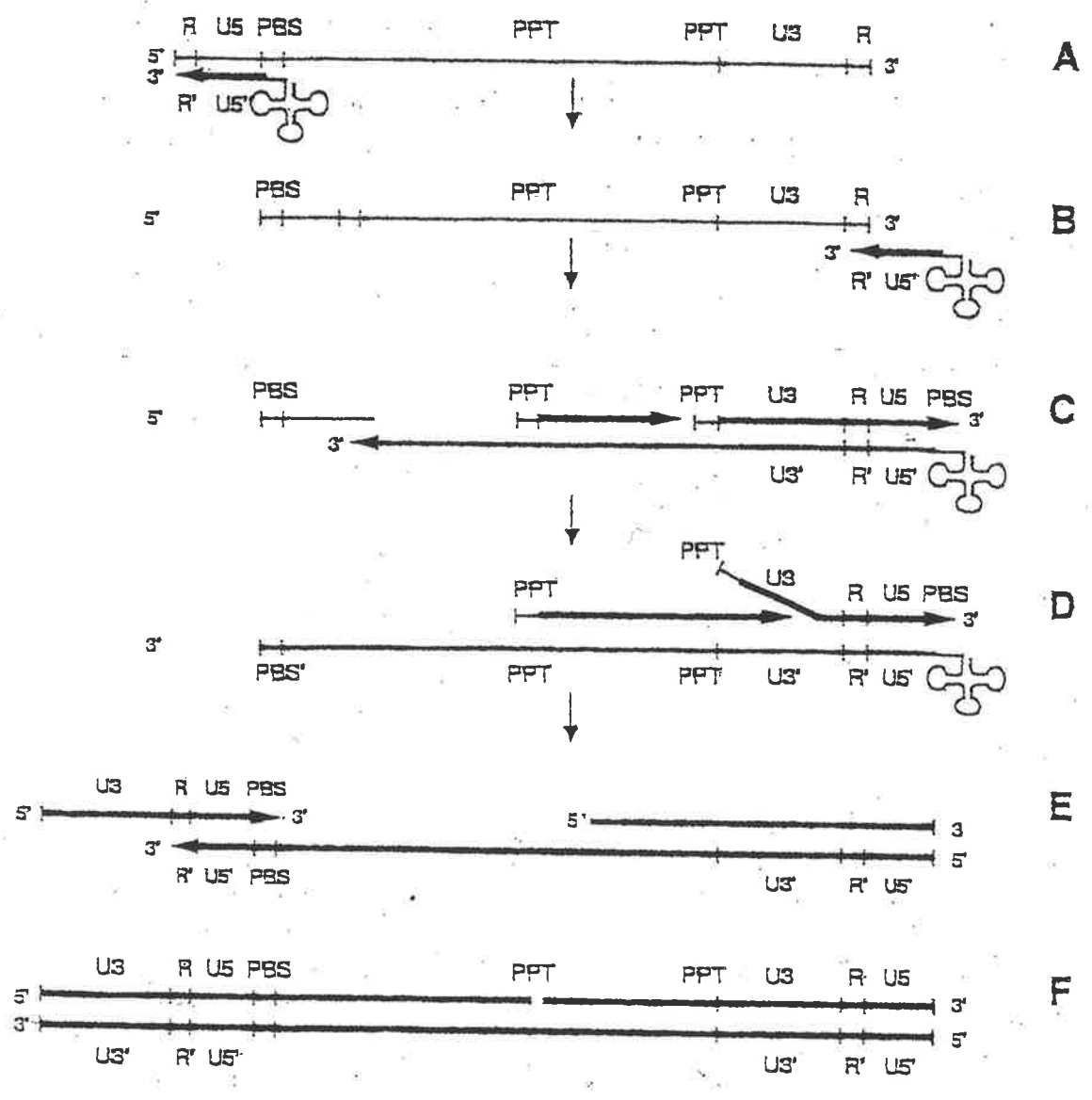
Figure 1.5 HIV-1 Reverse Transcription

A model of HIV-1 reverse transcription (see section 1.2.2 in text for a detailed description of the process). **A.** Extension of a host-derived tRNA₃^{lys} to produce minus-strand strong-stop DNA. RNase digestion of the viral genomic RNA accompanies synthesis of viral cDNA. **B.** Minus-strand cDNA synthesis following the first template switch. **C.** Priming and extension of RNase-resistant PPT viral RNA on the newly synthesised minus-strand cDNA. **D.** Displacement of newly synthesised plus-strand DNA. **E.** Annealing (via the complementary PBS region) and extension of displaced newly synthesised plus-strand DNA. **F.** Extension over gapped regions to generate the complete double-stranded HIV cDNA molecule.

PBS Primer Binding Site

PPT Polypurine Tract

Adapted from Karageorgos *et al.*, 1995 (131).



1.2.3 Transport of Viral DNA to the Nucleus

In order to integrate newly reverse transcribed DNA into the host cellular chromosome, viral DNA must gain access to the nucleus. In contrast to oncoviruses that require nuclear membrane breakdown during mitosis to access the nucleus, the HIV replication complex is actively transported across the nuclear membrane (23) allowing the virus to infect non-dividing cells (*eg.* macrophages). Transport of the replication complex to the nuclear membrane is accomplished by the karyopherin transporter pathway and is thought to occur shortly after initiation of reverse transcription (8, 90). Although potential nuclear localisation signal (NLS) sequences required for the interaction of the replication complex with components of the karyopherin pathway have been identified in the IN, MA and Vpr proteins (62, 63, 90), the contribution of each protein to the translocation of the replication complex into the nucleus remains unclear. Both the IN and Vpr proteins have been shown to accumulate in the nucleus of infected cells (63, 95, 194, 198), whereas the ability of the matrix protein to localise to the nucleus remains controversial (23, 91). Furthermore, the IN protein has been shown to readily form a ternary complex with importin- α and importin- β (key components of the karyopherin transport pathway) and mediate the efficient nuclear localisation of coupled marker proteins *in vitro* (95). The Vpr protein has demonstrated a strong interaction with importin- α (254) and can interact with nucleoporins *in vitro* and *in vivo* (91, 254). In addition to viral proteins, a triple-stranded viral cDNA intermediate of reverse transcription has recently been identified as an additional determinant that directs transport of the replication complex to the nucleus (275). Mutations destroying this triple DNA structure (termed the central DNA flap) abolish the ability of viral DNA to access the nucleus (275).

Taken together, these results suggest that the NLS present on the IN and Vpr proteins, and perhaps conformational requirements of the viral DNA molecule, primarily mediate an interaction of the replication complex with components of the transporter pathway, while the MA protein may act to both enhance this interaction (*eg.* by binding to an alternative site on importin- α) and also to facilitate the interaction of the replication/transporter protein complex with proteins of the nuclear pore. Once docked to the nuclear pore complex (NPC), the replication complex is transported across the

nuclear membrane in an ATP-dependent fashion (23). Evidence suggests that various proteins associated with the replication complex (such as the RT and MA proteins) may be lost during this translocation process (130).

1.2.4 Integration and Circularisation of Newly Synthesised Viral DNA

Once inside the nucleus, one of four main fates awaits the newly synthesised viral DNA: 1) integration in a colinear fashion into the cellular chromosome; 2) circularisation via recombination generating a circular form containing one long terminal repeat (1-LTR circle); 3) circularisation via direct head to tail ligation generating a circular form containing two long terminal repeats (2-LTR circle); 4) remaining as a linear viral DNA molecule (Fig. 1.6). Although each of these forms can be identified following infection in cell culture, in order for a productive infection to proceed, integration of the linear viral DNA into the host cell chromosomal DNA must occur. The linear DNA (lDNA) form is the immediate substrate for the integration reaction (20, 93), which principally occurs through the action of the virally encoded integrase (IN) enzyme within the context of the PIC. For a detailed review of the processes involved with both integration and circularisation of the viral DNA, refer to section 1.3.

1.2.5 Transcription of the Integrated Provirus

Once integrated, efficient transcription of the viral DNA is able to occur. Proviral transcription has been proposed to be an efficient process (as determined by the strong LTR promoter) and heavily influenced by the site of integration into cellular DNA (127, 246). However, integration into the chromosome is proposed to occur at random with respect to cellular sequence and therefore viral DNA may integrate into regions of highly condensed chromatin that are not suited to efficient transcription. Viral transcription is initiated from a promoter located in the R region of the 5' LTR and is regulated by a number of cis-acting elements up- and down-stream of this region (Fig. 1.7). The 9.7kb HIV genome (see Fig 1.2) is organised into three major genetic regions common to all retroviruses (GAG, POL and ENV) and has a total of nine open reading frames (ORFs). Three ORFs encode structural and/or enzymic proteins and six encode

Figure 1.6 HIV DNA Forms

Following nuclear import, one of 4 fates awaits newly synthesised HIV DNA: **A.**

Remaining as a linear DNA molecule. **B.** Circularisation via recombination to form a 1-

LTR circular form. **C.** Circularisation via direct end-to-end ligation to form a 2-LTR

circular form. **D.** Integration of the linear form into the host cell chromosome.

Extrachromosomal HIV DNA forms (**A.**, **B.** and **C.**) exist within structures derived from the replication complex and are subject to degradation over time. In contrast, integrated HIV DNA (**D.**) is highly stable within the cellular chromosome and is required for a productive infection to result.

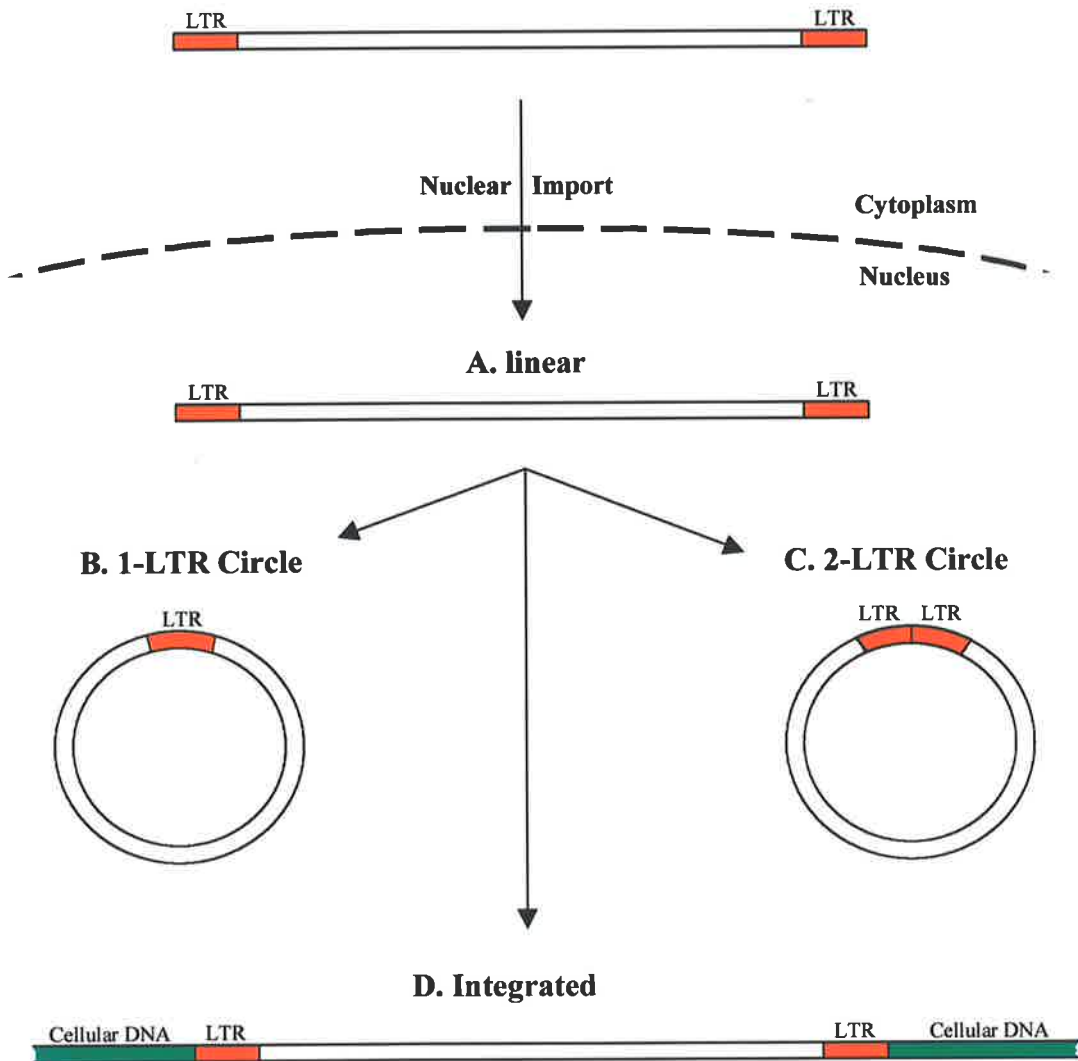
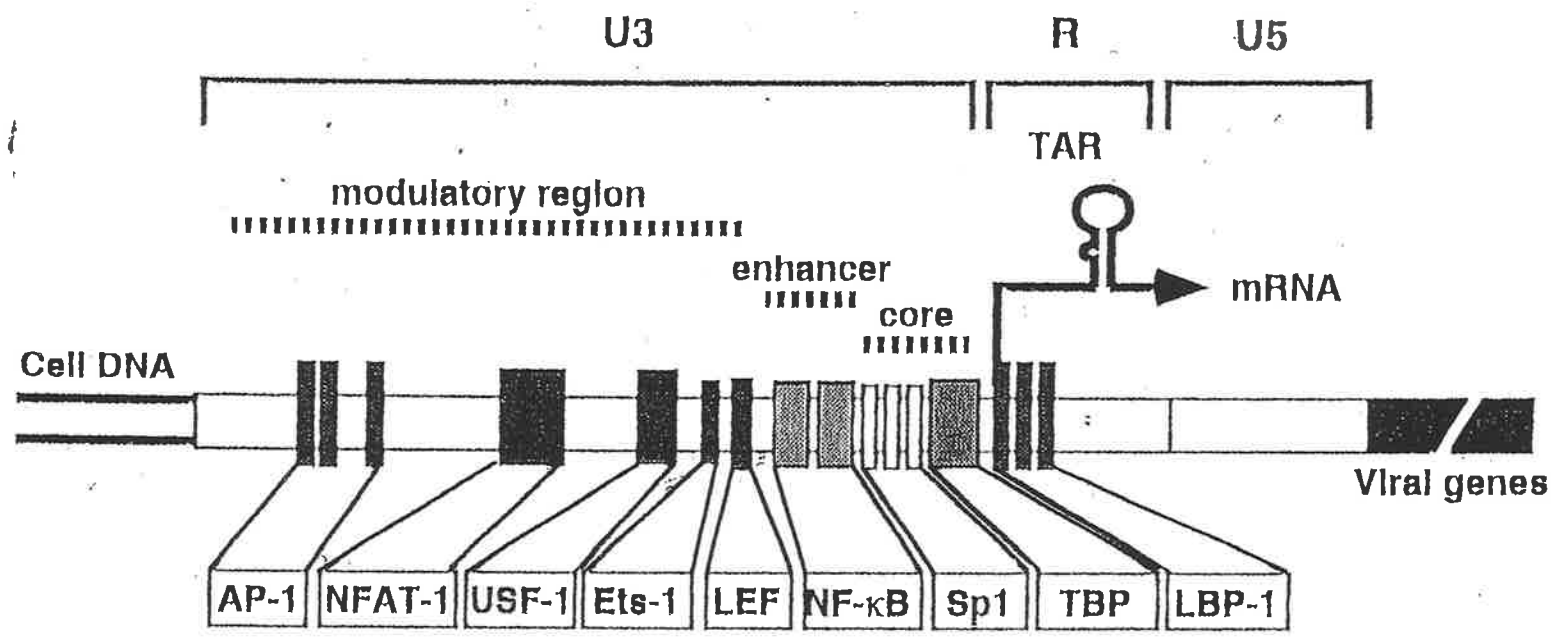


Figure 1.7 Regulatory Elements in the HIV-1 LTR

The HIV-1 LTR is divided into three functionally distinct regions designated U3, R and U5. Three transcriptional domains constitute the viral promoter in the U3 region: the core promoter, enhancer and modulatory region. The R region encodes the RNA sequence that forms the Tat transactivation response element, TAR. Binding sites of the various factors that bind to the 5' LTR are represented by shaded rectangles, superimposed onto the HIV-1 proviral genome.

Adapted from Luciw, 1996 (165).



regulatory proteins. However, multiple splicing patterns result in the production of more than 30 distinct RNA species. These have been divided into 3 main classes:

- 1) Multiply spliced 2kb transcripts. These are the first to appear (around 12h post infection) (60, 138) and encode the regulatory proteins Tat (**T**ransactivator of transcription), Rev (**R**egulator of virus expression) and Nef (**N**egative factor).
- 2) Singly spliced 4.5kb transcripts. These appear later in infection (16h-24h post infection) (60, 138) and encode the envelope protein gp160 which is subsequently cleaved by a host protease to produce the mature gp120 (**S**urface glycoprotein (SU)) and gp41 (**T**ransmembrane Protein (TM)) proteins. In addition, the accessory proteins Vif (**V**irion infectivity factor), Vpr (**V**iral protein **R**) and Vpu (**V**iral protein **U**) are encoded by these transcripts.
- 3) Unspliced 9.2kb transcripts. These appear late in infection (16h-24h post infection) (60, 138) and encode the gag (**g**roup-specific **a**ntigens) and gag-pol precursor polyproteins. Proteins p17 (**M**atrix (MA)), p24 (**C**apsid (CA)), p7 (**N**ucleocapsid (NC)), and p6 (important for incorporation of Vpr into the virion) are subsequently produced from the gag polyprotein, while the gag-pol polyprotein is cleaved by the viral protease enzyme (in an autocatalytic reaction) to produce the **P**rotease (PR), **R**everse transcriptase (RT) and **I**ntegrase (IN) proteins. In addition, unspliced transcripts are packaged as genomic RNA into assembling virions prior to budding.

The two virally encoded proteins Tat and Rev have been studied extensively and have been shown to act *in trans* to up-regulate transcription and regulate transport of transcripts to the cytoplasm respectively. Tat exerts its' effect through an interaction with a stem-loop structure, known as the TAR (**t**ransactivation **r**esponse) element, present in the nascent RNA chain. The interaction of Tat with this structure serves to enhance gene expression by increasing the rate of transcriptional initiation and enhancing transcriptional elongation (57, 153, 170, 232). Similarly, oligomeric Rev protein complexes interact with a 234 nucleotide RNA structure termed the Rev responsive element (RRE) present within the *env* gene to facilitate the switch from multiply-spliced transcripts encoding Tat, Rev and Nef proteins early in infection, to the singly and unspliced transcripts encoding structural proteins and genomic RNA later in infection. Although Rev has been shown to contain a nuclear export signal (NES) present in the C-terminal activation domain which enables it to be shuttled between the nucleus and the cytoplasm (176), the precise mechanism by which Rev achieves the

switch is unclear. Evidence suggests that Rev can directly enhance the export of unspliced and singly-spliced mRNAs (87) and/or inhibit the splicing of Rev-bound transcripts either by affecting spliceosome assembly or by the steric inhibition of the splicing reaction (38, 142, 143, 239).

1.2.6 Translation, Assembly and Budding

Following export to the cytoplasm, viral RNA molecules are translated by host cellular machinery. Early proteins (*eg.* regulatory proteins such as Tat and Rev) are translated from multiply spliced RNA molecules that appear from 12h after infection (60, 138), whereas later proteins (*eg.* structural proteins) are expressed from singly or unspliced mRNAs that appear between 16 and 24h after infection (60). The mRNA molecules encoding Gag and Gag-Pol proteins are translated into large polyproteins of 55 kD and 160kD in size, respectively. The 160kD Gag-Pol polyproteins are synthesised as a result of a frameshifting event that occurs during translation of *gag* mRNA at a frequency of approximately 1:20. These polyproteins are translated and imported into the endoplasmic reticulum (ER) and targeted to the cell membrane by virtue of an N-terminal myristoylation reaction (112, 179, 229). The assembly of the structural polyproteins at the cell surface and association with two strands of unspliced viral RNA (occurring as a result of packaging signal sequence located at the 5'-end of the viral RNA) allows the formation of a pre-core structure. Although there is some evidence to suggest that active RT can be found within infected cells (162), the complete cleavage of polyproteins to discrete individual proteins with enzyme activity (*eg.* RT and IN) and formation of the mature (condensed) core structure is thought to occur after pre-core assembly and virion budding from the cell (255, 265).

Env proteins are translated into the endoplasmic reticulum (ER) and, if not for the action of Vpu, would spontaneously associate with CD4 molecules present in this compartment (166). Vpu is thought to promote CD4 degradation within the ER, thereby freeing Env proteins and making them available to be incorporated onto the surface of budding virions. This action also serves to down-regulate the CD4 molecule on the surface of cells, thereby suppressing superinfection or re-infection of infected cells with

progeny virus. A second mode of CD4 down-regulation, in which the Nef protein promotes the endocytosis of CD4 at the cell surface, has also been described (168).

1.2.7 The Alternative Route: Cell-to-Cell infection

Like other viruses (*eg* Herpesviruses, Rhabdoviruses and other Retroviruses), some HIV strains are able to initiate infection via an alternative route known as cell-to-cell infection (104, 160). Cell-to-cell HIV infection is initiated when an infected cell expressing gp120/gp41 complex on its' surface contacts an adjacent, uninfected T cell. This interaction leads to fusion of the cellular membranes and the generation of large, short-lived, multinucleate cells known as syncytia. Following fusion of cellular membranes, infection is thought to proceed in a similar manner to that outlined above (see sections 1.2.1-1.2.6) for cell-free infection. The finding that cell-to-cell transmission can occur in the presence of neutralising antibody *in vitro* (92, 104, 196) indicates that the fusion of cell membranes may also occur through a mechanism other than simply the interaction of the gp120/gp41 complex with surface CD4. This is further supported by recent evidence demonstrating that a compound able to block cell-free infection by perturbing the redox chemistry of the CD4 molecule was unable to block cell-to-cell infection (Matthias LJ, Yam PTW, JiangX-M, Jaramillo A, Pourmbouris P, Vandegraaff NA, Li P, Donoghue N and Hogg PJ, *submitted for publication*).

Although conclusive evidence for cell-to-cell infection *in vivo* is lacking, syncytia have been observed in the brains of HIV positive individuals (225). Furthermore, the appearance of syncytium inducing (SI) strains *in vivo* has been correlated with the progression to AIDS (83). Due to the proximity requirements of cells for initiation of cell-to-cell infection to occur, this mode of transmission may be particularly important in areas of high CD4⁺ cell density such as the lymph nodes (T cells) and the brain (Macrophages). Seeding of these anatomical regions with SI strains might potentially contribute significantly to viral dissemination *in vivo* and perhaps even be a determining factor in the onset of AIDS in infected patients.

To date, a number of experiments evaluating the kinetics of viral nucleic acid synthesis have been performed using a cell-to-cell infection protocol (8, 130, 131, 158, 160-162). The cell-to-cell infection protocol allows an extremely efficient infection, abrogates the need to generate and store a viral stock, and bypasses the requirement for methods to enhance the uptake of free virus into cells. However, the presence of pre-existing viral DNA (and/or RNA) in infected (donor) cells can often make data difficult to interpret, particularly when viral nucleic acid accumulation and transcription is being monitored. Furthermore, the kinetics of early events following cell-to-cell infection has been proposed to occur at a faster rate than in cell-free infection (160). This was attributed to the bypass of the initial attachment/uncoating steps in the viral life cycle and transference of various intermediate components of the replication process from the persistently infected donor cell.

1.3 HIV Integration

1.3.1 Background

For a productive infection to occur, HIV must direct the integration of newly reverse transcribed DNA into the host cellular DNA (76, 151, 218, 238). Several groups have performed infection experiments using HIV-1 virions containing a mutation in the integrase gene which abolishes activity of the protein *in vitro* (see section 1.6.3.i below) (76, 151, 218, 237, 238). In the vast majority of cases, minimal RNA and protein expression was observed and in all cases a productive infection did not result. However, it was recently demonstrated that unintegrated viral DNA was able to direct the transcription and translation of the *tat* and *nef* genes in quiescent T cells (270). Taken together, these experiments indicated that although extrachromosomal DNA forms may direct synthesis of and contribute to the pool of viral proteins, integration is necessary for the production of infectious virus.

1.3.2 The Integration Reaction

A simplified integration reaction can be performed *in vitro* in the presence of divalent metal ions using purified IN protein and oligonucleotides mimicking the U3 regions of

the viral LTR (30, 56, 132) (see section 1.6.3.i); this reaction has been used to extensively characterise the mechanisms of retroviral integration. Although purified IN alone is able to facilitate a simplified integration reaction *in vitro*, a highly ordered positioning of both viral and cellular proteins within the PIC is likely to be required to facilitate the coordinated insertion of two viral DNA ends into host DNA in whole cells.

The process of HIV integration (see Fig. 1.8) can be divided into three main steps:

- 1) 3'-end processing, involving the removal of a di-nucleotide from the 3'-termini of the linear viral DNA molecule
- 2) strand-transfer, in which both 3'-ends of the viral DNA are covalently linked in a concerted fashion to cleaved host cellular DNA
- 3) gap repair, where the 5' ends of viral DNA are trimmed and ligated, and gapped regions generated by the strand-transfer reaction are filled.

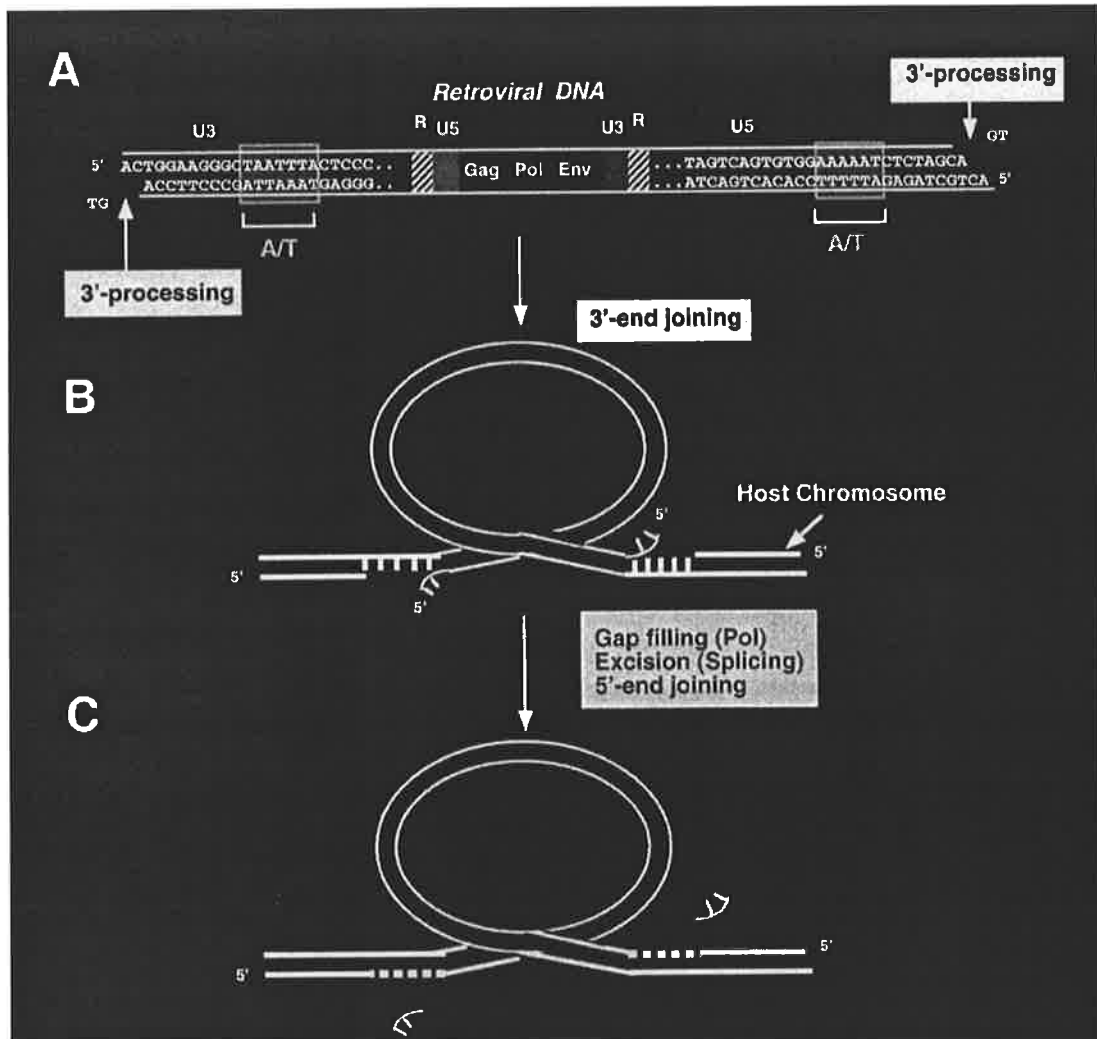
1.3.2.i 3'-end Processing

The 3'-end processing reaction is thought to occur as soon as synthesis of the ends of dsDNA is complete (227). This process results in the removal of the two terminal 3' nucleotides at either end of the linear viral DNA molecule generating staggered termini (see Fig. 1.8A) and may serve to remove non-templated nucleotides that are occasionally added to the termini of newly reverse transcribed DNA by the reverse transcriptase protein (178, 190). Integrase-mediated 3'-end processing *in vitro* has been shown by mutational analyses to be dependent on the presence of divalent metal ions (267) and a highly conserved 5'-CA-3' di-nucleotide situated on the viral DNA molecule immediately adjacent to the targeted nucleotides (29, 150, 227, 243, 244). Furthermore, mutation of various combinations of the four nucleotides immediately adjacent (internal) to the conserved CA di-nucleotide have been shown to affect the catalytic properties of IN *in vitro* (135, 136, 174, 252, 253). Taken together, these results suggest that the IN enzyme interacts with the terminal six or seven nucleotides of the linear viral DNA. Conserved sequences functionally analogous to this region (known as the attachment (*att*) site) to which the viral integrase enzyme is proposed to bind have been identified in other retroviruses (70, 128, 228).

Figure 1.8 Integration of HIV-1 DNA

Diagrammatic representation of the three (**A-C**) basic steps in HIV-1 integration. (**A**) The sequence of the termini of the HIV-1 (U455 strain) *att* sites is shown. Conserved A/T stretches are included in the grey boxes and labelled "A/T". In the 3'-end processing reaction, the viral integrase protein (IN) removes the last dinucleotide 3' to the conserved CA. (**B**) Concerted integration (3'-end joining or strand transfer) of both viral DNA termini into a host chromosome (note the five-nucleotide stagger). (**C**) Gap filling, excision, and 5'-end joining.

Adapted from Pommier and Neamati, 1999 (200).



1.3.2.ii Strand-transfer

The strand-transfer or "joining" reaction (Fig. 1.8B) describes the trans-esterification reaction that results in processed 3' viral cDNA termini being ligated to the 5'-ends of cleaved cellular DNA. Initially, integrase mediates the nicking of host cellular DNA on both strands at sites between four and six base pairs apart (75). This generates termini with staggered 5' overhangs to which the 3'-hydroxyl groups present on the underhangs of the processed viral DNA covalently couple. Like the 3'-end processing reaction, this process has been shown to require the presence of a divalent metal ion such as Mg^{2+} or Mn^{2+} *in vitro*. Both the 3'-end processing and the strand transfer reactions absolutely require the viral integrase enzyme (75).

1.3.2.iii Fill-in and Ligation

Following strand-transfer, the 3'-ends of the cellular DNA are extended over the single-stranded regions that are present as a result of proviral insertion, and ligated to the 5'-ends of the viral DNA (Fig. 1.8C). In addition, the mis-paired di-nucleotide present at the 5'-ends of the viral DNA molecule is removed. This process of gap repair generates short duplicated host sequences (5 bp) at either end of the fully integrated provirus. Whether a host enzyme or co-factor may mediate or assist the viral integrase protein in the process of gap repair in cells remains to be fully elucidated. Evidence has been obtained *in vitro* suggesting that IN possesses both a splicing activity (or disintegration activity) able to direct the cleavage then ligation of a target DNA molecule and a polymerase activity (2). However, PIC-mediated gap repair is inefficiently performed *in vitro* (107) supporting a role for host enzymes in this process *in vivo*. Furthermore, Daniel and colleagues demonstrated that murine scid cells defective for the DNA repair protein DNA-dependent protein kinase (DNA-PK), underwent apoptosis in response to infection with three different retroviruses (58). In contrast, infection with integrase-defective viruses did not result in cell death suggesting that the gapped intermediate DNA structure resulting from insertion of proviral DNA into cellular DNA is recognised as DNA damage and repaired by the DNA-PK-mediated pathway. However, Baekelandt and coworkers have recently challenged this conclusion by demonstrating

that cells deficient in the catalytic subunit of DNA-PK could be efficiently and stably transduced by an HIV-1-derived lentivirus vector.

1.3.3 The Integrase Protein

The IN protein is encoded by the 3' end of the pol gene and is translated as part of a large gag-pol polyprotein. This polyprotein is subsequently cleaved by the viral protease to release the composite proteins, which includes a 288 amino acid integrase protein. Proteolytic and mutational analyses have allowed the identification and characterisation of 3 distinct domains within the IN protein (Fig. 1.9A). Furthermore, although crystallisation of the full-length integrase has not been achieved, the crystal structures of each individual domain have been elucidated (40, 78, 256). Deciphering the structure of these domains (in particular the catalytic core domain, see Fig. 1.9B) has contributed greatly to our understanding of the structural and functional nature of the protein and, through co-crystallisation studies with inhibitors of IN, has aided in processes of rational drug design targeting the integration process *in vivo*.

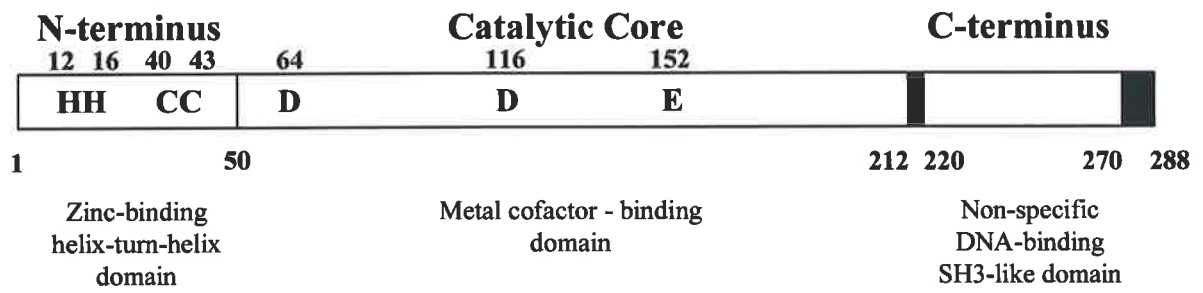
1.3.3.i The N-terminal Domain

The precise function of the small N-terminal domain (residues 1-50) remains unclear. This domain consists of an amino acid motif containing highly conserved histidine and cytidine residues (HHCC) which has been shown to bind a single zinc ion (25, 280). The binding of zinc ions to this N-terminal domain has been shown to enhance multimerisation *in vitro* of the full-length integrase protein, which is required for activity (156, 280). In addition to directing multimerisation, both structural analyses and cross-linking studies have indicated that this region might also confer the ability of integrase to bind to DNA (33, 113). However, conflicting evidence demonstrating that N-terminal deletion mutants of integrase can efficiently bind to DNA has been obtained (249). Furthermore, the region that has been proposed to bind to DNA based on similarities to other helix-turn-helix DNA binding domains (including that of Trp repressor), has been shown to form part of the dimer interface and would therefore not be available for binding to DNA (33).

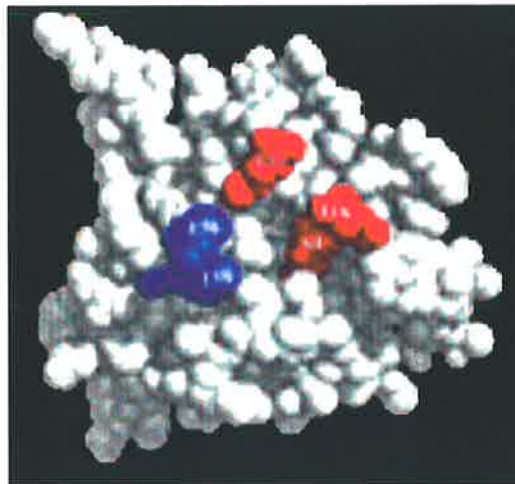
Figure 1.9 HIV-1 Integrase Protein

(A) Domain structure and conserved amino acid residues of the 288 amino acid HIV-1 integrase protein (IN). Conserved residues within the N-terminal (N), core (Catalytic Core) and C-terminal domains are indicated. (B) Model of the HIV-1 IN core domain. The amino acids comprising the active site of the enzyme (D₆₄, D₁₁₆ and E₁₅₂) are indicated (red, red and blue, respectively). Adapted from Esposito and Craigie, 1999 (78).

A.



B.



1.3.3.ii The Core Domain

The central domain (amino acids 50-212) was originally identified as a protease resistant region (73) and has since been shown to contain the catalytic site of the integrase protein. This domain consists of a triplet of amino acids (Asp-64, Asp-116 and Glu-152) that are conserved among all retroviruses with respect to both identity and position. These amino acids (known as the D,D(35)E motif) have also been observed in the transposase proteins of many mobile genetic elements (148) and their presence has been shown to be critical for the action of integrase. Correlations with other transposases have indicated that these residues coordinate a Mg^{2+} ion. Supporting this, the central domain of avian sarcoma virus (ASV) has been crystallised with a Mg^{2+} ion associated with it (21). Mutation of any one of these residues, or of amino acids known to stabilise the catalytic site, totally abolishes the catalytic activity of integrase in both cell and cell-free assay systems.

1.3.3.iii The C-terminal Domain

The C-terminal domain (amino acids 220-270) is the most diverged across retroviral integrase proteins. This domain possesses a sequence-independent DNA-binding ability and may bind target DNA during the strand-transfer (see above) reaction (74, 137, 167, 251). Upon dimerisation, a saddle-shaped cleft is formed that contains a number of positively charged amino acids which are proposed to interact with the phosphodiester backbone of DNA (137, 167, 251). Although the solution structure of this domain was shown to resemble the Src homology 3 (SH3) domains found in many signal transducing proteins (164), the role such a structure might play in the integration process remains unclear.

1.3.3.iv Multimeric Integrase

Studies showing that N-terminal deletion mutants of HIV integrase could be complemented by the addition of C-terminal deletion IN mutants *in vitro* indicated that the active form of HIV integrase is a multimer (71, 245). Consistent with this, IN proteins have been shown to spontaneously associate to form dimers, tetramers and

even higher order oligomers in solution (123, 126, 156, 280). As outlined in section 1.3.2.iii, integrase cleaves host DNA at two points separated by a distance of 5bp (249, 250). This distance is too small to extend over the region predicted to exist between the active sites of an IN homodimer. Taken together, these findings suggested, like the MuA transposase protein, the active form of the IN protein is a tetramer or higher order multimer.

1.3.4 Sequence Specificity

Integration of HIV DNA is generally considered to occur at random with respect to cellular chromosomal nucleotide sequence. A number of studies have assessed the sequence of the cellular DNA immediately adjacent to the integrated provirus and suggested that the G/C content and/or the presence of repeat elements known to exist throughout the human genome might influence target site selection (14, 88, 202, 236). However, the largest study undertaken to date (in which the DNA adjacent to 61 integrated proviruses sequences was analysed) found no bias in the G/C content of flanking DNA, and no preferential integration near sites containing human repeat sequences such as the *Alu* or LINE-1 elements (36). Although integration into host DNA seemingly occurs in a sequence independent fashion, preference for regions of distorted cellular DNA tertiary structure and/or tethering of the PIC to chromatin-bound proteins may influence the integration process *in vivo* (28, 31, 101, 133).

1.3.5 Host Proteins Involved in HIV DNA Integration

As indicated in section 1.2.2, a number of host proteins, including histones, the Barrier to Autointegration Factor (BAF) and the high-mobility group protein (HMG I(Y)), have been shown to be associated with the PIC (39, 80, 159). Although the presence of HMG I(Y) in the PIC was shown in *in vitro* assays to be essential for the integration process (see section 1.4), the exact role of this protein in the integration process remains unclear. However, studies using salt-stripped MoMLV PICs showed that the BAF protein was required to prevent intramolecular integration (or autointegration) (154, 155).

In addition to proteins directly implicated in facilitating integration of viral DNA into a target DNA substrate, a number of host proteins have been shown to considerably enhance the efficiency of the integration process. Ini-1 (a protein associated with the modification of chromatin structure to facilitate transcription) was shown to both bind tightly to integrase and enhance the integration efficiency 10-20 fold in *in vitro* assays (27). Similarly, the HMG-1 protein (distinct from HMG I(Y)) has been shown to enhance the activity of purified ASLV integrase 2- to 4-fold (4). Although shown to affect the efficiency with which integration occurs, neither of these two proteins have been identified within the HIV PIC.

In conclusion, the HIV PIC is a highly ordered nucleoprotein structure that is comprised of a variety of viral and host proteins. Furthermore, there is evidence to suggest that the composition of PICs may vary depending on their localisation within the cell. Specifically, PICs isolated from the nuclei of infected cells were smaller (as judged by sedimentation through a sucrose gradient) than those isolated from the cytoplasm (130). In addition to the constituent proteins of the PIC, additional interactions with cellular proteins associated with the host DNA may be required to ultimately coordinate the concerted integration of both viral DNA ends into the cellular chromosome.

1.3.6 *In Vivo* Significance of Integrated HIV DNA: Persistent Viral Reservoirs

Aside from the absolute requirement of integration in order for a productive infection to occur, the integration of viral DNA results in a very stable form of the virus that can persist within the cellular DNA for the life of the cell. This stability has major consequences for viral persistence in patients, including those undergoing Highly Active Antiretroviral Therapy (HAART).

Patients with high viral loads who commence HAART generally display a rapid decrease in plasma viraemia over the first few weeks (115, 193, 258). The initial, rapid drop in viral load is followed by a second phase in which levels of virus continue to decrease, but at a slower rate. This second phase has been attributed to the turnover of longer-lived populations of cells harbouring virus ($t_{1/2}$ =1-4 weeks) such as infected

macrophages (less susceptible to effects of killing), infected CD4⁺ T cells of low activation state and follicular dendritic cells that attract virions to their cell surface (12). Since two months of continuous therapy are generally required to reduce viral loads to undetectable levels in the blood, it was initially predicted that 3-4 years of continuous treatment would be required in order to eliminate the virus from the body entirely (193). However, this estimate was based on the assumptions that; 1) reservoirs of virus could not persist for the duration of treatment and 2) a complete block in viral replication could be achieved by the antiviral regimen used. The demonstration that patients on HAART who had exhibited undetectable serum viral loads for periods of greater than two years, displayed a rapid viral rebound upon cessation of therapy prompted investigators to look for long-lived sites of viral persistence in the body (45, 47, 49, 86).

Current antiviral regimens administered to patients attempt to specifically block various aspects of the viral replication cycle (namely, the processes of reverse transcription and proteolytic processing of viral proteins) and therefore target only actively replicating virus. Any virus able to persist within cells for the duration of therapy might therefore be able to direct the synthesis of progeny virus upon removal of therapy. To avoid immune surveillance for such a period, stably integrated viral DNA within infected cells was proposed to exist in a transcriptionally inactive form.

A prime candidate for a cell type harbouring integrated HIV DNA in a transcriptionally inactive form was proposed to be the memory T cell. Memory T cells are generated following T cell proliferation in response to antigenic stimulation (3, 12). A proportion of the proliferated cells then revert to a resting (or quiescent) memory phenotype that is able to quickly respond (by activation) to future encounters with the same antigen. Therefore, populations of activated T cells that escape the cytopathic effect of the virus and revert to a memory phenotype soon after infection with HIV might be expected to harbour persistent provirus. The quiescent nature (in which transcription is minimal) combined with the long-lifespan of these cells ($T_{1/2}$ of months to years (177)) was predicted to provide an environment in which integrated viral DNA could persist in the presence of HAART. The presence of stably integrated viral DNA was subsequently confirmed in populations of memory T cells isolated from patients on HAART with undetectable plasma virus levels (48). The ability to culture virus from these cells upon activation in cell culture demonstrated that the integrated viral DNA present existed in a

replication competent form (49). Further work demonstrated that there was little decrease of this viral reservoir over 2 years, as replication competent virus could be isolated from resting T cells sampled from patients on HAART who had exhibited undetectable blood virus levels for periods of up to 2.5 years (86, 94). Thus, a model for viral persistence was proposed in which a productively infected T cell was able to avoid immune surveillance long enough to return to a resting memory phenotype. In the absence of stimulating antigen, the long-lived nature of memory cells would be expected to allow the integrated viral DNA to persist within patients for the duration of therapy. Moreover, when these cells become activated in response to antigen in the absence of therapy, virus production would be expected to result. Based on our current knowledge of the *in vivo* lifespan of memory T cells, it was estimated that a period of anywhere between 10-60 years of continuous treatment would be required to eradicate virus entirely (278). However, this estimate has been further complicated by the demonstration that viral replication is not totally abolished in patients receiving HAART (65, 94, 226, 271, 278), a process that may allow the continual re-seeding of persistent viral reservoirs *in vivo*. Furthermore, additional sites of viral persistence may be identified in the future, providing further challenges for viral eradication *in vivo*.

It is worth noting however, that while virus could be cultured from memory cells isolated from patients on HAART, the frequency of culturable virus was much lower than the integrated HIV DNA levels observed in memory cells (44). This indicated that most of the integrated DNA detected by PCR may reside in a defective form. Although assessed in the absence of viral suppression, this conclusion was supported by a study that showed that defective genomes were present in a high proportion of DNA isolated from the PBMCs of chronically infected patients (220). Furthermore, a 48-week follow-up study on a cohort of patients placed on regimens of HAART showed a 5-fold decrease in the amount of total HIV DNA copy number, while levels of integrated HIV DNA over this time remained similar (119). These observations suggest that although measuring levels of integrated HIV DNA in patients on HAART may prove useful for the identification of sites of viral persistence *in vivo*, the ability of virus to be cultured from these cells may be a more appropriate way of assessing individuals for the extent of viral eradication.

1.4 Extrachromosomal HIV DNA Forms

As detailed in section 1.3.1, integration is one of four fates that awaits a newly reverse transcribed linear HIV-1 DNA molecule upon entering the nucleus. Viral DNA has also been shown to remain as a linear molecule, undergo recombination between the viral LTR regions to form a circular form with a single LTR or undergo direct end-to-end ligation (191) to generate a circular form with two LTR regions (see Fig. 1.6). Since the processes of recombination and end-to-end ligation require the presence of ligases present only in the nucleus (160), nuclear entry of viral DNA is presumed to occur before circularisation can proceed.

1.4.1 Transcriptional Activity of Extrachromosomal HIV DNA

Although there is substantial evidence to suggest that extrachromosomal viral DNA can direct limited transcription following infection, these forms are believed to be unable to mediate the production of mature virions capable of initiating infection (34, 76, 151, 160, 218, 237, 238). The limited transcription exhibited by these forms may reflect the complex structure in which they reside. Both the linear viral DNA molecule and the circular viral DNA molecules have been reported to reside within the PIC, or within a complex of proteins similar to the PIC but lacking IN (22), respectively. Since footprinting studies have demonstrated that proteins within the PIC interact with the viral LTR regions (178), a physical masking of elements responsible for transcriptional control may account for the inefficient transcription observed for these forms.

1.4.2 *In Vivo* Significance of Extrachromosomal HIV DNA

Extrachromosomal HIV DNA is generally considered to be actively degraded by nucleases within the nucleus and therefore relatively short-lived ($T_{1/2} \approx 2$ weeks (226)). Consequently, the successful PCR amplification of circular forms from cells isolated from patients has been used as a convenient marker of recent infection *in vivo* (226). This recent association of the circular forms with active viral replication also supports findings demonstrating that high levels of unintegrated viral DNA (in particular circular DNA) in PBMC are predictive of progression to AIDS (185, 211). Moreover, detection

of circular DNA molecules has indicated that ongoing viral replication occurs in patients on highly active anti-retroviral therapy (HAART) (226). It seems that while extrachromosomal DNA can be used as a useful diagnostic marker for virus replication, the biological contribution and significance (if any) of such viral DNA forms to the viral replication process *in vivo* remains unclear.

1.5 The Kinetics of Viral Nucleic Acid Metabolism in Cell Culture

Analysing the kinetics of HIV DNA synthesis following a one-step infection is critical to our understanding of basic HIV virology. Consequently, numerous attempts have been made to accurately monitor the accumulation of all major forms of HIV DNA following infection of cells in culture. Due to difficulties associated with macrophage culture and generating high titre preparations of macrophage-tropic HIV, our knowledge of the kinetics of HIV DNA accumulation has been largely limited to the analysis of DNA preparations made at various times following infection of T cells with T-tropic strains. Although there is substantial data available on the accumulation of total and circular viral DNA following infection, our understanding of the timing and extent of viral integration into the host cellular chromosome remains poor.

1.5.1 Total and Circular Viral DNA Accumulation

To date, the accumulation of total HIV DNA has been assessed following infection of a number of different T cell lines using a variety of virus strains and infection protocols (8, 131, 160, 238). The total HIV DNA complement at various times after infection has been monitored using either direct Southern hybridisation protocols on cellular DNA preparations or PCR-based techniques. PCR protocols designed to detect near full-length HIV DNA have generally utilised primers designed to anneal within the Gag or 5'U3-Gag regions of the viral genome (131). Such primers detect mid to later-stage reverse transcribed products at a stage in the reverse transcription process after the first template switch (see section 1.2.2). Following infection of T cells with free virus, near full-length DNA has routinely been observed between three and four hours post infection (h p.i.) (138, 157, 238). In two independent reports evaluating the HIV DNA

kinetics profile following cell-to-cell infection, reverse transcribed DNA was detected by 4h p.i. while the circular forms shown to first appear at 8h p.i. (8, 160). Furthermore, the 2-LTR species was shown to be a minor population compared to the 1-LTR and linear forms over the course of infection (8).

1.5.2 Integrated Viral DNA Accumulation

In contrast to the bulk of work assessing the accumulation of newly reverse transcribed linear and circular viral DNA forms, the accumulation of integrated DNA within infected cells following HIV infection has received limited study. This has been primarily due to the lack of an appropriate assay able to clearly distinguish between the integrated and extrachromosomal forms. Such an assay is required as chromosomal preparations following infection with HIV invariably contain significant contaminating extrachromosomal HIV DNA forms (8, 192, 238, 274).

In a cell-to-cell infection model, Barbosa and co-workers (8) indirectly estimated the amounts of integrated HIV DNA by subtracting the levels of extrachromosomal DNA from the total viral DNA complement. Using this approach, they found that the ratio of integrated DNA to unintegrated DNA at 24h p.i. was approximately 1:3 (34%). Furthermore, although acknowledging the poor transfer efficiency to Southern blots of high molecular weight DNA, these researchers demonstrated the presence of HIV DNA associated with high molecular weight chromosomal DNA by 12h p.i.. However, at the time this study was commenced, there had been no direct assessment of the accumulation of integrated HIV DNA over time following either cell-free or cell-to-cell transmission of HIV.

Recently however, Butler and colleagues reported a study that monitored the accumulation of integrated HIV DNA following infection of cultured T cells with pseudotyped HIV virions (capable of one round of infection) (32). Integrated viral DNA was directly monitored in their study by a modified nested-*Alu* PCR protocol (see section 1.7.3 below) similar to the assay presented in section 4.10 of this thesis. Although the low degree of assay sensitivity achieved in their hands may have resulted in the inability to detect integrated HIV DNA at earlier time points, integrated HIV

DNA was first detected at 7h p.i. and ultimately accounted for approximately 10% of the total amount of viral DNA reverse transcribed following infection. Although the efficiency of HIV integration appears lower than that observed by Barbosa and coworkers (8), the direct method used by Butler and colleagues to monitor integrated viral DNA is likely to be a more accurate predictor of integration levels than the indirect methods used in the former study.

Since an appropriate assay for the detection of integrated HIV DNA was unavailable at the time this project was initiated, and very little data had been collected on the efficiency of proviral integration following HIV infection in cell culture, one of the major aims of the study presented in this thesis was to establish a novel PCR-based assay designed to directly determine levels of integrated HIV DNA within cells. Such an assay was expected to not only allow the accurate assessment of integrated HIV accumulation following infection of target cells, but also to be a useful tool for the evaluation of inhibitors of HIV integration both in cell culture and *in vivo*.

1.6 Inhibition of HIV Integration

1.6.1 Antiviral Treatment of HIV Infected Individuals

As of 1999, 60 drugs were in use and greater than 100 new medicines for the treatment of AIDS, and AIDS-related diseases were being tested. Currently, combinations of antiretroviral agents are used to treat HIV-infected individuals. The exact combination of drugs used for this treatment (known as Highly Active Antiretroviral Therapy or HAART) differs between individuals, but generally consists of three or more compounds each designed to inhibit either the viral reverse transcriptase protein or the viral protease protein. These two viral enzymes are unique to, and essential for, the HIV replication process. The use of multiple drugs targeting different stages of the viral replication process (as opposed to the use of single drug regimens) has greatly reduced the frequency with which viral escape mutants displaying drug resistance arise. Although combinations of these drugs have been shown to decrease viral loads to undetectable levels indefinitely in patients (provided a rigorous adherence to the drug regimen administration is maintained) (103, 105), the numerous side effects (17, 35,

212) and the inability of the drugs to entirely block viral replication (278) have hampered the successful treatment of many patients. Furthermore, when patients on long-term treatment with undetectable viral loads are removed from therapy, virus levels rapidly rebound (86, 269). This has been determined to result from both the re-activation of virus from within persistent viral reservoirs (see section 1.3.6), and the outgrowth of actively replicating virus present at low levels in the presence of treatment. While the existence of sites in which viral replication can occur in the presence of therapy have been proposed (12), the extent and distribution of such sites remains unclear.

Clearly, additional drugs, that are designed to inhibit alternative aspects of the HIV replication cycle and able to freely access all sites of ongoing replication, will be required if complete viral eradication from an infected individual is to occur. Consequently, a number of approaches for intervention in key steps in the HIV replication cycle are being investigated, including the identification of compounds able to inhibit the HIV integration process.

1.6.2 HIV Integration as a Target for Intervention

In addition to the RT and PR proteins, the IN enzyme is another prime candidate against which antiviral drugs can be designed. This protein has no known cellular counterparts (125) and is absolutely required for the virus to replicate (see section 1.3.1). To date, few drugs specifically targeting the action of the IN protein have been identified. Two main reasons are responsible for this. Firstly, the structural characterisation (and hence, functional characterisation) of the active integrase protein complex has been slowed in part due to the inability of the protein to form crystals. However, individual domains have since been crystallised, which together with mutational analysis data, has allowed the detailed characterisation of the protein. Although the crystal structures of individual domains have been obtained, not knowing the number of proteins involved in the active integrase multimer and their modes of interaction remains a major obstacle for drug design. This is in contrast with both the RT and PR viral proteins, for which the complete crystal structures were obtained relatively early after their initial purification (11, 59, 121, 234, 266). Secondly, assays able to accurately assess the extent of

integration *in vivo* (both in cells and in patients) by allowing the specific detection of integrated HIV DNA species have only recently been developed (see section 1.7). Although some candidate inhibitors of integration have been shown to inhibit syncytia formation and/or virion release (as measured by P24 levels and/or reverse transcriptase activity in culture supernatants) following infection of cells in culture, demonstration of a direct effect on the integration process has been thus far limited to cell-free assay systems. Therefore, a major aim of the work presented in this thesis was to establish a cell-based infection model, that when used in conjunction with a PCR-based assay that specifically detects integrated HIV DNA, could allow the screening of potential inhibitors of integration in cell culture (see chapter 6).

1.6.3 Current Screening and Characterisation of Inhibitors of Integration

To date, the broad-range screening of potential integrase inhibitors has been primarily performed in *in vitro* systems using either purified integrase alone or within the context of a partially purified preintegration complex (26, 82, 84, 109, 110, 118, 163). Since these assays can be designed to test for inhibition of either the formation of the initial stable complex, 3'-end processing, strand transfer or disintegration (the reverse of strand-transfer), they can not only rapidly identify potential inhibitors but can also provide preliminary evidence as to their mode of action. However, inhibitors of integrase protein and/or the preintegration complex identified in this manner may not specifically inhibit HIV integration in infected cells, and are frequently cytotoxic or do not exhibit antiviral activities when tested in cell culture (200).

1.6.3.i In Vitro Assays Using Purified Integrase

The most commonly *in vitro* assays used to screen for integrase inhibitors assess the efficiency with which the 3'-end processing and 3'-end joining (or strand transfer) reactions occur using integrase purified from *Escherichia coli* expression systems (68, 75, 200). Although both assays require the presence of either Mn²⁺ or Mg²⁺ (69, 268), the preferred metal ion utilised was shown to depend on the reaction conditions used (72). Due to the relative concentration in cells, it is likely that the divalent magnesium

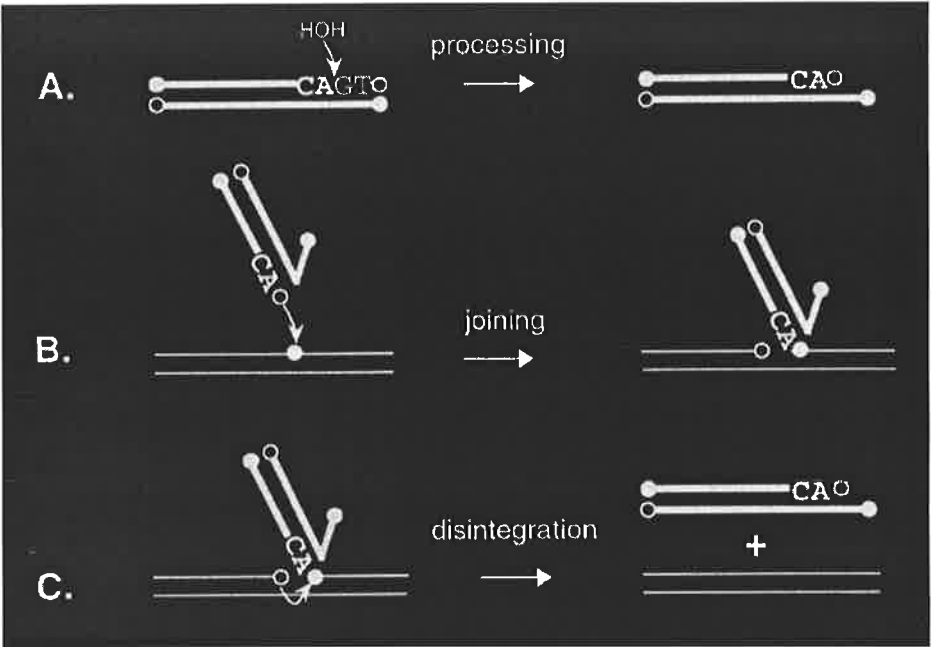
ion serves to facilitate this reaction *in vivo* (7). The 3'-end processing reaction (Fig. 1.10A) is used to assess the efficiency with which a di-nucleotide is removed from the end of a short, labelled DNA target mimicking the 5'-HIV LTR region. DNA products reduced in length by 2 nucleotides are distinguished from unprocessed targets by gel electrophoresis. The efficiency with which strand transfer reactions occur is assessed by monitoring the integrase-mediated integration of substrate oligonucleotides (mimicking the 3'-end processed form of the 5'-HIV LTR region) into target DNA molecules (Fig. 1.10B). Y-shaped integration products are resolved from target substrates by electrophoresis. Compounds exhibiting anti-integration abilities in these assays are often more efficient at inhibiting the strand transfer than the 3'-end processing reaction (184). This observation may however reflect the relative efficiencies of each reaction *in vitro*. The strand transfer reaction has been shown to be approximately one order of magnitude less efficient than the 3'-end processing reaction *in vitro*, and may therefore be more susceptible to inhibition.

An additional integrase-mediated reaction, known as disintegration, has been identified *in vitro* (43) and has been used to further characterise integrase inhibitors with respect to their site of action within the integrase protein. Disintegration is essentially a reversal of the strand transfer reaction, and can be achieved using only the core domain of the integrase protein (43). This reaction has been shown to be dependent on the presence of an intact active site within integrase, as mutation of any of the amino acids responsible for active site integrity (for example, Asp⁶⁴, Asp¹¹⁶ or Glu¹⁵²) abolishes the ability of integrase to catalyse this reaction (66, 73). Therefore, inhibitors that can be demonstrated to interfere with the disintegration reaction are likely to exert their effects at, or near to, the integrase active site. The process of disintegration is assessed by monitoring the ability of integrase to facilitate the conversion of a Y-shaped substrate DNA molecule to two separate products (Fig. 1.10C). Although useful for characterising the site of inhibitor action, disintegration reactions are not used extensively for the high throughput screening of compounds due to the relative inefficiency of this reaction *in vitro* compared to the 3'-end processing and strand transfer reactions, and the inability to detect inhibitors targeting other areas of the integrase protein.

Figure 1.10 Schematic of In Vitro Activities of IN

Internal target and 5'-phosphates are shown as closed circles, and 3'-OHs are shown as open circles. Curved arrows denote the coupled nucleophilic cleavage and joining that is common to all actions catalysed by IN. (A) Processing, depicted as a water OH group attacking the target phosphate that follows the CA dinucleotide (shown in boldface) to release the terminal dinucleotide (in this case, GT, as is found at the HIV-1 DNA termini). (B) Joining, showing the recessed 3'-OH attacking various phosphates on target DNA to yield a set of longer products. (C) Disintegration, the reversal of the joining reaction to regenerate viral and target DNA.

Adapted from Katzman and Katz, 1999 (134).



A number of adaptations of the assays outlined above have been employed to facilitate the high throughput screening of compounds for anti-integration properties. Although rapid and convenient, these assays only partially mimic the integration reaction as it proposed to occur *in vivo*. The inability of purified integrase alone to direct the coordinated integration of two viral ends into target DNA molecules together with the absence of proteins normally present within the PIC limit the ability of these assays to identify true inhibitors integration (29, 30). Consequently, the majority of drugs identified as possessing anti-integration activity in these assays do not demonstrate antiviral properties in cells. Therefore, *in vitro* assays using PICs, that more precisely re-create the *in vivo* integration reaction while retaining the convenience and speed of the assays using purified integrase, were developed.

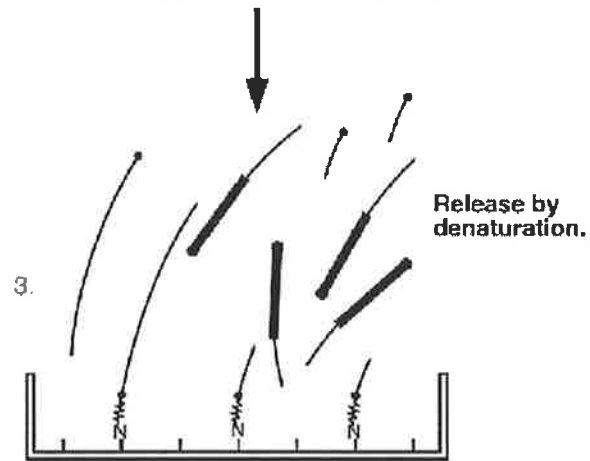
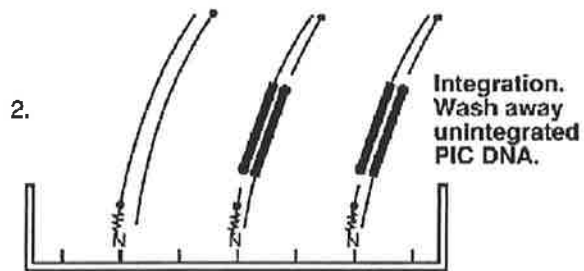
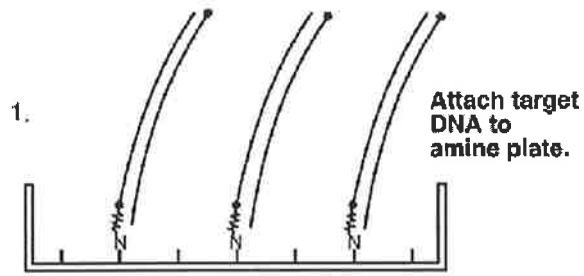
1.6.3.ii Purified PICs

As an alternative to purified integrase, assays using preintegration complexes (PICs) isolated from freshly infected cells have been developed to assess integration *in vitro* (82, 107, 110). PICs have been shown to direct the concerted integration of both viral DNA ends into target DNA substrates (107). This is in contrast to assays performed using purified integrase that direct only partial reactions involving one cDNA end. Consequently, PICs have been proposed to give a more authentic response to inhibitors *in vitro* than purified recombinant integrase proteins. Supporting this, AZT (a reverse transcriptase inhibitor) has been reported to inhibit the action of purified integrase (172), but an inhibitory effect was not demonstrated against PICs (82). Until recently, the practical complexity of assaying integration *in vitro* using PICs meant that they were not suited for use as a routine screening tool to identify inhibitors of integration. However, Hansen and colleagues (1998) have developed a simplified technique in which HIV PICs were added to microtitre wells containing immobilised target DNA molecules (see Fig. 1.11, panel 1) (107). Following incubation to allow integration and then washing to remove unintegrated viral DNA sequences (Fig. 1.11, panel 2), integrated HIV DNA is released from the wells by denaturing the viral-target DNA integrants (Fig. 1.11, panel 3) and detected by PCR amplification. Release of integrated viral DNA from the immobilised target DNA molecules was possible due to single-stranded gaps (that occur in the absence of gap repair) present on one strand at each

Figure 1.11 In Vitro Assay for Integration Using PICs

Lysates containing HIV vector-derived PICs are added to plate wells containing immobilized target DNA and incubated to permit integration. The 3' ends of viral cDNA become attached to target DNA due to integration (part 2). The 5' ends of viral cDNA do not become covalently attached in reactions with PICs. Unintegrated viral cDNA is washed away. Viral sequences are then released by denaturation, as the structure of the product is such that vector cDNA is linked to the plate by hydrogen bonding only (part 3). Released viral sequences are then quantified to determine the extent of integration.

Adapted from Hansen *et.al.*, 1999 (107).



Detect released HIV sequences.

viral-target DNA junction. This method is considerably more convenient than alternative assays using PICs, and can rapidly screen large numbers of potential inhibitors.

1.6.3.iii Cell-based assays used to indirectly measure integration

To date, few assays have been described that can precisely evaluate inhibitors of HIV integration in cell culture. The main advantages associated with cell-based assays are that they assess the ability of compounds to not only gain access to the cell, but also their ability to interact with the integration machinery in the context of the intracellular environment. Furthermore, the cytotoxicity of potential inhibitors of integration can be evaluated, allowing the selective nature of such chemicals for the viral integration process in cells to be assessed. Generally, cell-based assays have been used to determine whether the anti-integration properties of specific compounds observed *in vitro* translate to an antiviral effect in culture. For example; (i) Antiviral potency can be measured by monitoring the ability of virus to direct a productive infection in the presence of such compounds (as judged by release of either P24 or RT into the culture supernatant). (ii) A single-cycle assay following acute infection has been used extensively to assess whether integration is occurring in cells (139). This assay (termed the "MAGI" assay) involves the HIV infection of CD4⁺ HeLa cells that have been transfected with the β -galactosidase reporter gene under the regulation of HIV-1 LTR promoter elements. Integration of viral DNA and the resulting expression of the viral tat protein within cells, induces β -galactosidase expression which can then be detected using a chemiluminescent substrate. (iii) Inhibition of syncytia formation following infection (with a syncytia-inducing (SI) strain) performed in the presence of drugs can also be used to indicate antiviral activity. Although these techniques are used routinely to characterise anti-integration compounds, none allow the direct and quantitative assessment of proviral insertion. Furthermore, the inhibition of productive infection and/or proviral formation may reflect a block (or blocks) to steps of the HIV replication cycle prior to integration. Therefore, unless integrated and extrachromosomal viral DNA forms are directly measured in infected cells, a specific effect of compounds on integration cannot be determined.

Another cell culture-based technique that AIDS in the characterisation of inhibitors of viral replication, is the sequence analysis of HIV strains that are resistant to the effects of the drug. When infected cells are continually passaged in the presence of inhibitor, the outgrowth of viral strains that gain either partial, or full resistance to the action of the drug may occur. The subsequent sequence analysis of key regions (notably the integrase, protease, reverse transcriptase and gp120 genes) of viral genomes isolated from such strains often leads to the identification of consistent genetic mutations that confer drug resistance. The genetic region within which mutations correlating with drug resistance accumulate, indicates the likely target of the inhibitor. This approach has been used to confirm the specific viral targets of potential integrase inhibitors such as AR177, chicoric acid and the diketo acid inhibitors L-731,988 and L-708,906 (79, 111, 141) (see section 1.6.4).

1.6.4 Inhibitors of Integration

Very few inhibitors to date have been shown to specifically inhibit the viral integration process both *in vitro* and in cell culture. A summary of published compounds shown to possess anti-integration properties *in vitro* is presented in Table 1.3. However, the vast majority of these compounds are either inefficiently transported into cells, cytotoxic, or do not demonstrate anti-integration or antiviral effects in culture.

To date, only two compounds, both comprising a diketo acid moiety, have been proposed to inhibit viral replication by specifically targeting the integration reaction (see Table 1.3, compounds L-731,988 and L-708,906; Fig. 1.12A and B). Both L-731,988 and L-708,906 displayed 50% inhibitory concentrations (IC₅₀'s) of 80 and 150 nM, respectively for the *in vitro* strand transfer reaction using PICs, and were shown in a single-cycle (MAGI) assay for acute infection (see section 1.6.3.iii) to inhibit proviral insertion with IC₅₀'s of 1 to 2 μM (111). Furthermore, the presence of elevated levels of circular DNA forms following infection of a T cell line indicated that reverse transcription, and nuclear import of the newly reverse transcribed viral DNA, was occurring. Taken together, these results suggested a specific block to the viral integration process was occurring when infections were performed in the presence of these drugs. However, a direct evaluation of whether the drugs inhibited the

Table 1.3 HIV-1 IN Inhibitors

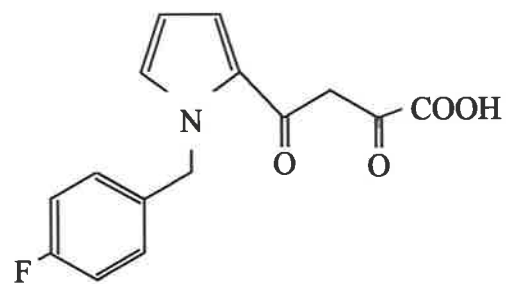
Adapted from Pommier and Neamati, 1999 (200).

Groups	Representative Drugs/Classes	IC₅₀ (μM)	Antiviral Activity	
DNA binders	Doxorubicin, mitoxantrone	1-10	No	
	Chloroquine	10	No	
	Phenanthroline-cuprous complex	1	No	
	Triple-helix forming oligonucleotides	0.05	No	
	Minor groove binders	0.1	Yes	
Nucleotides and Analogs	AZT nucleotides	100	Yes (also RT)	
	Mononucleotides	40	Yes (also RT)	
	Di- and tetranucleotides	3-300	No	
	Guanosine quartets (T30177)	0.05	Yes (also gp120?)	
Hydroxylated Aromatics	Aurintricarboxylic acid	<1	Yes (also gp120)	
	Cosalane derivatives	2	Yes (also protease)	
	Dihydroxynaphthoquinone	2	No	
	Anthraquinones (purpurin, alizarin)	10	No	
	Flavones (quercetin and derivatives)	0.1	No	
	CAPE and derivatives	3	Some derivatives	
	Biscatechols	0.5	No	
	Dicaffeoylquinic/chicoric acids	0.2	Yes	
	Arylamides	5	?	
	Hydrazides	0.2	?	
		Diaryl sulfones	2	No
		Aryl sulfonates	5	?
	Sulfones and Sulfonates	Suramin	0.2	Yes
Peptides	Hexapeptide (HCKFWW)	3	No	
	Anti-IN antibodies	0.5	?	
Diketo-acids	L-731,988 and L-708,906	0.08-0.15	Yes	

Figure 1.12 Inhibitors of Integration In Vitro

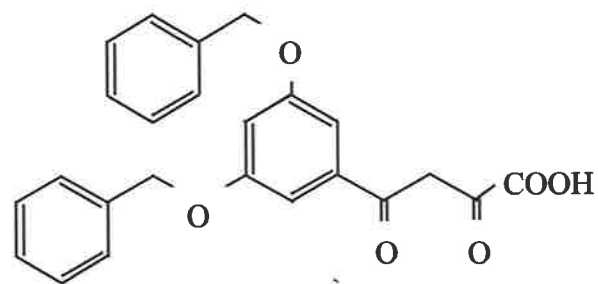
Four compounds shown to inhibit integration reaction at very low concentrations (nM) *in vitro*. Each of these compounds demonstrates cytoprotection from HIV infection in cell culture. However, recent analyses have indicated that only the diketo-containing compounds (**A** and **B**) specifically interfere with the process of integration in cells. (**A**) the diketo-acid containing compound L-731,988 (Merck) (**B**) the diketo-acid containing compound L-708,906 (Merck) (**C**) the oligonucleotide inhibitor AR177 (Zintevir) (**D**) L-chicoric acid.

A.



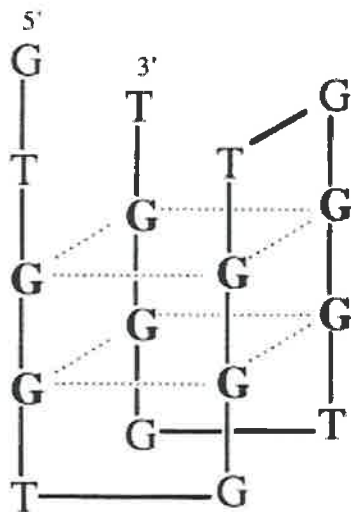
L-731,988

B.



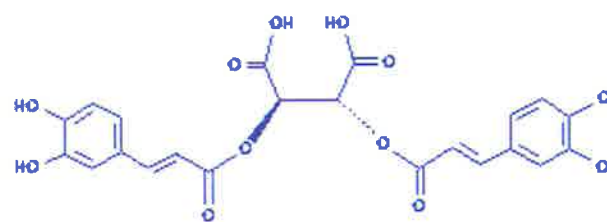
L-708,906

C.



AR177 5'-GTGGTGGGTGGGTGGGT-3'

D.



L-chicoric acid

accumulation of integrated HIV-1 DNA was not performed due to the lack of an assay that can specifically detect integrated HIV DNA.

Two structurally distinct drugs, AR177 (Fig. 1.12C) and *L*-chicoric acid (Fig. 1.12D), have been shown to inhibit *in vitro* integration reactions in the nM range (41, 163, 173, 215, 281). Furthermore, these compounds were shown to inhibit syncytia formation and productive infection in cell culture at non-toxic concentrations (79, 188, 215). AR177 is a G-quartet-containing oligonucleotide that forms highly stable intermolecular tetrad structures in solution (see Fig. 1.12C). The *in vitro* inhibition of the integrase protein has been proposed to be mediated by the interaction of the GTGT loop residues of the G-quartet inhibitor with the catalytic site of HIV-1 IN (124). In contrast to AR177, *L*-chicoric acid is a small molecule inhibitor of HIV-1 replication. However, like AR177, *L*-chicoric acid and a number of derivatives of *L*-chicoric acid (compounds containing two linked caffeoyl groups - see Fig. 1.12D) have been proposed to inhibit the *in vitro* activity of HIV-1 IN by binding within the catalytic site of the enzyme (214). Although activity against the viral integrase protein was demonstrated for both of these compounds *in vitro*, recent studies have suggested that the primary target of these drugs is the viral gp120 protein (79, 199). AR177 was shown to interfere with the binding of a monoclonal antibody raised against the V3 loop of gp120 (79). In addition, mutations that conferred viral resistance to both chicoric acid and AR177 in cell culture mapped to residues within the loop regions of gp120 protein, implying that the inhibition of infection observed may be due to a block in the viral entry process (79, 199). It was unlikely that AR177 directly inhibited the integration process within cells as the AR177-resistant viruses were shown not to possess any significant mutations within the integrase protein. However, in one study, mutations conferring resistance to chicoric acid were shown to map within the integrase gene (141). Although this conflicts with other reports (199), this finding might be interpreted as suggesting that chicoric acid targets both the processes of viral entry and integration. These studies on AR177 and chicoric acid underscore the importance of fully characterising compounds identified *in vitro* with respect to the precise viral and/or cellular target responsible for the antiviral activity observed in cell-based assays.

1.7 Whole Cell Assays for the Direct Detection of Integrated HIV DNA

The lack of a robust assay allowing the quantitative assessment of integrated HIV DNA in a population of cells has proved to be an obstacle for the comprehensive characterisation of drugs targeting the integration process, and for the identification of reservoirs of viral persistence in patients on HAART. Such an assay is required as chromosomal preparations following infection with HIV invariably contain significant amounts of contaminating extrachromosomal HIV DNA (8, 192, 238, 274). However, protocols allowing levels of integrated HIV DNA to be specifically assessed have been explored. These have either relied on the physical separation of extrachromosomal from integrated HIV DNA, or (more recently) have aimed to selectively amplify integrated viral DNA using novel applications of the polymerase chain reaction (PCR).

1.7.1 Electrophoresis/Electro-elution

Initial attempts to remove extrachromosomal viral DNA from preparations of chromosomal DNA involved the physical separation of the free viral DNA forms based on size (146, 238). The largest extrachromosomal HIV DNA form (the linear or 2-LTR species; see Fig. 1.6) is approximately 10kb in length. In contrast, chromosomal DNA fragments isolated by the HIRT method (114) (or other methods of extracting chromosomal DNA) are generally between 20kb and 150kb in length (204). Electrophoresis techniques were shown to significantly reduce the amount of free viral DNA contaminating chromosomal DNA preparations (146, 238). However, significant levels of mitochondrial DNA (which fractionates with extrachromosomal DNA) was still detected by PCR indicating that there remained substantial contamination of the chromosomal DNA with extrachromosomal DNA forms. Therefore, these methods could not totally eliminate error associated with the presence of free viral DNA forms, particularly when the highly sensitive polymerase chain reaction is employed. Moreover, since these methods are laborious, the chance of contamination of DNA preparations is significantly increased. The above drawbacks associated with these methods prompted the development of alternative methods to selectively detect integrated HIV DNA.

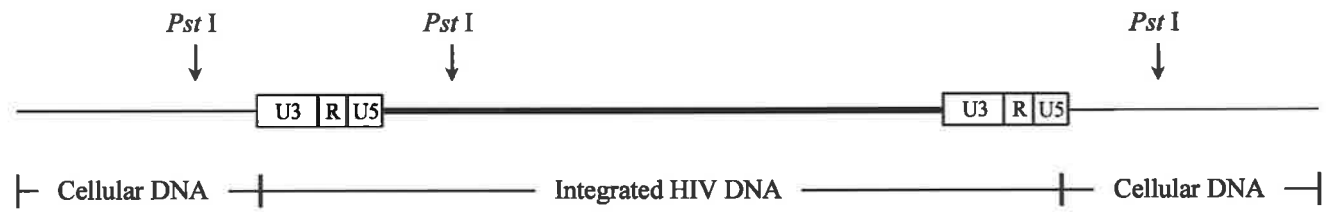
1.7.2 Inverse PCR

The inverse PCR technique (44) for the selective detection of integrated HIV DNA is outlined in Fig. 1.13. This approach involved a combination of restriction digestion and ligation reactions that ultimately generated a DNA template derived specifically from integrated HIV DNA that could be amplified by a tailored PCR procedure. Briefly, chromosomal preparations were initially digested with the restriction enzyme *Pst* 1. The *Pst* 1 restriction enzyme recognises a 6bp site and therefore cleaves the random cellular DNA sequence with an average frequency of approximately once every 4kb. The resulting linear fragments were subjected to ligation under conditions that favour intramolecular ligation (large ligation volumes). Under these ligation conditions, the DNA fragments resulting from the *Pst* 1 digestion of the 1-LTR circles, 2-LTR circles and the right-side of the integrated DNA, circularise in such a way that greater than 8kb of DNA is present between the primers to be used in a subsequent PCR (see Fig. 1.13, primers A and B). Linear extrachromosomal DNA does not circularise since a cohesive termini is not present at the 5'-end of this form. In contrast, circularisation of the left side (5'-end) of digested integrated HIV DNA yields a variable length region (1-4kb) of DNA between primers A and B. Since the *Taq* polymerase used had a defined extension rate, by limiting the extension time in the subsequent PCR, Chun and co-workers were able to restrict the amplification of products greater than 3.5kb in length. Using this technique, up to 85% of integrated DNA forms can theoretically be detected. The reported detection sensitivity of this assay was 16 copies of integrated HIV DNA per 10^6 cells when replicate samples were performed (44). However, uncertainties with respect to the variable levels of intermolecular ligation (rather than intramolecular ligation) that would be expected to occur between digested DNA fragments (regardless of ligation volume), and the varying extension efficiencies of long amplicons (up to 3.5kb) in a mixed population of integration events, has meant that the use of this assay has been limited to one study (44). Consequently, the preferred assay for the detection of integrated HIV DNA to date has been the nested *Alu*-PCR procedure (see below).

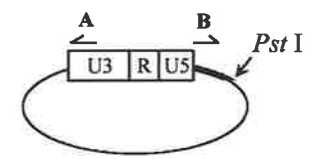
Figure 1.13 The Inverse-PCR Method for the Detection of Integrated HIV-1 DNA

Sample DNA is initially digested with *Pst* I and then subjected to ligation under dilute conditions. Inverse PCR is then performed using outward-directed LTR and *gag* primers (indicated as primers A and B, respectively). A nested PCR (using *gag* primers C and D) is then performed on an aliquot of the inverse PCR product to generate a final product of defined size which is then detected by Southern hybridisation using an internal probe (E).

Adapted from Chun *et.al.*, 1997 (44).



↓ Inverse PCR



↓ Nested PCR



1.7.3 Nested-*Alu* PCR

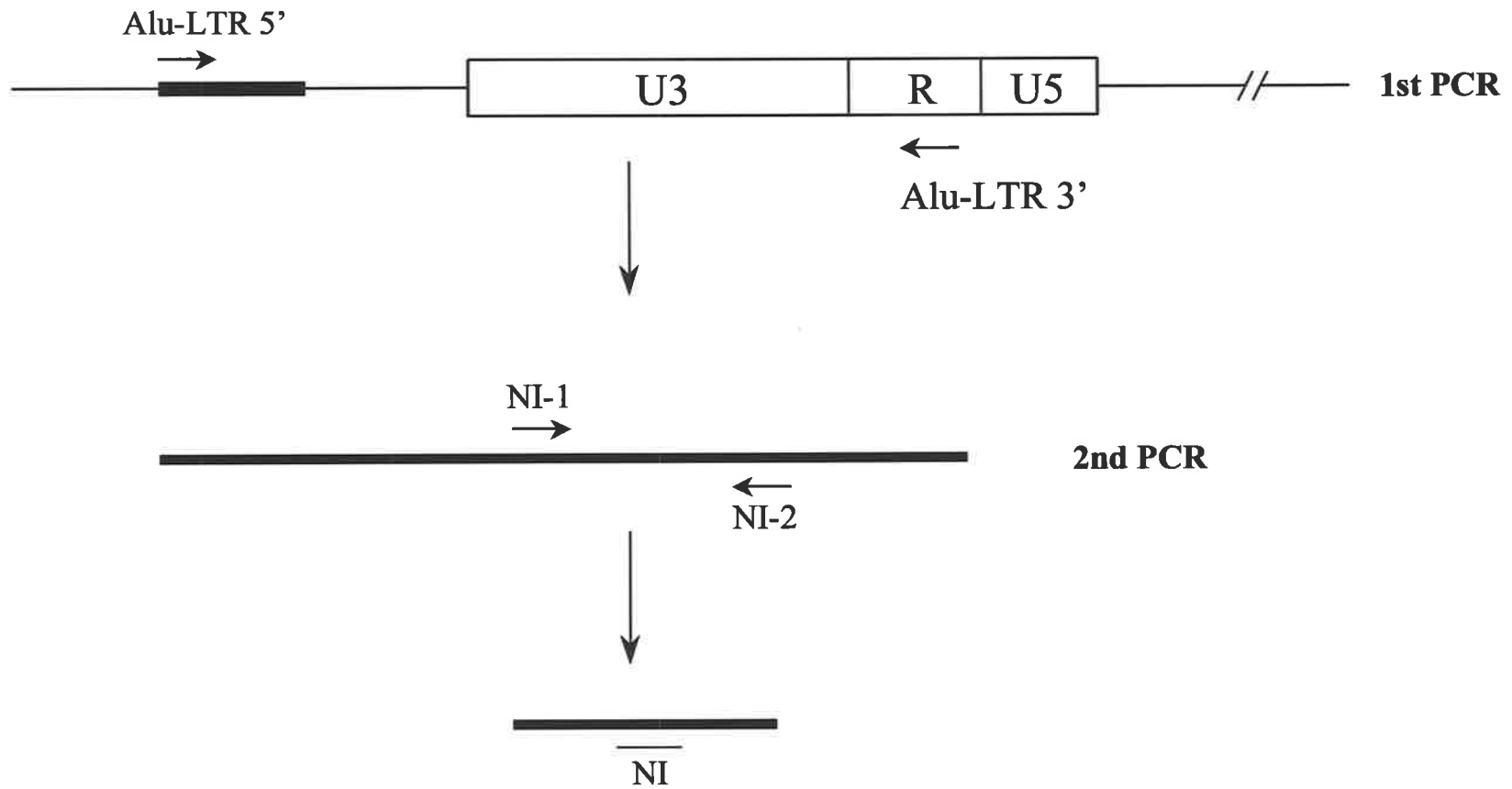
To date, a nested-*Alu* PCR protocol has been the preferred method for the detection of integrated HIV DNA. The demonstration that this assay allowed the highly sensitive and reproducible detection of integrated HIV DNA within the ACH-2 and U1 cell lines led to its extensive use for quantifying reservoirs of persistent infection in HIV positive patients undergoing HAART (49). This technique (based on a protocol initially described by Sonza and colleagues (1996) for the qualitative detection of integrated HIV DNA, (231)) utilised the presence of highly conserved *Alu* repeat elements that exist throughout the human genome (see Fig. 1.14). While *Alu* elements are not present in HIV DNA, approximately one million of these 300bp elements have been reported to exist per human genome (221, 235). This figure translates to an average frequency of one *Alu* element every 4kb of cellular DNA sequence. Since these elements are highly conserved with respect to sequence, a primer designed to anneal within this sequence was used in conjunction with an HIV-specific primer to amplify one end of the integrated HIV DNA sequence and adjacent cellular sequence (Fig. 1.14). In a mixed population of integrants, this amplification procedure generates products of varying lengths depending on the proximity of the nearest *Alu* element. Therefore, a nested PCR was performed on a dilution of the 1st-round PCR products to allow the visualisation and quantification of a discrete band corresponding to amplified integrated HIV DNA (49).

Although this technique has been shown to detect integrated HIV DNA within both the ACH-2 and U1 cell lines, and within chromosomal DNA extracted from cells of HIV positive patients, this protocol would be expected to underestimate randomly integrated viral loads when dilutions of the ACH-2 (or U1) cellular DNA are used as standards for the following reasons: (i) *Alu* elements may be clustered in particular regions of chromosomes, or be absent from certain loci. Either of these scenarios would serve to significantly increase the effective average distance between *Alu* elements and may exclude the amplification of HIV DNA integrated into *Alu*-deficient regions of cellular DNA. (ii) Although the *Taq* polymerase used has been shown to facilitate the qualitative amplification of long DNA sequences, the varying extension efficiencies of long amplicons in a mixed population of integration events has not been evaluated.

Figure 1.14 The Nested-Alu PCR Method for the Detection of Integrated HIV-1 DNA

DNA is subjected to 1st-round PCR using a 5' primer designed to anneal within a conserved *Alu* repeat sequence and a 3' primer designed to anneal within the HIV-1 LTR sequence (primers *Alu*-LTR 5' and *Alu*-LTR 3', respectively). This reaction amplifies both cellular DNA upstream of the integration site and integrated HIV-1 LTR DNA. An aliquot of the 1st-round PCR product is then further subjected to a 2nd-round of PCR using nested primers (primers NI-1 and NI-2) to generate a product of defined length that is then detected by Southern hybridisation using an internal probe (NI).

Adapted from Chun *et.al.*, 1997 (49).



Human Alu repeat sequences = 90 000/haploid genome
 Human Haploid genome size = 3.3×10^9 bp
 Average Distance between Alu copies = 4.0 kb

Inefficiencies associated with the amplification of longer amplicons would be reflected as a reduction in the amount of integration events detected. Furthermore, *Alu* elements can exist throughout the human genome in either orientation (221). This effectively reduces the number of *Alu* elements able to participate in the *Alu*-PCR by half and increases the average distance between these elements throughout the genome to $\approx 8\text{kb}$. (iii) Of further concern is the presence of a 5' \rightarrow 3' exonuclease activity associated with the *rTth* DNA polymerase used in this procedure (183). This activity would be expected to direct the digestion of annealed DNA (including primers) ahead of nascent DNA chains in each PCR cycle and reduce the efficiency with which many integrants are detected. A more detailed explanation of this is outlined in section 4.10.2. As a consequence of these observations, we have proposed that nested-*Alu* PCR procedure is restricted in its ability to detect the amplification of integrated HIV DNA and, unless used with copy number standards containing randomly integrated HIV DNA, is an inappropriate assay for quantifying levels of proviral DNA following infection either in cell culture or *in vivo*.

1.8 Quantification using the Polymerase Chain Reaction (PCR)

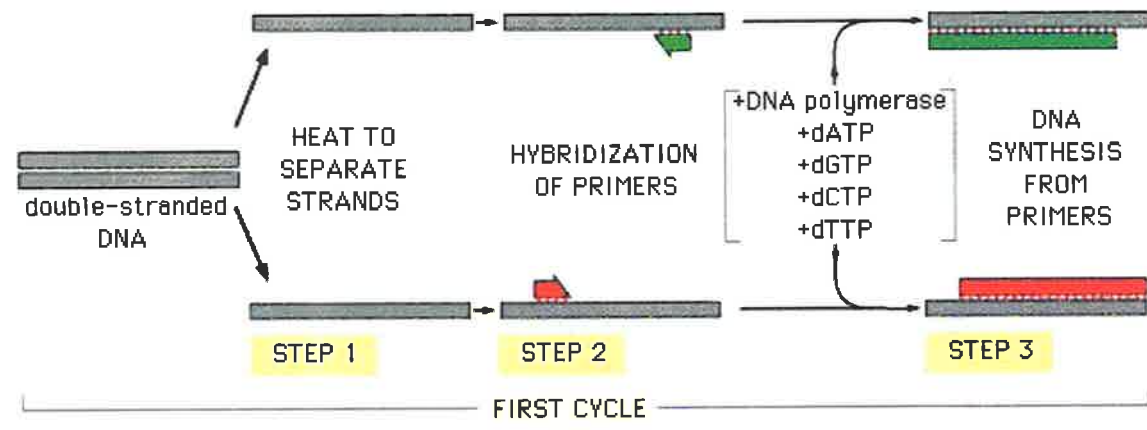
1.8.1 Background

The polymerase chain reaction (PCR) was first developed in 1986 (181, 182) and is today used extensively across all fields of biological science. This technique facilitates the amplification of target DNA sequences by cycling a set of three temperature-dependent reactions in the presence of a thermostable DNA polymerase (Fig. 1.15). Briefly, samples containing the targeted sequences are initially denatured by heating to a temperature ($T^{\circ}\text{C}_m$) in excess of 92°C . Oligonucleotide primers (of generally between 15 and 40 bases in length that are designed to be complementary to regions that flank the target DNA) are then allowed to anneal to complementary sequences within the sample DNA by reducing the temperature to a pre-determined level ($T^{\circ}\text{C}_{\text{anneal}}$). Finally, the reaction temperature is elevated to a level ($T^{\circ}\text{C}_{\text{ext}}$) generally below that of the denaturation but above the annealing temperature, allowing the efficient extension of

Figure 1.15 PCR Amplification of DNA

PCR amplification of target DNA sequences occurs by repeating a set of 3 sequential reactions (indicated as STEP 1, STEP 2 and STEP 3). Generally, template denaturation (STEP 1) is performed at temperatures $>92^{\circ}\text{C}$, primer annealing to denatured template (STEP 2) at temperatures between 50°C and 65°C and primer extension by *Taq* polymerase in the presence of dNTP (STEP 3) at 72°C .

Diagram downloaded at <http://usit.shef.ac.uk/~mba97cmh/> (University of Sheffield)



bound primers by the thermostable DNA polymerase in the presence of dNTPs. Often protocols include an additional denaturation step prior to the cycled reactions to ensure that all sample DNA exists in a single-stranded form prior to the initial primer annealing. Furthermore, a prolonged extension step is generally performed following the final extension step of the cycled reactions to allow the completion of any unfinished amplicons (264).

Assuming 100% reaction efficiency at each of the steps, cycling of these three temperature-dependent reactions allows the amplification of a target sequence in a logarithmic fashion (see Fig. 1.16 A and B) according to the equation below:

$$\mathbf{a = b \times 2^c} \quad \mathbf{(1)}$$

where: **a** is the number of amplified target DNA molecules

b is the initial number of target DNA molecules

c is the number of cycles used

The specific annealing of primers to their complementary sequence is highly dependent on the levels of salts present in the reaction mix and the annealing temperature used. Low annealing temperatures and high salt concentrations serve to decrease the stringency associated with the annealing of primers to target DNA, and result in high levels of non-specific binding throughout the sample DNA. Conversely, salt levels below, and temperatures above that required for efficient annealing of the primers to the template DNA, result in the highly specific, but inefficient amplification of the targeted DNA sequence. Therefore, for the highly specific and efficient PCR amplification of a target sequence, the optimum annealing temperature and MgCl₂ concentration within the reaction mix should be titrated for each primer pair used. These parameters will vary according to the length and G/C content of the primer pair used. For a detailed outline of primer design, see PCR Protocols: Current Methods and Applications (264).

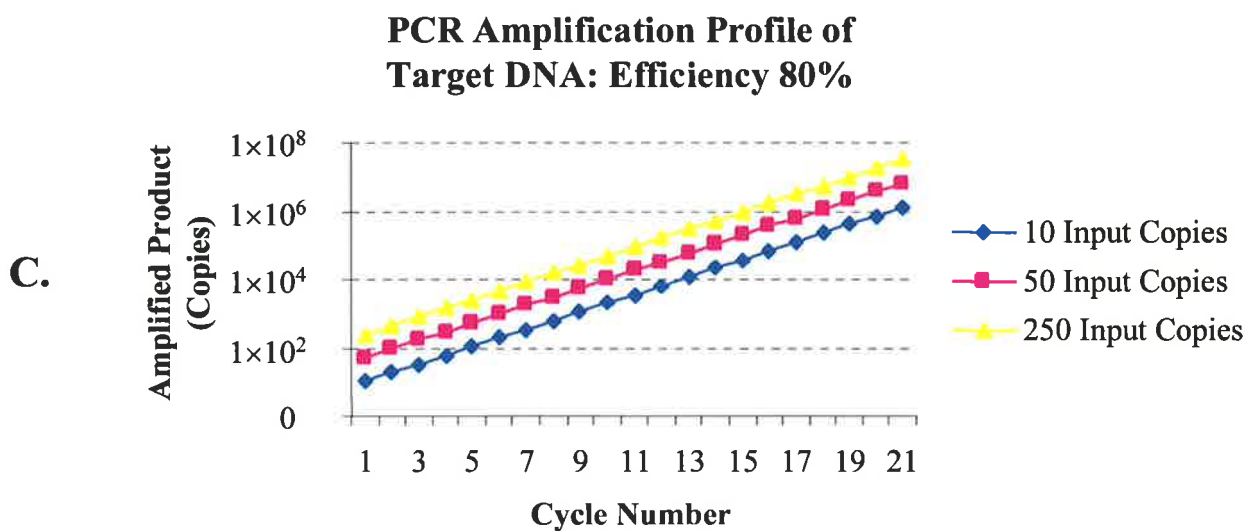
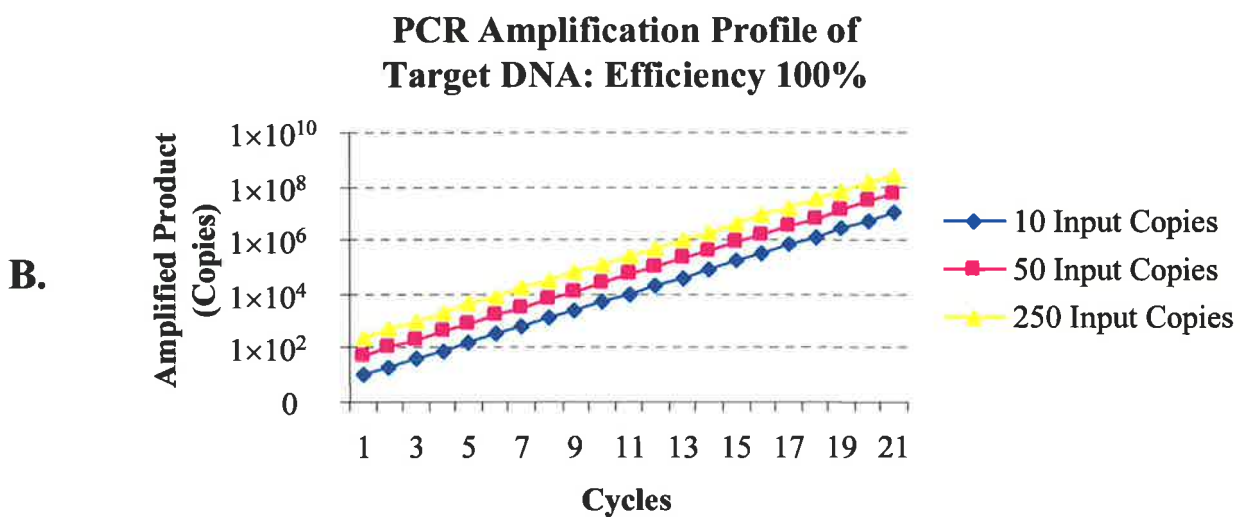
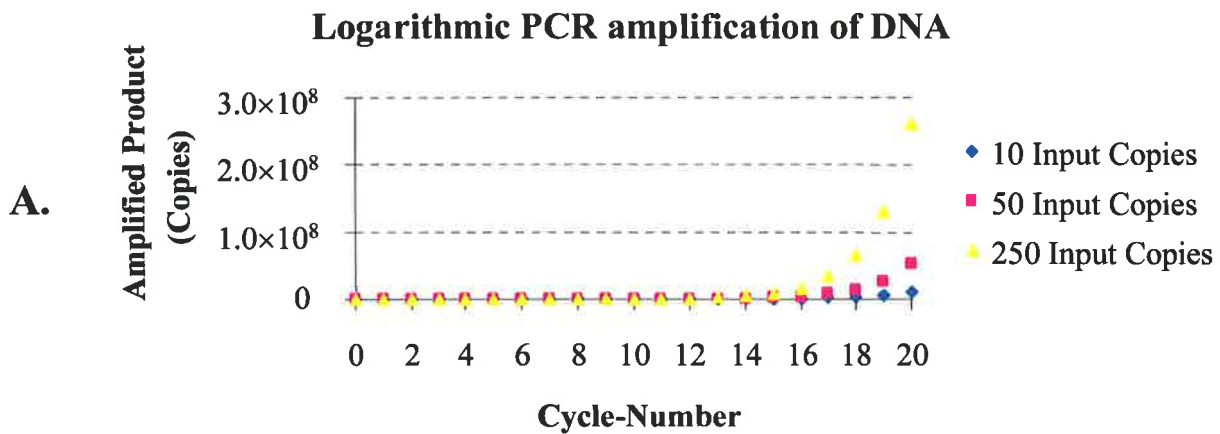
1.8.2 PCR as a Quantitative Tool

As outlined in section 1.8.1, PCR amplification of target DNA molecules occurs in a logarithmic fashion and can be modelled using a simple mathematical formula (equation (1)). However, the application of this formula to determine the number of amplified

Figure 1.16 PCR: Relationship of Amplified Product and PCR Cycle-Number

Modelling PCR-mediated amplification of three different starting amounts of target DNA (Input Copies) over 20 cycles. Amplification product (y axis) accumulates in a logarithmic fashion with each cycle (x axis) provided the efficiency of the procedure remains constant. **(A)** Accumulation of amplification product assuming a PCR efficiency of 100% (linear y axis). **(B)** Accumulation of amplification product assuming a PCR efficiency of 100% (logarithmic y axis). **(C)** Amplification of target DNA when the efficiency of the PCR is 80% (logarithmic y axis).

See Appendix 1.1 for source data.



target DNA molecules requires the assumption that the reaction efficiency of each step is 100%. A number of factors (including the frequency with which specific primer extension is initiated and the processivity of the thermostable DNA polymerase used) serve to limit the efficiency with which target DNA is amplified. Therefore, assuming a conserved efficiency of target DNA amplification, the copy-number of amplified product is more precisely expressed as follows (205, 209):

$$\mathbf{a} = \mathbf{b} \times (\mathbf{1} + \mathbf{E})^{\mathbf{c}} \quad (2)$$

where: **a** is the final number of target DNA molecules

b is the initial number of target DNA molecules

E is the % efficiency of the PCR expressed as a decimal (*e.g.* 80% = 0.8)

c is the number of cycles used

Although a reduced amplification efficiency will result in lower amounts of the final amplified product, if the efficiency (**E**) with which target DNA is amplified *remains constant* throughout the PCR procedure (regardless of the level of efficiency attained), a linear relationship will exist when the cycle-number is plotted against the log of the amplified DNA product (compare Fig. 1.16 B and C; see Appendix 1.1 for source data). Furthermore, a linear relationship will exist between the levels of amplified product and the initial input DNA copy-number (Fig. 1.17). However, a variety of factors can affect the constant, logarithmic amplification of target DNA. Fortunately, these factors tend to exert their effect at later stages of the PCR procedure and in most cases can be avoided where problematic (see below). The predominant cause of an ongoing reduction in amplification efficiency is the late-stage depletion of components within the PCR reaction mix. As a PCR progresses and the templates available for amplification increase, competition for various constituents of the reaction such as primers, nucleotides, MgCl₂ and (most importantly) molecules of the polymerase, increases (180, 222, 223). Furthermore, it has been suggested that the accumulation of by-products (such as DNA, degraded proteins or pyrophosphate) in the later-stages of PCR cycling may serve to lower PCR efficiency (120, 180). As competition for reagents (or by-product inhibition) increases, the corresponding decrease in PCR efficiency eventually leads to the loss of the linear relationship (*ie.* $R^2 < 0.95$) between amplified product and the amount of initial input target DNA (Fig. 1.18; see Appendix 1.2 for source data).

Figure 1.17 PCR: Relationship of Amplified Product to Input Copy-Number

For a fixed cycle number, the relationship between the accumulation of amplified product and input copy-number is maintained as long as the PCR efficiency remains constant. In this example, the efficiency of PCR is 100% .

The source data is available in Appendix 1.1.

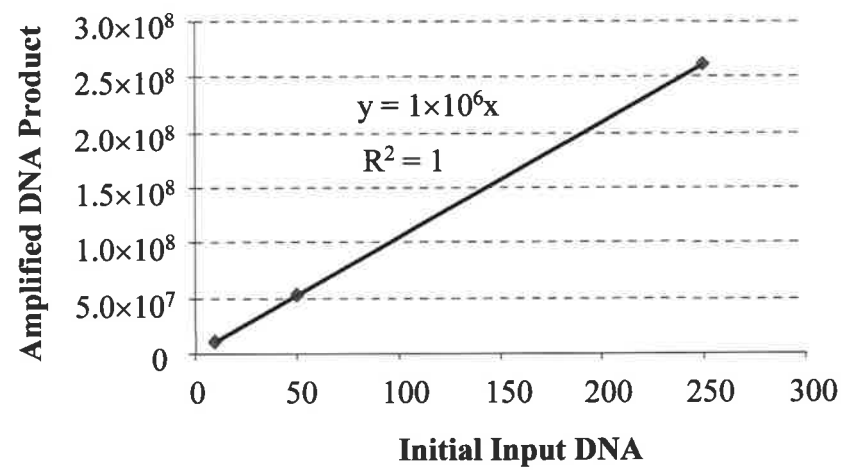
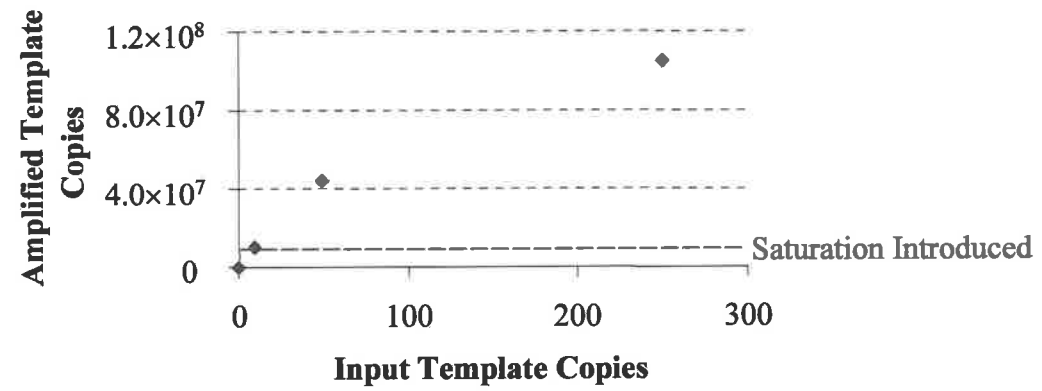


Figure 1.18 Effect of Saturation on PCR Amplification

Saturation of components within the PCR mix (such as primers, nucleotides, MgCl₂ and molecules of polymerase) will occur when the number of templates available for amplification reaches a certain level. In this example, the PCR efficiency has been sequentially decreased (in a logarithmic manner) over ensuing cycles after amplified product reaches 10⁷ copies (dashed, red line). Note that in this example, the regression value (R² value) of the curve is <0.95 (R² = 0.9209) indicating that the relationship of the accumulated product to the input DNA is no longer linear (R²>0.95).

The source data is available in Appendix 1.2.

**Logarithmic saturation introduced at
 10^7 amplified copies**



Maintaining the efficiency of PCR amplification at a constant rate throughout cycling is critical for quantification purposes (205, 223) and can be achieved in two ways: 1) limitation of the PCR cycle-number such that reagents, within the sample containing the highest input DNA copy number, do not become limiting and, 2) reducing the input copy numbers such that reagents do not become limited when a fixed cycle-number is used. A combination of these two methods (*i.e.* low input copy-numbers and limited cycles) may be required in order to ensure the logarithmic amplification of target DNA. However, for the generation of data presented in this thesis, we have primarily chosen to limit the PCR cycle-number of reactions rather than reduce input DNA copy numbers to minimise error associated with the sampling of small numbers of target DNA molecules.

In the laboratory, the use of external standards containing known numbers of a target DNA sequence can be used to confirm that amplification is proceeding in a constant, logarithmic manner (see Fig. 1.17) (205, 209). In addition, if a linear relationship (regression value (R^2) > 0.95) is demonstrated to exist between the levels of amplified product and the initial input copy number of the DNA standard, the copy number contained within an unknown sample (simultaneously amplified in a separate but otherwise identical PCR) can be determined. This is achieved using standard linear regression analyses as follows:

$$\mathbf{x} = (\mathbf{y} - \mathbf{d})/\mathbf{m} \quad (3)$$

where: **x** is the unknown number of target DNA molecules

y is the amount of amplified product (as measured by PCR product quantification techniques (see below)

d is the linear constant (y-axis intercept)

m is the slope of the linear graph

Today, numerous techniques are available allowing the specific detection and accurate quantification of PCR products (209). The amount of amplified product (**y**, equation (3)), synthesised in the PCRs presented in this thesis, was calculated by quantifying the hybridisation signals following Southern analyses of PCR products using ImageQuant software (see Chapter 2). Recently, technology allowing the amount of amplified product over the course of the PCR to be monitored in a real-time manner (as measured

by the rate at which a fluorescence-labelled primer is incorporated into newly synthesised DNA) has become available. However, the use of this real-time quantitative PCR format to detect amplified product is significantly more expensive than the Southern techniques outlined above and was consequently not considered for use in this study.

1.9 Aims of this thesis

The aims of this study were, (i) to establish a robust assay to quantify integrated HIV-1 proviral DNA in a population of cells, and apply this assay to evaluate the kinetics of integrated HIV DNA accumulation following a one-step infection of T cells; (ii) to establish a cell-based assay for assessing compounds that have previously been shown to both inhibit HIV integration *in vitro* and protect cells from HIV infection in culture, for their ability to inhibit the accumulation of integrated HIV DNA following HIV infection of T cells.

At the time this study was initiated, a PCR-based assay for integrated HIV DNA was unavailable. Consequently, there was limited information regarding the kinetics and efficiency of HIV-1 integration in cell culture. Although the efficiency of integration has been estimated indirectly following a single round of cell-to-cell infection (8), there had been no direct assessment of viral integration following either cell-to-cell or cell-free infection, primarily due to the lack of an assay that could distinguish between the integrated and extrachromosomal DNA forms. Since cell-free infection is the predominant route by which HIV is proposed to disseminate *in vivo*, evaluating integration following this mode of infection is critical for our understanding of HIV pathogenesis. Furthermore, a robust assay for the detection of integrated HIV DNA can not only be applied to evaluate the kinetics of integration, but can also be used to evaluate compounds for their ability to specifically inhibit the accumulation of integrated HIV DNA following infection of T cells.

This thesis describes:

- 1) the development of two novel PCR-based methods (DP-PCR and LP-PCR) designed to allow the specific amplification of integrated HIV DNA
- 2) the comparison of LP-PCR with a modified nested *Alu*-PCR protocol for the detection of integrated HIV DNA
- 3) the application of LP-PCR to study the kinetics of integrated HIV DNA accumulation following a one-step infection of T cells
- 4) the use of the modified nested-*Alu* PCR procedure to assess potential inhibitors of HIV integration for their ability to abolish the accumulation of integrated DNA in cell culture.

Chapter 2

Materials and Methods

2.1 Materials

2.1.1 Cells and Cell Culture

HuT-78 cells are a CD4⁺ lymphoblastoid cell line obtained from the NIH AIDS Research and Reference Reagent Program (98). The H3B cell line is a laboratory clone of H9 cells (201) persistently infected with the HTLV-III_B (HIV_{HXB2}) strain of HIV (160). These cells are CD4⁺, secrete approximately 0.01 TCID₅₀ units of virus/hour and are >95% HIV P24 positive by immunofluorescence. H3B cells contain two copies of HIV-1 provirus per cell and undetectable levels of unintegrated viral DNA by Southern blot hybridisation (160). ACH-2 and 8E5 clonal cell lines are T cell lines persistently infected with HIV-1 (53, 89) obtained from the NIH AIDS Research and Reference Reagent Program.

Cells were maintained in RPMI 1640 medium supplemented with 10% heat-inactivated foetal bovine serum (CSL), 0.2M L-glutamine, 1.2µg/ml penicillin and 16µg/ml gentamycin at 37°C and in 5%CO₂. For each experiment, cells were subcultured at a density of 1×10⁶ cells/ml, 12h prior to use, to ensure that cells were in the log phase of growth.

2.1.2 Virus Stocks

The virus inoculum consisted of H3B cell culture supernatant after clarification to remove cells and debris. Briefly, H3B cells were maintained at a density of 5×10⁷ cells per/ml with a medium change every hour. The culture medium was harvested hourly, chilled on ice and centrifuged at 2800g for 10min to pellet cells and debris before storing the supernatant at -70°C. The infectivity of the viral stock was determined by scoring for syncytia 10 days after the addition of serial dilutions of virus to HuT-78

cells. The TCID₅₀ (the highest dilution causing syncytia in 50% of inoculated cultures) was then calculated to be 3.16×10^6 TCID₅₀/ml according to a published method (102).

2.1.3 Bacterial Culture

Escherichia coli (*E. coli*) strain DH5 α was propagated in 1 \times LB or on plates containing 1 \times LB and 15g/L bacto-agar at 37°C for 12-18h. DH5 α containing pBluescript KS (pBS) was grown in 1 \times LB containing 100 μ g/ml ampicillin (Boehringer Mannheim) or on 1 \times LB plates containing 100 μ g/ml ampicillin. Bacterial cultures were permanently stored in 15% glycerol at -70°C.

Competent DH5 α were prepared by growing cells to log phase growth (OD₅₅₀ \approx 0.3) in 100ml 1 \times SOB. Bacteria were chilled rapidly on ice for 15min and the bacterial pellet isolated by centrifugation at 3000g for 5min at 4°C. The bacterial pellet was then resuspended in 32ml transformation buffer 1 {100mM RbCl, 50mM Mn Cl₂, 30mM potassium acetate pH 7.5, 10mM CaCl₂, 15% glycerol} and 200 μ l aliquots frozen, using liquid nitrogen, and stored at -70°C until use.

2.1.4 Plasmid Vectors

The phagemid pBluescript KS(+) vector (pBS) (Stratagene) was used for all cloning work presented in this study. Small-scale preparation of plasmid DNA was achieved using the alkali lysis method (219). Larger-scale preparations were performed using a Qiagen Plasmid Extraction Kit (Qiagen) according to the manufacturers' protocol.

2.1.5 Oligonucleotide Sequences

All oligonucleotides were synthesised by Geneworks. Dried oligonucleotide pellets were resuspended in autoclaved water to a final concentration (stock) of 500pmol/ μ l before further dilutions to working concentrations (generally 25pmol/ μ l). The sequences and nucleotide positions of the oligonucleotide primers used in this study are presented in Table 2.1.

Table 2.1 Sequences and position of oligonucleotides used in this study

Primer Name	Sequence	Sequence Coordinates
β -glo 1	5'-CAACTTCATCCACGTTCCACC-3'	nt 938-918 ^a
β -glo 2	5'-GAAGAGCCAAGGACAGGTAC-3'	nt 671-690 ^a
M1	5'-GACGTTAGGTCAAGGTGTAG-3'	nt 1320-1340 ^b
M2	5'-GGTTGTCTGGTAGTAAGGTG-3'	nt 1715-1695 ^b
GAG-P1(+)	5'-GAGGAAGCTGCAGAATGGG-3'	nt 1408-1426 ^c
GAG-III(-)	5'-CTGTGAAGCTTGCTCGGGTC-3'	nt 1722-1703 ^c
PBS-659(-)	5'-TTTCAGGTCCCTGTTCGGGCGCCAC-3'	nt 659-635 ^c
R7	5'-GGGTCTCTCTGGTTAGACC-3'	nt 454-472 ^c
U3NV	5'-GGCTTCTTCTAACTTCTCTGGCTC-3'	nt 179-156 ^c
LPNV	5'-TCATGATCAATGGGACGATCACATG-3'	same as B101 ^d
U3PNV	5'-GGTACTAGCTTGTAGCACCATCC-3'	nt 151-129 ^c
U3-106(-)	5'-CCTGGCCCTGGTGTGTAGTTC-3'	nt 106-86 ^c
AC+1-21	5'-ACTGGAAGGGCTAATTCCTCCC-3'	nt AC+1-21 ^c
INT-2	5'-CAGCATTATCAGAAGGAGCC-3'	nt 1307-1326 ^c
Primer B	5'-CTGCTAGTTCAGGGTCTACTTGTGTGC-3'	nt 5350-5324 ^c
M13-20	5'-GTAAAACGACGGCCAGT-3'	nt 600-616 pKS(+) ^e
<i>Alu</i> 164	5'-TCCCAGCTACTCGGGAGGCTGAGG-3'	nt 164-187 ^f
U3.1(+)	5'-GGAAGGGCTAATTCCTCC-3'	nt 2-20 ^c
PBS(-)	5'-GGTCCCTGTTCGGGCGC-3'	nt 654-638 ^c
U3.2(-)	5'-GGTAGATCCACAGATCAAGG-3'	nt 57-38 ^c
GAG-FL2	5'-GTTCTAGGTGATATGGCCTG-3'	nt 1240-1221 ^c

^aHuman β -globin sequence genbank accession number L26462

^bHuman Mitochondrial sequence genbank accession number NC_001807

^cHuman Immunodeficiency Virus Type 1 (HXB2) genbank accession number K03455

^dWattel *et.al.*, 1995 (257)

^eStratagene

^f*Alu* consensus sequence (129)

2.1.6 Commonly used buffers and solutions

Ethidium bromide stock solution:

10mg/ml ethidium bromide (Sigma) dissolved in DDW. Store at 4°C in a dark bottle.

Gel Loading Buffer (10×):

60% Glycerol; 100mM EDTA (pH 8.0); 100mM Tris (pH 7.5); Bromphenol blue;

Xylene cyanol

HIRT Solution 1 (1×):

5mM Tris pH 7.7; 10mM EDTA

HIRT Solution 2 (1×):

10mM EDTA; 5mM Tris pH 7.7; 1.2% SDS

LB (1×):

10g/L bacto-tryptone, 5g/L bacto-yeast extract, 10g/L NaCl

Non-denaturing polyacrylamide gel solution:

8% acrylamide {acrylamide (Bio-Rad) : N,N'-methylenebisacrylamide (National Diagnostic); 33:0.9}, 1×TBE. Gels were polymerised by the addition of ammonium persulfate (Bio-Rad) to 0.1% and TEMED (N,N,N',N'-Tetramethylethylenediamine; Bio-Rad) to 0.075%.

PBS:

140mM NaCl, 3mM KCl, 1mM KH₂PO₄, 8mM Na₂HPO₄

Phenol:

Tris-equilibrated Phenol was prepared according to suppliers instructions (Sigma)

SOB (1×):

20g/L bacto-tryptone, 5g/L bacto-yeast extract, 0.6g/L NaCl, 0.5g/L KCl, filter sterile
10mM MgSO₄ (added before use)

SSC(1×):

150mM NaCl, 15mM trisodium citrate, pH 8.0

STE:

100mM NaCl, 10mM Tris pH 7.6, 1mM EDTA

TAE (1×):

40mM Tris-actetate, 1mM EDTA

TBE (1×):

90mM Tris, 90mM Boric acid and 2.4mM disodium EDTA

TE (1×):

10mM Tris, 1mM EDTA, pH 8.0

2.2 Methods

2.2.1 Cell-free HIV Infection

HuT-78 cells were routinely subcultured 16h prior to infection to ensure cells were at a similar stage of growth, and then incubated with virus inoculum (0.5 TCID₅₀ units per cell for drug infection experiments (sections 4.9.1.ii and chapter 6) and 1 TCID₅₀ unit per cell for the viral DNA integration kinetics study (see chapter 5)) at 4°C for 30min. Cells and virus were then spun at 2500×g for 1hr at 37°C after which cells were allowed to recover in pre-warmed fresh media for 15min at 37°C. Under these conditions of centrifugal enhancement, the rate of infection has been reported as being increased by a factor of 10 (116, 160, 197). Infected cells were then washed 3 times in culture media to remove unbound virus and plated in a 48-well tray at a density of 1×10⁶ cells/ml.

2.2.2 Detection of HIV P24 Antigen

Viral release was monitored over time using a commercially available kit (NEN) by measuring the P24 concentrations in 1/50, 1/200 and 1/500 dilutions (in PBS) of culture supernatant inactivated by the addition of Triton-X (to a final concentration of 0.5%).

2.2.3 HIRT DNA Extraction

Chromosomal and extrachromosomal DNA fractions were separated by the HIRT method (114) of DNA extraction (as the HIRT pellet and HIRT supernatant, respectively) in the presence of 0.5mg/ml of Proteinase K (Merck) and each fraction was then purified independently. Briefly, cells were pelleted by low-speed centrifugation in a bench microfuge and then washed twice with PBS. Pelleted cells were then carefully resuspended in 160µl of HIRT solution 1 (see section 2.1.6) prior to the addition of 20µl of Proteinase K (10mg/ml; Merck). HIRT solution 2 (see section 2.1.6) was then added (200µl) and tubes carefully (to minimise chromosomal DNA shearing) inverted 5 times to mix reagents. Following incubation of samples for 30min at 37°C (waterbath), 100µl of 5M NaCl was added and tubes carefully inverted 5 times to mix reagents. Samples were then stored at 4°C overnight before centrifugation (17000g for 45min at 4°C) to separate the extrachromosomal (HIRT supernatant) from the chromosomal (HIRT pellet) DNA.

HIRT supernatant DNA fractions (containing extrachromosomal DNA) were subjected to phenol/chloroform/isoamylalcohol extraction (24:25:1) and then ethanol precipitation in the absence of additional salt (*ie.* 3M Sodium acetate pH 5.2 was not added). Washed and dried pellets were resuspended in autoclaved DDW (at ≈ 5000 cells/µl) and stored at -20°C until use.

HIRT pellet DNA fractions (containing chromosomal DNA) were heated to 70°C for 20min (vortexing occasionally) in 500µl of autoclaved DDW to fully resuspend DNA. An equal volume of phenol/chloroform/IAA (500µl) was added and samples vortexed thoroughly before being centrifuged for 5min at full-speed in a microfuge. The supernatant was removed to tubes containing 50µl 2M NaCl (0.2M final) and 1µl of glycogen (10mg/ml). Ethanol (2.5 volumes) at room temperature was then added and tubes inverted before centrifugation at full-speed (bench microfuge) for 30min at room temperature. Washed and dried pellets were resuspended (at ≈ 5000 cells/µl) in autoclaved DDW by incubation (with occasional mixing) at 70°C for 30min and stored at -20°C until use. It is worth emphasising that after phenol/chloroform/isoamylalcohol

extraction (24:25:1) of HIRT pellets, care was taken to minimise SDS contamination of DNA by performing all ethanol precipitations and washes at room temperature. Furthermore, NaCl was used as the salt component instead of sodium acetate to further minimise SDS precipitation.

2.2.4 DNA Standard Preparations

2.2.4.i HA8

An integrated proviral DNA standard (designated HA8) was produced by mixing 5×10^5 , 1×10^6 and 1×10^6 cells of the H3B, ACH-2 and 8E5 cell lines, respectively and preparing chromosomal DNA by the method of HIRT (114). These cell lines contain 2, 1 and 1 copies of the integrated HIV proviral DNA, respectively with little or no detectable extrachromosomal forms (53, 89, 160). Rather than a single cell line, a mixture of clonal cell lines was used as the integrated standard to account for potential variations associated with the nested *Alu*-PCR- or the LP-PCR-mediated amplification of different integrants. DNA cell equivalents were calculated based on the average signal obtained after performing β -globin PCR (see below) on six chromosomal DNA extractions from cells counted independently. These standards were used as copy number controls in the total HIV DNA (1.2 HIV copies/cell), integrated HIV DNA (1.2 HIV copies/cell) and β -globin PCR (2 copies/cell) assays. Wherever required, HuT-78 chromosomal DNA was used as background DNA.

2.2.4.ii linear (*lin*) construct (section 4.11.1)

The full length HIV linear control construct (*lin*) used in Chapter 4 (section 4.8) was generated in a multi-stage manner by Dr. Raman Kumar. Firstly, amplification of the first 5350 nucleotides of HIV-1_{HXB2} from plasmid pHXB2(kleen) (Australian National Centre in HIV Virology Research, HIV&AIDS Research Reagents Catalogue) was achieved with primers AC+1-21 and Primer B (Table 2.1) using *Pfu* DNA polymerase (Stratagene). Amplification of the remaining HIV-1 sequence was achieved using the same plasmid template and primers INT-2 and M13-20 (Table 2.1). Cleavage of both products with *Pst* I resulted in four fragments which were then subjected to

electrophoresis through a 0.5% TAE agarose gel. The two fragments, which on ligation would produce a full-length linear HIV-1 DNA molecule, were then eluted (see section 2.2.10) and ligated. The full-length linear fragment was then gel-purified (see section 2.2.10) after electrophoresis through agarose. Following estimation of the construct copy number (see section 2.2.4.iv), the HIV copy numbers within this control construct were established by comparative PCR amplification (see GAG PCR conditions) against the HA8 standard mix using primers GAG-P1 and GAG-III(-) (Table 2.1) in the presence of background DNA. DNA was stored in aliquots at -70°C in siliconised tubes until use.

2.2.4.iii 2-LTR Construct

The 2-LTR control construct was generated by Dr. Raman Kumar. Briefly, PCR amplification of HIRT supernatant samples taken from a cell-free infection of HuT-78 cells at 26h p.i. using primers R7 and U3NV (see Table 2.1) was performed using *Pfu* polymerase (Stratagene). The resulting 2-LTR junction product was gel purified and then ligated into the *Eco* RV site of pBluescript KS(+) vector (Stratagene). Following estimation of the construct copy number (see section 2.2.4.iv), the copy number of the 2-LTR construct stock solution was normalised by PCR against the HA8 chromosomal mix using primers U3.1(+) and U3PNV (Table 2.1) in the presence of appropriate background DNA. DNA was stored in aliquots at -70°C in siliconised tubes until use.

2.2.4.iv Estimating construct copy number

The A_{260} of 1/50 dilutions of the stock control constructs was measured (by spectrophotometer) allowing the DNA concentration of each sample to be calculated as follows:

$$[\text{DNA}] = A_{260} \times 50 \times 50 \mu\text{g/ml}$$

where: [DNA] is the concentration of DNA ($\mu\text{g/ml}$)

50 corresponds to the dilution factor used

50 $\mu\text{g/ml}$ is the concentration of DNA that gives an absorbance (A) at 260nm of 1.0 when using a 1cm path cuvette

Concentrations of DNA in stock solutions were converted to copy numbers according to the formula below:

$$\text{copy number}/\mu\text{l} = m/M \times \text{AvC}$$

where, m = mass (in g) of DNA per μl

M = molecular weight of construct (= $660\text{Da} \times \text{number of bp}$)

AvC = Avagadros Constant (6.02×10^{23})

2.2.5 Standard PCR Procedures

All PCRs were performed in a Perkin-Elmer GeneAmp PCR system (9600). All dNTPs used were obtained from Promega. Where “standard condition” are mentioned in the text, PCRs were performed in $1 \times$ PCR Buffer II (Perkin-Elmer), 2.5mM MgCl_2 , and 0.2mM dNTPs (Promega). 25pmol of each primer and 2.5U of AmpliTaq DNA polymerase was used. Reactions were cycled as follows: 94°C 3min; X cycles (X = cycle-number and is given in the text) of 94°C 30s, 58°C 30s, 72°C 45s; and a final extension of 72°C 10min.

2.2.5.i β -globin PCR

PCR amplification of the single-copy human β -globin gene was used to estimate the DNA content of the chromosomal DNA preparations made. PCRs ($25\mu\text{l}$) were performed using ≈ 50 cell-equivalents of chromosomal DNA in $1 \times$ PCR Buffer II (Perkin-Elmer), 2mM MgCl_2 , 0.2mM dNTPs (Promega), 25pmol β -glo 1 and 25pmol β -glo 2 primers (Table 2.1) using 2.5U AmpliTaq DNA Polymerase. Reactions were cycled as follows: 94°C 3min; 25 cycles of 94°C 45s, 58°C 30s, 72°C 45s; and a final extension of 72°C 10min.

2.2.5.ii Mitochondrial PCR

Mitochondrial DNA was amplified and used to standardise the cell-equivalent amounts of DNA extracted in each HIRT supernatant fraction. PCRs (20µl) were performed using ≈100 cell-equivalents of HIRT supernatant extractions in 1×PCR Buffer II (Perkin-Elmer), 2.5mM MgCl₂, 0.2mM dNTPs, 25pmol M1 and 25pmol M2 primers (Table 2.1) using 1U AmpliTaq DNA Polymerase. Reactions were cycled as follows: 95°C 5min; 20 cycles of 95°C 45s, 59°C 30s, 72°C 35s; with a final extension of 72°C 15min.

2.2.5.iii GAG-PCR

Extrachromosomal HIV DNA forms were detected by amplification of the GAG region from 1000 cell-equivalents of purified DNA estimated from HIRT supernatants. GAG PCR amplifications were performed in 25µl reactions consisting of 1×PCR Buffer II (Perkin-Elmer), 2.5mM MgCl₂, 0.2mM dNTPs, 25pmol GAG-P1(+) and 25pmol GAG-III(-) primers (Table 2.1), and 2.5U AmpliTaq DNA Polymerase. Reactions were cycled as follows: 94°C 3min; 20 cycles of 94°C 30s, 55°C 30s, 72°C 45s; and a final extension of 72°C 10min.

2.2.5.iv Nested-*Alu* PCR

The nested-*Alu* PCR for detection of integrated viral DNA was performed on 1000 cell-equivalents of chromosomal DNA essentially as published (49). Briefly, a primer (*Alu*164; see Table 2.1) designed to anneal within highly conserved human *Alu* repeat elements (approximately 9×10^5 elements/haploid genome) is used with an HIV-specific primer (PBS-659(-); see Table 2.1) to mediate amplification of the left-side viral LTR and upstream chromosomal sequence. First-round PCRs were set up in a 2-stage process: Firstly, 25pmol of each primer (*Alu*164 and PBS-659(-); see Table 2.1) were mixed with dNTPs (0.2mM final) and water (to a final volume of 10µl) in PCR tubes and a wax bead (PCR Gem 50; Perkin-Elmer) placed over the reagents. Tubes were heated to 75°C for 1min and then cooled to 4°C in the PCR machine allowing a solid

wax layer to form over the reagents. Secondly, reactions were then made up to 50 μ l (including lower phase) by water following the addition of sample DNA, 3.3 \times PCR Buffer II (Perkin-Elmer) to a final concentration of 1 \times , 1.2mM (final) Magnesium acetate (Perkin-Elmer), and 1.6U of rTth DNA Polymerase. Reactions were cycled as follows: 94 $^{\circ}$ C 3min; 22 cycles of 94 $^{\circ}$ C 30s, 66 $^{\circ}$ C 30s, 70 $^{\circ}$ C 5min; and a final extension of 72 $^{\circ}$ C 10min.

Following 1st-round amplification, 1/1000th (instead of 1/400th as previously published (49)) of amplified products were then re-amplified using a set of nested primers. Nested PCR amplification was performed in 25 μ l reactions consisting of 1 \times PCR Buffer II (Perkin-Elmer), 1.25mM MgCl₂, 0.2mM dNTPs, 25pmol NI-1 and 25pmol NI-2 primers (Table 2.1), and 2.5U AmpliTaq Gold DNA Polymerase. Reactions were cycled as follows: 94 $^{\circ}$ C 12min; 20 cycles of 95 $^{\circ}$ C 30s, 63 $^{\circ}$ C 30s, 72 $^{\circ}$ C 1min; and a final extension of 72 $^{\circ}$ C 10min.

2.2.5.v LP-PCR

Chromosomal DNA was first digested with 20 U of *Bgl* II (New England Biolabs) in 2 \times OPA Plus buffer (Pharmacia) for 3h at 37 $^{\circ}$ C in a final volume of 20 μ l. The sample buffering conditions were then altered to final concentration of 1 \times OPA Plus, 20mM Tris-Acetate pH 7.9, 0.1mg/ml BSA (New England Biolabs) and 1mM dithiothreitol (Boehringer Mannheim) prior to the addition of 10 U of *Nla* III (New England Biolabs) and incubation for 3h at 37 $^{\circ}$ C in a final volume of 40 μ l. Two nucleotides (G and A) of the *Bgl* II overhang generated by digestion were “filled in” with 5 U of Klenow (3’-5’ exo⁻) (New England Biolabs) after modification of the buffering conditions to final concentrations of 7.5mM DTT, 0.25mM dGTP (Promega) and 0.25mM dATP (Promega), and incubation at 37 $^{\circ}$ C for 30min in a final volume of 50 μ l. Samples were then extracted once with phenol/chloroform/isoamylalcohol (25:24:1), ethanol precipitated in the presence of 2mg/ml glycogen (Boehringer Mannheim) and washed in 70% ethanol prior to resuspension of the pellet in water. Linker ligation reactions in 1 \times ligation buffer (New England Biolabs) and 50pmol (vast excess) of the oligonucleotide LPNV (Table 2.1) were heated to 60 $^{\circ}$ C for 10min and snap-cooled to

minimise inter- and intra-molecular ligation of *Nla* III fragments, followed by the addition of 400U of T4 DNA ligase (New England Biolabs) and incubation overnight at 16°C. First-round PCR was performed in 1×PCR Buffer II (Perkin-Elmer), 2mM MgCl₂, 0.2mM dNTPs (Promega) using 150pmol of LPNV, 100pmol of U3NV (Table 2.1) and 5 U of Amplitaq Gold DNA Polymerase in a final volume of 100µl. PCR was cycled as follows: 95°C 12min; 22 cycles of 94°C 30s, 58°C 30s, 72°C 1min; 72°C 10min. Nested PCR was performed on 1/100th of the 1st-round PCR product in 1×PCR Buffer II (Perkin-Elmer), 1.5mM MgCl₂, 0.2mM dNTPs (Promega) using 25pmol each of primers U3.1(+) and U3-106(-) (Table 2.1) and 2.5 U of AmpliTaQ DNA Polymerase (Perkin-Elmer) in a final volume of 25µl. PCRs were cycled as follows: 94°C 3min; 22 cycles of 94°C 45s, 58°C 30s, 72°C 45s; 72°C 10min.

2.2.5.vi 2-LTR PCR

Quantification of the 2-LTR circular DNA forms was carried out by performing PCR on 500 cell-equivalents of total DNA in 1×PCR Buffer II (Perkin-Elmer), 1.5mM MgCl₂, 0.2mM dNTPs using 25pmol each of primers R7 and U3PNV (Table 2.1) and 1.5 U AmpliTaQ Gold DNA Polymerase. Reactions were cycled as follows: 94°C 12min; 26 cycles of 94°C 15s, 58°C 30s, 72°C 45s; 72°C 10min.

2.2.6 Southern Transfer and Hybridisation Techniques

After cycling, 10µl of each PCR product was subjected to electrophoresis on 8% polyacrylamide gels (Hoeffer SE600) and then Southern transfer (Bio-Rad electroblot apparatus) onto Hybond N+ nylon filters (Amersham). Electroblotting was performed by sequentially layering pre-soaked Whatman paper (2 sheets), nylon filter and the gel onto the base of the transfer apparatus. An additional layer of Whatman paper was then placed on top (2 sheets). The Whatman paper and nylon filter used was pre-soaked in 0.3×TBE. Gels were transferred for 90min at 50mA. Following denaturation and fixation for 20min using 0.4M NaOH, the filters were washed in 2×SSC, pre-hybridised and then hybridised at 42°C using the relevant probes (see Table 2.2) in UltrahybTM solution (Ambion) according to the manufacturers instructions in an hybridisation

incubator (Robbins Scientific). All washes were performed at 55°C. When an end-labelled oligonucleotide probe was used, filters were washed 3 times (20min) in prewarmed 5×SSC/0.5%SDS. When double-stranded DNA probe was used, filters were initially washed for 15min in 5×SSC/0.5%SDS and then washed twice (20min) in prewarmed 0.5×SSC/0.1%SDS. Filters were exposed to phosphorous-coated screens (Molecular Dynamics) for between 30min and 16h depending on signal strength.

2.2.6.i Quantification of band intensity

The intensity (quantification) of bands following Southern hybridisation was performed using PhosphorImager Image-Quant software (Molecular Dynamics). Briefly, an object (rectangle) was initially drawn around the band of largest intensity. This object was then copied and placed over each band to be quantified to ensure that the same area was quantified for each band. The pixel-number was then quantified within each box (using the “integrate volume” function) and transferred to Microsoft Excel for analysis. Pixel-counts in lanes in which no amplified DNA was present served to control for background and were subtracted in Microsoft Excel from sample counts. Where multiple PCRs were run for the same sample, results were averaged in Microsoft Excel prior to either graphing as a function of input copy-number to generate a standard curve (DNA standards), or conversion to copy-number by regression analysis (experimental samples).

2.2.7 Preparation of Probes for Southern hybridisation

Oligonucleotides used for probing were obtained from Geneworks. All double-stranded DNA probes were generated by PCR amplification of target sequences as outlined below. A list of the probes used in this study is given in Table 2.2.

Table 2.2 Sequence and position of probes used in this study

Probe Name	Sequence Position	
Glo	Flanked by primers β -glo 1 and β -glo 2	nt 671-938 ^a
Mit	Flanked by primers M1 and M2	nt 1320-1715 ^b
GAG	Flanked by primers GAG-P1(+) and GAG-III(-)	nt 1408-1722 ^c
U3-106	Nucleotides 2-106 of the HIV _{HXB2} genome	nt 2-106 ^c
T7	5'-GTAATACGACTCACTATAGGGC-3'	nt 625-646 pKS(+) ^d

^aHuman β -globin sequence genebank accession number L26462

^bHuman Mitochondrial sequence genebank accession number NC_001807

^cHuman Immunodeficiency Virus Type 1 (HXB2) genebank accession number K03455

^dStratagene

2.2.7.i U3-106 probe

The U3-106 probe was used to detect products resulting from the nested *Alu*-PCR, the LP-PCR and the 2-LTR PCR (see section 2.2.5). This probe was generated by a 30-cycle PCR amplification of DNA within the pHXB2(kleen) plasmid (100ng) using primers U3.1(+) and U3-106(-) under reaction conditions identical to that described in section 2.2.5.v (nested PCR). The resulting product was separated on a 2% TAE agarose gel and purified using the Qiagen quick-spin column as outlined by the supplier. DNA was eluted with DDW and stored at -20°C until use.

2.2.7.ii β -glo probe

The Glo probe was used to detect products arising from the β -globin PCR (see section 2.2.5.i). This probe was generated by a 30-cycle PCR amplification of DNA within purified HIRT pellet DNA (see section 2.2.3) extracted from 1×10^4 HuT-78 cells under reaction conditions identical to that described in section 2.2.5.i. The resulting product was separated on a 2% TAE agarose gel and purified using the Qiagen quick-spin column as outlined by the supplier. DNA was eluted with DDW and stored at -20°C until use.

2.2.7.iii Mit probe

The Mit probe was used to detect products arising from the mitochondrial PCR (see section 2.2.5.ii). This probe was generated by a 30-cycle PCR amplification of DNA within HIRT supernatant DNA (see section 2.2.3) extracted from 1×10^4 HuT-78 cells under reaction conditions identical to that described in section 2.2.5.ii. The resulting product was separated on a 2% TAE agarose gel and purified using the Qiagen quick-spin column as outlined by the supplier. DNA was eluted with DDW and stored at -20°C until use.

2.2.7.iv GAG probe

The GAG probe (Table 2.2) was used to detect fragments resulting from GAG PCR analysis (see section 2.2.5.iii) of DNA preparations. This probe was generated by a 30-cycle PCR amplification of DNA within the pBS(HXB2) plasmid using primers GAG-P1 and GAG-III(-) under reaction conditions identical to that described in section 2.2.5.iii. The resulting product was separated on a 2% TAE agarose gel and purified using the Qiagen quick-spin column as outlined by the supplier. DNA was eluted with DDW and stored at -20°C until use.

2.2.7.v Labelling of oligonucleotides

Oligonucleotides or pUC19/*Hpa* II markers were labelled at the 5' termini with [γ - 32 P]ATP using T4 polynucleotide kinase (PNK) (Boeringer Mannheim) in the following reaction conditions (219): a 20 μ l reaction containing 100ng oligonucleotide DNA (or pUC19/*Hpa* II marker DNA), 1 \times T4 PNK buffer {50mM Tris-HCl pH7.6, 10mM MgCl₂, 1mM DTT, 1mM β -mercaptoethanol}, 5U T4 PNK, and 50mCi [γ - 32 P]ATP (3000Ci/mM Geneworks) was incubated at 37°C for 45min. The labelled DNA was purified through a G-25 Sephadex spun column (section 2.2.7.vii) and stored at -20°C before use.

2.2.7.vi Random primer labelling of DNA

Double-stranded DNA probe fragments were labelled with [α - 32 P]ATP using the Megaprime DNA Labelling Kit (Amersham), according to the manufacturers' instructions. Reaction mixtures contained 120ng denatured DNA template, 5 μ l of denatured random nonamer primers in aqueous solution, 5 μ l reaction buffer, 4 μ l of each dGTP, dTTP and dCTP, 5 μ l of [α - 32 P]ATP (10 μ Ci/ml; Geneworks) and 2 μ l DNA polymerase (Klenow fragment), and were incubated for 45min at 37°C. Probes were then purified through a G-25 Sephadex spun column (see section 2.2.7.vii) and stored at -20°C until use.

2.2.7.vii Purification of probes by spun-column chromatography

Spun-column chromatography was used to purify [32 P]-labelled DNA from unincorporated [32 P]-nucleotides contained in labelling reactions. The column matrix was composed of Sephadex G-25 (Pharmacia Biotech) equilibrated in STE buffer. Briefly, Sephadex bead suspension was compacted by centrifugation (350g for 4min (Heraeus Biofuge 17RS; swing-out rotor 2147)) into 1ml syringes (Terumo) plugged with glass microfibre filters (Whatman). Labelled DNA within the PNK reaction mix was volume adjusted with DDW to 100 μ l, loaded onto the top of the column and the column centrifuged as before. The eluate containing purified DNA was collected in an eppendorf and stored at -20°C until use.

2.2.8 Sequence analysis

All DNA sequencing was performed in the Sequence Centre by Arthur Mangos at the Institute of Medical and Veterinary Science, using the dye terminator method. Reagents were provided in the ABI Prism Dye Primer Cycle Sequencing Ready Reaction Kit and reactions were performed as recommended by the supplier (Perkin-Elmer).

2.2.9 Phenol extraction and ethanol precipitation of DNA

Phenol (Sigma) for use in DNA extractions was prepared according to the manufacturers recommendations. An equal volume of the 25:24:1 phenol/chloroform/isoamylalcohol mix was added to a crude preparation of nucleic acid, vortexed and centrifuged at 13000rpm for 5min. Generally, the DNA (recovered in the aqueous phase) was precipitated by adding 10% volume of 3M sodium acetate pH5.2, 5µg of glycogen (Boerringer Mannheim) and 2.5×volume absolute ethanol and then left at -20°C for at least 1h. The nucleic acid was then pelleted by centrifugation at 13000 rpm for 25min, washed once in 70% ethanol (followed by centrifugation at 13000 rpm for 5min) and then air-dried. DNA was finally resuspended in autoclaved water.

2.2.10 Agarose gel electrophoresis and purification of DNA from agarose

For separation of DNA fragments of varying size, 1-2% DNA grade agarose (Progen) gels were made up in 1×TAE buffer. The agarose solution was melted and allowed to set in a gel mould for 30min. Gels were run in 1×TAE buffer for 1-3h at 120V and stained in 100µg/ml ethidium bromide solution for 5min. A 10min destain in 1×TAE buffer was performed prior to visualisation of bands under UV light. Photographs were taken either using Polaroid 667 film (Kodak) or the Spot digital camera system.

DNA fragments to be purified from the agarose gel were visualised and excised under low-strength UV light. DNA was then purified from the agarose plugs using the QIAquick® gel extraction kit (Qiagen) according to the manufacturers' protocol.

2.2.11 Restriction enzyme digestions

All restriction enzymes used were obtained from NEB and, unless otherwise stated, used with the buffers and under the conditions recommended by the supplier. Typically, restriction digests were incubated with 1µl of neat enzyme (5-10U) for 1-3h.

2.2.12 Drugs and Cytotoxicity Assays

The compounds 5,8 dihydroxynaphthoquinone and quercetin dihydrate were obtained from Aldrich. L-708,906 and 3TC were kind gifts from David Bourke, Department of Medicinal Chemistry, Victorian College of Pharmacy, Australia. L-731,988 and an additional sample of L-708,906 were obtained from the Department of Antiviral Research, Merck Research Laboratories, West Point, PA 19486, USA and L17 was synthesised in the Laboratory of Medicinal Chemistry, Division of Basic Sciences, National Cancer Institute, Bethesda, Maryland, 20892-4255, USA. AR177 was synthesised locally (Geneworks) and AZT was obtained from Sigma. With the exception of AR177, all drugs were made up to 10mM stocks in DMSO and then diluted further in serum free RPMI 1640 media to the working concentration. AR177 was dissolved and diluted in phosphate buffered saline. Working concentrations of all drugs used except quercetin dihydrate were based on concentrations shown to inhibit viral release following infection of T cells (54, 111, 140, 161, 188). Quercetin dihydrate was used at 50 μ M, a concentration approximately four-fold higher than that shown to inhibit strand transfer in cell-free systems (12 μ M).

Cell cytotoxicity experiments were performed in triplicate by incubating 2×10^5 HuT-78 cells with concentrations of drugs ranging approximately 5-fold below and above that used in the infection experiments. After 24h and 48h in the presence of drugs, cultures were assessed for cell death by trypan blue exclusion and increase in cell number. Drugs were considered non-toxic if there was <5% inhibition of HuT-78 cell growth over 48h (LD₅) compared to drug-free cultures.

2.3 List of Suppliers

Amersham, Amersham International plc, Little Chalfont, Buckinghamshire, UK

ABI, Applied Biotechnology Inc., Division of Perkin-Elmer Corporation

BDH, BDH Chemicals Australia Pty Ltd, Kilsyth Vic, Australia

Bexkman, Beckman Instruments Inc., Fullerton CA, USA

Bio-Rad, Bio-Rad Laboratories, Hercules CA, USA

Boeringer, Boeringer Mannheim, Mannheim, Germany

Geneworks, Geneworks Pty Ltd, Adelaide SA Australia

Kodak, Eastman Kodak, Rochester NY, USA

Merck, Merck Biochemica, Darmstadt, Germany

Molecular Dynamics, Inc., Beverly MA, USA

Perkin-Elmer Corporation, Roche Molecular Systems, Inc., Branchberg NY, USA

Pharmacia, Pharmacia Biotech, Uppsala, Sweden

Promega, Promega Corporation, Madison WI, USA

Sigma, Sigma Chemical Company, Sigma-Aldrich Pty Ltd, St. Lois MO, USA

Stratagene, Stratagene Cloning Systems, La Jolla CA, USA

Terumo, Tokyo, Japan

Whatman, Whatman International Ltd., MAIDStone, UK

Chapter 3

Development of an Assay for Integrated HIV DNA: Degenerate Primer Polymerase Chain Reaction (DP-PCR)

3.1 Introduction

3.1.1 Background

As outlined in section 1.5.2, research into the kinetics of accumulation of integrated HIV DNA following infection has been restricted by the extensive contamination of cellular chromosomal DNA preparations with extrachromosomal HIV DNA. At the time this study was designed, the protocols used to obtain highly purified chromosomal DNA preparations were laborious and failed to consistently remove all contaminating free viral DNA forms (146, 238). This has prompted researchers to investigate methods that specifically detect integrated proviral DNA.

Selection procedures require the identification and capitalisation of unique characteristic(s) associated with the target. The most prominent feature distinguishing integrated from extrachromosomal HIV DNA forms is the presence of cellular sequence flanking the proviral DNA molecule. However, HIV integrates at random with respect to cellular chromosomal sequence (see section 1.3.4). Consequently, the sequences immediately flanking integrated HIV DNA are essentially random with respect to sequence and have, until recently, been of little aid in allowing the selective detection of integrated HIV DNA in a quantifiable manner.

Currently, the most widely used technique detecting integrated HIV DNA (first described by Sonza and co-workers (231)) exploits the presence of conserved *Alu* repeat elements within human chromosomal DNA (see section 1.7.3 and Fig. 1.14). However, the ability of this protocol (and variations thereof) to direct the amplification of

randomly integrated HIV DNA is likely to be restricted by both the varying extension efficiencies of different amplicons and the presence of a 5'→3' exonuclease activity associated with most thermostable polymerases (see section 1.7.3 and 4.10.2). Here, I present a serious attempt to develop an alternative PCR proviral assay (degenerative primer PCR or DP-PCR) that utilises the presence of the random cellular DNA immediately adjacent to the 5'-end of the integrated provirus.

3.1.2 DP-PCR Principle

3.1.2.i DP-PCR Amplification of the 5'-end of Integrated HIV DNA

The basic principle of DP-PCR is outlined in figure 3.1. Briefly, a pool of semi-degenerate primers (DP; composed of an HIV-specific 3'-end and a degenerate 5'-end) was designed such that they could anneal to the 5'-end of the integrated HIV DNA and upstream cellular chromosomal sequence. The length of the HIV-specific region of the DP was limited to ensure that complementarity over this short region alone was insufficient to facilitate priming in a PCR. Therefore, the annealing conditions could be manipulated so that partial binding of the DP degenerate 5'-end to sequences immediately upstream of the proviral integration site was required in order to stabilise the primer on the template and facilitate PCR amplification. When used in conjunction with an HIV-specific primer (PBS-659(-); see Table 2.1) in a PCR, primers within the degenerate primer (DP) pool were expected to mediate the amplification of the 5'-end of the integrated provirus but not from free HIV DNA forms (see section 3.1.2.iii).

To ensure that degenerate primers based on the 5'-end of the HIV_{HXB2} sequence could also anneal to this region in other isolates, a sequence comparison was performed between 50 clade B isolates (the major clade in which viruses affecting patients in Australia and the USA are classed). It is worth noting that the first 10 bp of the HIV LTR region were expected to be highly conserved across the majority of isolates due to sequence requirements of the integration machinery (see section 1.3.2.i). When compared, 44/51 isolates within this clade exhibited 100% homology over this region, with the remaining isolates differing by a single base (Fig. 3.2). Furthermore, the consensus sequences over this region for strains within other clades were highly

Figure 3.1 DP-PCR Principle

DP-PCR amplification of the HIV 5'LTR region using a population of degenerate primers (indicated as DP; designed to anneal over the 5'U3/chromosomal DNA junction) and primer PBS(-).

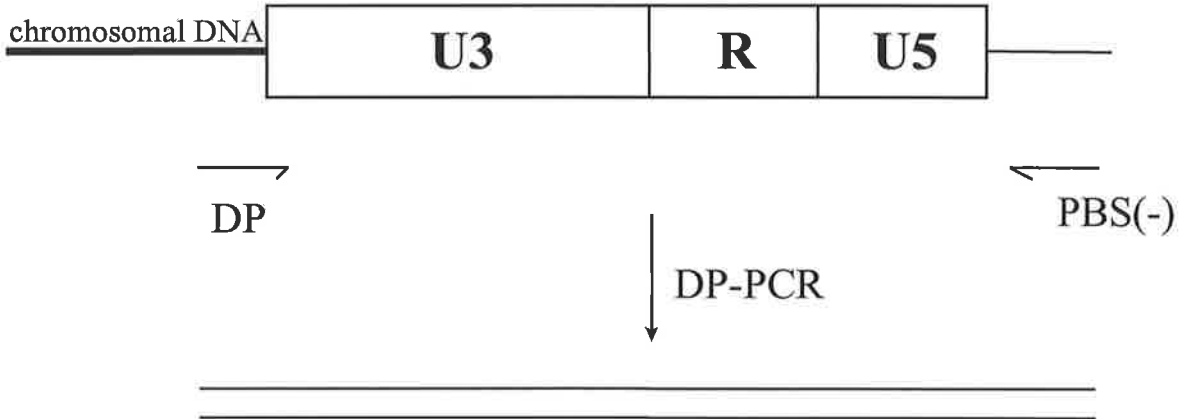


Figure 3.2

Sequence conservation of the first 10 bases of the 5'U3 region across isolates within clade B. The consensus sequences (CONS:) of other clades over this region are also shown (bottom right).

CONS: B **TGGAAGGGCT**
B_SF2 TGGAAGGGCT
B_SF2B13 TGGAAGGGCT
B_LAI TGGAAGGGCT
B_HXB2 TGGAAGGGCT
B_N1TE TGGAAGGGCT
B_NY5 TGGAAGGGCT
B_NL43 TGGAAGGGCT
B_MN TGGATGGGTT
B_BRVA TGGAAGGGCT
B_SC TGGAAGGGCT
B_BAL1 TGGAAGGGCT
B_JRCSF TGGAAGGGCT
B_JRFL TGGAAGGGCT
B_OYI TGGAAGGGCT
B_SF162 TGGAAGGGTT
B_CAM1 TGGAAGGGAT
B_CDC41 TGGAAGGGCT
B_SF33 TGGAAGGGCT
B_HAN TGGAAGGGTT
B_P896 TGGAAGGGCT
B_SWB884 TGGAAGGGCT
B_RF TGGATGGGCT
B_YU2 TGGAAGGGCT
B_BCSG3C TGGAAGGGCT
B_E90NEF TGGAAGGGCT
B_YU10 TGGAAGGGCT
B_3202A12 TGGAAGGGCT
B_3202A21 TGGAAGGGCT
B_LTRAA TGGAAGGGCT

CONS: B **TGGAAGGGCT**
B_LTRAB TGGAAGGGCT
B_LW123 TGGAAGGGCT
B_P50069-100 TGGAAGGGCT
B_P50192-201 TGGAAGGGCT
B_P50219-11 TGGAAGGGCT
B_P50346-307 TGGAAGGGCT
B_P50349-11 TGGAAGGGCT
B_P50363-402 TGGAAGGGCT
B_P50442-500 TGGAAGGGTT
B_P50495-10 TGGAAGGGCT
B_P50864-10 TGGAAGGGCT
B_E81NEF TGGAAGGGCT
B_E88NEF TGGAAGGGCT
B_F12CG TGGAAGGGCT
B_MCK1 TGGAAGGGCT
B_PM213 TGGAAGGGCT
B_PV22 TGGAAGGGCT
B_TH475A TGGAAGGGCT
B_896 TGGAAGGGCT
B_AD8 TGGAAGGGCT
B_C18MBC TGGAAGGGCT
B_WEAU160 TGGAAGGGCT

CONS: A TGGATGGGTT
CONS: D TGGAAGGGCT
CONS: AE TGGA ?GGGCT
CONS: O TGGA ?GGGTT
CONS: CPZ TGGAAGGGTT

homologous (differing by no more than 2 bases) to the clade B consensus sequence. Taken together, these results suggested that the DP-PCR procedure would be effective both using laboratory HIV strains and in clinical settings.

It is worth emphasising that each of the degenerate primers evaluated in this study are actually a pool of primers containing thousands of potential variants (see calculations, Fig. 3.7). Although a proportion of these variants would be expected to exhibit high levels of specificity for the 5'HIV/cellular DNA junction sequence, the remaining primers within the pool will be able to anneal to this region with varying efficiencies depending on the nucleotide composition of their 5'-end. It was therefore expected that the amplification efficiency of target DNA would be high during early cycles of PCRs performed using degenerate primers and will decrease as the highly specific primers are incorporated into product and depleted from the reaction mix. To partially overcome this, a vast excess of degenerate primers were used in all reactions (see section 3.2.3).

3.1.2.ii Choice of Downstream Primer

The PBS sequence was initially chosen as the downstream priming site rather than the LTR region. Primers annealing at this site [eg. primer PBS(-) (Table 2.1)] should not allow amplification of the HIV 3'-LTR present in the linear DNA form and therefore contribute to the specificity for integrated sequences.

3.1.2.iii Selection Against Amplification of Extrachromosomal HIV DNA

The two major extrachromosomal HIV forms that are likely to be present in HIRT pellet (chromosomal) DNA preparations are the linear and the single-LTR circular DNA forms (8). HIV linear DNA (Fig. 3.3A) is thought to predominate early after infection while levels of the circular form increase as infection progresses (8). Since the linear viral DNA does not contain any sequence adjacent (upstream) to the 5'-U3 region to which the degenerate portion of a DP is required to anneal, primer annealing to this region (and as a consequence, amplification) would not be expected to occur. Furthermore, amplification of the 3'-LTR would not be possible because of the absence of a downstream PBS priming sequence.

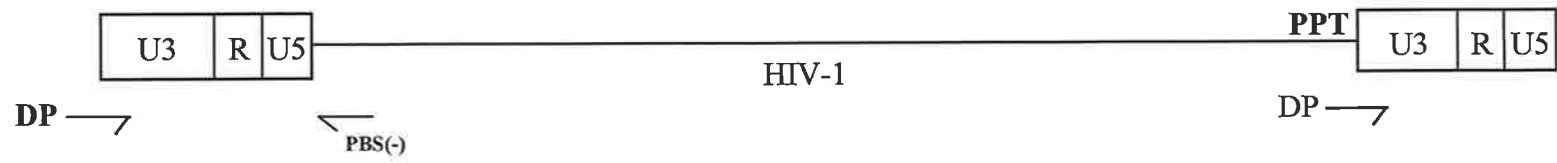
Figure 3.3 DP-annealing to Extrachromosomal HIV DNA

Diagrammatic representation of the 3 extrachromosomal HIV DNA forms showing potential annealing sites of primers within the DP primer-pool. **A.** Linear extrachromosomal HIV DNA **B.** 1-LTR HIV DNA **C.** 2-LTR HIV DNA. The positions at which the PBS(-) primer is designed to anneal are also shown.

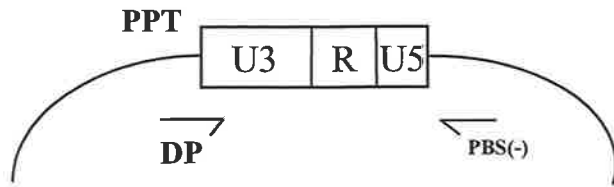
DP Degenerate Primer

PPT Polypurine tract

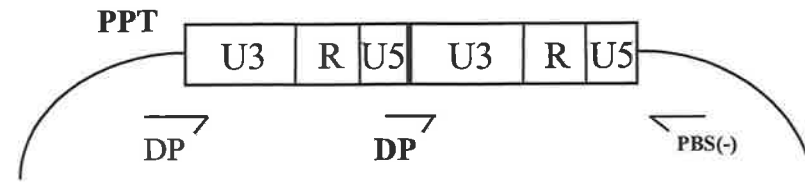
A.



B.



C.



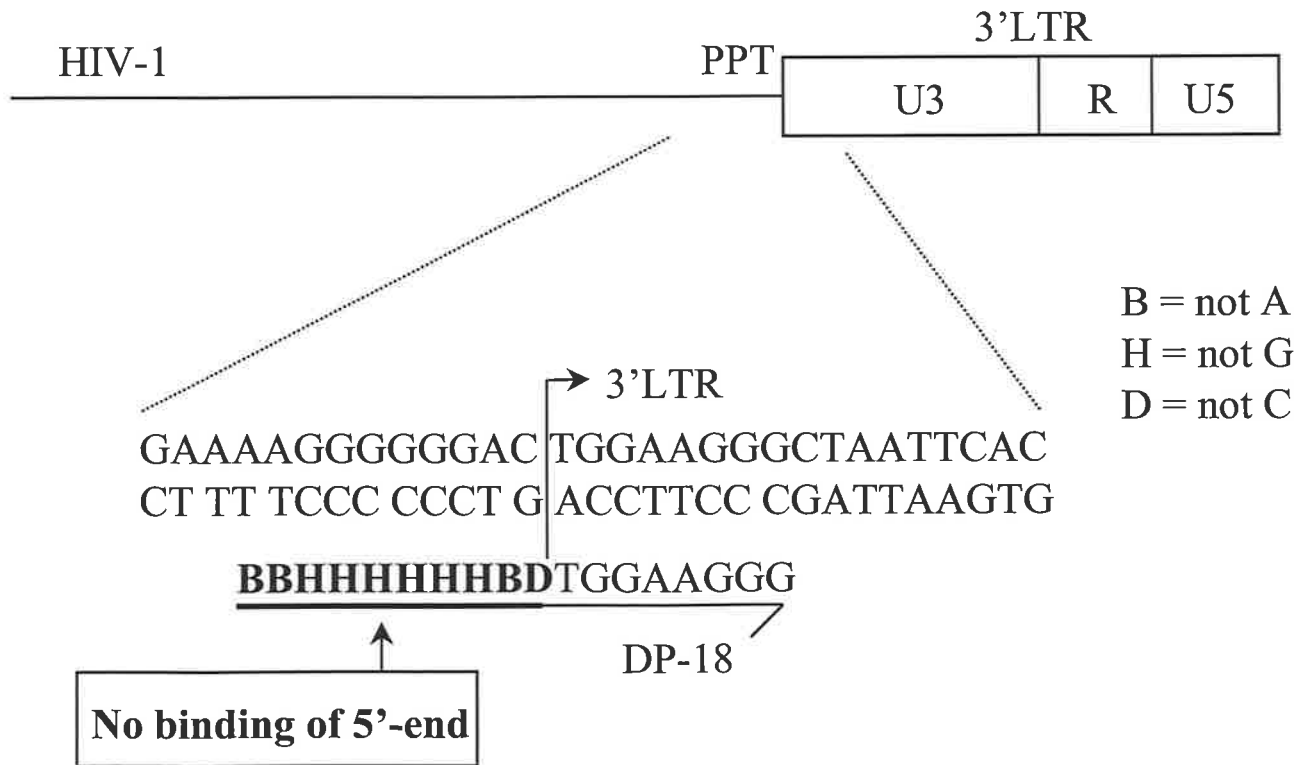
However, DP-mediated amplification of the HIV 3'-LTR region of 1-LTR circular HIV DNA could occur in a similar fashion to that outlined for amplification of DNA at the 5'-insertion site of the integrated provirus (see Fig. 3.3B). Since amplification over this region is not specific for integrated HIV DNA, steps were undertaken to abolish amplification over this site and thereby inhibit detection of extrachromosomal HIV DNA forms.

The sequence of DNA immediately upstream of the 3'-LTR is known as the **polypurine tract (PPT)** and is highly conserved. This region within the viral RNA genome is highly resistant to the action of RNase H during the reverse transcription process and remains bound to the newly synthesised '-' sense DNA, serving as a primer to initiate '+' sense DNA synthesis (see section 1.2.2). The highly conserved nature of the PPT allowed the 5'-ends of DP's to be designed such that they contained nucleotides unable to pair within this region (Fig. 3.4). This was achieved by limiting the degeneracy of the 5'-tail region to nucleotides that were non-complementary to the corresponding bases in the PPT sequence. In this way, annealing at the PPT (and hence PCR amplification over this region) was expected to be abolished.

Later experiments (see section 3.4) incorporating the use of the restriction enzyme *Bsr* 1 (which cleaves at the PPT/3'U3 junction effectively removing all upstream sequence), were undertaken to further ensure that priming did not occur at this region. Taken together, these design features were expected to minimise the frequency with which free viral DNA forms were amplified. Binding of the DP could still potentially occur over the LTR junction region in the 2-LTR form (in the absence of *Bsr* 1 digestion; see Fig. 3.3C), however this form is a very minor species and was not expected to significantly contribute to the signal arising from integrated HIV.

Figure 3.4 Degenerate Tail Design: Discouraging Primer-Annealing to the HIV PPT Region

Diagram showing the aligned sequence of both the 3'U3/PPT junction and the degenerate primer DP-18.



3.2 Demonstrating the DP-PCR Principle

3.2.1 Generating constructs mimicking integrated and linear HIV DNA

3.2.1.i *The Integrated HIV DNA Construct (cc)*

To generate a construct representing integrated HIV DNA, the entire genomic sequence of HIV_{HXB2} was cloned into the *Xba* I and *Sac* II sites of pBluescript (Dr. Hairong Peng, unpublished). The resulting construct (Fig. 3.5A; pHXB2(kleen) or cc) served to mimic integrated HIV DNA.

3.2.1.ii *The linear HIV DNA Construct (lin2)*

Similarly, a construct mimicking linear extrachromosomal HIV DNA was generated. This form predominates early after infection and has been found to be in approximately equal amounts to the single-LTR form later in the infection cycle (8). For all initial experiments, the use of this construct allowed the levels of amplification over the 5'-LTR region in the absence of upstream DNA to be measured.

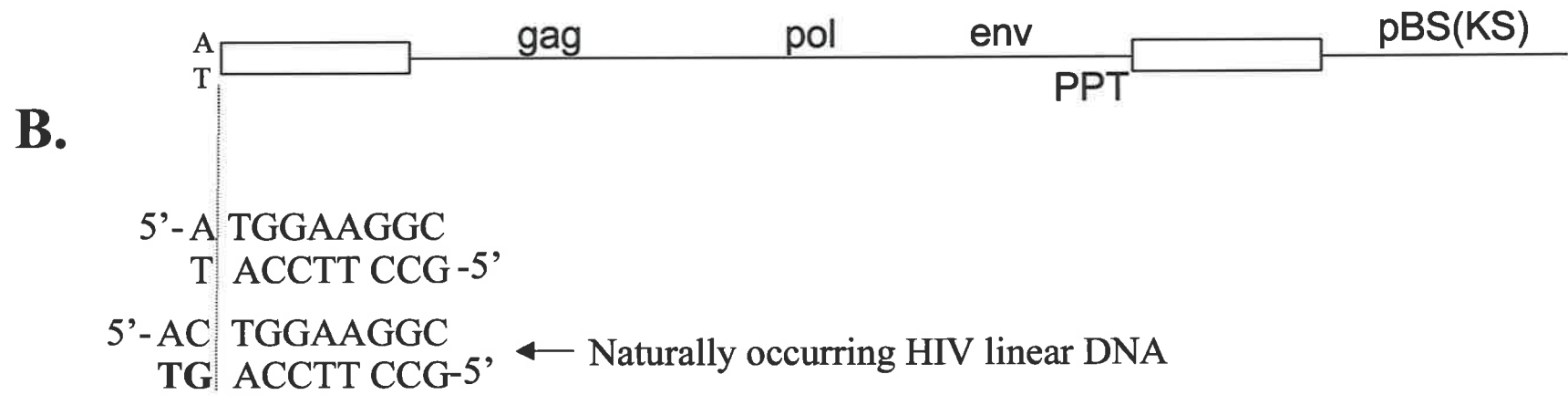
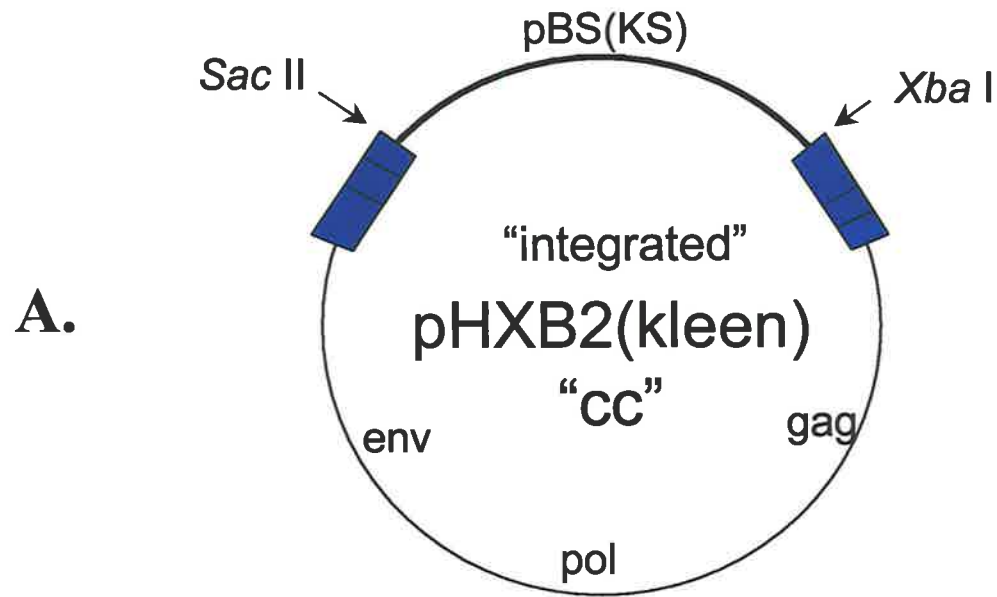
To generate the linear construct (lin2), pHXB2(kleen) (Fig. 3.5A) was cleaved with the restriction enzyme *Xba* I and then treated with Mung Bean Nuclease to remove the 5'-overhang generated upon digestion. The resulting linear molecule was then gel purified to remove any contaminating undigested construct. Although the final purified linear construct (Fig. 3.5B) closely mimicked the naturally occurring HIV linear molecule, the 5'-end of the construct was not absolutely identical with respect to sequence. Naturally occurring HIV linear molecules exist in two forms with respect to sequence at the 5'-end (and 3'-end): 1) Unprocessed, in which the terminal 3' di-nucleotides of both strands of HIV DNA are present, and 2) 3'-processed, in which the 3' dinucleotide (TG) has been removed from each strand of HIV DNA in preparation for integration (see section 1.3.2.1). Due to the nature of the *Xba* I cleavage reaction, the linear construct lacked either of the above termini, and instead contained a single additional base-pair (A:T) rather than the dinucleotide. Therefore, the overall difference between the naturally

Figure 3.5

Structure of the pHXB2(kleen) (cc; **A.**) and the lin2 (**B.**) constructs used to mimic the integrated and linear extrachromosomal HIV DNA forms, respectively. The position of the major genomic regions encoded by HIV in each construct is indicated (gag, pol and env). The TG dinucleotide removed in the 3'-end processing reaction during the integration process is indicated in boldface.

PBS(KS) pBluescript sequence (Stratagene)

PPT Polypurine Tract



occurring linear HIV DNA molecule and the linear construct was the presence in the former of a single C≡G base pair immediately downstream of the terminal A=T base pair (see Fig 3.5B). For the purposes of demonstrating the ability of a variety of degenerate primers to direct the selective amplification of the integrated construct, this linear construct was deemed sufficient to control for the amplification of linear HIV DNA.

3.2.1.iii Quantification of Standard Constructs

The copy numbers of stock samples of both the integrated (cc) and linear (lin2) were initially determined by spectrophotometric analysis (see section 2.2.4.iv). These stock solutions were then aliquotted and dilutions then precisely assessed for relative copy number by performing a 25-cycle PCR using primers U3.1(+) and PBS(-) (see Table 2.1) under standard reaction conditions (see section 2.2.5). Dot blot analysis of products using the internal ³²P-labelled U3PNV oligonucleotide demonstrated that slightly higher DNA copies were present in the stocks of the linear construct compared to the cc construct (Fig. 3.6). However, the cc and lin2 stocks were used in subsequent experiments as if they contained equal copy-numbers in equal volumes. This ensured that any differences observed in signal intensities after Southern or dot blot analysis of DNA DP-PCR products reflected the selective nature of the assay, and not initial differences in the amounts of template DNA.

3.2.2 Designing a Degenerate Primer: RDM#2

Initially, a single DP (RDM#2), was tested under various PCR conditions (different [MgCl₂] and T°C_{anneal}) with the PBS(-) primer for the ability to facilitate the specific amplification of cc in the absence of background DNA. This degenerate primer comprised a 9-base, HIV-specific, 3'-end and an 8-base semi-degenerate 5'-end (Fig. 3.7). The semi-degenerate part was designed such that guanosine bases were absent within this region of the primer. Since the '+' DNA sequence immediately upstream of the 3' viral LTR (within the PPT region) consists primarily of guanosine bases, it was expected that minimal annealing of this semi-degenerate tail region to the '-' sense PPT region would occur.

Figure 3.6 Equating the pHXB2(kleen) and lin2 Constructs

Various dilutions (indicated top) of the pHXB2(kleen) and lin2 constructs were subjected to PCR amplification of DNA within the HIV LTR region (see section 3.2.1.iii for details).

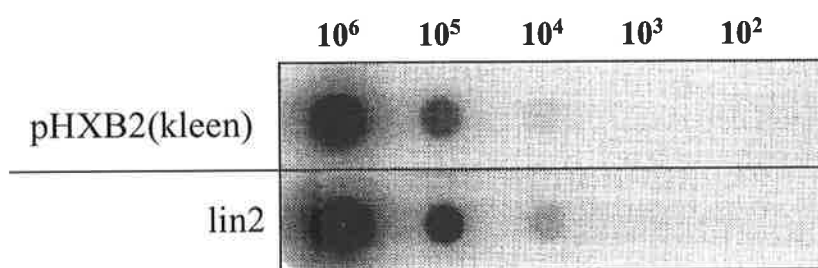
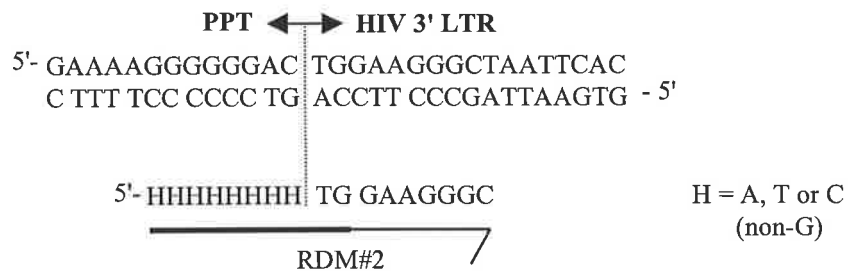


Figure 3.7

Sequence identity and alignment of the RDM#2 primer with the 3'U3/PPT junction DNA sequence. The total number of primer species within the RDM#2 degenerate primer pool is calculated below.



Total primer combinations within the a degenerate primer pool:

$$= X_1 \times X_2 \times \dots \times X_y$$

where, X = The number of potential nucleotides able to occupy a given position (indicated by the subscript number) in the primer sequence

y = the length of the degenerate primer

∴ For primer pool RDM#2,

Given: H can be one of 3 nucleotides (A, T or C)

$$= 3 \times 3 \times 3 \times 3 \times 3 \times 3 \times 3 \times 3 \times 1 \times 1 \dots$$

= **6561 primer combinations**

3.2.3 Confirming the Basic Principle of DP-PCR

Initially, a 30 cycle-PCR was performed using primers PBS(-) and RDM#2, a nominal annealing temperature of 56°C and a final MgCl₂ concentration of 2.5mM on 2×10⁵ copies of either the integrated (cc) or linear (lin2) constructs. Although background chromosomal DNA was not present, a vast excess of RDM#2 was used (100pmol) to minimise the chance of non-specific priming events depleting selected variants within the available pool of DPs. After subjecting PCR products to Southern analysis using the U3.2(-) oligonucleotide probe (see Table 2.1), preferential amplification of the integrated construct compared to the linear construct was seen (Fig. 3.8). However, since amplification of the linear construct was evident (albeit inefficiently), under these PCR conditions, annealing of RDM#2 to the 5'-termini of the linear construct was still able to occur. Therefore, various reaction conditions were investigated to increase the annealing stringency of the PCR.

3.2.4 Optimising Selectivity: [MgCl₂] and Annealing Temperature (T°anneal)

Lower MgCl₂ concentrations are routinely used to increase the specificity of primers for their corresponding target DNA templates (see section 1.8.1). Therefore, MgCl₂ concentrations were sequentially reduced by 0.5mM increments while maintaining other PCR conditions identical to those outlined in section 3.2.3. Lowering MgCl₂ levels reduced both the overall assay-sensitivity and the selectivity of DP-PCR for the integrated construct (Fig 3.9A). Conversely, when MgCl₂ levels were sequentially increased (in increments of 1mM) in otherwise identical DP-PCR reactions, there was a corresponding elevation of signal intensity (Fig. 3.9B). However, the selective nature of the DP-PCR for the integrated construct was not significantly altered by MgCl₂ concentrations under these reaction conditions.

Although preferential amplification of the integrated construct was occurring, there was still significant amplification of the linear construct. The assay was therefore repeated using an annealing temperature of either 60°C or 63°C over a range of MgCl₂ concentrations to establish whether specific combinations of MgCl₂ concentration and a

Figure 3.8 Demonstrating the DP-PCR Principle

PCR amplification of equivalent copies (2×10^5) of the pHXB2(kleen) (shown as cc) and the lin2 constructs using the RDM#2 primer pool and the PBS(-) primer. Details of the amplification procedure used are given in the text (see section 3.2.3). Amplified products were detected by Southern hybridisation using the U3.2(-) oligonucleotide probe (see Table 2.1). A control reaction involving amplification of water alone (H_2O) is indicated. DNA mass markers (pUC19/*Hpa* II) are shown.


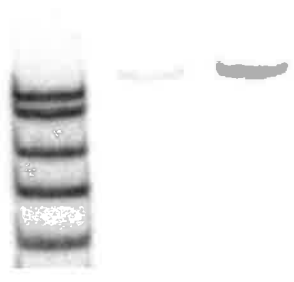
	H₂O
	2×10⁵ cc
	2×10⁵ lin2
	pUC19/<i>Hpa</i> II

Figure 3.9 Impact of [MgCl₂] on DP-PCR Sensitivity and Selectivity using RDM#2

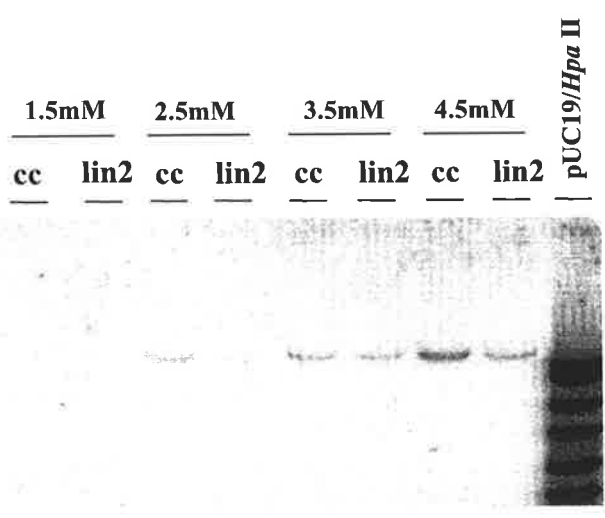
The effect of decreasing (A.) and increasing (B.) MgCl₂ concentrations within the reaction mix on DP-PCR amplification of equivalent copies (2×10^5) of the pHXB2(kleen) (indicated as cc) and the lin2 constructs using the RDM#2 primer pool and the PBS(-) primer. DNA mass markers (pUC19/*Hpa* II) are shown.

pUC19 <i>Hpa</i> II	2.5mM [MgCl ₂]		2.0mM [MgCl ₂]		1.5mM [MgCl ₂]		1.0mM [MgCl ₂]	
	lin2	cc	lin2	cc	lin2	cc	lin2	cc

A.



B.



higher annealing temperature would increase the selectivity. No signal was obtained from any samples after a 2h exposure of probed filters to phosphorous screens (Fig. 3.10, A and B). However, upon overnight exposure of filters, signals were obtained with the integrated but not the linear construct at both annealing temperatures using a MgCl_2 concentration of 4.5mM (Fig 3.10, C and D). Interestingly, signal intensity within the integrated construct samples was more pronounced at the higher annealing temperature of 63°C. This might have resulted from reduced background priming events and/or reduced inter- and intra-molecular annealing of primers within the degenerate primer pool at the higher temperature.

Although in this experiment amplification of the linear construct was below the level of detection, limited amplification of the linear construct was evident in future experiments using identical PCR conditions (*eg.* see Fig. 3.11). However, the difference in signal intensities when identical amounts of the integrated and linear construct were amplified was consistently greatest when a high MgCl_2 concentration and high annealing temperature was used. Similar results were obtained in later experiments performed by Dr. Raman Kumar using a number of different DPs (data not shown).

3.3 Background Chromosomal DNA and DP-PCR

Optimisation

3.3.1 Background HuT-78 DNA

To assess the impact background cellular DNA had on the sensitivity of the DP-PCR procedure, 1×10^5 copies of either integrated or linear constructs were spiked onto 2.5×10^5 cell-equivalents (1.5 μg) of HuT-78 chromosomal DNA. PCRs were performed using an annealing temperature of 63°C exactly as described above (see section 3.2.4) and products subjected to dot-blot analysis using the U3.2(-) probe (see Table 2.1 for the oligonucleotide sequence). The addition of background DNA significantly reduced the signal obtained following amplification (Fig. 3.11). It was noted that the dot-blot procedure did not allow the detection of 1×10^5 copies of input DNA (Fig. 3.11, no PCR

Figure 3.10 Impact of Annealing Temperature on DP-PCR Sensitivity and Selectivity using RDM#2

DP-PCR amplification of known copies (3.3×10^4) of the pHXB2(kleen) and lin2 construct using an annealing temperature of either 60°C (A and C) or 63°C (B and D) and a range of buffer MgCl₂ concentrations. Short exposures (2h; A and B) and Long exposures (overnight; C and D) are shown. An arrow indicates the position of the desired band.

1. pHXB2(kleen); 0mM MgCl₂
2. lin2; 0mM MgCl₂
3. pHXB2(kleen); 3mM MgCl₂
4. lin2; 3mM MgCl₂
5. pHXB2(kleen); 3.5mM MgCl₂
6. lin2; 3.5mM MgCl₂
7. pHXB2(kleen); 4mM MgCl₂
8. lin2; 4mM MgCl₂
9. pHXB2(kleen); 4.5mM MgCl₂
10. lin2; 4.5mM MgCl₂
11. pHXB2(kleen); 5mM MgCl₂
12. lin2; 5mM MgCl₂
13. pUC19/*Hpa* II DNA mass markers

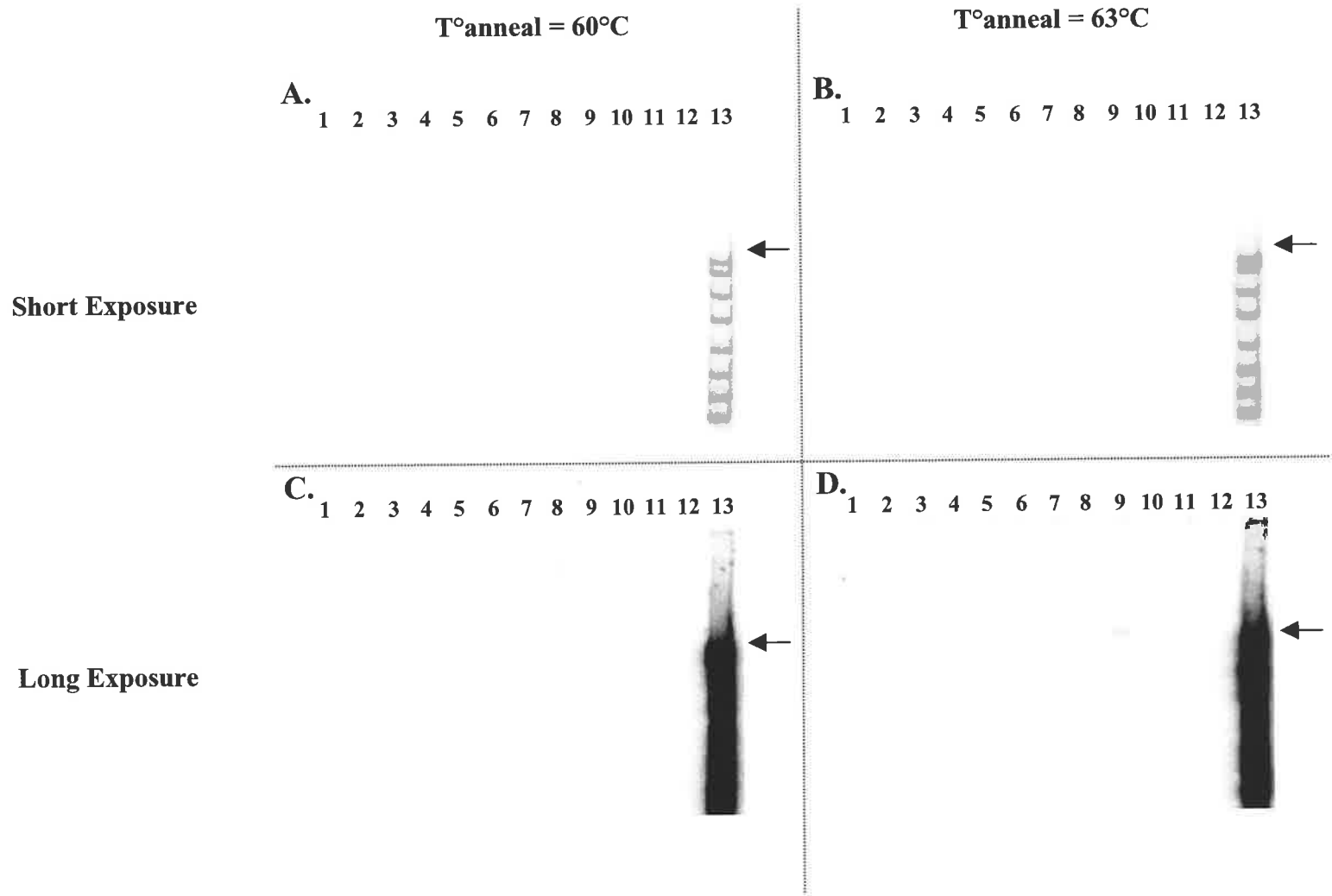


Figure 3.11 Impact of Background Chromosomal DNA on DP-PCR

DP-PCR amplification of known copies (1×10^5) of the pHXB2(kleen) (cc) and lin2 construct in the presence (+HuT-78) or absence (-HuT-78) of 2.5×10^5 cell-equivalents ($1.5 \mu\text{g}$) of HuT-78 chromosomal DNA. Control reactions in which 1×10^5 copies of each construct were spotted directly onto the hybridisation membrane are shown (no PCR).

lin2 + HuT-78

lin2 - HuT-78

cc + HuT-78

cc - HuT-78

lin2 (no PCR)

cc (no PCR)



samples), confirming that all signals observed were due to DP-PCR-mediated amplification of DNA.

The reduction in signal intensity obtained in the presence of background DNA may have resulted from the annealing of the degenerate primer to background chromosomal sequence. Since the HIV-specific region of primer RDM#2 is nine bases in length, the primer was predicted to bind in a specific manner to random chromosomal sequence at a frequency of approximately 2×10^4 times/diploid genome (see calculation below).

Size of single genome = 2.8×10^9 bp

Frequency a nine base sequence occurs in random DNA sequence = $(1/4)^9$
 = 1/262144

∴ this sequence exists $2.8 \times 10^9 / 262144$ times per genome

= 10681/single genome

= **21362/diploid genome**

Therefore, when 1×10^5 cell equivalents of background DNA is present within samples, there are $\approx 2 \times 10^9$ ($21362 \times 1 \times 10^5$) potential sites of priming in addition to sites of HIV DNA integration. Since the average distance between these sites in background DNA is ≈ 250 kb, the specific annealing of primers within the DP-pool to these regions would not be expected to result in exponential amplification of DNA sequences. However, if two sites exist in close proximity ($\leq \approx 2$ kb) to each other, or additional false priming events occur nearby, exponential amplification of background DNA could theoretically occur. Furthermore, the ability of both the degenerate primer and the HIV-specific primer to anneal to background chromosomal sequence may have been enhanced due to the high concentrations of $MgCl_2$ used in the PCR procedure.

3.3.2 Optimising Degenerate Primer Composition and Length

To increase the efficiency of DP-PCR (particularly in the presence of background chromosomal DNA), a systematic approach was undertaken to determine components of the degenerate primer necessary to maximise primer-annealing to desired target sequences. Therefore, primers were altered with respect to the length of both the 5'-

semi-degenerate region and the 3'-HIV-specific region and assessed for their ability to amplify both the integrated and linear constructs (see Fig. 3.12).

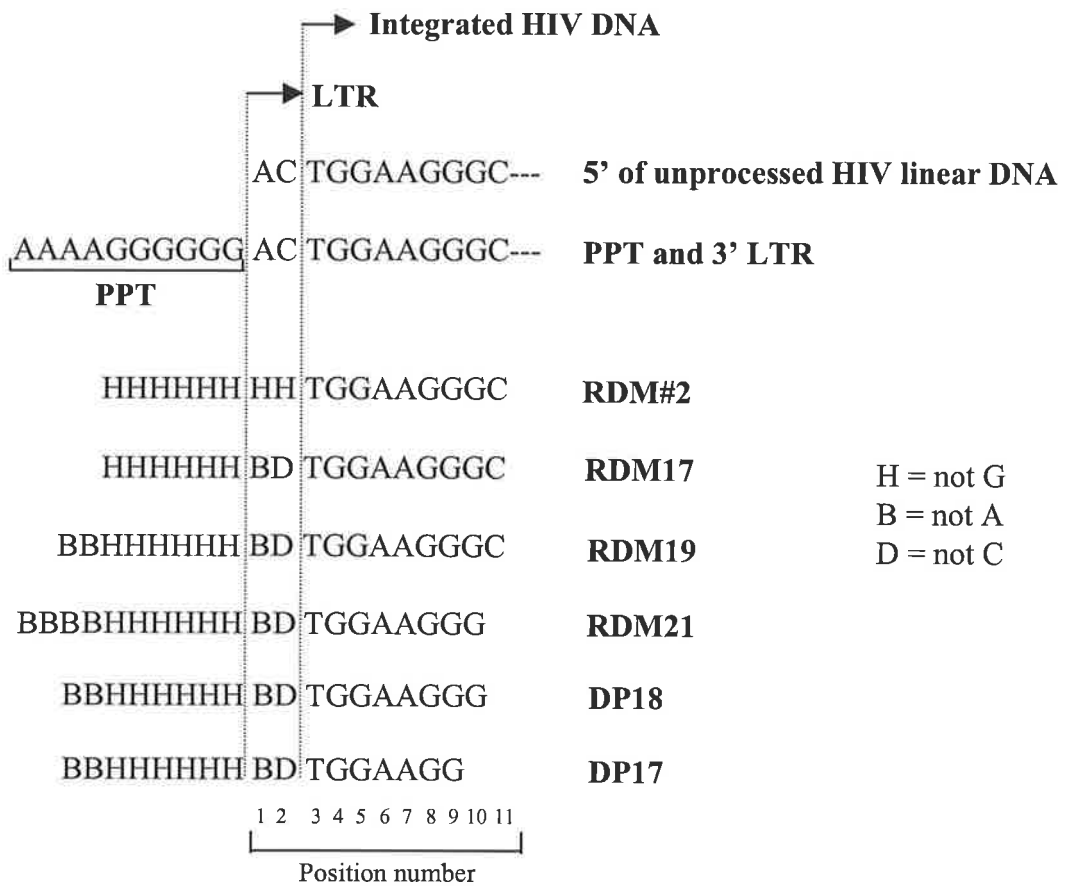
3.3.2.i Varying the composition and length of the degenerate region

To specifically examine the effect length of the DP degenerate region had on the sensitivity of the DP-PCR procedure, a number of primers were prepared based on the structure of RDM#2. RDM#2 was initially modified by substituting the HH di-nucleotide present immediately upstream of the HIV-specific region of RDM#2 (see section 3.2.2) with a BD di-nucleotide (not A and not C, respectively) to produce RDM17 (see Fig. 3.12). This modification was expected to abolish annealing of bases within the degenerate region to the AC di-nucleotide present at the 5'-end of unprocessed linear extrachromosomal DNA and the 3'-LTR region of all HIV DNA forms. Since the terminal 5'AC di-nucleotide is removed upon integration of HIV DNA into cellular DNA, amplification of the integrated provirus was expected to remain unaffected.

In addition, the degenerate region of RDM17 was sequentially lengthened by 2 nucleotides to generate primers RDM19 and RDM21 (Fig. 3.12). In all cases, the 9mer HIV-specific region (complementary to the first nine bases within the 5'-U3 region of the integrated provirus) was retained. The amplification efficiency (and relative selectivity) of all newly synthesised primers was assessed by amplifying both the cc and lin2 constructs using an annealing temperature of 63°C in the presence of 1×10^5 cell equivalents of HuT-78 background DNA over a range of MgCl₂ concentrations. All DPs were able to direct the highly selective amplification of the cc construct at a MgCl₂ concentration of 4.5mM (Fig. 3.13) supporting the earlier observation that stabilisation of such DPs to the template was heavily dependent on high MgCl₂ levels (see section 3.2.4). Although both RDM17 and RDM19 were able to amplify the integrated construct with comparable efficiency and selectivity, the RDM19 primer pool appeared to direct a more robust amplification with respect to MgCl₂ concentration and was therefore selected for further analysis.

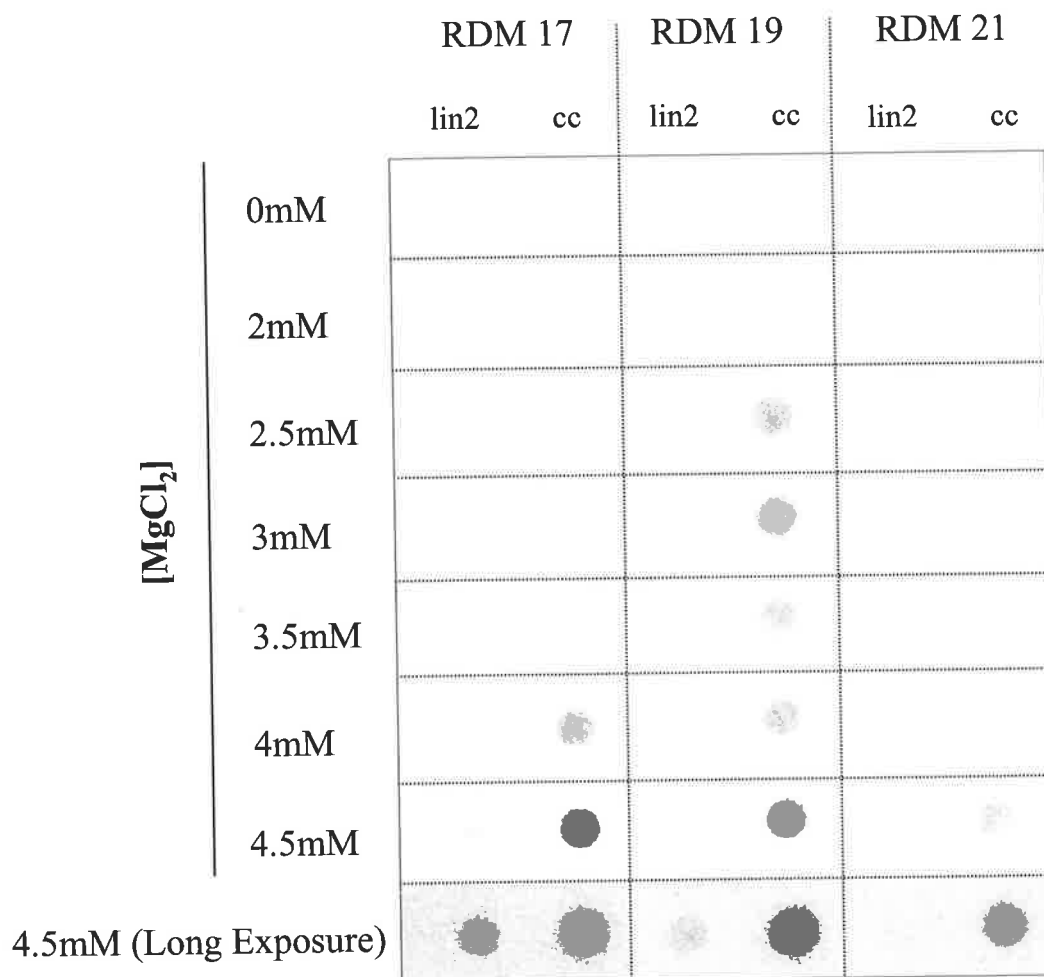
Figure 3.12

Alignment of various degenerate primers with the 5' sequence of the unprocessed linear HIV extrachromosomal form and the 3'U3/PPT junction sequence. Position numbers (shown at bottom) refer to nucleotides of the DP's that can bind to the HIV LTR region.



***Figure 3.13 Optimising Degenerate Primer Length and Composition:
Lengthening the 5' Degenerate Region***

Equivalent amounts (1×10^5 copies) of each construct (pHXB2(kleen) (cc) or lin2) spiked into 1×10^5 cell-equivalents of HuT-78 amplified using three different degenerate primers (RDM 17, RDM 19 and RDM 21) over a range of MgCl_2 concentrations. Each of the degenerate primers used differs in the length of the 5'-degenerate region (see Fig. 3.12). A longer exposure of the signals obtained using a MgCl_2 concentration of 4.5mM is shown.





To establish both the level of sensitivity and selection for the integrated construct achieved using RDM19, dilutions of the cc and linear constructs spiked onto 1×10^5 cell equivalents of HuT-78 background DNA were subjected to DP-PCR. The assay allowed the detection of $\approx 10^4$ copies of the cc construct (Fig. 3.14). In addition, the assay was between 10 and 100-fold more specific for the integrated form. Since negligible signal was expected to arise as a result of primer-annealing over the 3'-PPT/LTR region (as there is no downstream HIV-specific primer (analogous to the PBS(-) primer) with which to amplify the signal), the majority of the signal observed with the linear construct was assumed to result from the low-level binding of RDM19 to the 5'-LTR region. This was particularly surprising since the HIV-specific 9mer region of RDM19 was not expected to anneal at all to the 5'-terminal 9 bases of the linear viral DNA at 63°C (see also Fig. 3.16). The ability of RDM19 primers to bind (albeit inefficiently) to the 5'-end of the linear construct at this temperature may have been due to an unknown stabilising effect of the degenerate tail (see section 3.8.2).

3.3.2.ii Varying the length of the HIV-specific region

Although at very low levels, amplification of the lin2 construct using the RDM19 primer-pool was still observed (see Fig. 3.13 and Fig. 3.14). To assess whether reducing the length of the HIV-specific region could further increase the selectivity for the integrated construct, primer RDM19 was further modified to produce primers DP17 and DP18 (Fig. 3.12). These two primers comprised a degenerate region identical to that of RDM19, together with a 7mer and 8mer HIV-specific region, respectively. Use of these primers in the presence of 1×10^5 cell equivalents of background HuT-78 chromosomal DNA indicated that the efficiency with which the integrated construct was amplified decreased as the length of the HIV-specific component of the DP was decreased (Fig. 3.15A and B). However, although amplification of the integrated construct was slightly more efficient when RDM19 was used, DP18 appeared more specific than RDM#19 for the integrated construct. Even after an overnight exposure of the dot blot to a phosphorous screen, there was little evidence of DP18-mediated amplification of the linear construct (Fig. 3.15B).

Figure 3.14 DP-PCR Sensitivity Using RDM 19

Known amounts of either pHXB2(kleen) (cc) or lin2 were spiked into 1×10^5 cell equivalents of HuT-78 background chromosomal DNA and subjected to DP-PCR using the RDM 19 and PBS(-) primers. Amplified products were detected by dot-blot hybridisation using a ^{32}P -labelled oligonucleotide (U3.2(-)). Control reactions in which HuT-78 chromosomal DNA alone (HuT-78) or water alone (H_2O) were amplified are shown.

	HuT-78
	10^3 cc
	10^4 cc
	10^5 cc
	10^6 cc
	10^6 lin2
	10^5 lin2
	H ₂ O

***Figure 3.15 Optimising Degenerate Primer Length and Composition:
Shortening the 3' HIV-Specific Region***

A. Equivalent amounts (1×10^5 copies) of each construct (pHXB2(kleen) (cc) or lin2) spiked into 1×10^5 cell-equivalents of HuT-78 amplified using three different degenerate primers (DP17, DP18 and RDM19). Each of the degenerate primers used differs in the length of the 3'-HIV-specific region (see Fig. 3.12). **B.** A longer exposure of the signals obtained using the DP18 and RDM19 primers is shown.

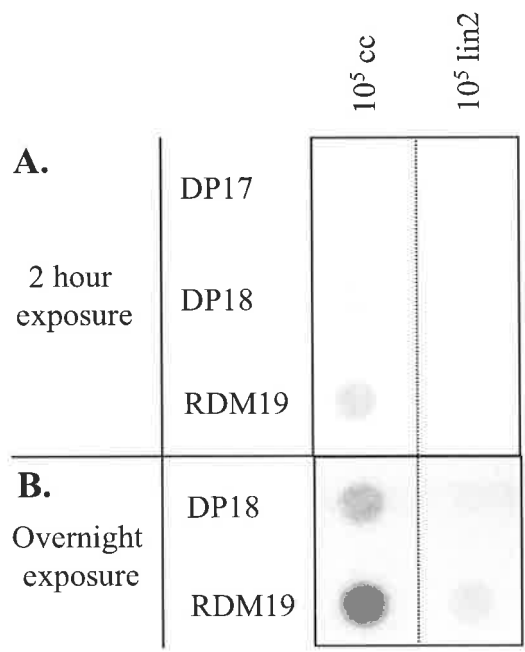
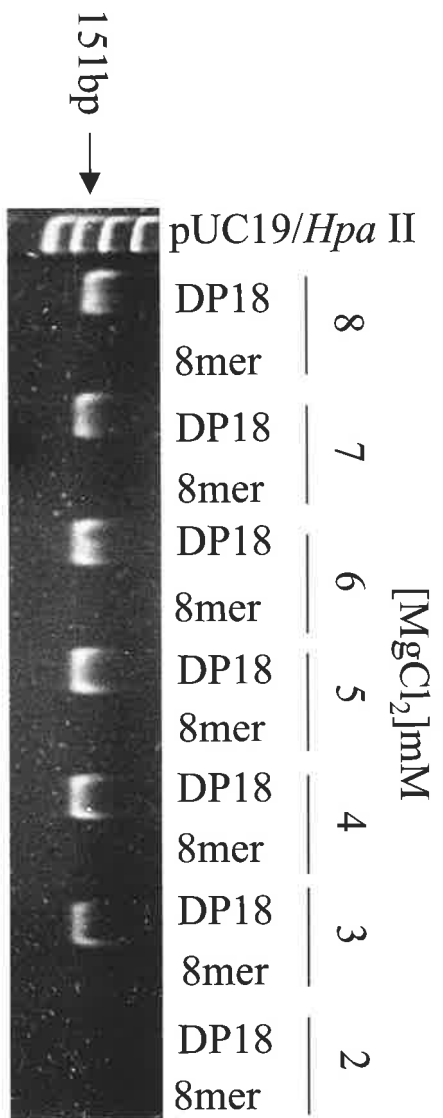


Figure 3.16 Confirming the Requirement for the 3' Degenerate Region

Amplification of $\approx 1 \times 10^7$ copies of the integrated construct using primer U3PNV (see Table 2.1) and either DP18 or an 8mer oligonucleotide corresponding to the HIV-specific region of DP18. Reactions were performed over a range of MgCl_2 concentrations. Amplified products were detected by ethidium bromide staining following electrophoresis through an 8% PAGE gel. DNA mass markers (pUC19/*Hpa* II) and the size of the expected band (arrow, left side) are shown.



A subsequent experiment in which DP18 was assessed for its' ability to direct amplification of $\approx 1 \times 10^7$ copies of the integrated construct over a range of MgCl_2 concentrations confirmed that optimal amplification was achieved at MgCl_2 concentrations $> 4\text{mM}$ (Fig. 3.16). Since lower MgCl_2 concentrations were expected to reduce the amount of annealing to background chromosomal sequence, the original MgCl_2 concentration of 4.5mM (established for RDM#2; see section 3.2.4) was used in all future experiments involving DP18. In addition, the complete absence of signals when an 8mer primer corresponding to the HIV-specific region of DP18 was used instead of DP18 (Fig. 3.16) conclusively demonstrated the requirement of the degenerate portion for primer stability on the DNA template.

3.4 Enhancing the Selectivity of DP-PCR: Incorporation of *Bsr* 1

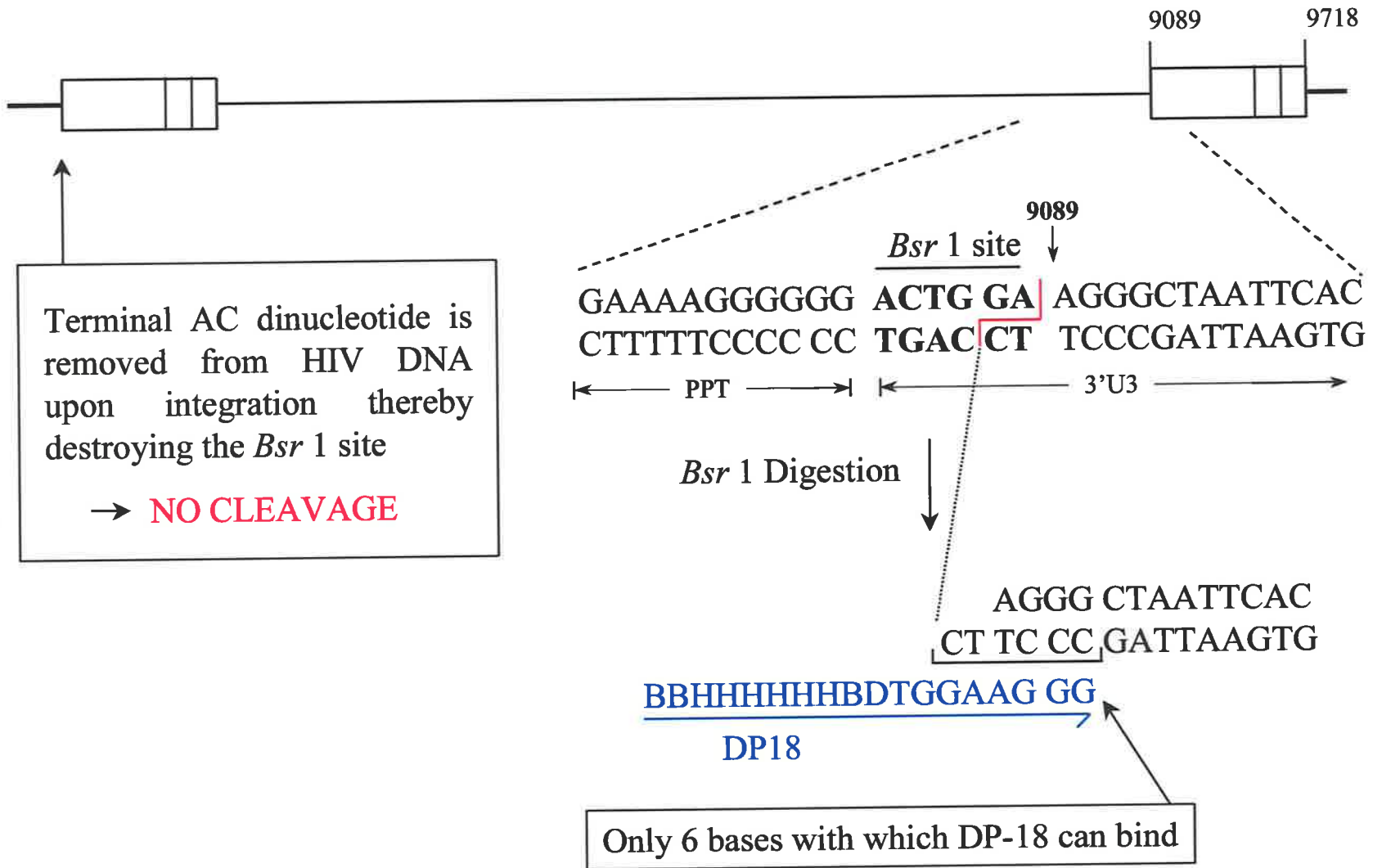
Although the sensitivity of the assay at this stage was poor regardless of the degenerate primer used (*eg.* 10^4 cc copies; see section 3.3.2.i), emphasis was initially placed on optimising the selective nature of the procedure. It was expected that alternative techniques (such as nested PCR procedures) could be employed in later experiments to increase the sensitivity of the assay.

3.4.1 The Principle of *Bsr* 1 Digestion

In an attempt to further increase the specificity of DP-PCR for integrated HIV DNA, cleavage of target DNA with the restriction enzyme *Bsr* 1 was investigated. The 6-base recognition site corresponds to the first 6 nucleotides at the left-hand end of the HIV U3 sequence (Fig. 3.17 and 3.18). This enzyme therefore cleaves the 3'-U3 region from upstream PPT sequences, and cleaves at the junction of the two LTR regions in 2-LTR circles. However, cleavage at the junction of the 5'-U3 and cellular sequence in the integrated form is unlikely since a di-nucleotide is removed upon integration of HIV DNA (see section 1.3.2.i) that destroys the recognition site (see Fig. 3.17). However, integrated viral DNA will be cleaved if it is inserted immediately adjacent to an identical di-nucleotide (a CA di-nucleotide) in the cellular sequence, thereby restoring the *Bsr* 1 site. The frequency with which this event can be expected to occur in random

3.17 Bsr 1 Digestion of Integrated HIV DNA

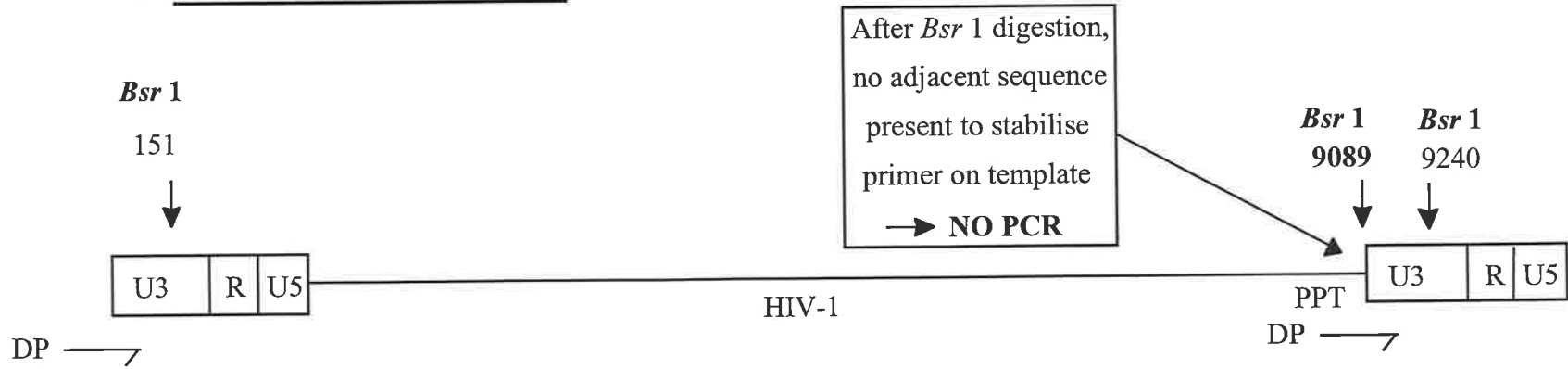
Bsr 1 digestion of integrated HIV DNA at the junction of the 3'U3/PPT regions. The *Bsr* 1 recognition site is shown in boldface and the pattern of cleavage is shown as a red line. The U3 DNA fragment resulting from *Bsr* 1 cleavage is shown aligned to the DP18 primer sequence (shown in blue). Cleavage at the analogous site at the 5'U3 region (5'U3/chromosomal DNA junction) with *Bsr* 1 does not occur since a terminal dinucleotide (AC) is removed upon integration that destroys the *Bsr* 1 recognition sequence.



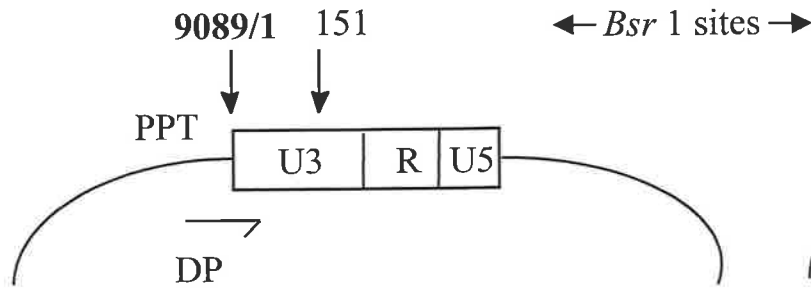
3.18 Bsr 1 Digestion of Extrachromosomal HIV DNA

Diagram showing the key positions at which *Bsr* 1 cleaves in the linear (**A.**), 1-LTR (**B.**) and 2-LTR (**C.**) HIV extrachromosomal DNA forms. Sites at which *Bsr* 1 cleaves to remove sequence to which DP18 is expected to bind are shown in boldface. The position of the polypurine tract (PPT) is also shown.

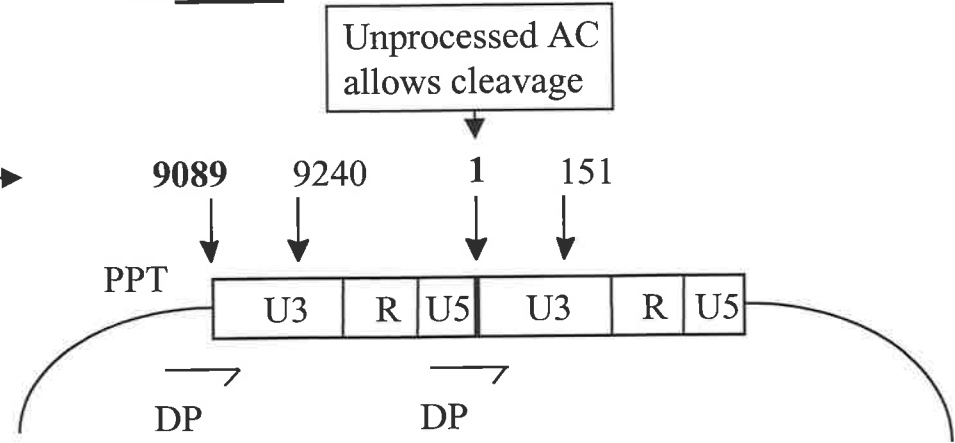
A. Linear Extrachromosomal



B. 1-LTR



C. 2-LTR



DNA sequence is one every 16 integration events. Theoretically therefore, 94% of all integration events can be detected by DP-PCR when *Bsr* 1 digestion of sample DNA prior to amplification is employed.

Although expected to be only a minor contaminant of HIRT pellet (chromosomal DNA) preparations, the chance that the 2-LTR circular HIV DNA form will be amplified by the DP-PCR protocol is significantly reduced by *Bsr* 1 digestion. Circularisation to form the HIV 2-LTR form involves the direct end-to-end ligation of LTR regions present in the linear viral DNA molecule (see section 1.4). Sequence analysis of LTR-junction regions has revealed that the terminal U3 region AC dinucleotide (removed upon integration) is retained in these forms (191). Therefore, a complete *Bsr* 1 recognition site exists at the junction of the LTR regions in the 2-LTR circular HIV DNA form. Cleavage with *Bsr* 1 will therefore remove sequence upstream of the two U3 regions present within 2-LTR HIV DNA (the PPT/U3 junction and the U5/U3 junction; see Fig. 3.18C) to which the degenerate region of DP18 might bind.

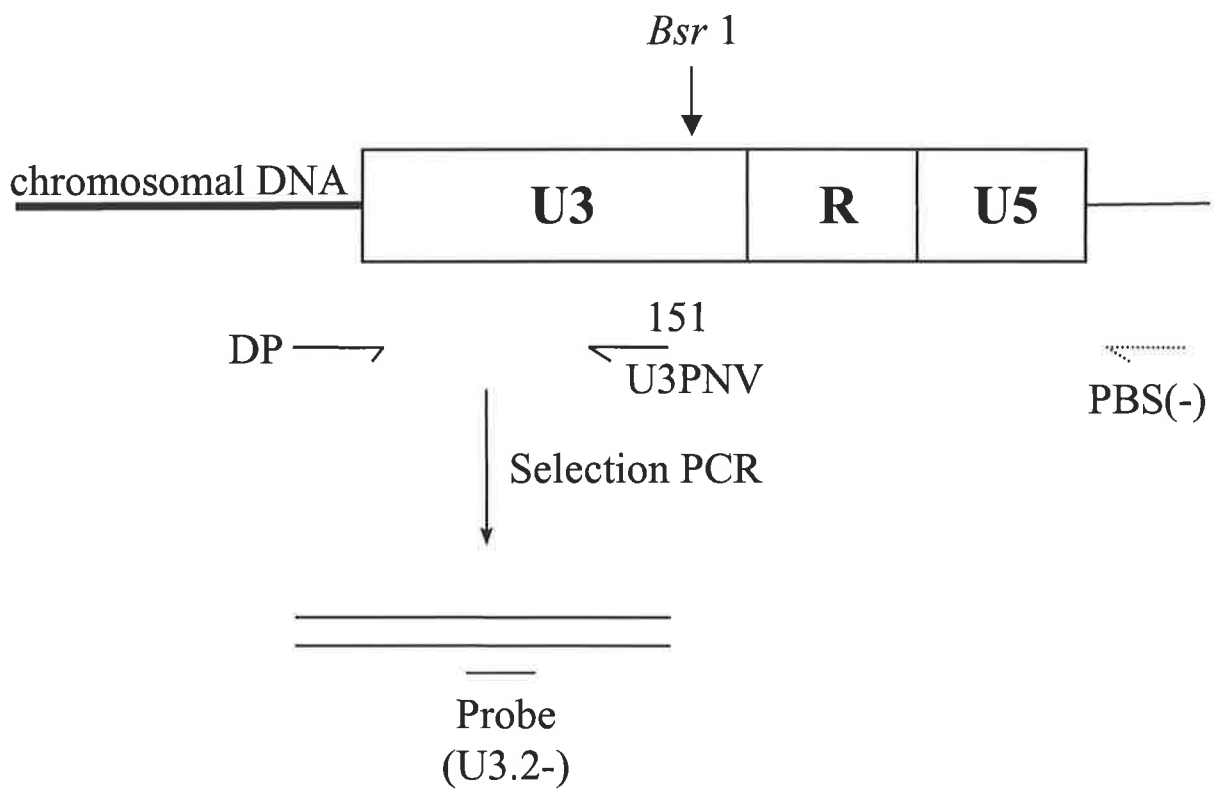
Since additional *Bsr* 1 sites are present in the HIV sequence prior to the annealing site of the PBS(-) primer, this primer could not be used in the subsequent DP-PCR amplification of integrated HIV DNA. Instead, the U3PNV primer (see Table 2.1) was used whenever *Bsr* 1 pre-digestion of target DNA was employed. An outline of the revised DP-PCR reaction is given in figure 3.19.

3.4.2 Benefit of *Bsr* 1 digestion in DP-PCR

Prior to performing the modified DP-PCR protocol, *Bsr* 1 digestion conditions (in a 10 μ l volume) were optimised by both a gel electrophoresis protocol and a PCR-based protocol using primers flanking the *Bsr* 1 site at the HIV PPT/U3 junction (data not shown). Digestion of target DNA sequences with *Bsr* 1 (in a 10 μ l volume) prior to performing DP-PCR (in a final volume of 20 μ l) was subsequently shown to substantially reduce the efficiency with which the lin2 construct was amplified (Fig. 3.20, compare 10⁶ lin2 lane in rows 1 and 2), without affecting amplification of the integrated construct (Fig. 3.20, 10⁶ cc lane in rows 1 and 2). Furthermore, U3PNV-mediated amplification of undigested samples (both integrated and linear constructs)








3.19 Modification of the DP-PCR Strategy to Incorporate Bsr 1

Diagrammatic representation of the revised strategy used for DP-PCR amplification of integrated HIV DNA following digestion with *Bsr* 1. The positions of the DP, the new HIV-specific primer (U3PNV) and the old HIV-specific primer (PBS(-); dashed arrow, right-side) are shown.



3.20 Bsr 1 Digestion Enhances DP-PCR Selectivity

DP-PCR amplification of 1×10^6 copies of either pHXB2(kleen) (cc) or the lin2 construct spiked into 5×10^4 cell-equivalents of background DNA in the presence (row 1) or absence (row 2) of *Bsr* 1 digestion. DP-PCR amplification in the absence of *Bsr* 1 digestion using primers DP18 and PBS(-) was performed (row 3) to allow a comparison of the PBS(-) and U3PNV primers. Control reactions using LTR-specific primers (U3.1(+) and U3PNV) to control for the input copy-number of each construct are shown in row 4. For details regarding the amplification conditions used in each reaction, see section 3.4.2.

Row #	Description	10 ⁶ cc	10 ⁶ lin2
1	<i>Bsr</i> 1 digested; PCR in 20μl		
2	Undigested; PCR in 20μl		
3	Undigested; PCR in 20μl (primers: PBS(-)/DP-18)		
4	Undigested; U3PNV/U3.1(+)		

were comparable (if not more efficient) than PBS(-)-mediated amplification (compare Fig. 3.20, rows 2 and 3). This indicated that the substitution of the U3PNV primer for the PBS(-) primer did not affect the efficiency of DP-PCR amplification of target sequences. In addition, comparable signals arose in samples of integrated (cc) and linear (lin2) constructs subjected to PCR using common LTR-specific primers (Fig. 3.20, row 4) confirming that similar amounts of each construct were present in all samples analysed. In conclusion, sample digestion with *Bsr* 1 prior to performing DP-PCR using the DP18 and U3PNV primers significantly enhanced the selective nature of the assay.

3.5 Increasing the Sensitivity of DP-PCR

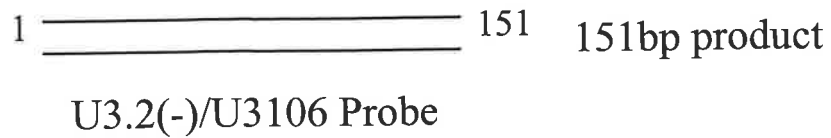
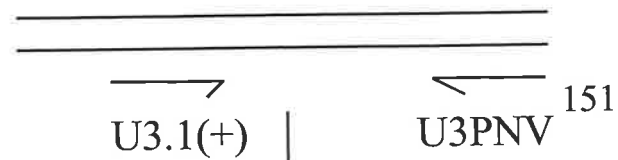
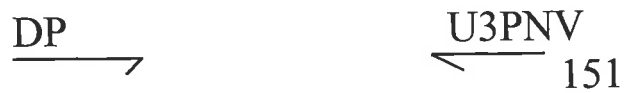
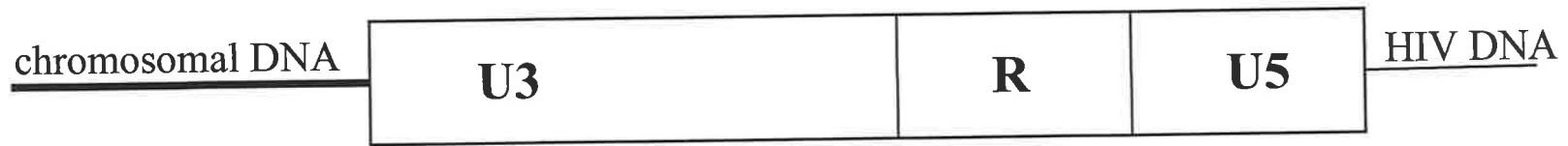
Routinely, 5×10^4 cell-equivalents (300ng) of HuT-78 DNA had been used in previous experiments as background DNA. In the context of this background, the DP-PCR procedure using the RDM19 primer was shown to be able to detect 10^4 copies of the integrated construct (see section 3.3.2.i). Since the sensitivity of assay was not expected to increase dramatically when DP18 was used, attempts were made to increase the sensitivity of the procedure by the incorporation of a nested PCR protocol (Fig. 3.21).

3.5.1 Optimising the Nested PCR

To ensure that the nested PCR amplification of 1st-round PCR products gave a linear dose-response, first-round PCR products were sequentially diluted to a point where a nested PCR (of fixed cycle-number) did not display evidence of saturation. Equally, alteration of the nested PCR cycle-number using a fixed amount of 1st-round PCR product could have been used to achieve the same goal. However, this approach would have required numerous independent PCRs to be performed, whereas the use of the first approach allowed the sub-saturating conditions required for the nested PCR to be established in a single PCR. Consequently, 1×10^5 copies of either the integrated or linear constructs were amplified by 1st-round PCR (28 cycles) and then various dilutions of product subjected to a 25-cycle nested PCR. The nested PCR procedure (using primers U3.1(+) and U3PNV) was performed essentially under the same conditions as those used to establish the copy number of integrated and linear control constructs (see

Figure 3.21 Incorporating a Nested PCR

Diagram showing the nested DP-PCR amplification strategy used to detect integrated HIV DNA. Nested PCR products were detected using either the U3.2(-) oligonucleotide (Table 2.1) probe or the U3-106 (Table 2.2) probe.



section 3.2.1.iii). Since primer U3PNV was common to both 1st-round and 2nd-round PCRs, this 2nd-round PCR is more correctly termed a semi-nested PCR.

Nested PCRs performed on more than 1/500th of the 1st-round PCR product (*ie.* 1/10th, 1/50th and 1/100th) gave a plateau of maximal signal intensity for the integrated construct (see Fig. 3.22A, B and C). Although difficult to definitively conclude without performing additional nested PCRs on lower amounts of template, amplification of 1/500th and 1/1000th of the 1st-round product appeared to be on the linear portion of the dose response curve (Fig. 3.22C). At both these dilutions of 1st-round product, the signals observed following amplification of the *lin2* construct were very low.

3.5.2 Sensitivity and Selectivity of the Nested DP-PCR Procedure

The sensitivity of the nested DP-PCR procedure was subsequently assessed by amplification of various dilutions of the integrated construct spiked onto 1×10^5 cell-equivalents of HuT-78 DNA. Concurrently, 1×10^5 copies of the linear construct was amplified in a similar fashion to determine the selectivity of the procedure for amplification of the integrated construct. All DNA samples were digested with the restriction enzyme *Bsr* 1 prior to the 1st-round PCR. The results showed that the nested DP-PCR procedure allowed the successful detection of 100 copies of the integrated construct (Fig. 3.23A). Signals arising from amplification of the integrated construct were quantified and plotted against cc copy-number to generate a curve of the amplification profile (Fig. 3.23B). The curve generated was linear ($R^2 > 0.997$) indicating that amplification was proceeding in a logarithmic fashion. The signal arising from amplification of 1×10^5 copies of the linear construct corresponded to approximately 3×10^3 copies of the integrated construct (see graph, Fig 3.23B) indicating that the assay was ≈ 33 -fold more specific for detection of the integrated construct.

To apply this assay to a preparation of infected cell DNA, an extract of 8E5 cell DNA (1 HIV copy/cell (89); 1×10^5 cell equivalents) was also amplified by the nested DP-PCR procedure. The signal obtained corresponded to 9.7×10^4 copies (expected = 1×10^5) when read from figure 3.23B.

Figure 3.22 Optimising the Nested PCR

Various amounts ($1/10^{\text{th}}$ to $1/1000^{\text{th}}$) of first-round DP-PCR products were subjected to a 25-cycle PCR using primers U3.1(+) and U3PNV. First-round PCRs were performed for 28-cycles on 1×10^5 copies of either pHXB2(kleen) (cc) or the lin2 constructs using primers DP18 and U3PNV. Nested PCR products were detected by Southern hybridisation using the U3.2(-) oligonucleotide probe and signals quantified using ImageQuant software. **A.** Signals following a 1h exposure of blots to a phosphor screen. **B.** Signals following a 4h exposure of blots to a phosphor screen. **C.** Signal-intensities (Counts; y-axis) of spots in **B.** (cc) graphed against the dilution of 1st-round product used in the nested PCR (x-axis).

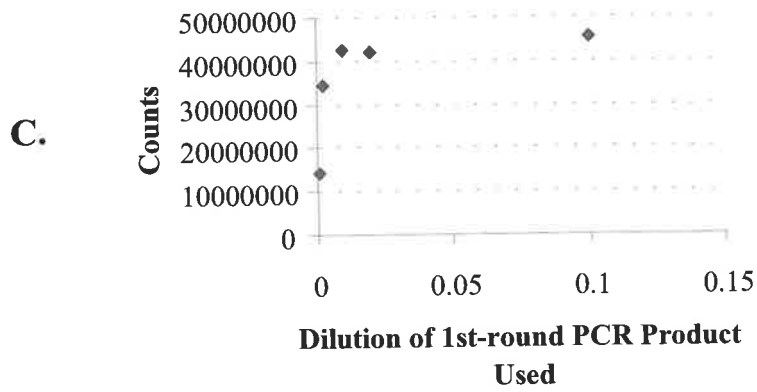
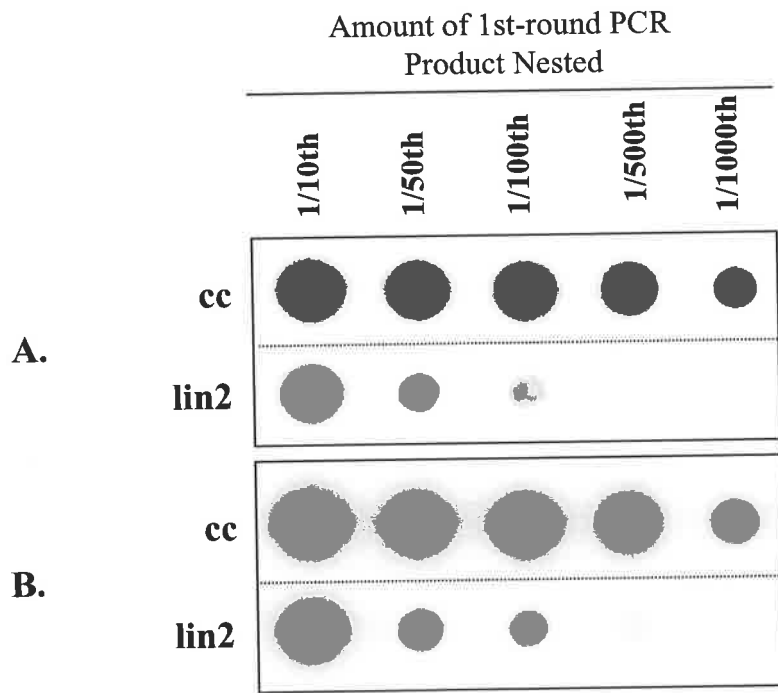
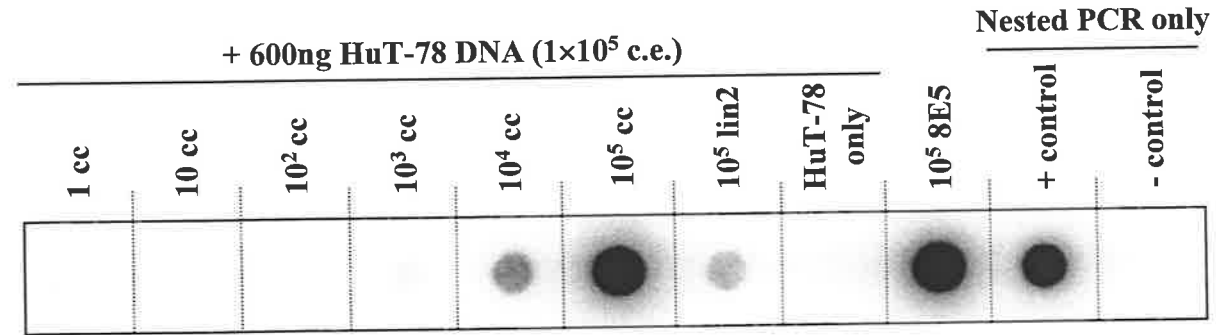


Figure 3.23 Sensitivity and Selectivity of DP-PCR

A. Nested DP-PCR amplification of known amounts of the pHXB2(kleen) (cc) and the lin2 constructs spiked into 1×10^5 cell-equivalents (c.e.) of HuT-78 chromosomal DNA using the DP18 degenerate primer. Amplification of 1×10^5 cell-equivalents of chromosomal DNA extracted from 8E5 cells (1 integrated HIV copy/cell) is also shown (10^5 8E5). A reaction involving amplification of HuT-78 chromosomal DNA alone (HuT-78 only) is shown. Control reactions in which direct amplification of either 1×10^5 copies of the cc construct or water alone was performed using the nested PCR conditions are shown on the right side (+control and -control, respectively). **B.** Signal intensities resulting from the nested DP-PCR amplification of cc dilutions were quantified using ImageQuant software (Counts) and plotted against the input copy-number (Copies of cc). The equation ($y=$) and the regression value (R^2) of the resulting curve is shown.

A.



B.

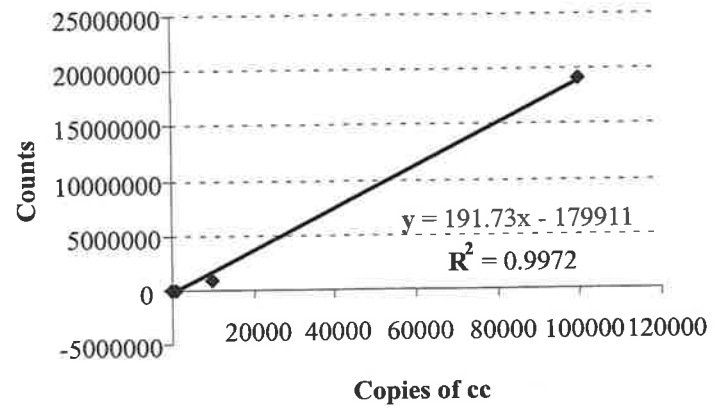
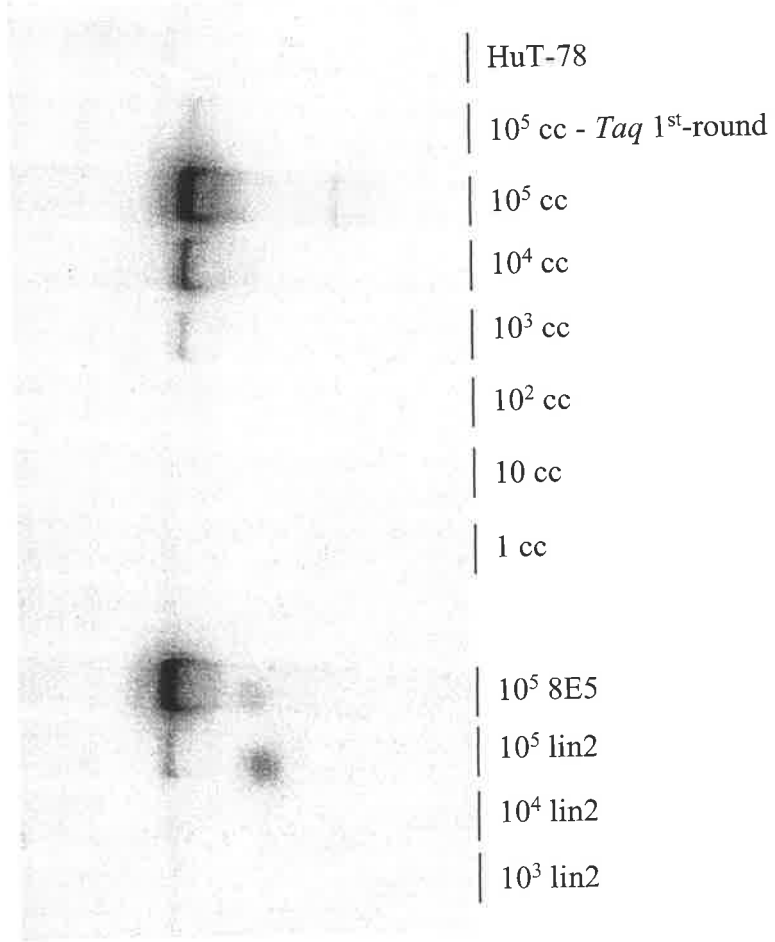


Figure 3.24 Nested DP-PCR Using a ³²P-Labelled Nested PCR Primer

Nested DP-PCR amplification of known amounts of the pHXB2(kleen) (cc) construct, the lin2 construct and 8E5 chromosomal DNA using the DP18 degenerate primer. The cc and lin2 constructs were spiked into 1×10^5 cell-equivalents (c.e.) of HuT-78 chromosomal DNA. First-round DP-PCRs and nested PCRs were performed exactly as outlined for Fig. 3.23 (see also section 3.5.2) except that nested PCRs were performed using ³²P-labelled U3.1(+) primer. Nested PCR products were run through an 8% polyacrylamide gel, which was subsequently dried and exposed to a phosphorous screen to allow visualisation of bands. Control reactions in which first-round DP-PCR amplification was performed on HuT-78 chromosomal DNA alone (HuT-78) or on 10^5 copies of the cc construct in the absence of *Taq* polymerase (10^5 cc - *Taq* 1st-round) are shown.

149bp →



In an attempt to further increase the sensitivity of the assay, the nested PCR was performed using primer U3.1(+) that had been labelled with ^{32}P using T4 polynucleotide kinase. Nested PCR products were run down an 8% polyacrylamide gel, which was subsequently dried, and exposed directly to a phosphorous screen. A single distinct band of PCR product was obtained but no increase in sensitivity was achieved (Fig. 3.24).

3.6 Characterisation of DP-PCR with Additional Control Constructs

3.6.1 Preparation of HA8 Integrated HIV DNA Standards

In order to confirm that the DP-PCR procedure could detect a number of different integration events, a mixture of three clonal cell lines containing known amounts of integrated HIV DNA was prepared (section 2.2.4.i.). Briefly, the clonal cell lines H3B (2 integrated HIV DNA copies per cell (160)), ACH-2 (1 copy of integrated HIV DNA per cell (53)) and 8E5 (1 copy of integrated HIV DNA per cell (89)) were mixed in ratios of 0.5:1:1 (H3B:ACH-2:8E5) and chromosomal DNA prepared (see section 2.2.3). Therefore, this mixed integrated HIV DNA control standard (HA8) contained 4 distinct HIV DNA integration events. It is worth noting, that these cell lines have each been shown to harbour very little extrachromosomal HIV DNA (53, 89, 160).

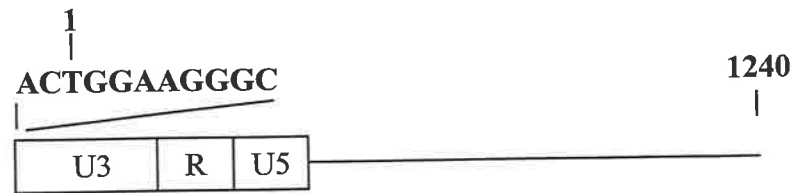
3.6.2 Preparation of Additional Constructs Mimicking Extrachromosomal HIV DNA

To determine whether the unwanted signal consistently observed following amplification of the linear control construct was derived from amplification over the 5'-U3 or 3'-U3 regions, a control construct containing only the terminal 5'-LTR region (termed the LHS construct, Fig. 3.25A) was generated by PCR amplification (30 cycles) of the cc construct using primers AC+1-21 and GAG-FL2 (see Table 2.1) and standard reaction conditions (see section 2.2.5). The resulting PCR product, containing the entire 5'-LTR (complete with the terminal AC di-nucleotide) and downstream sequence to

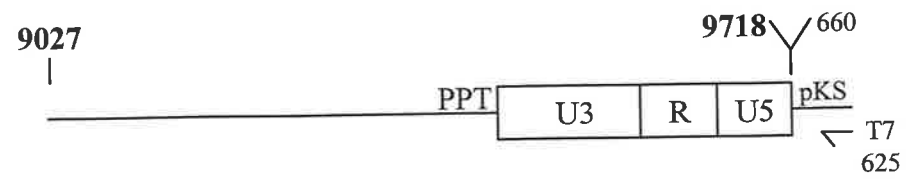
Figure 3.25

Constructs mimicking the left side of linear extrachromosomal HIV DNA (**A.**), the right side of linear extrachromosomal HIV DNA (**B.**) and the 2-LTR junction region (**C.**). pBluescript sequence is indicated as pKS. Numbers in boldface represent base-pair positions of the Human Immunodeficiency Virus Type 1 (HXB2) DNA sequence (genebank accession number K03455). Non-boldface numbers represent base-pair positions of the pBluescript(KS) vector sequence (Stratagene).

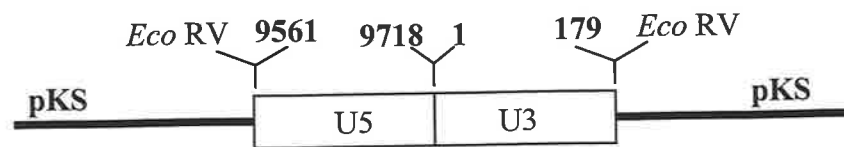
A.



B.



C.



position 1240 in the HIV DNA sequence, was gel purified and equated for LTR copy-number with the HA8 standard (see below).

Similarly, a control construct containing only the 3'-LTR region and surrounding HIV sequence (termed the RHS construct; see Fig. 3.25B) was generated using primers 9027(+) (Table 2.1) and T7 (designed to anneal within pBluescript(KS); see Table 2.1). The resulting PCR product, containing the entire HIV sequence from position 9027 onwards (see the HIV DNA sequence, genbank accession number K03455) was gel purified and equated for LTR copy-number with the HA8 standard (see below).

A 2-LTR control construct (Fig. 3.25C) prepared by Dr. Raman Kumar (see section 2.2.4.iii) was also equated for LTR copy-number with the HA8 standard (see below) to assess the ability of the DP-PCR procedure to detect circular HIV DNA containing a double LTR region. Although this form was not expected to be present at significant levels in HIRT chromosomal preparations, it was considered necessary to establish the efficiency with which this form was amplified by DP-PCR.

3.6.3 Equating New Control Standards

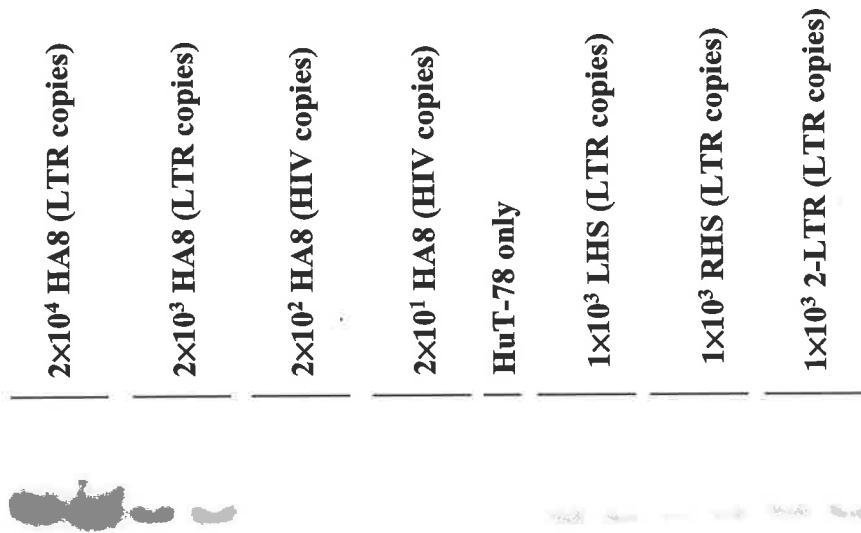
Stock solutions of the LHS, RHS and 2-LTR control constructs standards were diluted to approximately 200 copies/ μ l based on their DNA content by spectrophotometry (see section 2.2.4.iv). Samples (5 μ l) of each diluted solutions were mixed with 1×10^4 cell-equivalents of background HuT-78 DNA and subjected to a 23 cycle standard PCR (see section 2.2.5) using primers U3.1(+) and U3PNV (Table 2.1). Following Southern analysis using probe U3-106 (Table 2.2), the signals were compared to those obtained after amplification of various dilutions of the HA8 integrated standard (also in background DNA).

The signal intensities of the HA8 standards demonstrated a linear dose response ($R^2 > 0.997$; Fig. 3.26A and B). Since integrated copies present within the HA8 standard contained two copies of the viral LTR region, the copy number used for calculations referred to the number of LTR copies present in the sample (see Fig. 3.26A). Signal intensities from duplicate samples of LHS, RHS and 2-LTR samples were averaged and

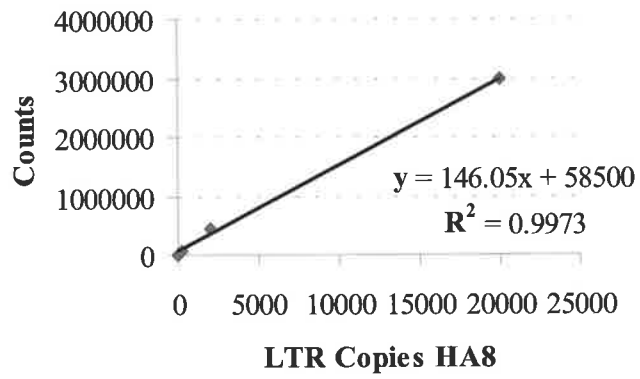
Figure 3.26 Equating the LHS, RHS and 2-LTR constructs with the HA8 Integrated HIV DNA Standard

A. Amplification of known amounts of the HA8 integrated HIV standard to establish the copy-number within stock solutions of the LHS, RHS and 2-LTR control constructs. PCR was performed using primers U3.1(+) and U3PNV (Table 2.1) as described in section 3.6.3. The LHS, RHS and 2-LTR constructs were spiked into 1×10^4 cell-equivalents (c.e.) of HuT-78 DNA. HuT-78 chromosomal DNA was added where necessary to samples containing HA8 DNA such that 1×10^4 c.e. of DNA (final) was present. A control reaction in which HuT-78 chromosomal DNA alone (1×10^4 c.e.) was amplified is shown (HuT-78 only). **B.** Signal intensities (Counts; see Appendix 3.1) resulting from amplification of the HA8 integrated HIV DNA standards plotted against input copy-number (LTR copies HA8). The equation ($y=$) and regression value (R^2) of the resulting curve is shown.

A.



B.



converted to LTR content using the standard curve generated by the HA8 standards (see Fig. 3.26B and Appendix 3.2). Values obtained were equivalent to 880, 340 and 1140 HA8 LTR copies, respectively. Based on these results, the stock concentrations of each construct were calculated.

3.6.4 Characterising DP18 using New Controls

To determine whether the sensitive and selective amplification of DNA present within the HA8 integrated control standard was occurring, duplicate samples containing 100 copies of integrated DNA, and samples containing 100 copies of either LHS, RHS or 2-LTR control constructs, were subjected to the nested DP-PCR procedure. The PCR conditions used were identical to those used in section 3.5.2. In this experiment, 1.2 μ g (2×10^5 cell equivalents) of background HuT-78 chromosomal DNA was added to each sample prior to PCR. Following Southern analysis of nested PCR products using the U3-106 probe (Table 2.1), strong signals arising from amplification of integrated HIV DNA present within the HA8 standards were observed (Fig. 3.27A). In contrast, light bands of approximately equal intensity were observed in only one lane (of duplicate lanes) corresponding to amplification of 100 copies of either LHS or RHS constructs. This suggested that the unwanted amplification of the linear control construct (lin2) observed in previous experiments may have been occurring at either the 5'-U3 or the 3'-U3 regions. Furthermore, additional experiments performed on the LHS and RHS control constructs failed to consistently produce equivalent signals within duplicate samples, suggesting that chance priming events occurring in early cycles of the 1st-round PCR may significantly influence the extent to which either control construct was amplified (data not shown).

To compare the signals between each of the components of the HA8 integrated standards, individual preparations of 5×10^4 cell-equivalents of H3B, ACH-2 and 8E5 were subjected to DP-PCR. The resulting signals differed between the preparations (Fig. 3.27B). The signal corresponding to the 8E5 sample (5×10^4 HIV copies) was slightly lower than that observed after amplification of 1×10^5 copies of the integrated construct (cc) indicating that the efficiency of the DP-PCR procedure was comparable to that observed previously (see section 3.5.2). However, integrated DNA present within the

Figure 3.27 Nested DP-PCR Amplification of the HA8 Integrated HIV DNA Standard

A. Known amounts of the HA8 integrated HIV DNA standard (100 HIV copies) and the LHS, RHS and 2-LTR control constructs (100 LTR copies) subjected to the nested DP-PCR protocol using the DP18 degenerate primer in the presence of 2×10^5 c.e. of HuT-78 chromosomal DNA. Nested PCR products were detected by Southern hybridisation using the U3-106 probe. **B.** Independent DP-PCR amplification of the cc construct and integrated HIV DNA within the H3B, ACH-2 and 8E5 cell lines. Known amounts of the cc construct (1×10^5 copies) and chromosomal DNA isolated from the H3B, ACH-2 and 8E5 cell lines (5×10^4 cell equivalents (c.e.)) were subjected to nested DP-PCR as described in section 3.5.2. Control reactions in which first-round DP-PCR amplification was performed on HuT-78 chromosomal DNA alone (HuT-78) or on 10^5 copies of the cc construct in the absence of *Taq* polymerase (Nested PCR only) are shown.

A.

100 HA8
100 LHS
100 RHS
100 2-LTR
H ₂ O
H ₂ O nested PCR

B.

	1x10 ⁵ cc (+ 5x10 ⁴ c.e. HuT-78)
	5x10 ⁴ c.e. 8E5 (5x10 ⁴ HIV copies)
HuT-78	5x10 ⁴ c.e. ACH-2 (5x10 ⁴ HIV copies)
Nested PCR only	5x10 ⁴ c.e. H3B (1x10 ⁵ HIV copies)
Blank	

ACH-2 cell line was amplified with greater efficiency than that present in either the 8E5 or H3B cell lines. Furthermore, the signal observed following amplification of integrated HIV DNA present within 5×10^4 cell equivalents of the H3B cell line (a total of 1×10^5 integrated HIV DNA copies) was well below that observed following amplification of 5×10^4 cell equivalents of the 8E5 cell line (containing a total of 5×10^4 integrated HIV DNA copies). The differences in the amplification profile of integrated DNA present in each of the three cell lines suggested that the identity of the cellular sequence immediately adjacent to the 5'-U3 region of the integrated provirus greatly affected the amplification efficiency.

3.6.5 Sequence Analysis of the H3B, ACH-2 and 8E5 Cell Lines

To establish whether the identity of sequence to which the degenerate portion of DP18 was designed to bind could affect the efficiency of DP-PCR, the sequence of cellular DNA immediately upstream of the integrated provirus in each cell line (previously obtained by Dr. Raman Kumar, see (248)) was analysed. Although the H3B cell line contained two integrated proviruses (160), only a single sequence was obtained. Alignment of the three sequences (see Fig. 3.28) confirmed that the sequence to which the HIV-specific region of DP18 was designed to bind was extremely well conserved between the isolates present in each of the cell lines. In addition, in each case extensive annealing of bases within the degenerate tail of DP18 (a minimum of 7/10) could theoretically have occurred. However, the theoretical ability of primers within the DP18 primer pool to bind to the cellular sequence adjacent to the integrated provirus did not correlate with the amplification efficiency of DP-PCR. For example, despite possessing only one nucleotide in the upstream cellular sequence to which the degenerate tail of DP18 could not bind (Fig. 3.28, H3B sequence, boldface and underline), the integrated DNA present within the H3B cell line was amplified with the lowest efficiency by DP-PCR. In contrast, cellular DNA adjacent to integrated proviral DNA present in the ACH-2 cell line contained nucleotides at three positions in which mis-matching of the degenerate primer would be expected to occur. However, the efficiency with which DP-PCR-mediated amplification of integrated HIV DNA occurred in this cell line was significantly greater than that observed for either the H3B (1 mismatch), 8E5 (3 mismatches) or cc (1 mismatch) cell line. Furthermore, sequentially mismatched

Figure 3.28

Sequence alignment of the DP18 primer with the cellular DNA/5'HIV U3 junction sequences obtained from the H3B, ACH-2 and 8E5 cell lines and the pBluescript(KS)/5'HIV U3 junction sequence of the pHXB2(kleen) (cc) construct. The efficiency with integrated HIV DNA within each standard is amplified by the DP-PCR protocol is indicated (DP-PCR column) as follows:

- + low amplification efficiency
- ++ medium amplification efficiency
- +++ high amplification efficiency

Nucleotide positions that would be expected to form a mis-match with the DP18 primer are indicated by boldface and underline.

DP18	HHHHHHHHBD	TGGAAGGG	DP-PCR
H3B	C TTC CACTT <u>C</u>	TGGAAGGG	+
ACH-2	<u>G</u> CTT T <u>G</u> TAA <u>A</u> T	TGGAAGGG	+++
8E5	<u>G</u> <u>G</u> C CTCTT T <u>C</u>	TGGAAGGG	++
cc	TAG <u>G</u> TTCTAGA	TGGAAGGG	++

H = non G
 B = non A
 D = non C

nucleotides (3 or more in a row), that might be expected to confer primer instability to the template DNA, were not present. Taken together, mismatching of nucleotides present in the cellular DNA with bases in the degenerate region of DP18 could not explain the variable nature of DP-PCR.

3.7 Detection of Integrated DNA Following HIV Infection of HuT-78 T Cells

The variation of DP-PCR in amplifying integrated DNA within individual cell lines suggested that similar variability might occur in amplifying integrated HIV DNA molecules present in cells following infection. Additionally, although the HA8 construct consisted of 4 individual integration events, it was possible that this standard would not accurately reflect the efficiency with which multiple integrated HIV DNA molecules inserted at random into cellular sequence were amplified. Therefore, chromosomal DNA preparations of Hut-78 T cells acutely infected with HIV_{HXB2} were assayed to determine the efficiency with which the DP-PCR procedure amplified randomly integrated HIV DNA.

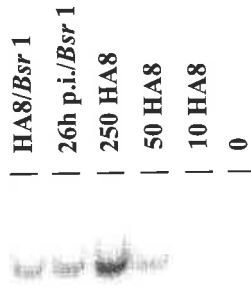
The infected HuT-78 T cell chromosomal DNA used in this experiment (26 hours post infection) had previously been assayed for the presence of integrated HIV DNA by a modified nested Alu-PCR procedure (see section 4.10.1). Using this technique, infected HuT-78 chromosomal samples were shown to contain 2-5 copies of the integrated provirus per cell (see Fig. 6.5; 26h p.i., No Drug). Therefore, a sample containing 1000 cell-equivalents of infected HuT-78 chromosomal DNA was expected to contain >2000 copies of integrated provirus.

Samples of the infected HuT-78 chromosomal DNA and HA8 integrated standard DNA digested with *Bsr* 1 were initially assayed by PCR for β -globin content (see section 2.2.5.i) to confirm that equivalent amounts of cellular chromosomal DNA were present in each case (Fig. 3.29A). Amplification of 1×10^3 cell-equivalents of infected HuT-78 chromosomal DNA and HA8 DNA were then subjected to 30 cycles of 1st-round DP-PCR (30 cycles). In addition, 1000 copies of a mixture of the LHS, RHS and 2-LTR constructs (330 copies of each) spiked into 1000 cell-equivalents of uninfected HuT-78

Figure 3.29 DP-PCR Amplification of Randomly Integrated HIV DNA

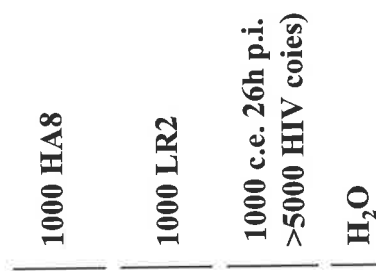
Infected HuT-78 T cell chromosomal DNA was harvested 26h p.i. following infection of HuT-78 T cells at an MOI of 1 TCID₅₀ units/cell using centrifugal enhancement (see section 2.2.1). This preparation was shown to contain 2-5 copies of the integrated HIV provirus per cell using the nested-*Alu* PCR protocol. **A.** β -globin PCR analysis of *Bsr* 1-digested chromosomal DNA isolated from HA8 (HA8/*Bsr* 1) and HuT-78 cells infected with HIV_{HXB2} (26h p.i./*Bsr* 1). Known amounts of undigested HA8 amplified in a similar manner served as copy-number standards. **B.** Nested DP-PCR amplification (in duplicate) of 1000 cell-equivalents (c.e.) of *Bsr* 1-digested chromosomal DNA isolated from HA8 (HA8/*Bsr* 1) and HuT-78 cells infected with HIV_{HXB2} (26h p.i./*Bsr* 1). Control reactions involving DP-PCR amplification of a mixture of the LHS, RHS and 2-LTR constructs (1000 copies each; LR2) and water alone (H₂O) are shown.

A.



B.

149bp →



chromosomal DNA were also assayed. The absence of a 151bp PCR product following amplification of the infected HuT-78 chromosomal DNA sample suggested that amplification of integrated HIV DNA within this sample was extremely inefficient (Fig. 3.29B). In contrast, amplification of integrated HIV DNA within the HA8 sample was evident. Repeated experiments gave similar results (data not shown). This suggested that for reasons we do not fully understand, the reaction conditions determined for the DP-PCR procedure had favoured the amplification of certain integration events. Consequently, it was concluded that unless it was substantially remodelled, the DP-PCR procedure was unlikely to allow the robust amplification of randomly integrated HIV DNA and was therefore not pursued further.

3.8 Discussion

In this chapter, I have described an attempt to develop a PCR-based assay that specifically amplifies integrated HIV DNA. This assay was based on the ability of primers within a semi-degenerate primer pool to anneal to the terminal 5'U3 region of HIV and the upstream sequence present only in the integrated HIV proviral form. Initial experiments in which a variety of degenerate primers were constructed (differing with respect to the length of the HIV-specific region, and the length and composition of the degenerate region) identified the DP18 primer pool (10-base degenerate region and 8-base HIV-specific region) as the outstanding degenerate-primer candidate. Furthermore, after *Bsr* I digestion and nested PCR, the DP-PCR procedure was shown to be consistently 10-100 times more specific for an integrated construct (pHXB2(kleen)) than a construct mimicking linear HIV DNA (lin2). However, DP-PCR was repeatedly unable to mediate the efficient amplification of integrated HIV DNA present within cellular chromosomal DNA preparations extracted following infection of T cells with HIV.

3.8.1 Sensitivity of DP-PCR

The success of the DP-PCR protocol (regardless of the DP pool used) in early experiments appeared to be heavily dependent on high concentrations of MgCl₂ in the PCR mix. MgCl₂ levels within reaction mixes are known to influence the stringency

with which primers bind the template (target) DNA. Low MgCl_2 concentrations have the effect of increasing the specificity of the primer for the template (*ie.* there is a more stringent interaction of primers with their complementary sequences) and therefore serve to reduce the amount of non-specific primer-annealing events occurring (see section 1.8.1). In contrast, a high level of MgCl_2 within reaction mixes serves to decrease the stringency requirements for successful annealing of primers. Consequently, elevated amounts of background priming events occur due to primers binding to sub-optimal (non-100% complementary) sequences within sample DNA. Therefore, the high MgCl_2 (low stringency) environment used in the DP-PCR procedure may have increased the chance of successful binding and extension events occurring over the cellular DNA/5'U3 junction by lending less significance to any mis-pairing of the degenerate region of the DP within the adjacent chromosomal sequence. Supporting this, the melting temperature ($T^{\circ}\text{m}$) of an oligonucleotide with 2 mis-matches annealed to a target sequence is significantly higher in the presence of 4.5mM MgCl_2 than 2.5mM MgCl_2 (see Fig. 3.30A). Since the melting temperature of an oligonucleotide is defined as the temperature at which 50% of the oligonucleotides in solution are in a duplex formation (224), it can be used as a measure of the strength of a primer-template interaction. Therefore, the number of successful annealing events were likely to be limited under conditions of high stringency (low salt) to those primers within the DP pool that consisted of a degenerate region that was precisely complementary to the cellular sequence immediately upstream of the 5'U3 region.

The early experiments demonstrating successful DP-mediated amplification of integrated HIV DNA were performed using an integrated DNA control plasmid (pHXB2(kleen)). However, in contrast to the diversity of integrants expected to be present within chromosomal DNA samples isolated following infection, this construct represented a single integration event. Although later experiments showed that integrated HIV DNA within chromosomal DNA isolated from a mixture of persistently infected cell lines (HA8) was efficiently detected using the DP18 primer pool, the efficiency with which each of the integrants within this standard was detected was shown to vary considerably (see section 3.6.4). In addition, integrated HIV DNA within chromosomal DNA isolated 26h following acute infection with HIV was unable to be detected using the DP18 primer pool under the established PCR conditions. Taken

together, these results suggested that the conditions established for DP-PCR using the DP18 primer population might have favoured amplification of particular integration events.

Although the mechanism remains unknown, a number of explanations can be proposed to explain the varying efficiencies with which different HIV integrants were amplified using the DP18 primer pool. It is worth emphasising that the DP18 primer pool theoretically consists of 3^{10} (59049) primer-types, each consisting of a different degenerate 5'-region. For each integration event, only one primer-type will be able to anneal with 100% complementarity to the cellular sequence immediately adjacent to the 5'U3 sequence. The remaining primer combinations display various affinities for the cellular/5'U3 DNA sequence depending on the composition of the DP degenerate region.

An unknown factor affecting the ability of the DP-PCR protocol to direct amplification of integrated HIV DNA, was the number of nucleotides of the DP degenerate region required to interact with sequence upstream of the 5'U3 region in order to stabilise the DP on the target sequence. Our results suggested that although amplification could occur when 3 mis-matches (*ie.* 7/10 potentially paired nucleotides) within the degenerate region of DP were present, the position at which the mis-matches occurred affected the amplification efficiency of the target sequence (see Fig. 3.28, compare ACH-2 and 8E5). Annealing of oligonucleotides with their target sequence has been shown to be heavily influenced by the extent of mis-matching that occurs, the strength of the successfully paired nucleotides and the positions at which mis-pairing occurs (5, 6, 195, 263). Supporting this, a primer within the DP pool mis-paired at the 5'-end with target DNA was predicted to mediate a more stable primer-template interaction than if mis-pairing had occurred more centrally (Fig. 3.30B; compare integrants #3 and #4). This occurs because the stability of a single base-pair is largely influenced by the stability of neighbouring pairing events (5, 6, 195). Furthermore, higher proportions of $G \equiv C$ pairing events over the DP18 degenerate region mediate a stronger interaction of the primer with the template compared to high proportions of $A = T$ pairing events (Fig. 3.30C; compare integrants #5 and #6). If we assume that both integrants #5 and #6 (see Fig. 3.30C) are present within a sample, and that an excess of the two primer-types exist

Figure 3.30 Modelling DP18 Primer Interactions

These modelling experiments were performed using MeltCalc software (Schutz and von Ahsen, 1999) that can be downloaded at <http://www.meltcalc.com>. The melting temperature ($T^{\circ}\text{C}_m$) of the best matched primer within the DP18 primer-pool is calculated for each example template (Integrand) and reflects the strength of primer/template interaction. Positions at which mis-matches between the DP18 primers and target sequences are indicated by boldface. The source data for these modelling experiments is available in Appendix 3.2. **A.** The effect of MgCl_2 concentration on the T°_m of a primer containing 2 mis-matches with the template sequence (Integrand #1). **B.** The effect of no mis-matches (Integrand #2), mis-matching of three terminal bases (Integrand #3) and mis-matching of three central bases (Integrand #4) on the T°_m of a primer within the DP18 primer pool. **C.** The effect on primer T°_m of multiple $\text{C}\equiv\text{G}$ (Integrand #5) vs $\text{A}=\text{T}$ (Integrand #6) interactions of bases within the degenerate region of the DP with the target sequence.

A. 5' - H H H H H H H H B D T G G A A G G G **DP18**
5' - C A T G A T C T A T T G G A A G G G **Integrant #1**

T^{°m} = 32.3 @ 2.5mM MgCl₂
T^{°m} = 37.1 @ 4.5mM MgCl₂

B. 5' - H H H H H H H H B D T G G A A G G G **DP18**
5' - C A T C A T C T C T T G G A A G G G **Integrant #2**

T^{°m} = 59.4

5' - H H H H H H H H B D T G G A A G G G **DP18**
5' - G G G C A T C T C T T G G A A G G G **Integrant #3**

T^{°m} = 47.1

5' - H H H H H H H H B D T G G A A G G G **DP18**
5' - C A T C A T C G A C T G G A A G G G **Integrant #4**

T^{°m} = 42.4

C. 5' - H H H H H H H H B D T G G A A G G G **DP18**
5' - C C C C C C C C G T G G A A G G G **Integrant #5**

T^{°m} = 75.0

5' - H H H H H H H H B D T G G A A G G G **DP18**
5' - A A A A A A A T A T G G A A G G G **Integrant #6**

T^{°m} = 52.7

H = non G
B = non A
D = non C

within the DP18 primer pool that pair perfectly over the degenerate sequence, the differences in the melting temperature between these two primer-types is 22.3°C based on G/C content alone. This difference in primer-template stability would be expected to result in differences in the efficiency with which each integrant was amplified.

Another potential explanation describing why certain integrants were inefficiently amplified involved the relative proportions of each primer-type within the pool of degenerate primers. Since only a small proportion of primers within the DP18 primer pool would be expected to be capable of directing amplification of an individual integration event, variation in the primer pool composition may contribute to the inability of the DP-PCR protocol to amplify some integrants. Supporting this, we obtained evidence (data not shown) that strong biases of certain primer-types existed within a pool of primers consisting of theoretically random 5'-regions (Dr. Raman Kumar, unpublished). This suggested that synthesis of the DP degenerate region may not have been occurring in an absolutely random fashion. Furthermore, results obtained in early experiments could not be subsequently repeated with newly synthesised stocks of DP18, indicating substantial batch variation in the manufacture of degenerate primer solutions (data not shown). Taken together, these observations suggested that the DP18 primer pool (and other DP primer pools) was unlikely to have contained an equal distribution of all the possible primer combinations. Therefore, the presence of certain primer combinations in relatively low proportions may have resulted in the inefficient amplification of selected integration events.

A further complicating factor associated with the successful detection of certain HIV integrants, was the presence of the DP HIV-specific sequence (present at the 3'-end of the DP) throughout the cellular chromosome. The 8-base oligonucleotide sequence at the 3'-end of DP18 was expected to occur approximately once every 66kb ($1/4^8$) in random sequence. Therefore, there were $\approx 8.5 \times 10^4$ sites throughout a single cellular chromosome (assuming that the diploid human T cell genome is 5.6×10^9 bp in length) to which primers within the DP18 population could potentially anneal and direct DP-PCR amplification. During DP-PCR, amplification at each of these sites would be competing with amplification at the site at which the HIV provirus was inserted, leading to the rapid depletion of various primer-types and saturation of components within the

reaction mix. Integrants within clonal cell lines, in which the cellular sequence immediately adjacent to the 5'U3 region do not bare significant resemblance to the sequences present immediately upstream of randomly occurring 8mer (DP18 3'-region) sequences in the cellular chromosome, might be expected to be amplified with reasonable efficiency, as competition for primer-types would likely be minimal. This scenario may have contributed to our observations that clonal HIV integrants within the pHXB2(kleen) construct or certain cell lines (such as ACH-2) could be detected by DP-PCR, but amplification of random HIV integrants present within cellular chromosomal DNA isolated following infection could not be achieved.

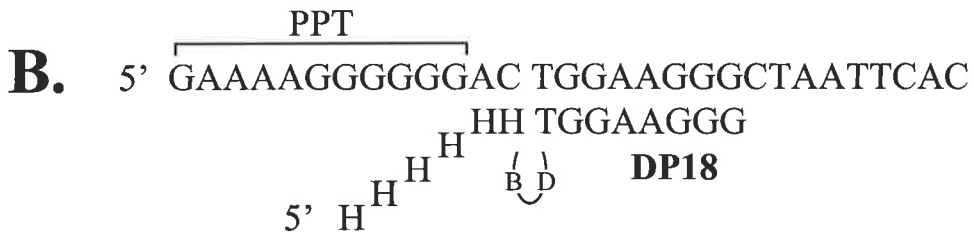
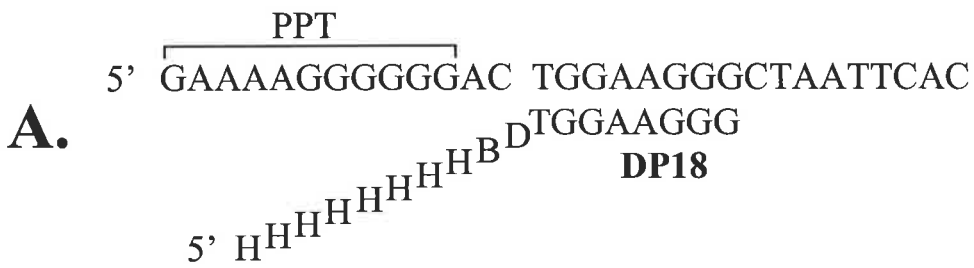
Finally, due to the **semi**-degenerate nature of the 5'-region of oligonucleotides within the DP18 primer population, integration events in which the upstream cellular sequence (present immediately adjacent to the 5'U3) resembles the HIV PPT region would be inefficiently amplified. However, such integration events would be expected to be a rare occurrence (a precise match of the PPT: one in every $\approx 1.0 \times 10^6$ integrants; 1 mis-match: one in every $\approx 3.5 \times 10^5$ integrants; 2 mis-matches: one in every 1.2×10^5 integrants, *etc*) and would therefore represent a minority of integrants following infection.

3.8.2 Selectivity of DP-PCR

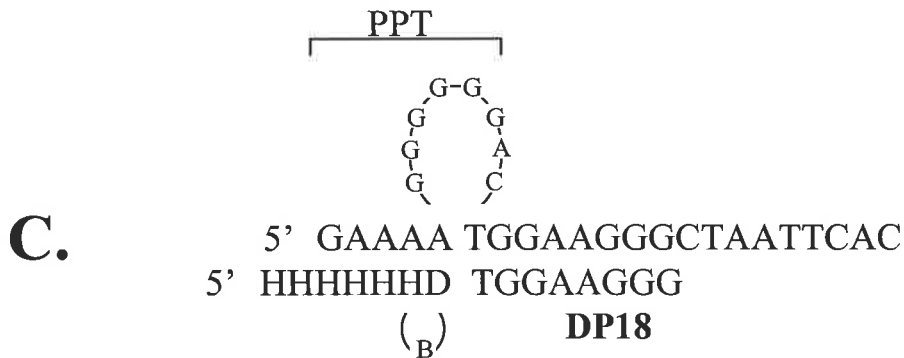
Although the degenerate regions of DP's were designed to be unable to anneal to the HIV PPT region, DP18-mediated amplification (albeit inefficient) of a construct mimicking the linear HIV extrachromosomal form (in the absence of *Bsr* 1 digestion) was routinely observed. A likely explanation accounting for this observation involves the ability of single-stranded DNA to loop essentially conferring the ability of primers to shift through reading frames on target DNA sequences. This ability of primers to loop at mis-paired sites is the basis for many PCR-based site-directed mutagenesis protocols (120, 264). In the context of the DP-PCR procedure, DNA looping may potentially occur within the 5'-degenerate region of the DP (Fig. 3.31B), the PPT sequence (Fig. 3.31C) or both (see also Fig. 3.31C). Each of these scenarios could potentially increase the extent of base-pairing between DP18 primers and the PPT/3'U3 region compared to an unlooped scenario (Fig. 3.31A) and thereby increase the chance of successful amplification over this region. Furthermore, the extent to which amplification of

Figure 3.31

Potential scenarios in which DNA looping may increase the chance of primer-annealing over the HIV 3'U3/PPT region. **A.** Unlooped primer and template **B.** Looping within the degenerate region of DP18 **C.** Looping of both the degenerate region of DP18 and the PPT region.



H = non G
 B = non A
 D = non C



extrachromosomal HIV DNA proceeded was expected to reflect the time at which chance annealing of primers within the DP18 pool to the 5' or 3' U3 regions occurred in these forms. If priming (and extension) occurs at these sites during the early cycles of the PCR, then successive priming events on the resulting product will ensure that significant amplification of extrachromosomal DNA will occur (see Fig. 3.32A and B for proposed mechanisms). Therefore, the degree to which amplification of extrachromosomal HIV DNA forms were abolished was likely to be heavily dependent on the frequency with which priming events occurred early in the PCR procedure.

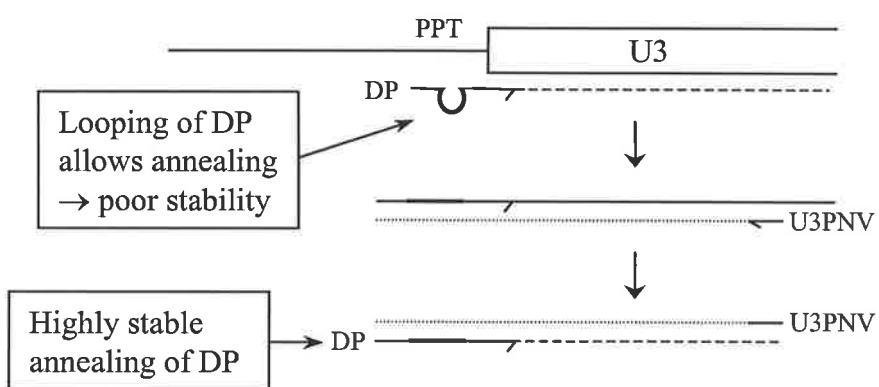
Interestingly, in addition to the amplification of the RHS construct, DP-mediated amplification of HIV DNA present within the *Bsr* 1-digested lin2 construct and the LHS construct was also observed (see Fig. 3.27A). Since the HIV-specific 3'-end of degenerate primers alone was unable to mediate amplification of HIV DNA within the pHXB2(kleen) construct (see section 3.3.2.ii and Fig. 3.16), a stabilising effect associated with the presence of the degenerate region seemed to allow binding of DPs to the 5'U3 termini. Although the mechanism by which the degenerate region was able to confer a stable primer-template interaction at this site remains unclear, the existence of oligonucleotides able to form stable triplex structures, particularly at purine- or pyrimidine-rich stretches of DNA sequence, has been well documented (37, 240)). Triple-helix forming oligonucleotides (TFO's) are oligonucleotides that are able to bind in a sequence-specific fashion to the major groove of double-stranded DNA to form a triplex structure, and have been (and continue to be) widely investigated for potential applications in gene therapy. Indeed, 11mer TFO's designed to form triple helical structures at the U5 LTR region have been shown to be able to inhibit HIV integration *in vitro* (18, 19). Since the HIV sequence to which degenerate primers were designed to bind is highly pyrimidine-rich, we therefore speculate that the unbound DP degenerate region may be able to loop back on itself to form a stable triplex structure (see Fig. 3.33 A and B). Although expected to be heavily dependent on the sequence of the degenerate region and therefore probably occurring at low frequencies, partial or full triplex formation at the 3'-U3 termini may confer sufficient primer-template stability to allow the subsequent extension of the DP by *Taq* polymerase.

An additional effect that may have contributed to the ability of both extrachromosomal HIV DNA and integrated HIV DNA to be amplified involves the inherent potential for

Figure 3.32

The effect of chance priming events resulting from looping (**A.**) and mis-matching (**B.**) of DPs on primer annealing in subsequent PCR cycles. In both cases, an exact copy of the bound primer (including any looped or mis-matched bases) is made in the next cycle to which primers within the DP primer pool would be expected to then bind with high affinity. Mis-matches are shown as crosses.

A.



B.

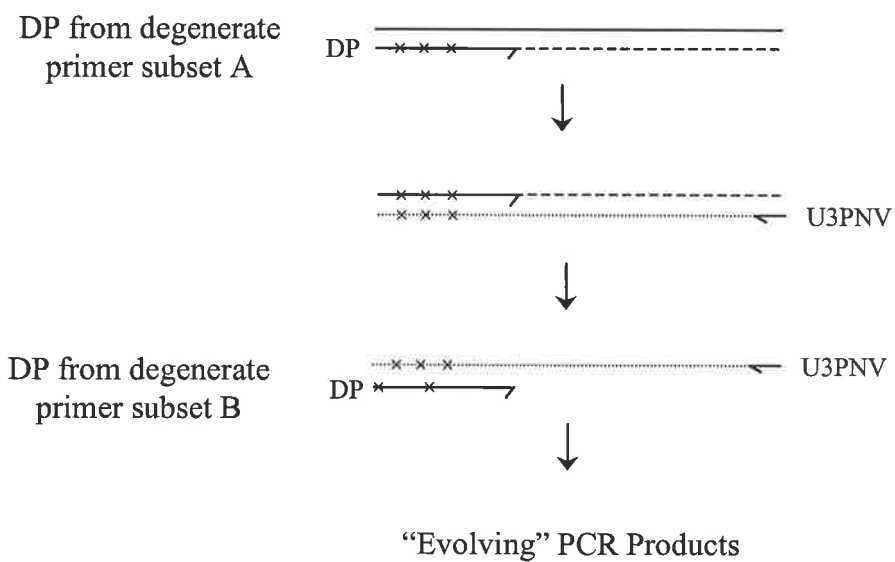
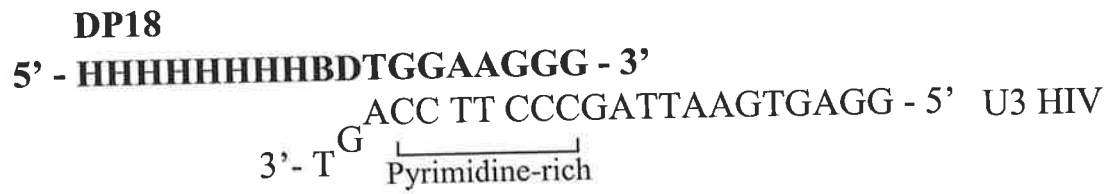


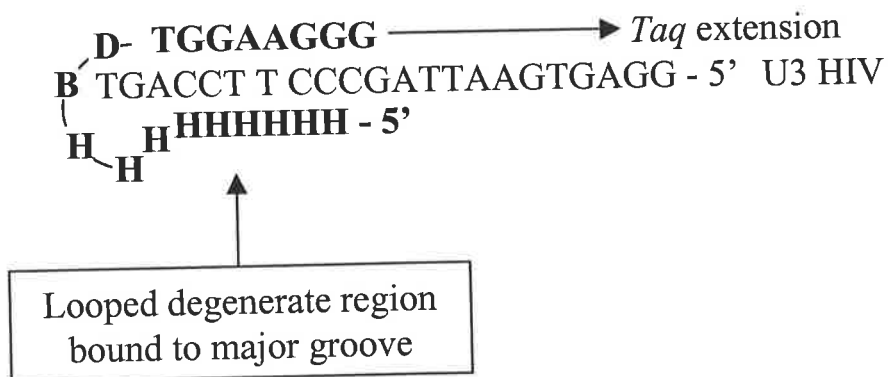
Figure 3.33

Triple-helix formation of DP18 (boldface) with the 5'U3 region of the linear extrachromosomal HIV DNA form. The pyrimidine-rich region of the U3 HIV region is shown. **A.** Unlooped DP18 annealing to the terminal 5'U3 sequence. **B.** Looping of the DP18 degenerate region allowing triple-helix formation at the termini of the 5'U3 sequence.

A. Unstable template-primer interaction



B. Stable template-primer interaction



newly synthesised PCR products to be amplified by different subsets of primer-types, if mis-matching within the DP degenerate region is occurring. Successful priming involving a mis-matched degenerate primer type (*eg.* 3 mis-matches within the degenerate region) would lead to the incorporation of the degenerate region of the DP at the 5'-end of the resulting product (Fig. 3.32B). Subsequent extension over this region in the next extension phase of the PCR by the HIV-specific primer (*eg.* U3PNV) would generate an exact complementary copy of the degenerate region (including mis-matched bases) to which further degenerate primers may bind. Since the subset of degenerate primers able to mediate amplification is defined by the sequence to which binding within the degenerate region occurs, the spectrum of degenerate primers able to anneal to the copied sequence will have changed due to the incorporation of the mis-matched nucleotides. Thus, there would appear to exist the potential for an "evolution" of PCR products (driven by the annealing and extension of mis-matched degenerate primers over template sequences) that could be amplified by changing subsets of degenerate primers. Another likely scenario (also exacerbated by the occurrence of mis-matched priming events) is the occurrence of extensive primer-primer interactions (primer dimerisation) that would be expected when using degenerate primers. These effects, coupled with the extensive background annealing of DPs expected to occur throughout cellular chromosomal DNA, would likely result in the very early saturation of DP-PCR.

3.8.3 Conclusion

The discussion presented above highlights the highly complex and diverse nature of potential DP degenerate region interactions with both sequences adjacent to HIV proviral DNA and throughout the cellular chromosome. In this study, we have demonstrated the ability of semi-degenerate primers to mediate the amplification of specific HIV integrants in a reasonably selective fashion. However, due to largely unknown mechanisms, semi-degenerate primers were ultimately shown to be unable to mediate the amplification of integrated HIV DNA in a mixed population of integration events. Therefore, this procedure was discarded as a potential candidate assay for assessing the kinetics of HIV DNA integration following infection in cell culture and a larger emphasis was placed on the development of an alternative assay (Linker-Primer PCR or LP-PCR) for the specific detection of integrated HIV DNA (see Chapter 4).

Chapter 4

Development of a Quantitative HIV-1 Integrated Proviral Assay: Linker-Primer Polymerase Chain Reaction Method (LP-PCR)

4.1 Introduction

4.1.1 Background

Previous attempts to quantify integrated HIV DNA in a sample of human chromosomal DNA have been hampered by the presence of extrachromosomal HIV DNA. At the time this study was initiated, only extensive electrophoresis of sample DNA over a prolonged time period could remove most of the extrachromosomal forms (146, 238). However, even after electrophoretic purification of sample DNA, contaminating free DNA forms remained that could subsequently be detected, particularly when the highly sensitive PCR amplification was employed. Since then, two techniques have been published that specifically detect integrated HIV DNA *viz* the nested-*Alu* PCR (49) and the inverse PCR method (44) (see section 1.7). Here, the development of an alternative PCR-based assay termed the linker-primer PCR (LP-PCR) is presented.

4.1.2 The Principle of the Linker-Primer PCR Assay: Amplification of Integrated HIV DNA

The linker-primer PCR technique was inspired by a PCR-based assay used to characterise the clonal nature of populations of cells harbouring the retrovirus HTLV-1 among patients (257). This assay was extensively modified to selectively amplify integrated HIV DNA in a manner that was readily detected and quantifiable after gel analysis. Briefly, the assay can be divided into 4 major steps:

Digestion of the DNA sample(s) with the restriction enzyme *Nla* III to create cellular-5'HIV U3 junction sequences with unique overhanging ends

Ligation of oligonucleotide linkers to the ends of the DNA fragments created by *Nla* III digestion

Selective PCR amplification of the above DNA fragments mediated by primers designed to anneal at the linker sequence and within the HIV U3 region

Detection of PCR products via Southern hybridisation techniques

The basic principle is presented in figure 4.1. Briefly, addition of a linker (LPNV) to restriction enzyme (*Nla* III) cleaved chromosomal DNA upstream of the viral 5'-LTR region creates a template to which an oligonucleotide primer (LPNV) can anneal. When used in conjunction with a second primer designed to anneal within the viral 5'-U3 region (U3NV) in a subsequent PCR, amplification of the 5'HIV LTR region and upstream chromosomal DNA results. The PCR products are then detected by standard Southern hybridisation techniques.

Initially, target DNA samples are digested with the restriction endonuclease *Nla* III which cleaves regularly throughout the chromosome generating 4-base 3' overhangs (Fig. 4.2, parts 1. and 2.). The resulting 4-base overhangs are of sufficient length to allow efficient binding of the oligonucleotide linker LPNV that bears sequence complementary to the overhang sequence at its 3'-termini. In the presence of T4 DNA ligase, the linker ligates to the 5'-termini of the *Nla* III fragments (Fig. 4.2, part 3.) essentially creating additional sequence from which priming can occur in a subsequent PCR reaction (Fig. 4.2, part 4). Following template denaturation, the first cycle of the PCR involves extension of the HIV-specific primer (U3NV) from its annealing site over adjacent chromosomal DNA and the additional linker sequence (Fig. 4.2, part 4, broken blue line). In this way, a complementary copy of the linker primer is generated that can then act as a template from which LPNV can prime in subsequent cycles (Fig. 4.2, part 4, broken red line).

Since HIV is known to integrate at random with respect to sequence (and hence with respect to *Nla* III sites) in chromosomal DNA (see section 1.3.4), the products resulting from the LP-PCR procedure will be of various lengths. Digestion of the amplified products with the restriction enzyme *Bst* YI releases a common fragment of 129bp

Figure 4.1 LP-PCR Procedure

Diagrammatic representation of the LP-PCR procedure. Chromosomal DNA is initially digested with the restriction enzyme *Nla* III generating DNA fragments with cohesive termini to which linkers (LPNV) are ligated. A PCR is then performed using a primer specific for the HIV U3 region (U3NV) and the LPNV primer. Since HIV integration is random with respect to cellular sequence, PCR products are digested with the restriction enzyme *Bst* YI to generate a product of defined size that can be detected by Southern hybridisation using an internal probe.

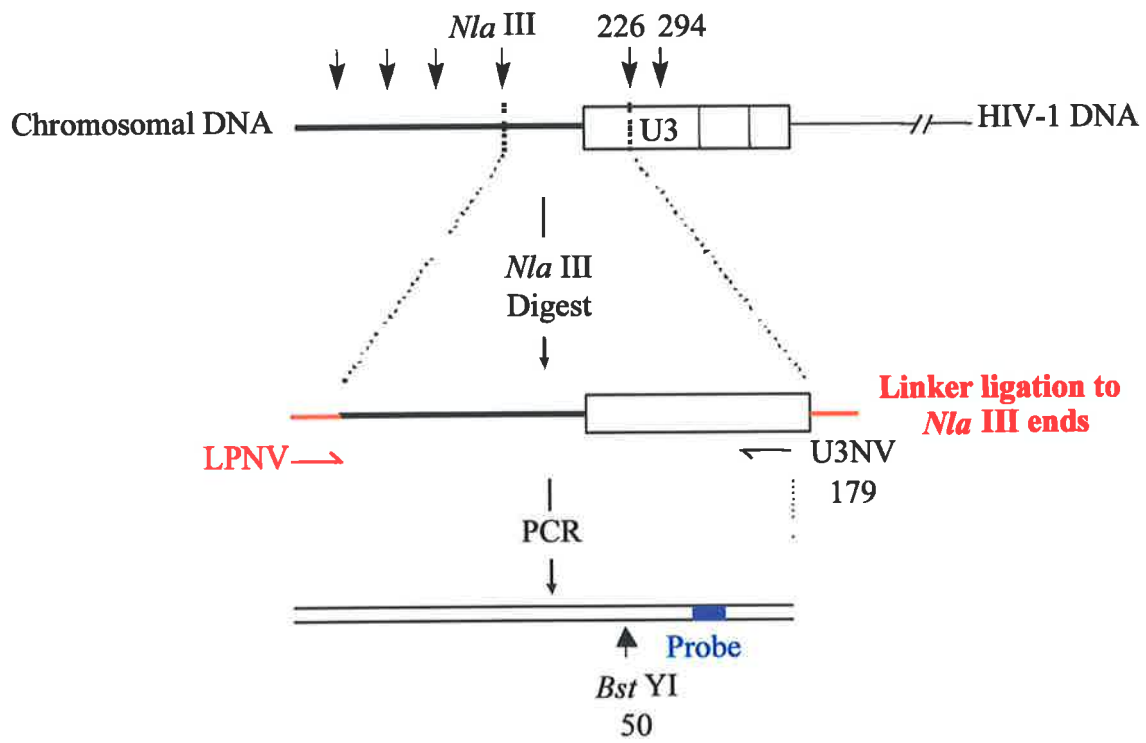
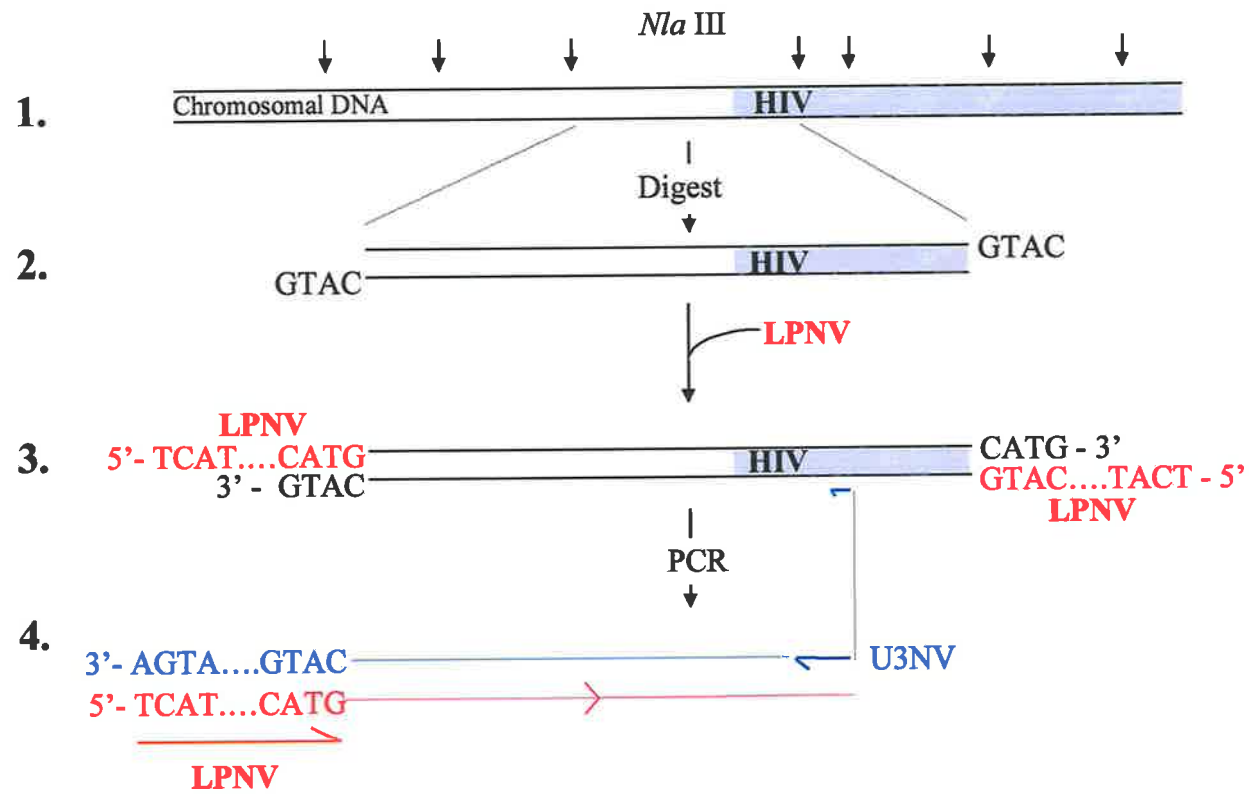


Figure 4.2 LPNV Amplification of Integrated HIV DNA

The oligonucleotide LPNV acts as both the linker and the primer in the subsequent PCR. The 5' and 3' markers represent the orientation of DNA strands by the presence of phosphate and hydroxyl groups, respectively. **1.** The average distance between *Nla* III sites in chromosomal DNA is 256bp. **2.** *Nla* III cleaves DNA to produce 4-base 3'-overhangs. **3.** Linker (LPNV)-ligation to *Nla* III-digested DNA. The 3'-end of the LPNV sequence is complementary to the 4-base 3'-overhang produced upon *Nla* III digestion of target DNA. **4.** Extension from the annealed U3NV primer in the first-round of PCR cycling over the ligated linker generates complementary sequence to which LPNV can anneal (and therefore act as a primer) in subsequent cycles.



within the HIV coded sequence which can then be detected and quantified by Southern analysis using an internal ^{32}P -labelled probe (see Fig. 4.1).

4.1.3 The Principle of the Linker-Primer PCR Assay: Selection Against Amplification of Extrachromosomal HIV DNA

Unless modified, the strategy outlined above for the detection of integrated HIV forms will also amplify the three extrachromosomal HIV forms (see section 1.4). Digestion of all extrachromosomal HIV forms with *Nla* III (the “assay” enzyme) releases a fragment spanning the PPT-3’U3 junction region between nucleotides 8917 and 9310 in the HIV sequence (see Fig. 4.3). In addition, digestion of the 2-LTR circular HIV form will release an additional fragment from nucleotides 9378 to 226 (see Fig. 4.3C.). Both of these fragments can be amplified using the LP-PCR protocol. To prevent this happening, the restriction endonuclease *Bgl* II (the “selection” enzyme). *Bgl* II cleaves DNA generating 5’-overhangs of 4 bases in length to which LPNV cannot ligate. Since the additional potential templates generated through *Nla* III digestion of the extrachromosomal forms contain internal *Bgl* II sites (see Fig. 4.3), internal cleavage of these fragments will prevent any subsequent linker-mediated PCR amplification.

Another potential template for amplification exists over the 5’-LTR region in the linear form should the linker be able to ligate to the terminal 5’-end. Since the terminal 5’-end of the HIV linear species exists in either a blunt or 3’-recessed form (see section 1.4), oligonucleotide linker ligation to this site is highly unlikely.

4.1.4 Relative Restriction Frequencies of *Nla* III and *Bgl* II

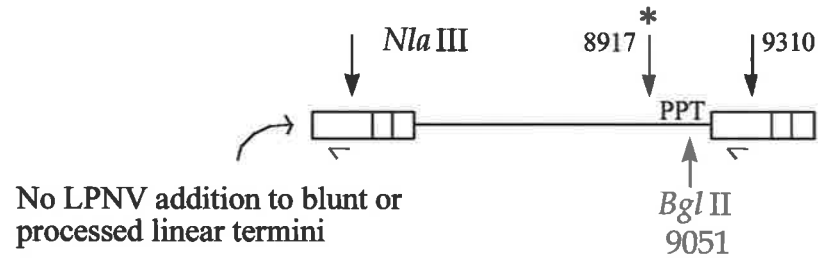
Since the selection enzyme acts by cleaving at a site within specific *Nla* III fragments in the extrachromosomal sequence thereby preventing amplification, care needed to be taken in the assay design to ensure this would not interfere with detection of integrated HIV DNA. Therefore, a restriction enzyme requiring a 4bp site was selected for the assay enzyme (*Nla* III) and a restriction enzyme recognising a 6bp site was chosen for the selection enzyme (*Bgl* II). Due to the smaller recognition site, *Nla* III will cut far more frequently (on average once per 256bp) than *Bgl* II (on average once per 4096bp)

Figure 4.3 Selection Against Amplification of Extrachromosomal HIV DNA

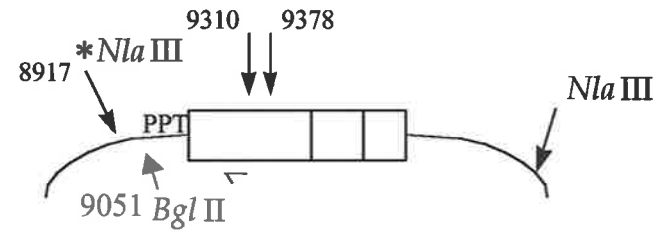
Bgl II digestion abolishes linker-ligation to HIV 3'-LTR (**A.**, **B.** and **C.**) and 2-LTR junction (**C.**) *Nla* III fragments containing the U3NV primer-annealing sequence.

Consequently, extrachromosomal HIV DNA is not amplified by LP-PCR. The scenarios in each of the three major extrachromosomal HIV DNA forms (**A.**, **B.** and **C.**) are shown.

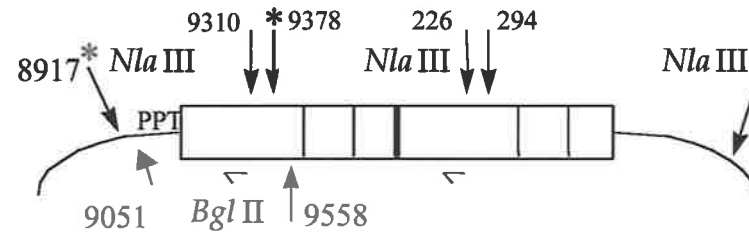
A. Linear



B. Single-LTR Circular



C. Double-LTR Circular

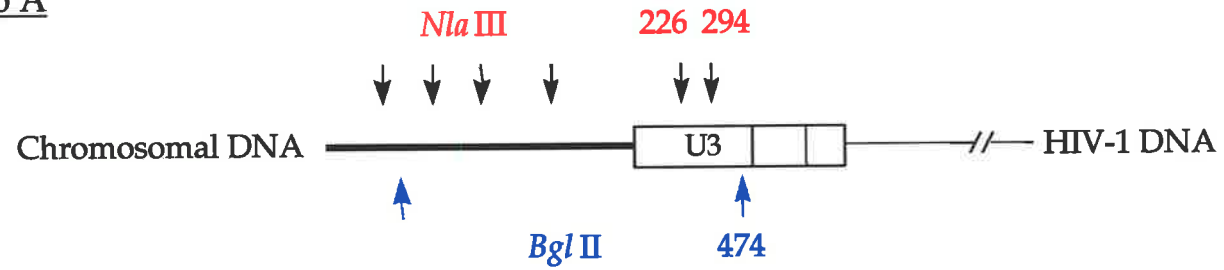


* - indicates where LPNV ligation would occur in the absence of *Bgl* II digestion
◁ - indicates the position of primer U3NV

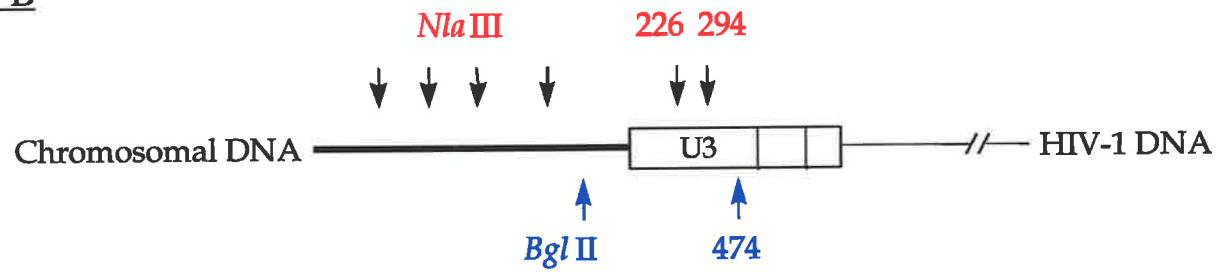
Figure 4.4 Relative Restriction Frequencies of Nla III and Bgl II

In order for successful LP-PCR amplification of integrated HIV DNA sequence to occur, an *Nla* III site is required to be present prior to a *Bgl* II site in the cellular DNA immediately upstream of the integrated provirus (Scenario A). However, if a *Bgl* II sites exists prior to an *Nla* III site in the upstream DNA (Scenario B), LP-PCR amplification of the integrated 5'-end of the HIV provirus will not occur. However, due to the relative sizes of the restriction enzyme recognition sequences for *Nla* III and *Bgl* II (4bp and 6bp, respectively), the chance that a *Bgl* II sites exists prior to an *Nla* III site in the cellular sequence immediately upstream of the integrated proviral DNA (Scenario B) is low (≈ 6 times per 100 integration events or 6%; see calculations section 4.1.4 in text). By implication, 94% of all HIV integration events can theoretically be detected using the LP-PCR protocol.

Scenario A



Scenario B



throughout the chromosomal DNA, ensuring that the majority of integrated HIV forms will be detected using this method (Fig. 4.4, Scenario A). The chance that *Bgl* II will cut before *Nla* III in the chromosomal sequence adjacent to the integrated HIV genome (Fig. 4.4, Scenario B) is calculated below:

$$\begin{aligned}
 \% \text{ Chance Scenario B} &= \frac{\text{Frequency of } Bgl \text{ II cleavage}}{\text{Frequency of } Nla \text{ III cleavage}} \times 100\% \\
 &= \frac{1/4096}{1/256} \times 100\% \\
 &= 6\%
 \end{aligned}$$

Thus, $(100-6)\% = \underline{\underline{94\% \text{ of integrated forms can be detected}}}$

4.2 Optimisation of Digestion Conditions

Since both the basic principle and the selectivity of the assay is dependent on complete restriction endonuclease digestion of target DNA, it was necessary to rigorously optimise the digestion conditions for both enzymes involved. In particular, optimisation and assessment of cleavage efficiency of chromosomal DNA with *Nla* III and complete digestion at the site positioned at nucleotide 9051 in the HIV sequence by *Bgl* II were deemed critical to the success of the assay. Since the assay was ultimately to be performed on chromosomal DNA preparations, optimisation of these reaction conditions were performed in the presence of background human lymphocyte (HuT-78) chromosomal DNA extracted and purified by the HIRT protocol (114) mixed with Qiagen preparations of the HIV-containing plasmid, pHXB2(kleen) (see Fig. 3.5A).

4.2.1 Optimisation of *Bgl* II Digestion

4.2.1.i Gel Analysis

Since *Bgl* II recognises a site 6 base-pairs in length, this enzyme was expected to cleave random chromosomal sequence on average once every 4096 base-pairs (see calculations section 4.1.5). Digestion of HuT-78 chromosomal DNA in either 1×NEBuffer3 or 2×OPA buffers produced a smear of DNA fragments from <1.5kb to >8.5kb in length (Fig. 4.5, panel A). A similar result was seen in the plasmid (pHXB2(kleen)) digestions (Fig. 4.5, panel B) in which little or no undigested plasmid DNA was observed with these buffers when compared to the undigested control. In contrast, digestions performed in 1×OPA (One-Phor-All PLUS; Pharmacia Biotech) or 1×ThermoPol buffer (NEB) did not proceed to completion (Fig. 4.5, panels A and B).

4.2.1.ii PCR Analysis

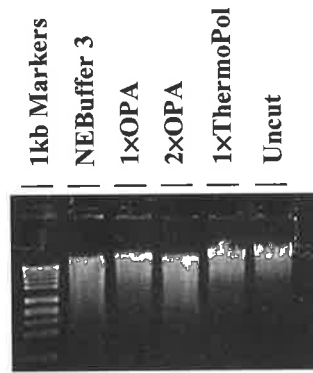
To further assess the extent of digestion of episomal HIV forms, particularly at nucleotide 9051 in the HIV sequence, PCR was performed to determine whether digestion with *Bgl* II could abolish amplification of DNA flanking this site. Initially, aliquots (1/1000th) of the digested plasmid samples (section 4.2.1.i) were subjected to 23 cycles of a PCR (standard conditions, see section 2.2.5) using primers U3PNV and 8918+ (Table 2.1). The absence of amplified product from samples in which *Bgl* II digestion was performed in either NEBuffer 3 or 2×OPA buffer confirmed that digestion neared completion in these samples (Fig. 4.6A).

Optimisation of *Bgl* II digestion in 2×OPA buffer was then further optimised by manipulating reaction volumes and enzyme concentrations. In contrast to the results presented above (see Fig. 4.6A), the increased PCR cycle-number (28 cycles) and the highly sensitive Southern detection protocol used to visualise DNA in this experiment, ensured that amplified product could be detected from all samples (except the water control) regardless of the digestion conditions used (Fig. 4.6B). This result was expected since restriction enzymes will rarely cleave to completion (100%) even under

Figure 4.5 Optimising Bgl II Digestion: Gel Analysis

Chromosomal digests (20µl) were performed at 37°C for 3hrs and then subject to electrophoresis through a 0.9% TAE agarose gel. Bands were stained with ethidium bromide and visualised under UV light. Plasmid digests (20µl) were performed at 37°C for 3hrs and then subject to electrophoresis through a 1% TAE agarose gel. Bands were stained with ethidium bromide and visualised under UV light. Uncut controls (right side) and DNA mass markers (1kb; left side) are marked. **A:** Digestion of 2µg of HIRT-purified HUT-78 chromosomal DNA (500ng loaded) with *Bgl* II in various buffers. **B:** Digestion of 2µg of pHXB2(kleen) plasmid DNA (500ng loaded) with *Bgl* II.

A.



B.

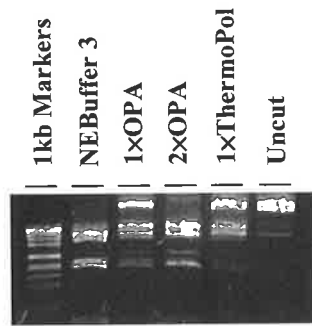
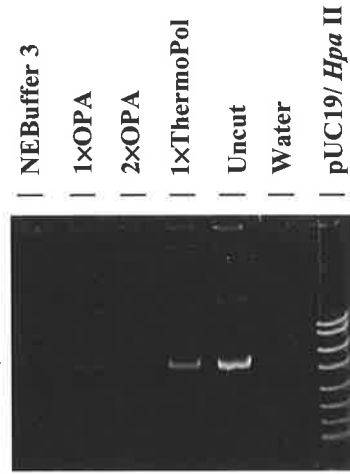


Figure 4.6 Optimising Bgl II Digestion: PCR Analysis

PCR evaluation of *Bgl* II digestion under different buffering conditions (**A.**) and different reaction-volumes/enzyme-concentrations (**B.**) using primers (U3PNV and 8918+) flanking the *Bgl* II restriction site at position 9051 in the HIV nucleotide sequence (accession number K03455). The sizes of expected bands (arrows, right side), DNA mass markers (pUC19/*Hpa* II) and water controls (Water) are indicated. **A.** Aliquots equivalent to approximately 2ng of pHXB2(kleen) plasmid digested in 2×OPA buffer with *Bgl* II (see section 4.2.1.i and Fig. 4.5 for reaction conditions) were subjected to a 23-cycle PCR (standard conditions, see section 2.2.5). **B.** Aliquots equivalent to approximately 2ng of pHXB2(kleen) plasmid digested in 2×OPA buffer with *Bgl* II were subjected to a 28-cycle PCR (standard conditions, see section 2.2.5). The reaction-volumes/enzyme-concentrations (μl/U) are indicated.

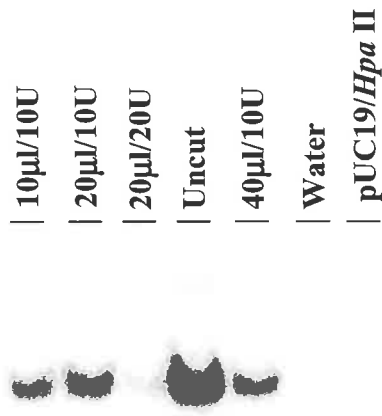
A.

322bp →



B.

322bp →



ideal conditions. However, *Bgl* II digestion in 2×OPA buffer was most efficient when performed using 20U of enzyme in a final volume of 20μl (see Fig. 4.6B).

4.2.2 Optimisation of *Nla* III Digestion

4.2.2.i Gel Analysis

The restriction endonuclease *Nla* III was employed to cleave chromosomal DNA containing the integrated provirus in order to provide cohesive termini to which a specifically designed oligonucleotide linker (LPNV) could be ligated. *Nla* III was chosen over other 4bp cutters as it gives a terminal 4bp 3' overhang after digestion (as opposed to shorter overhangs) thus favouring a more stable interaction with the linker primer.

Since the *Nla* III recognition site is 4 base-pairs in length, this enzyme was expected to cleave random chromosomal sequence on average once every 256 base-pairs (see calculations section 4.1.5). Following digestion of HuT-78 chromosomal DNA, the majority of resulting fragments were less than ≈400bp in length (Fig. 4.7A) indicating that digestion was nearing completion in each of the 6 buffers tested. This result was confirmed by showing that very little uncut pHXB2(kleen) construct DNA remained following digestion in each of the buffers (Fig. 4.7B).

4.2.2.ii PCR Analysis

To further assess the extent of chromosomal DNA digestion, a PCR was performed to determine whether digestion with *Nla* III could abolish amplification of a region within the β-globin gene (Fig. 4.8A). Digested chromosomal samples were subjected to 30 cycles of amplification using primers (β-glo1 and β-glo2) flanking an *Nla* III site present within the β-globin gene. Amplified product was clearly observed in lanes corresponding to *Nla* III digestion performed in the 1×OPA and NEBuffer 3 buffers (and the uncut control reaction) indicating that digestion was incomplete under these buffering conditions (Fig. 4.8B). In addition, faint signals were observed when *Nla* III

Figure 4.7 Optimisation of Nla III Digestion: Gel Analysis

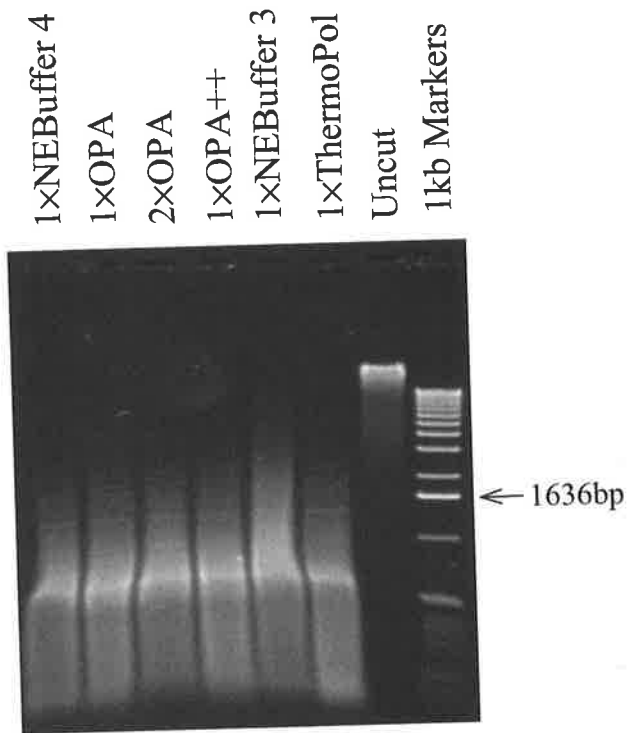
Chromosomal digests (20µl) were performed at 37°C for 3hrs and then subject to electrophoresis through a 2% TBE agarose gel. Bands were stained with ethidium bromide and visualised under UV light. Plasmid digests (20µl) were performed at 37°C for 3hrs and then subject to electrophoresis through an 8% polyacrylamide gel. Bands were stained with ethidium bromide and visualised under UV light. Uncut controls (Uncut), DNA mass markers (1kb and pUC19/*Hpa* II) and the sizes of selected bands (for reference) within DNA mass markers (arrows) are marked. **A:** digestion of 2µg of HIRT purified HUT-78 chromosomal DNA (500ng loaded) **B:** digestion of 2µg of pBSHXB2(cc) plasmid DNA (500ng loaded).

1×OPA++ Buffer = OPA (Pharmacia Biotech) with Tris-acetate, BSA and DTT to final concentrations of 20mM, 100µg/ml and 1mM, respectively.

All other buffers were obtained from NEB.

NB. Due to the short half life of this enzyme, digestion times were not extended beyond 3hrs.

A.



B.

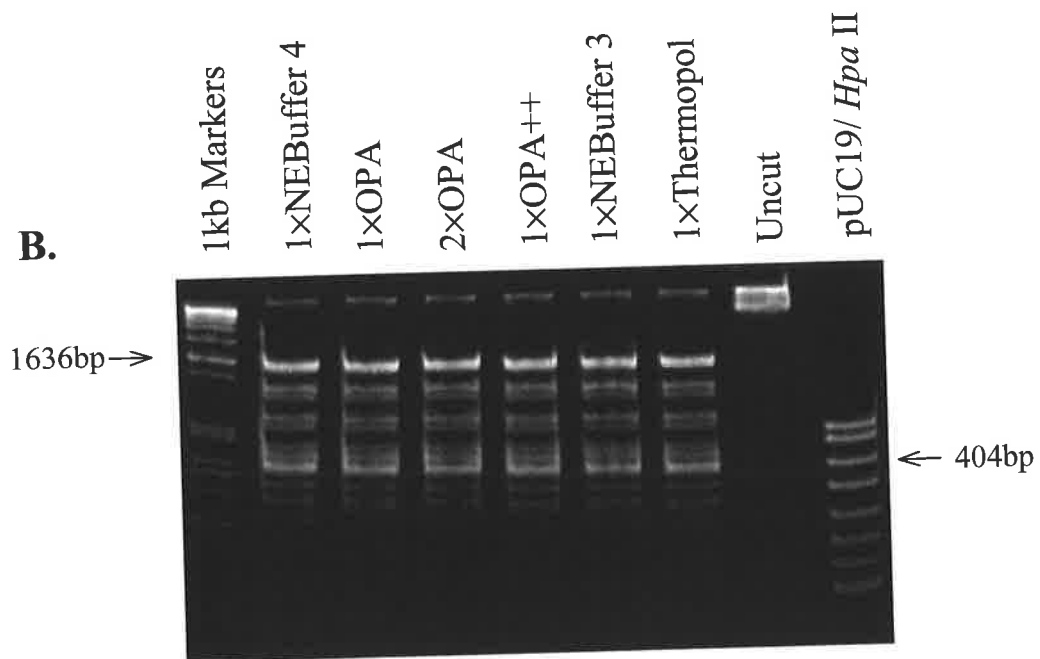


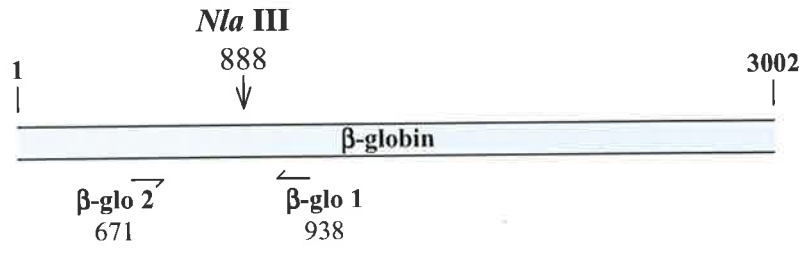
Figure 4.8 Optimisation of Nla III Digestion: PCR Analysis

PCR evaluation of *Nla* III digestion under different buffering conditions using primers (β -glo1 and β -glo2) flanking an *Nla* III restriction site within the human β -globin gene sequence (accession number L26462). The size of expected band (arrows, left side), DNA mass markers (pUC19/*Hpa* II) and water control (H₂O PCR Control) are indicated. **A.** Relative positions of the β -glo1 and β -glo2 primers within the human β -globin gene sequence. **B.** Aliquots equivalent to approximately 2ng of HuT-78 chromosomal DNA digested in various reaction buffers with *Bgl* II were subjected to a 30-cycle PCR (β -globin PCR conditions, see section 2.2.5.i).

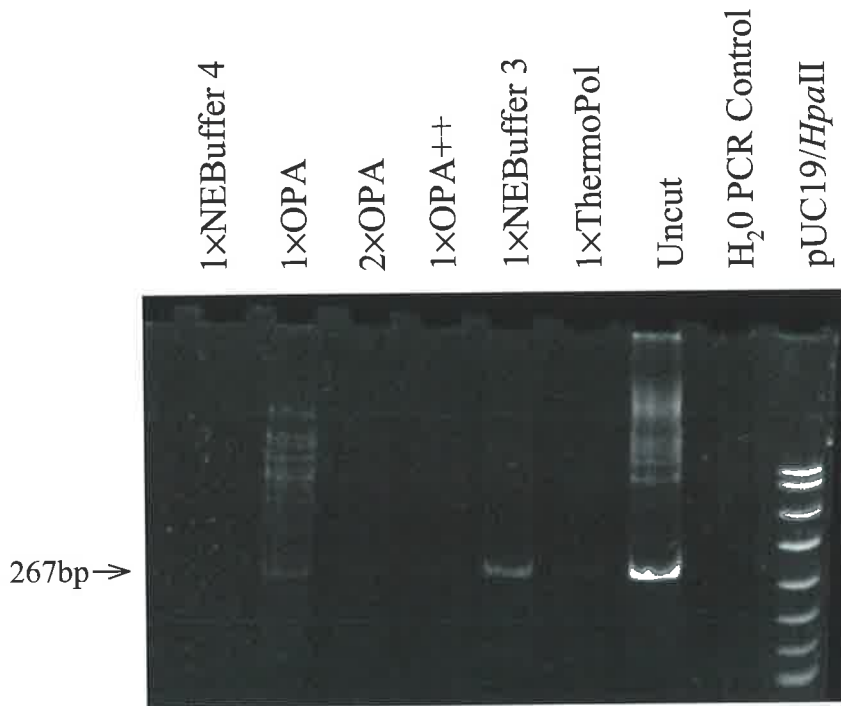
1×OPA++ = OPA (Pharmacia Biotech) with Tris-acetate, BSA and DTT to final concentrations of 20mM, 100 μ g/ml and 1mM, respectively

All other buffers were obtained from NEB.

A.



B.



digestion was performed in either the 2×OPA or 1×Thermopol buffer. In contrast, digestion appeared to near completion when either the NEBuffer 4 (the supplied buffer) or the 1×OPA++ (OPA with Tris-acetate, BSA and DTT to final concentrations of 20mM, 100µg/ml and 1mM, respectively) buffering conditions were used.

4.2.3 Final Digestion Conditions

In order to minimise assay variation resulting from multiple DNA handling steps, it was important to perform both digests without having to remove the buffer between digestions by phenol/chloroform/IAA or any other means. Although buffering conditions did not allow for a concurrent double digest, sequential enzymic digestion of DNA was achieved by the addition of supplementary solutions to the initial restriction enzyme reactions. Since *Nla* III digestions in 1×OPA++ were reproducibly comparable to digestions performed in the supplied buffer, after initial *Bgl* II digestion in 2×OPA, the conditions were manipulated by the addition of Tris-acetate, BSA and DTT (final concentrations of 20mM, 100µg/ml and 1mM, respectively) to generate conditions ideal for *Nla* III digestion (1×OPA++).

4.3 Proof of Principle

To establish whether PCR amplification of HIV DNA using the LP-PCR approach could occur, initial experiments were performed using known amounts (see section 2.2.4.iv for calculations) of the pHXB2(kleen) plasmid construct in the absence of background DNA. The digestion of sample DNA was performed as detailed in section 4.2.3 and subsequently phenol/chloroform/IAA extracted and ethanol precipitated (to remove the buffer prior to linker ligation) as described in section 2.2.9. Ligation reactions were performed using 20U of T4 DNA ligase (NEB) in the presence of a vast excess of the LPNV linker (50pmol). In order to minimise DNA loss associated with transferring aliquots between tubes, ligations were performed in 0.5ml PCR tubes and the entire reaction used in a subsequent PCR. PCR was performed in a final volume of 50µl using primers LPNV and U3NV under standard reaction conditions (see section 2.2.5). Reactions were cycled as follows: 30 cycles of 94°C 1min, 58°C 1min, 72°C

1min; and a final extension of 72°C 5min. Since HIV DNA was situated a defined distance away from a known upstream *Nla* III site in the pHXB2(kleen) plasmid construct, the PCR product generated by extension of primers LPNV and U3NV was of known size. Therefore, digestion with the restriction enzyme *Eco* RV or *Bst* YI was not performed prior to gel analysis.

The presence of bands (Fig. 4.9, heavy arrow) following Southern analysis of amplified products using the U3PNV end-labelled oligonucleotide probe confirmed that the LP-PCR assay was able to amplify HIV DNA within the pHXB2(kleen) construct in the absence of background DNA. Furthermore, amplification appeared to occur in a relatively quantitative manner. Although not displaying a particularly high degree of sensitivity (1000 copies of the pHXB2(kleen) construct), this experiment confirmed the principle of linker-primer-mediated amplification of HIV DNA within the context of flanking DNA.

Later experiments (see section 4.8) indicated that the second band (Fig. 4.9, thin arrow) probably resulted from amplification of sequence between position 8917 in the HIV_{HXB2} sequence and U3NV binding at position 9263 within 3'LTR (see Fig. 4.10B).

Amplification over this region generates a 367bp product and might be expected if there is incomplete digestion with *Bgl* II or if re-ligation of *Bgl* II sites during the linker-ligation step occurs. Methods addressing re-ligation of *Bgl* II sites are addressed in section 4.8.

4.4 Optimisation of Linker Ligation

The efficiency of LPNV ligation to the termini of *Nla* III fragments was expected to be critical to the sensitivity of LP-PCR. Therefore, optimisation of the ligation reaction was performed using the pHXB2(kleen) construct alone in the absence of background DNA.

In these experiments, two main issues were considered. Firstly, reaction conditions needed to be optimised for a highly efficient ligation reaction. Secondly, ligation of linkers to *Nla* III fragments rather than inter- and intra-ligation of *Nla* III fragments

Figure 4.9 Demonstrating the Principle of LP-PCR

LP-PCR performed on various dilutions of the pHXB2(kleen) construct. PCR conditions are given in the text (see section 4.3). The expected band is indicated with an heavy arrow. An additional band, likely resulting from linker-mediated amplification over the HIV 3'LTR/PPT region (see section 4.3 for details), is indicated by a thin arrow. Selected bands within the DNA mass ladder used (pUC19/*Hpa* II) are marked for reference (right side).

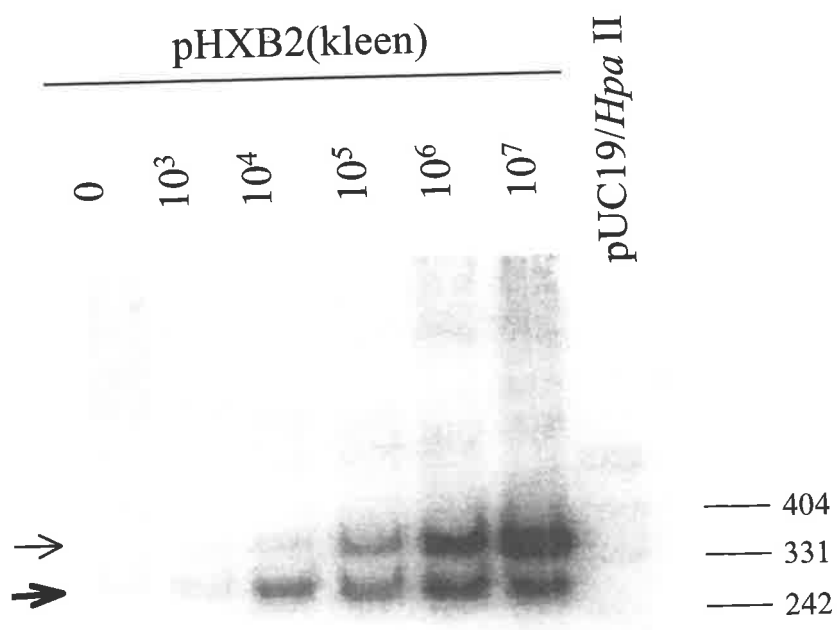
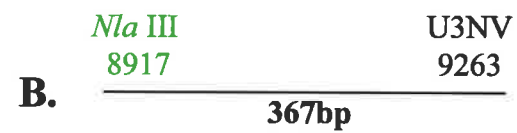
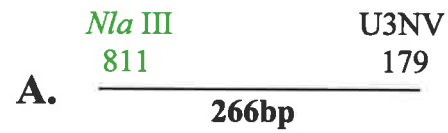
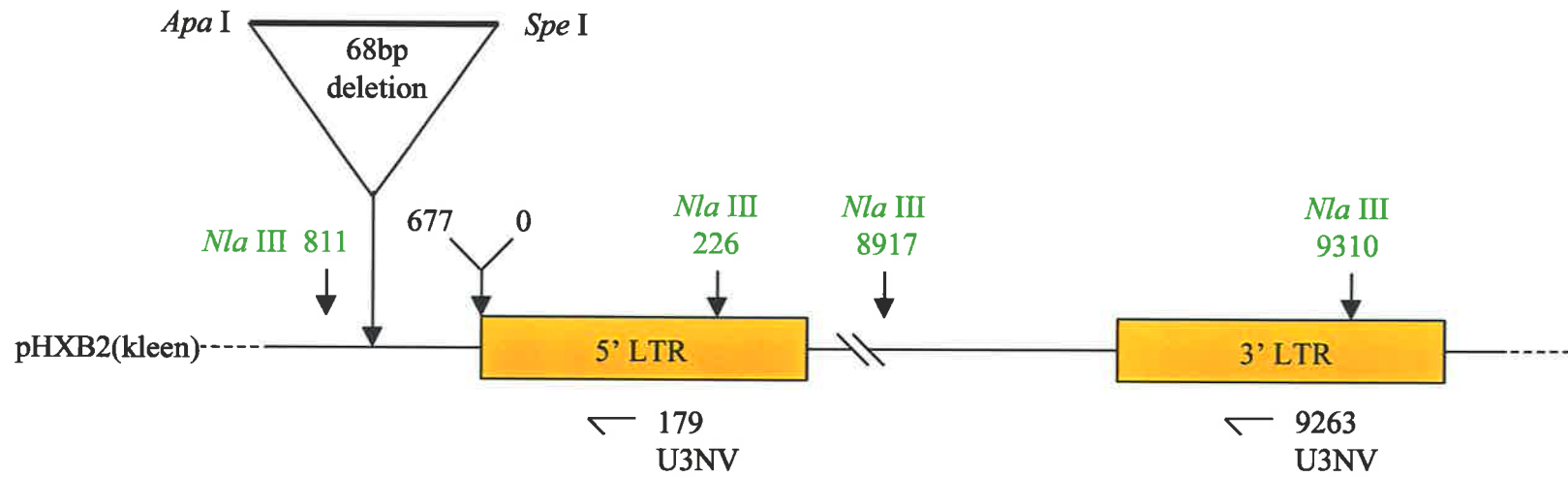


Figure 4.10 LP-PCR Amplification at the HIV 5' and 3' LTR Regions

A. The expected 266bp product resulting from LP-PCR amplification over the HIV 5'LTR and upstream (pBluescript) sequence in plasmid pHXB2(kleen). **B.** The expected 367bp product resulting from LP-PCR amplification over the HIV 3'LTR and upstream (internal HIV) sequence in plasmid pHXB2(kleen). A detailed map of the HIV genome and surrounding pBluescript plasmid sequence is shown. In addition, the positions of relevant *Nla* III restriction sites and the positions at which the U3NV primer was expected to anneal are shown.



needed to be favoured. To address the second issue, all ligation reactions were heated and then snap-cooled in the presence of a **vast excess** of LPNV in order to both maximise the chance of LPNV molecules interacting with *Nla* III sticky ends, and to out-compete the inter- and intra-molecular ligation events within and between *Nla* III fragments. The samples were initially heated in order to dissociate any *Nla* III fragments that may have already annealed via their complementary sticky 4bp termini.

The use of calf intestinal phosphatase (CIP) treatment of samples to inhibit circularisation (intra- molecular ligation) of *Nla* III fragments was considered. However, since linker addition occurs via ligation of the 3'-termini of LPNV to the 5'-termini of the *Nla* III fragment, the phosphate group on the *Nla* III fragment is required in order to complete the phosphodiester bond formation.

To establish whether different preparations of T4 DNA ligase displayed significant differences with respect to their ability to ligate LPNV to *Nla* III fragments, two preparations of T4 DNA ligase were compared as judged by LP-PCR amplification of pHXB2(kleen) (see Fig. 4.11 for the amplification and detection strategy used). Following *Nla* III and *Bgl* II digestion of plasmid samples, ligations were performed in the presence of either T4 DNA ligase (NEB) or T4 DNA ligase (USB) in a final volume of 20 μ l using equivalent amounts of LPNV and the recommended buffering conditions. Using this amplification protocol, no significant differences in ligation efficiency were observed (Fig. 4.12A). However, due to the slightly increased overall signal in the 10³ copy number sample, the ligase obtained from New England Biolabs (NEB) was used in all subsequent experiments.

In order to further optimise ligation efficiencies, the effect ligation volume and ligase concentration on the reaction was investigated. The results demonstrated that there was little difference in ligation efficiency associated with using either 400U or 800U of T4 DNA ligase (Fig. 4.12B). In addition, there was no appreciable difference in the ligation efficiencies observed when a 2-fold increase in the final reaction volume was used. Therefore, most subsequent experiments (see sections 4.9 onwards) involved ligation reactions performed in a final volume of 40 μ l using 400U (NEB Units) of T4 DNA

Figure 4.11 Incorporating Bgl II into the LP-PCR Procedure

Diagrammatic representation of the LP-PCR procedure. Chromosomal DNA is initially digested with the restriction enzymes *Nla* III and *Bgl* II. Linkers (LPNV) are ligated to only those DNA ends with the *Nla* III-digested phenotype (4-base, 3'-overhang). A PCR is then performed using a primer specific for the HIV U3 region (U3NV) and the LPNV primer. Since HIV integration is random with respect to cellular sequence, PCR products are digested with the restriction enzyme *Bst* YI to generate a product of defined size that can be detected by Southern hybridisation using an internal probe.

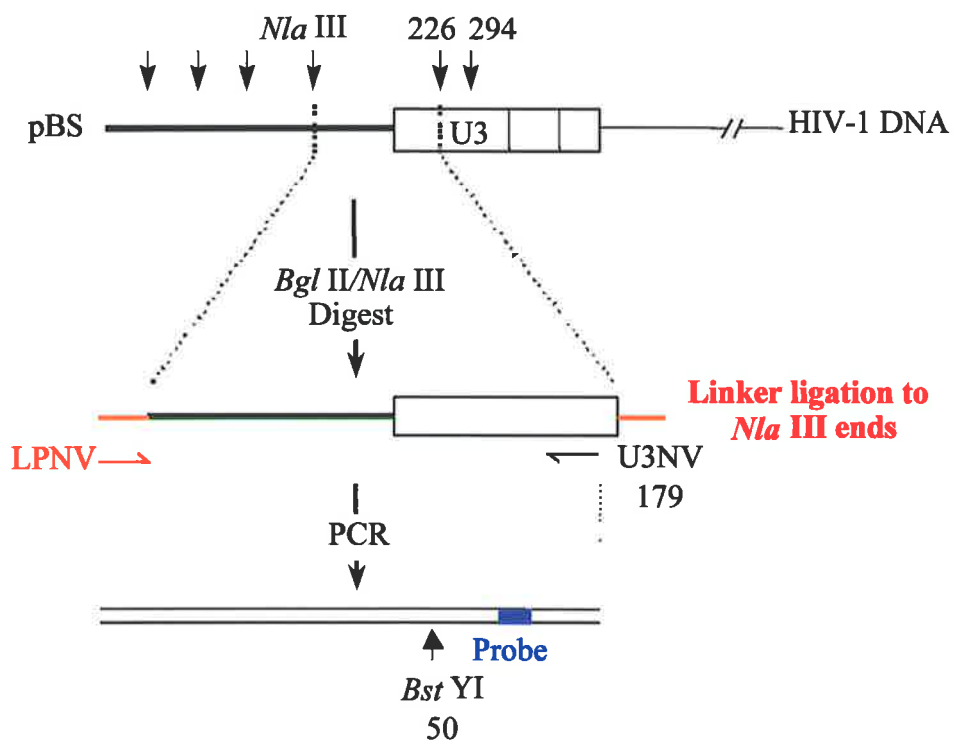
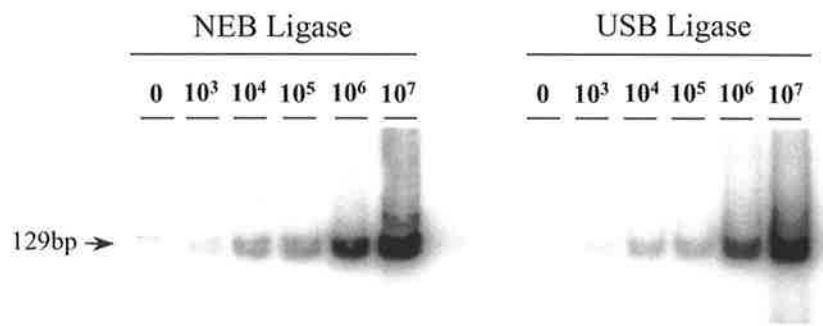


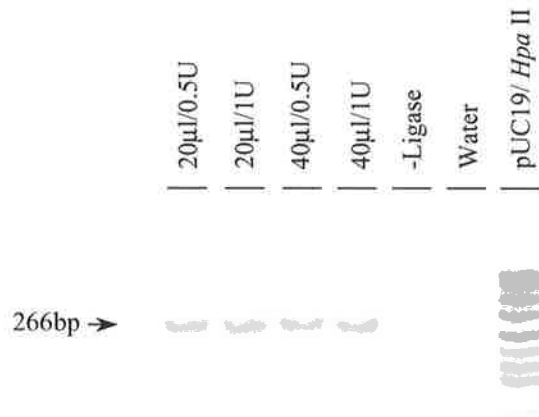
Figure 4.12 Optimisation of Linker Ligation Conditions

A. Relative efficiency of linker ligation using two different T4 DNA ligase enzymes as judged by LP-PCR amplification of various dilutions of pHXB2(kleen). The identity of each T4 DNA ligase used (NEB or USB) and the size of the expected band is indicated (left side). Ligations were performed under conditions recommended by the suppliers in a final volume of 20µl in the presence of 50pmol of LPNV. LP-PCR was performed as described in section 4.3 and products (digested with *Bst* YI) detected using the U3PNV oligonucleotide probe (see Table 2.1). **B.** Assessing the effect of reaction-volume and enzyme-concentration on the efficiency of linker ligation as judged by LP-PCR amplification of 10⁵ copies of pHXB2(kleen). Ligations were performed using NEB T4 DNA ligase. The reaction-volumes/enzyme-concentrations used in each reaction are indicated (µl/U). Control reactions performed in the absence of T4 DNA ligase (-Ligase) or pHXB2(kleen) (Water) are shown. LP-PCR and detection of amplified products (not digested with *Bst* YI) were performed as in **A.**

A.



B.



ligase. However, earlier experiments (see sections 4.4-4.8), in which the final PCR volumes were less than 100 μ l, involved ligation reactions performed in 20 μ l.

4.5 Detection of HIV DNA in the Presence of Background Human Lymphocyte DNA

4.5.1 Standard PCR Technique

In the absence of background chromosomal DNA, amplification of the pHXB2(kleen) construct achieved a sensitivity of detection of approximately 1000 copies (see Fig. 4.9). To assess the impact of background chromosomal DNA on this sensitivity, varying amounts of pHXB2(kleen) were spiked into 3 μ g ($\approx 5 \times 10^5$ cell-equivalents) of HuT-78 chromosomal DNA and a 30-cycle LP-PCR performed (as described in section 4.3). PCR products were analysed directly (no *Bst* YI digestion) by gel electrophoresis and hybridisation with the U3PNV probe (Table 2.2). The results indicated that the addition of the background DNA abolished signal entirely (Fig. 4.13A).

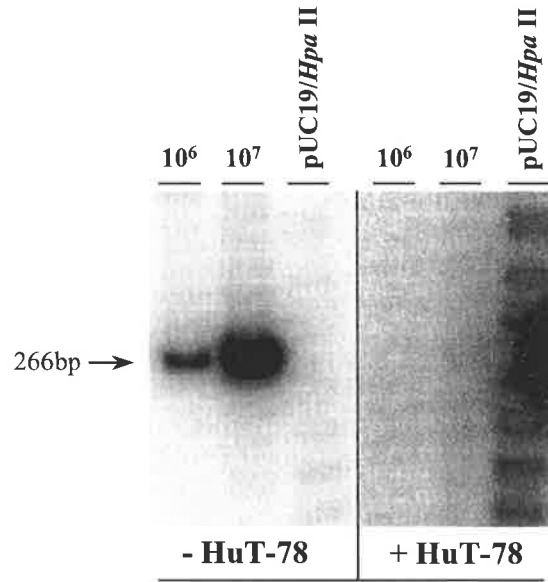
To further examine the effect of background DNA on LP-PCR-mediated amplification of integrated HIV-1 DNA, amplification of HIV DNA within H3B chromosomal preparations was attempted. Following LP-PCR-mediated amplification of 1.5 μ g ($\approx 2.5 \times 10^5$ cell-equivalents) of H3B chromosomal DNA (2 integrated HIV copies per cell), very little signal was observed (Fig. 4.13B.i. and ii.). The expected signal was a level slightly higher than that seen for amplification of 1×10^5 copies of the pHXB2(kleen) construct. However, the actual signal corresponded to a level of amplification equivalent to below 1×10^4 copies of the extrachromosomal construct.

It was considered likely that increased levels of non-specific priming events associated with the presence of background chromosomal DNA may have contributed to a reduction in assay-sensitivity. Therefore, approaches to minimise the amount of background (non-specific) priming throughout cellular chromosomal DNA were investigated.

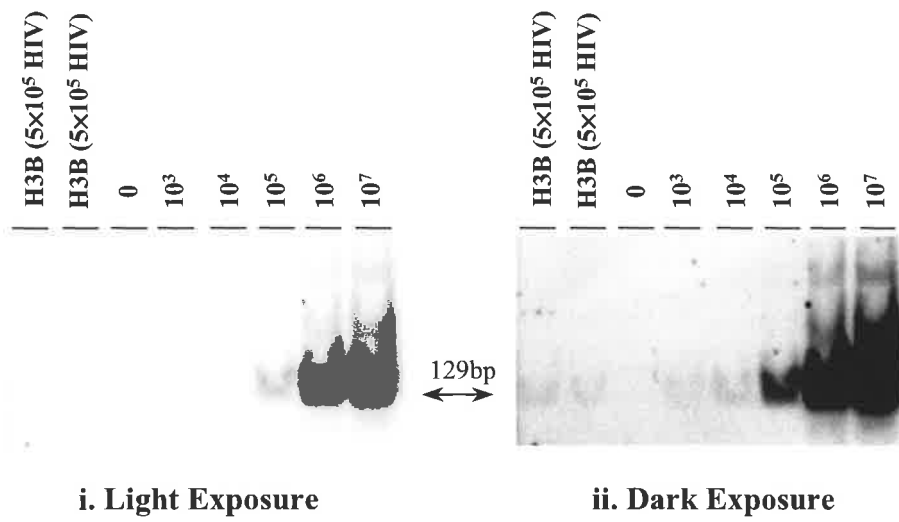
Figure 4.13 The Effect of Background Chromosomal DNA on LP-PCR

A. LP-PCR amplification of pHXB2(kleen) in the presence (+ HuT-78; right panel) or absence (- HuT-78; left panel) of 3 μ g of HuT-78 chromosomal DNA. DNA mass ladders (pUC19/*Hpa* II) and the expected size of the PCR product (not digested with *Bst* YI) are shown. **B.** LP-PCR amplification of 2.5 $\times 10^5$ cell-equivalents (5 $\times 10^5$ integrated HIV copies) of H3B chromosomal DNA. Various dilutions of the pHXB2(kleen) construct were amplified (in the absence of background HuT-78 DNA). Amplified products were detected by Southern analysis following digestion with *Bst* YI using the U3PNV probe. A 2h (i. Light Exposure) and an overnight (ii. Dark Exposure) exposure of the gel is shown.

A.



B.



4.5.2 Hot Start PCR Technique

The hot-start PCR procedure was employed to reduce non-specific priming and extension events occurring at sub-annealing temperatures within samples prepared for PCR in the time immediately prior to the initial denaturation step. Such priming and extension events can potentially occur at room temperature during sample transport to the thermocycler or over the temperature ramping period as the thermocycler heats to its initial denaturation temperature. Depending on the extent of such early non-specific priming and extension events, the templates available for amplification over the ensuing cycles may be dramatically increased resulting in elevated non-specific amplification. This can in turn significantly influence the efficiency with which the desired template is amplified.

In order to investigate whether a hot-start PCR protocol might enhance the efficiency of the integrated proviral assay procedure, the LP-PCR procedure was performed using *rTth Taq* polymerase in conjunction with wax beads (PCR gemsTM). The conditions in which PCRs were performed were altered from that described above (see section 4.3) to comply with the manufacturers' recommendations. PCRs (50 μ l) were performed using 50pmol each of the primers LPNV and U3NV (see Table 2.1) in 1 \times PCR Buffer II (Perkin-Elmer), 1.2mM Magnesium acetate (Perkin-Elmer) and 1.6U of rTth DNA Polymerase. Reactions were cycled as follows: 30 cycles of 94 $^{\circ}$ C 1min, 58 $^{\circ}$ C 1min, 72 $^{\circ}$ C 1min; and a final extension of 72 $^{\circ}$ C 5min. The use of wax beads allowed the segregation of primers and dNTPs from both the *rTth* DNA polymerase and target DNA by a wax layer that melts upon heating to temperatures in excess of 70 $^{\circ}$ C. Consequently, when temperatures during the initial denaturation step of the PCR exceeded 70 $^{\circ}$ C, reagents were free to mix allowing the PCR to proceed as normal.

A strong signal was obtained with 1 $\times 10^5$ copies of pHXB2(kleen) in the absence of HuT-78 background DNA, confirming that the hot-start PCR conditions used were able to mediate the amplification of HIV DNA (Fig. 4.14). However, signal was abolished when amplification was performed in the presence of 1 $\times 10^5$ c.e. of HuT-78 DNA. Furthermore, no signals were detected when the assay was performed on 1 $\times 10^5$ c.e. of ACH-2 chromosomal DNA (1 integrated HIV copy/cell). Taken together, these results

Figure 4.14 Hot-Start (PCR GemsTM) LP-PCR

Hot-start LP-PCR was performed using *rTth* DNA polymerase and PCR gemsTM on various amounts of pHXB2(kleen) (indicated on figure as **kleen**) in the presence (+ HuT-78) or absence (- HuT-78) of 1×10^5 cell equivalents (c.e.) of background HuT-78 chromosomal DNA. Amplification of 10^5 c.e. of ACH-2 chromosomal DNA (1 integrated HIV copy/cell) and HuT-78 alone are also shown. For a detailed explanation of the reaction conditions, see the text (section 4.5.2). Amplified products were detected (in the absence of *Bst* YI digestion) by Southern analysis using the U3PNV oligonucleotide probe (Table 2.1). The size of the expected band (266bp) and the DNA mass ladder used (pUC19/*Hpa* II) is shown.

266bp →



- | 10³ kleen - HuT-78
- | 10⁵ kleen - HuT-78
- | 10³ kleen + HuT-78
- | 10⁵ kleen + HuT-78
- | 10⁵ c.e. ACH-2
- | HuT-78
- | pUC19/*Hpa* II



indicated that this hot-start protocol did not significantly increase the sensitivity of the LP-PCR procedure and suggested that an alternative element (other than non-specific primer-annealing) was primarily responsible for the severe decrease in assay-sensitivity observed when background DNA was present.

4.6 Overcoming Background *Nla* III-fragment Amplification

4.6.1 Background *Nla* III-fragment Amplification

The loss of assay-sensitivity in the presence of background DNA was ultimately attributed to linker-mediated amplification of background *Nla* III fragments leading to early saturation of the PCR. Amplification of such fragments may occur if the recessed 3'-ends of DNA fragments containing a ligated linker are extended over the remaining 21 bases of the LPNV sequence (forming blunt-ends) prior to the initiation of the PCR cycling procedure (see Fig. 4.15). Such events might be expected to occur within reaction mixes at room temperature or during the ramping period as the thermocycler heats to its initial denaturation temperature and generate multiple templates (approximately 1×10^{13} per $3 \mu\text{g}$ of chromosomal DNA; see calculations below) for LPNV-mediated PCR amplification. In a situation where multiple templates for PCR exist, PCR saturation would be expected to occur well before the allocated 30 cycles due to competition and depletion of reagents such as primers, nucleotides and the thermostable DNA polymerase.

Calculating the number of potentially amplifiable *Nla* III fragments:

Assume:

5 600 000 000 bp per diploid genome

Nla III cuts once every 256bp \rightarrow 21875000 fragments/genome

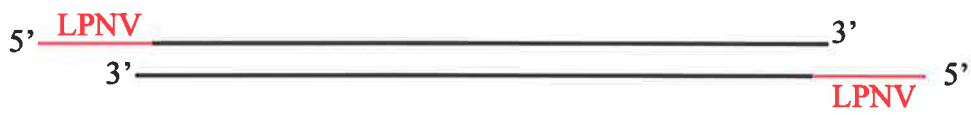
Bgl II cuts once every 4096bp \rightarrow 1367188 fragments/genome

Therefore amplifiable fragments = 21875000 - 1367188
 = 2×10^7 fragments/genome

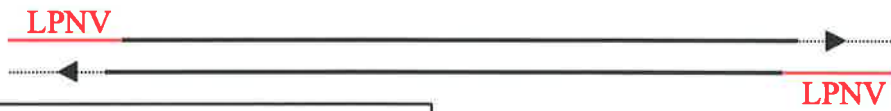
Therefore in $\approx 5 \times 10^5$ genomes there are $2 \times 10^7 \times 5 \times 10^5$
 = $\approx 1 \times 10^{13}$ fragments with *Nla* III termini

Figure 4.15 LPNV-mediated Amplification of Nla III Fragments

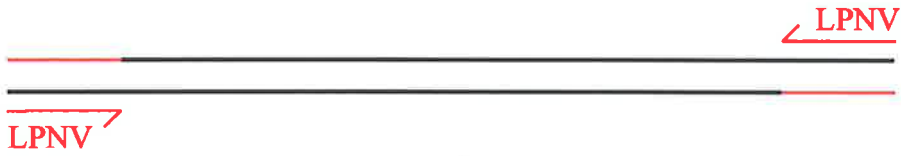
Proposed mechanism allowing LPNV-mediated amplification of linker-ligated *Nla* III fragments. Extension of 3' DNA ends (dashed black line with arrowhead) by *Taq* polymerase using the ligated linker (shown in red) as a template may occur immediately following addition of PCR reagents to samples. The resulting sequence is able to mediate binding of the LPNV primer in subsequent cycles of PCR.



↓
Addition of
PCR reagents



Taq-mediated extension over
remaining 21bp of LPNV sequence at
RT or during PCR ramp to 94°C



↓
LPNV-mediated amplification of *Nla* III fragment

Therefore, to address the issue of background *Nla* III-fragment amplification, a number of approaches were investigated to either increase the proportion of the desired template or to abolish the amplification of background *Nla* III fragments.

4.6.2 Incorporation of a Asymmetric or "Linear" PCR

To address the problem associated with the relative small amount of desired template compared to background templates, the use of asymmetric PCR was investigated. Asymmetric PCR using one primer only (U3NV) was performed to substantially increase the HIV-specific templates available for subsequent LP-PCR mediated amplification. In this way, a far higher proportion of the desired *Nla* III fragment containing the U3 region of HIV compared to the potential background *Nla* III fragments was expected to be generated. The asymmetric PCR process is outlined in figure 4.16. This reaction is essentially identical to a standard PCR procedure except it requires the use of a single primer and results in amplification of the target sequence at a linear rate rather than the exponential rate observed when a primer pair is used.

4.6.2.i Incorporation of a Standard Asymmetric PCR

Initial experiments employing the use of an asymmetric PCR (performed after ligation of linkers (LPNV) to digested DNA and prior to the LP-PCR) involved cycling for 99 cycles using AmpliTaq GoldTM (Perkin-Elmer) under standard reaction conditions (see section 2.2.5) in a final volume of 50 μ l. AmpliTaq GoldTM was used as it has a substantially longer half-life than other commercially available thermostable DNA polymerases and would better withstand the long cycling procedure. A 30 μ l aliquot of the asymmetric PCR product was then used as a template for further LP-PCR-mediated amplification (see section 4.3) using primers U3NV and LPNV in a final volume of 50 μ l.

Following amplification of serial dilutions of pHXB2(kleen) in the presence of 1×10^5 c.e. (600ng) of HuT-78 background DNA, strong signals were observed in lanes corresponding to amplification of both 10^7 and 10^6 copies of the integrated control

Figure 4.16 Asymmetric PCR Amplification of Target Sequences

Diagrammatic representation of the asymmetric PCR protocol used to enrich for the desired *Nla* III fragments containing the 5'U3 HIV sequence. Linkers (LPNV) are shown in red while HIV sequence is shown in blue.



x 99
cycles



LP-PCR on newly generated
single-stranded template

construct (Fig. 4.17A). In addition, signals were visible in lanes corresponding to amplification of 10^5 copies of H3B DNA (2×10^5 copies of integrated HIV proviral DNA) amplified in a similar manner (data not shown). Subsequent experiments varying the concentration of $MgCl_2$ (over the 0-3mM range) within the asymmetric PCR reaction mix did not affect the level of sensitivity attained (data not shown). Therefore, the incorporation of an asymmetric PCR prior to performing LP-PCR appeared to increase the assay sensitivity by between 10 and 100-fold. However, the sensitivity level achieved remained well below that required for future applications (see chapters 4 and 5).

4.6.2.ii Incorporation of a Hot-Start Asymmetric PCR

As described in section 4.5.2, the hot-start PCR procedure is employed to reduce the amount of non-specific annealing and extension events that can potentially occur immediately prior to PCR cycling. To increase the efficiency of the asymmetric PCR, attempts were made to amplify various amounts of the pHXB2(kleen) construct either in the presence or absence of HuT-78 DNA using a hot-start protocol. The asymmetric PCR (using primer U3NV) was performed in a final volume of 50 μ l using the *rTth* DNA polymerase enzyme and the commercially recommended buffering conditions. A 10 μ l aliquot of amplified product was then used in a subsequent LP-PCR (see section 4.3). Following Southern analysis of amplified products, signals were obtained for 8000 copies of the pHXB2(kleen) construct amplified in the absence of background DNA (Fig. 4.17B). However, the addition of background DNA completely abolished this signal. These results indicated that the use of this hot-start protocol (under the commercially recommended conditions) did not direct the efficient asymmetric amplification of HIV DNA when background DNA was present.

4.6.2.iii Ethanol Precipitation of the Asymmetric PCR Products

To assess whether factors associated with the large sample volume taken into the LP-PCR might be affecting the efficiency of amplification, asymmetric PCR products (performed as described in section 4.6.2.i) were subjected to phenol/chloroform/IAA extraction and ethanol precipitation prior to LP-PCR. Following re-suspension of pellets

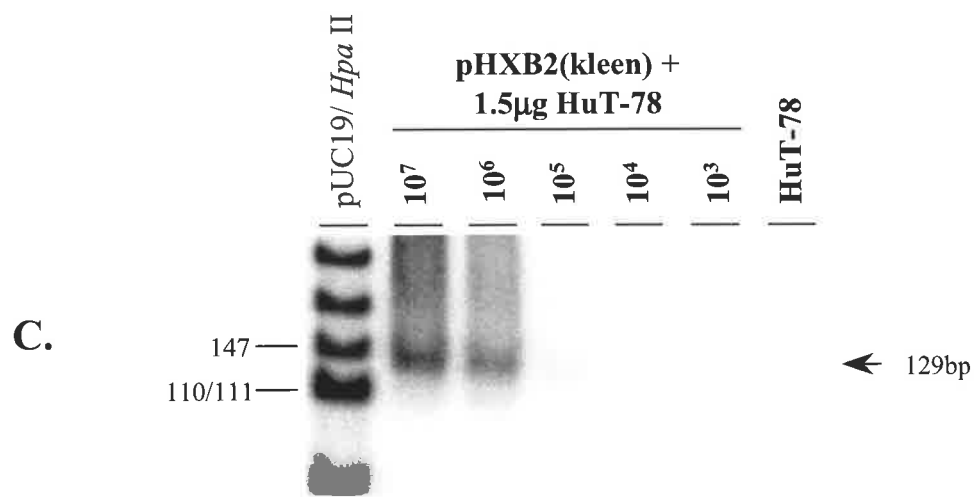
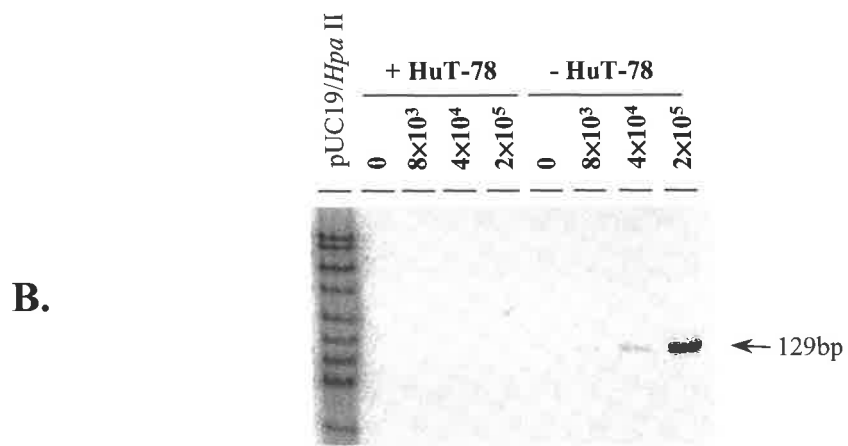
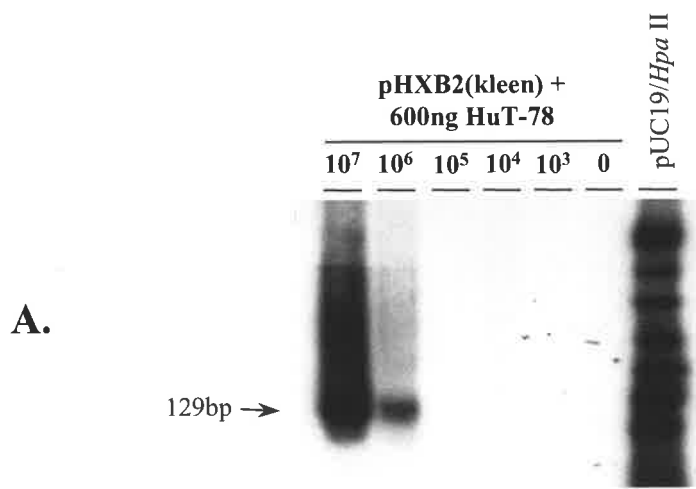
Figure 4.17 Asymmetric PCR Amplification: Effect on LP-PCR

Sensitivity

The effect of various asymmetric PCR protocols on the sensitivity of the LP-PCR procedure. PCR products were detected following digestion with *Bst* YI by Southern hybridisation using the U3PNV probe. The position of the expected band (129bp) is indicated in each case. DNA mass markers (pUC19/*Hpa* II) are shown. **A.** Asymmetric PCR performed on various copies of pHXB2(kleen) using AmpliTaq Gold™ DNA polymerase. An aliquot of the resulting reaction mix was then subjected to LP-PCR. For a detailed explanation of the reaction conditions used, see section 4.6.2.i. **B.**

Asymmetric PCR performed on various copies of pHXB2(kleen) in the presence (+ HuT-78) or absence (- HuT-78) of 600ng HuT-78 chromosomal DNA using *rTth* DNA polymerase. An aliquot of the resulting reaction mix was then subjected to LP-PCR. For a detailed explanation of the reaction conditions used, see section 4.6.2.ii. **C.**

Phenol/chloroform/IAA extraction and ethanol precipitation of asymmetric PCR products prior to performing LP-PCR. Asymmetric PCR was performed as described in **A.** For a detailed explanation of the reaction conditions used, see section 4.6.2.iii.



in water, the entire volume was used in a subsequent LP-PCR. Following *Bst* *YI* digestion, LP-PCR amplification of the ethanol precipitated DNA allowed detection of 1×10^4 copies of pHXB2(kleen) in a background of 2.5×10^5 cell-equivalents (1.5 μ g) of HuT-78 chromosomal DNA (Fig. 4.17C). Although representing a significant increase in the assay-sensitivity (compare with Fig. 4.17A), this level of sensitivity was still below that required for future applications.

4.6.2.iv Removal of LPNV

To address the possibility that the excessive amounts of LPNV present in the ligation reaction were contributing to premature saturation of the asymmetric PCR, steps were taken to remove residual linker from the ligated samples. Since the presence of LPNV alone was proposed to be capable of background *Nla* III fragment amplification, its presence in the asymmetric PCR mix could potentially allow both U3NV-LPNV (HIV-specific) and LPNV-LPNV (non-specific) amplification to occur. If occurring, non-specific amplification was proposed to have a profound effect on the efficiency with which the asymmetric PCR could direct the desired linear amplification of the target template.

In order to remove residual linker from linker-ligated samples prior to the asymmetric PCR, the ligation products were subjected to spin-column purification (both CL6B and qiaquickTM spin columns). Each of these techniques removes DNA fragments of less than 50 bases in length from sample preparations. However, the use of these techniques was shown to result in a highly variable and reduced yield of the ligation product (data not shown). In particular, the sample loss associated with the CL6B column was estimated to be approximately 50% (data not shown).

4.6.2.v Asymmetric PCR: Conclusions

Although these results show that an asymmetric PCR step prior to subsequent LP-PCR amplification was able to increase the assay sensitivity when background HuT-78 chromosomal DNA was present, the sensitivity levels achieved were less than that required for either the analysis of HIV integration kinetics over time or the evaluation of

drugs inhibiting the integration process in cell culture. Therefore, methods to further increase the levels of successful amplification of the targeted HIV sequence, and reduce background amplification of *Nla* III fragments, were investigated.

4.6.3 DynabeadTM selection of the HIV LTR DNA fragments

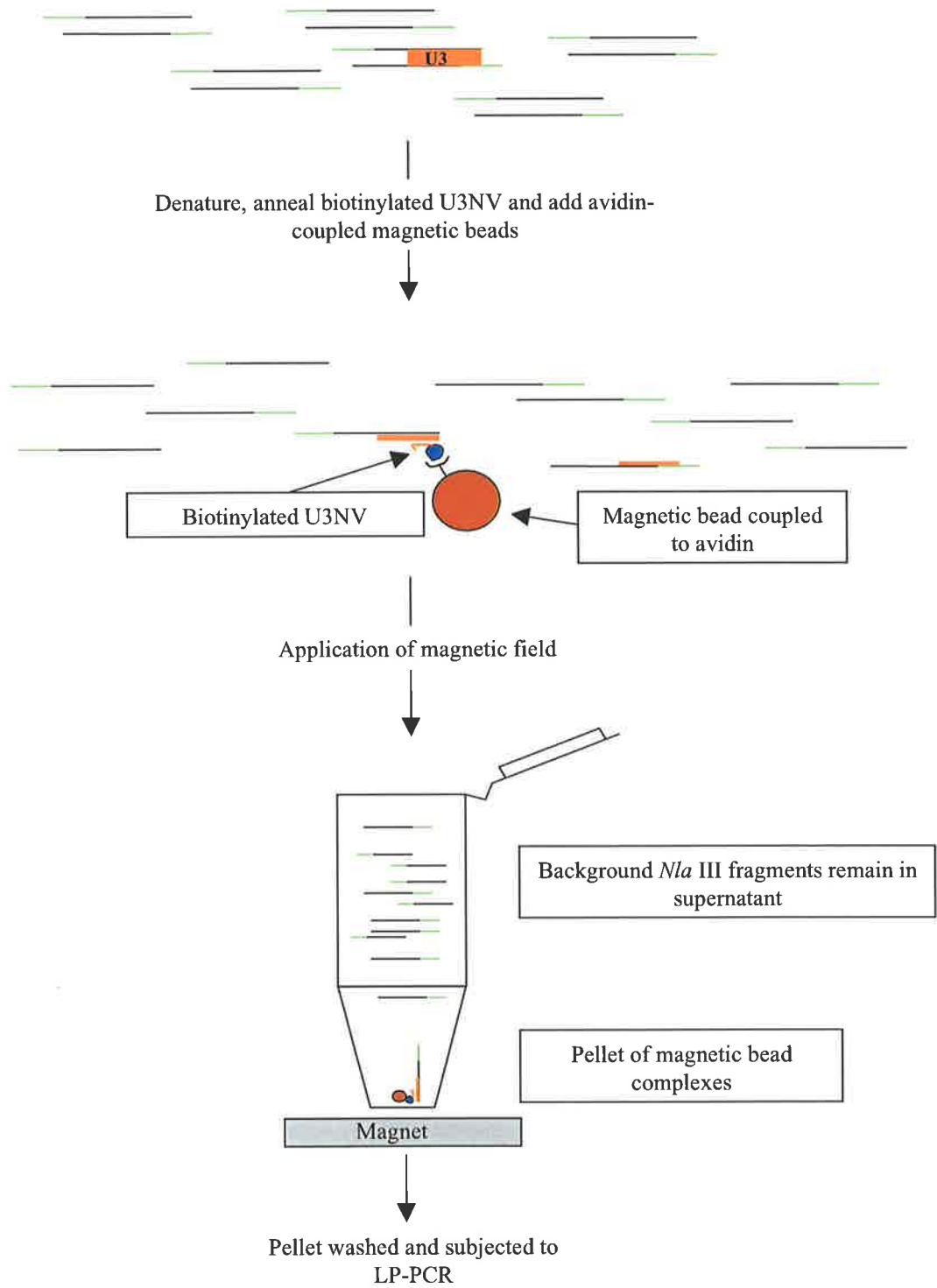
4.6.3.i The DynabeadTM Selection Principle

Since the incorporation of an asymmetric PCR was, on its own, unable to increase the sensitivity of the LP-PCR procedure to the levels required for future applications, an alternative technique involving the physical separation and purification of HIV-specific *Nla* III fragments from the non-specific fragments using DynabeadsTM was investigated. The removal of non-HIV-specific *Nla* III fragments from the linker-ligated sample was expected to substantially reduce the amount of linker-mediated background amplification occurring during the subsequent LP-PCR procedure. Consequently, we hoped to avoid saturation of LP-PCR, allowing the more specific and efficient LP-PCR-mediated amplification of HIV-specific *Nla* III fragments spanning the 5'U3 region and upstream cellular sequence.

The principle of active DynabeadTM selection for *Nla* III fragments containing the U3 HIV region is outlined in figure 4.18. Following ligation of LPNV to *Nla* III termini, DNA was denatured by heat. A biotinylated primer (U3NV-bio, see Table 2.1) was then allowed to anneal to complementary sequences within the denatured sample. Magnetic beads coupled to avidin molecules were then added and the sample incubated to allow interaction of the avidin and biotin molecules. By virtue of the strong biotin/avidin interaction, all fragments containing the sequence complementary to the U3NV primer were then selected by applying a magnetic field across the base of the tube holding the sample. DNA fragments not annealed to the U3NV primer (and therefore not involved in a biotin/avidin interaction) were removed by extensive washing of the magnetic pellet. The pellet was then resuspended in a volume of water and subjected to LP-PCR.

Figure 4.18 DynabeadTM Selection of HIV LTR DNA Fragments

Outline of the method used to enrich for *Nla* III fragments containing HIV U3 DNA sequence. Linkers (LPNV) are shown in green and the target U3 sequence is shown in orange. The biotinylated U3NV oligonucleotide and the avidin-coated magnetic beads are shown.



4.6.3.ii Assessing the DynabeadTM Selection Process

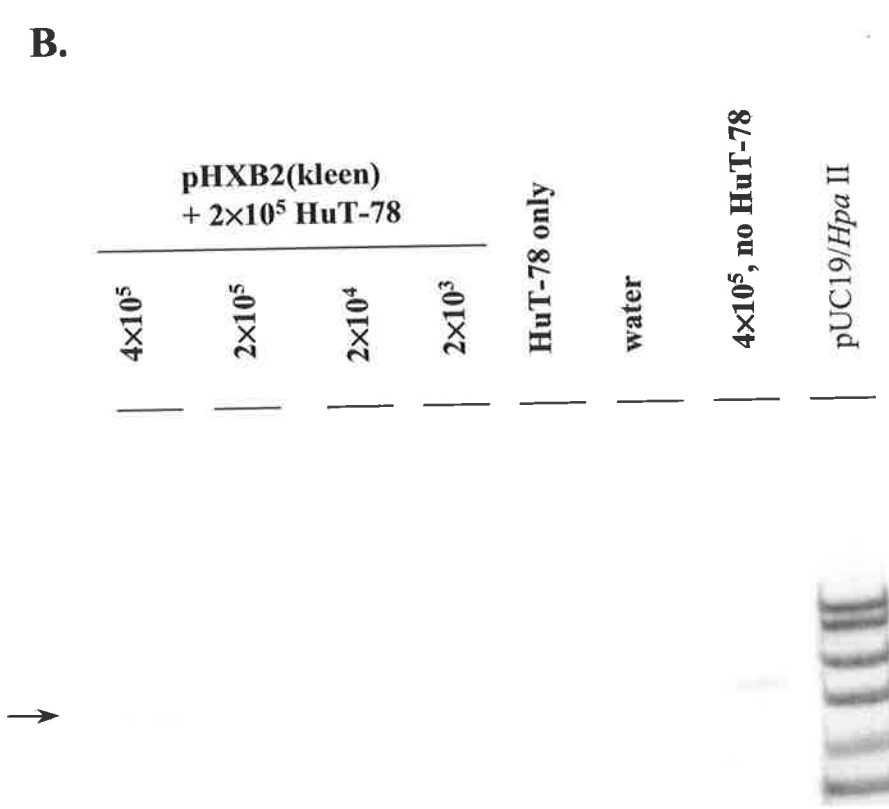
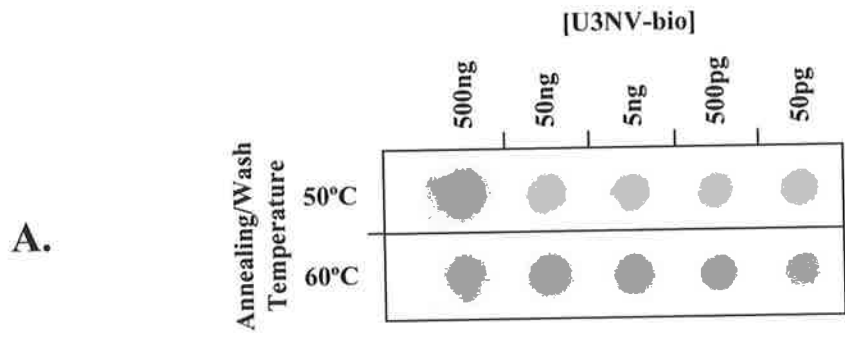
The DynabeadTM selection of target DNA sequences was essentially performed using the conditions recommended by the manufacturer. However, efforts to optimise the dynabead selection procedure using the U3NV-bio primer concentrated on establishing the concentration of biotinylated U3NV to be used and the temperatures of annealing and washing that allowed the greatest target sequence recovery.

Experiments were performed on 1×10^5 copies of pHXB2(kleen) spiked into 2×10^5 cell-equivalents of HuT-78 chromosomal DNA. DNA samples were digested with *Bgl* II and *Nla* III and then subjected to hybridisation with various amounts of biotinylated U3NV at either 50°C or 60°C. Dynabeads were then added and incubated with samples for 2h at room temperature. Two washes were carried out at temperatures identical to that in which hybridisation was performed (50°C or 60°C) using the buffer supplied by the manufacturer, and recovery of target DNA was assessed by a 25-cycle PCR (standard conditions; see section 2.2.5) using primers U3NV and U3.1(+). PCR products were then subjected to dot-blot analysis using the U3PNV oligonucleotide probe (Table 2.2). The results indicated that the efficiency of target sequence recovery did not differ significantly between the two annealing temperatures (50°C and 60°C) tested (Fig. 4.19A). However, there was a noticeable dependence of the final signal obtained on the amount of biotinylated primer used at both temperatures tested. Since the strongest signal was observed in the samples which involved the use of 500ng of biotinylated U3NV at an annealing temperature (and wash temperature) of 50°C, these conditions were used in subsequent experiments incorporating the use of LP-PCR.

To determine whether the magnetic bead selection process could increase the sensitivity of LP-PCR, an initial experiment was performed using various copy numbers of pHXB2(kleen) spiked onto 2×10^5 cell-equivalents of HuT-78 chromosomal DNA. Following linker ligation to digested DNA samples, DNA was heated to 94°C to facilitate strand separation and then cooled to 50°C in the presence of the biotinylated U3NV primer (500ng) to allow hybridisation. Following magnetic separation of biotin/avidin complexes, the pellet was washed twice at 50°C and LP-PCR was performed. The resulting products were subjected to Southern analysis. The absence of

Figure 4.19 The Effect of the Dynabead™ Selection Process on LP-PCR

- A.** Optimisation of the Dynabead™ selection protocol. The effects of various concentrations of biotinylated U3NV ([U3NV-bio]; horizontal labelling) and the temperatures at which U3NV-bio was annealed/washed (annealing/washing temperatures; vertical labelling) on the signal observed following dot-blot analysis of LP-PCR products. For a detailed explanation of the protocol used, see section 4.6.3.ii.
- B.** Sensitivity of the LP-PCR procedure following enrichment of U3-containing DNA fragments by the Dynabead™ selection protocol and subsequent asymmetric and LP-PCR. Dynabead™ selection was performed using 500ng of U3NV-bio annealed (and washed) at 50°C.



any detectable signal indicated that LP-PCR-mediated amplification of the U3 region and upstream cellular sequence had not occurred (data not shown). However, when an asymmetric PCR was performed (as outlined in section 4.6.2.i) on the washed pellet prior to LP-PCR, signal was clearly present in lanes corresponding to amplification of 4×10^5 and 2×10^5 copies of pHXB2(kleen) (Fig. 4.19B). This equated to an ability of the assay to pick up approximately one copy of integrated HIV DNA per cell when 2×10^5 cell-equivalents of input DNA were used. However, in the absence of the DynabeadTM selection procedure, detection of 1×10^4 copies of pHXB2(kleen) in a background of 2.5×10^5 cell-equivalents of background DNA was achieved (see section 4.6.2.iii). This implied that there was a measure of target sequence loss associated with the complex nature of the magnetic bead selection process. Alternatively, the presence of the magnetic beads and/or the biotin or avidin proteins may have been directly interfering with the LP-PCR amplification process.

4.6.3.iii DynabeadTM Selection: Conclusions

The incorporation of the DynabeadTM selection process led to a decrease in assay-sensitivity possibly due to a loss of target sequences. In addition, the complexity of this technique also provided potential areas in which variability might compromise the quantitative nature and reproducibility of the assay. Furthermore, in any system in which PCR is to be used to generate the final detectable product, there is a chance of contamination of samples with DNA sequences generating false positive results. To reduce the chance of sample contamination, all such assays should keep sample handling to a minimum. Therefore, since incorporation of the DynabeadTM selection procedure substantially increased the complexity of the assay without increasing detection sensitivity, this avenue of investigation was not pursued further.

4.6.4 Genuine Hot-start LP-PCR + Nested PCR

4.6.4.i Incorporating the Genuine Hot-start LP-PCR

An alternative method to address the problem of non-specific amplification of background *Nla* III fragments was next investigated. Since extension over the linker

sequence to generate sequence to which the linker can anneal (see Fig. 4.15) was proposed to be necessary in order for linker-mediated amplification of *Nla* III fragments to occur, methods to stop this process were expected to reduce background amplification.

To abolish extension over the linker sequence prior to the initial denaturation step in the LP-PCR procedure (either at RT during sample transport to the thermocycler or over the temperature ramping period as the thermocycler heats to its initial denaturation temperature), strand dissociation prior to the addition of the thermostable DNA polymerase was performed. Initially, this was achieved by heating the samples to 94°C for 5min in the thermocycler prior to the addition of 5U of AmpliTaq DNA polymerase and then resuming normal cycling. The PCRs were performed in a final volume of 100µl and cycled as follows: 94°C 3min; 30 cycles of 94°C 30s, 58°C 30s, 72°C 1min; and a final extension of 72°C for 10min. Using this method to amplify 1×10^5 copies of the pHXB2(kleen) construct spiked onto a background of 1.2µg (2×10^5 cell-equivalents) of HuT-78 chromosomal DNA, no signal was detected (Fig. 4.20A).

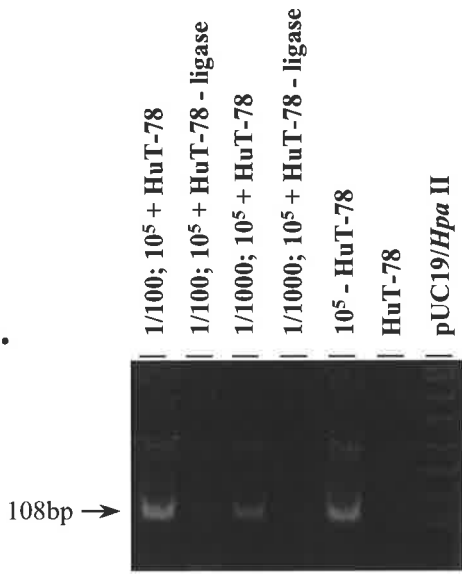
4.6.4.ii Incorporating the Nested PCR

Although signals were absent after first-round amplification using the genuine hot-start protocol, when a 22-cycle nested PCR (using primers U3.1(+) and U3-106(-) on either 1/100th or 1/1000th of the final LP-PCR reaction mix was performed (as described in section 2.2.5.v), clear signals were observed as determined by ethidium bromide staining (Fig. 4.20B). The DNA bands were then confirmed to be specific for an HIV probe (U3-106 probe; see Table 2.2) by Southern analysis (Fig. 4.20C). Since control reactions in which the ligation reaction was performed in the absence of ligase failed to produce signal after the nested PCR (even after Southern analysis), all signals arising from the assay were dependent on the addition of linkers to digested DNA. This control also demonstrated that the nested PCR amplification of asymmetric PCR products (mediated by the U3NV primer) arising during the LP-PCR were not contributing significantly to the final signal observed. Thus, the incorporation of a nested PCR allowed the detection of products produced from the 1st-round LP-PCR.

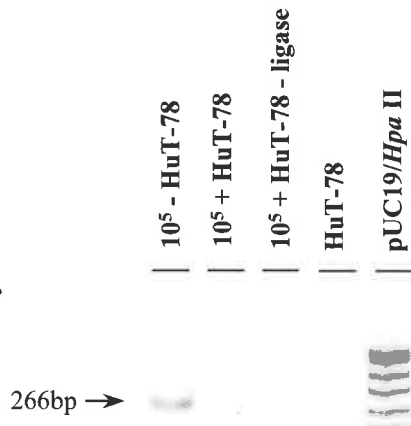
Figure 4.20 Incorporating a Genuine Hot-Start Protocol with a Nested PCR

A. Genuine hot-start PCR protocol using AmpliTaq DNA polymerase (see section 4.6.4.i for details) performed on 10^5 copies of pHXB2(kleen) in the presence (+ HuT-78) or absence (- HuT-78) of HuT-78 chromosomal DNA. Controls in which LP-PCR was performed on HuT-78 chromosomal DNA alone (HuT-78) or on 10^5 copies of pHXB2(kleen) in the absence of linker ligation (-ligase) are shown. DNA mass markers (pUC19/*Hpa* II) and the size (266bp) and position of the expected band are indicated. **B.** Nested PCR (22 cycles) performed using primers U3.1(+) and U3-106(-) on aliquots (either 1/100 or 1/1000 as indicated) of the products generated in A. PCR products are visualised by ethidium bromide. **C.** Southern analysis of products generated in B. using the U3-106 probe (Table 2.2).

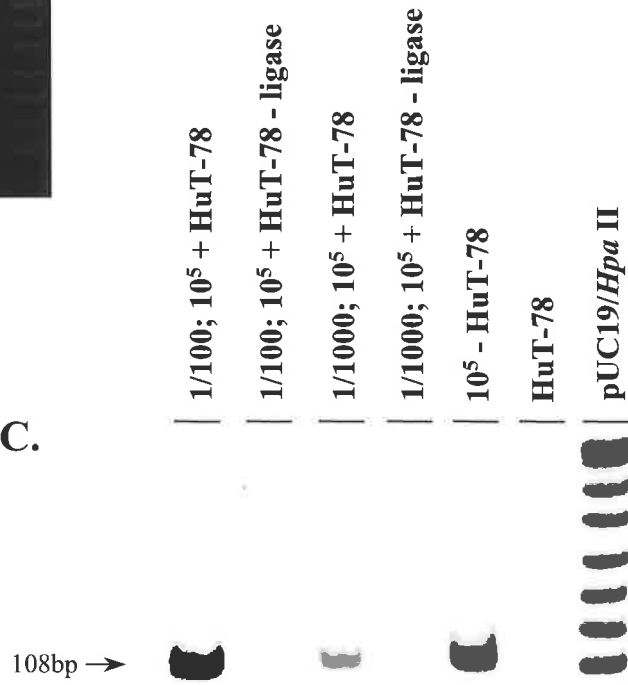
B.



A.



C.



Taken together, these results suggested that amplification of target sequences was occurring in the 1st-round LP-PCR, although the levels of amplified product were below that required for successful detection using Southern analysis. Furthermore, since a strong signal was detected following amplification of a 1/100th aliquot of the 1st-round PCR, this dilution (as opposed to 1/10th of the 1st-round PCR) was used in all future experiments to reduce the potential that PCR saturation might occur.

4.6.4.iii Incorporating AmpliTaq Gold™

In order to simplify the hot-start process, the use of AmpliTaq Gold™ DNA polymerase was investigated. AmpliTaq Gold™ DNA polymerase is initially present as a proenzyme which when exposed to high temperatures (above 92°C) converts to the active polymerase form. Thus, at temperatures below that required for DNA strand dissociation, AmpliTaq Gold™ is inactive. The use of AmpliTaq Gold™ was therefore expected to abolish any linker-mediated amplification of background *Nla* III fragments as the enzyme is only active after strand dissociation has occurred.

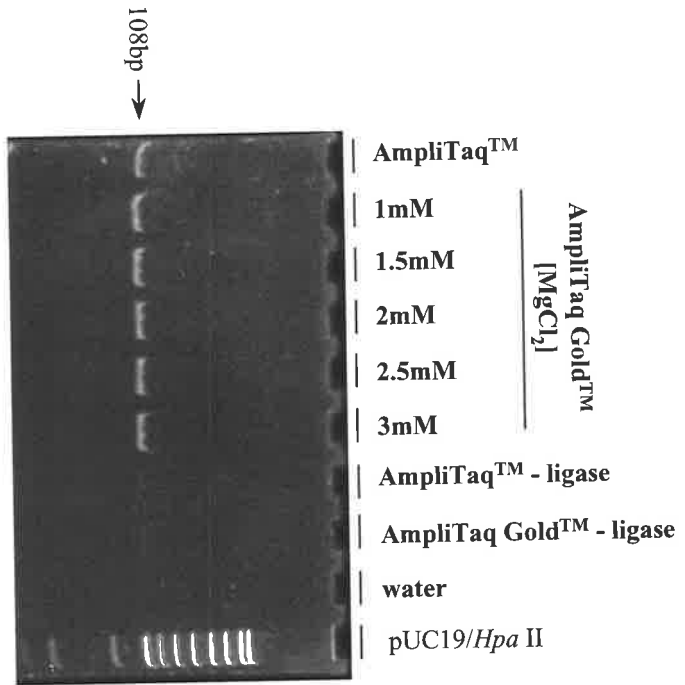
Following nested LP-PCR amplification of 1×10^5 copies of the pHXB2(kleen) construct spiked onto 1.2 µg of HuT-78 chromosomal DNA using either AmpliTaq™ DNA polymerase (with heat dissociation of sample as described above) or AmpliTaq Gold™ DNA polymerase over a range of MgCl₂ concentrations, little difference in the amplification efficiencies was observed (Fig. 4.21). However, since the use of AmpliTaq Gold™ DNA polymerase greatly simplified the hot-start LP-PCR procedure, this enzyme was used (with a reaction MgCl₂ concentration of 1.5mM) for all future LP-PCR experiments.

4.6.5 Amplification of Integrated HIV DNA Present in Cell Lines

Previous experiments had indicated that the pHXB2(kleen) construct may amplify more efficiently by LP-PCR than equivalent copies of HIV (as judged by GAG PCR) within chromosomal DNA (see section 4.5.1). Therefore, to assess the assay sensitivity, a mixture of three persistently infected cell lines [H3B (2 copies of integrated provirus),

Figure 4.21 Optimising the Hot-Start Protocol: The effect of [MgCl₂] on AmpliTaq Gold™

Genuine hot-start PCR protocol using either AmpliTaq™ DNA polymerase (used as described in Fig. 4.20) or AmpliTaq Gold™ DNA polymerase performed on 10⁵ copies of pHXB2(kleen) in the presence of 1.2µg of HuT-78 chromosomal DNA. PCR cycling of reactions with AmpliTaq Gold™ DNA polymerase involved an extended initial denaturation step (94°C for 12min) to ensure activation of the polymerase. A subsequent nested PCR was performed on 1/100th of the 1st-round PCR. Control reactions performed on water alone (water) or 10⁵ copies of pHXB2(kleen) in the absence of linker ligation (-ligase reactions) are shown. DNA mass markers (pUC19/*Hpa* II) and the size (108bp) and position of the expected nested PCR product is indicated.



ACH-2 (1 copy of integrated provirus) and 8E5 (1 copy of integrated provirus)] was prepared (the HA8 integrated standard, see section 2.2.4.i).

To confirm that a *Bgl* II recognition site was not present prior to an *Nla* III site in the chromosomal sequence upstream of the integrated HIV molecule in each cell line, sequence analysis was performed. This was performed by Dr. Raman Kumar. Although sequence data for upstream chromosomal sequence at one integration site within the H3B sequence was unable to be obtained, the remaining three sequences all indicated the presence of an *Nla* III site immediately adjacent to the site of integration without an intervening *Bgl* II site (Fig. 4.22).

Various amounts of the HA8 standard were spiked into 1.2 μ g (2×10^5 cell-equivalents) of HuT-78 chromosomal DNA and subjected to nested LP-PCR. When nested PCR was performed on 1/100th of the 1st-round LP-PCR, the HA8 standard equivalent to 10 copies of HIV DNA gave a PCR product that was clearly visualised by ethidium bromide staining when run through an 8% polyacrylamide gel (Fig. 4.23). In addition, signal was absent in a control reaction in which the linker ligation reaction was performed in the absence of ligase. It is also worth noting that using this LP-PCR protocol, there appeared to be little difference in the amplification efficiencies of either the HA8 or the pHXB2(kleen) control standards.

4.7 Optimising the 1st-Round PCR

In order to further optimise the 1st-round LP-PCR procedure, the number of PCR cycles required for amplification and the optimal temperature for primer annealing was investigated using the HA8 standards.

No significant decrease in final sensitivity was seen when the 1st-round PCR cycle-number was reduced from 30 to 22 (compare Fig. 4.24A. and B.). Furthermore, in this experiment a strong signal was observed following the 30-cycle amplification of 1000 copies of HA8 prepared in the absence of ligase. This signal likely resulted from the nested amplification of asymmetric PCR products generated by extension of U3NV in the 1st-round PCR and was abolished when the 1st-round cycle-number was reduced to

Figure 4.22

Diagram showing the relative distances of *Nla* III sites upstream of the integrated HIV 5'LTR in the H3B, ACH-2 and 8E5 cell lines. The sequence upstream of each integrated provirus was obtained by Dr. Raman Kumar (unpublished).

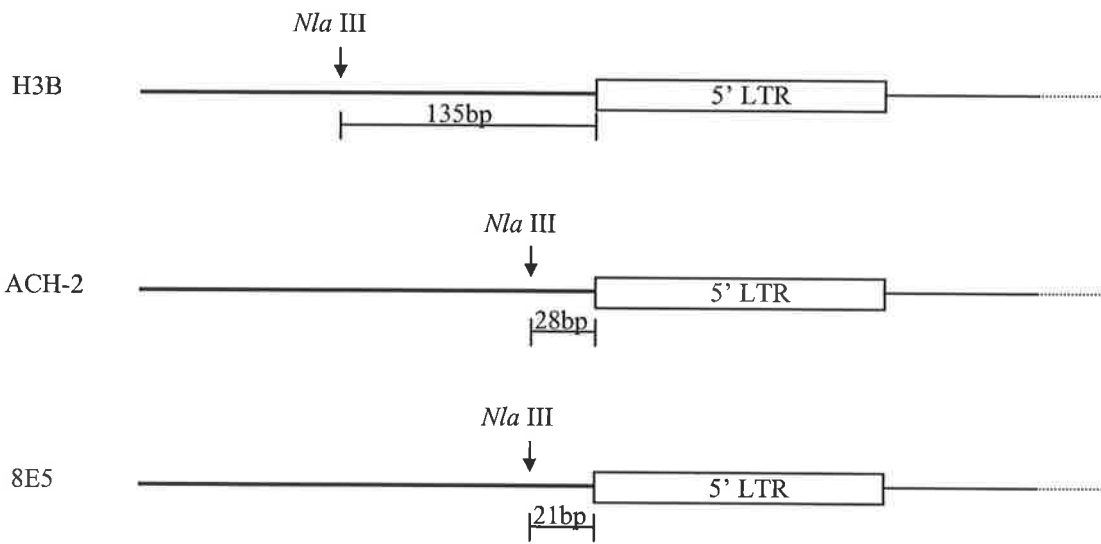
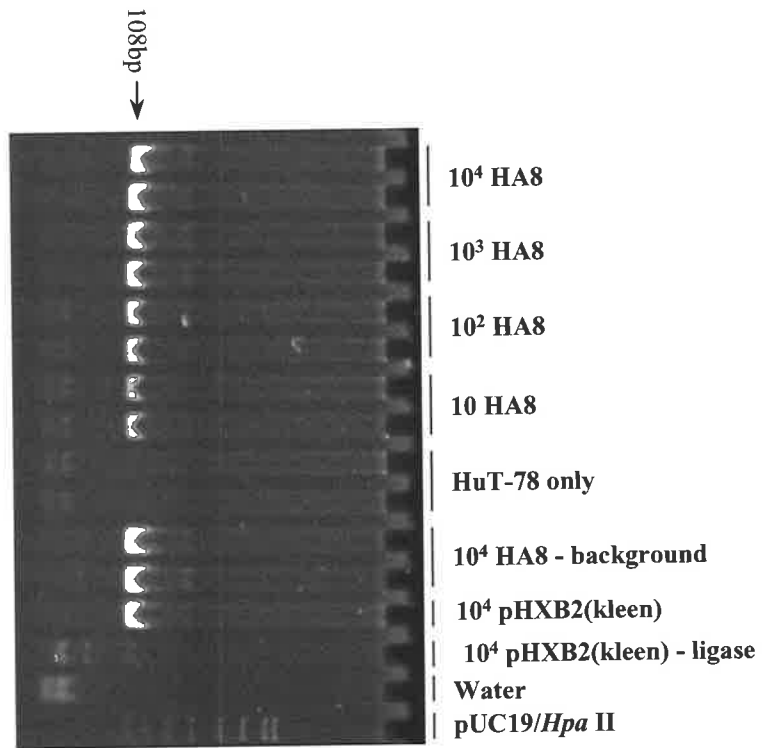


Figure 4.23 LP-PCR Amplification of Integrated HIV DNA: The HA8 Integrated Standard

Nested LP-PCR amplification of integrated HIV DNA present within the HA8 integrated standard. All samples (except –background control) contained 1.2µg of HuT-78 chromosomal DNA. Control reactions in which 10⁴ copies of pHXB2(kleen) were amplified following linker ligation in the presence or absence (-ligase) of ligase are shown. DNA mass markers (pUC19/*Hpa* II) are shown. First-round LP-PCR (30 cycles) was performed on various amounts of the HA8 standards (based on integrated HIV copy-number) as described previously (see section 4.6.4.iii). Nested PCR (22 cycles) was performed on 1/100th of first-round PCRs (see section 4.6.4.ii). Bands were visualised by ethidium bromide staining.



22 (compare Fig. 24A. and B., -ligase samples). When the 1st-round cycle number was further reduced from 22 to 18, the final signal obtained dropped considerably (data not shown). Taken together, these results indicated that the 1st-round LP-PCR amplification had reached a plateau when cycle numbers above 22 were used, resulting in no further amplification of target sequences. This was in part attributed to unavoidable background amplification mediated by non-specific annealing of either the LPNV or U3NV primer to chromosomal DNA (see Fig. 4.25A), or by non-specific interactions between *Nla* III fragments (containing linkers) at temperatures of annealing and/or extension thereby creating templates that could subsequently be amplified by pairs of LPNV primers (see Fig. 4.25B). In either case, subsequent PCR from these templates could deplete the reaction of PCR precursors.

In addition, 1st-round LP-PCR was performed using an annealing temperature of either 58°C (used in experiments above) or 63°C. The use of an elevated annealing temperature was investigated in an attempt to decrease the amount of non-specific primer annealing occurring in the LP-PCR. However, increasing the temperature of primer annealing significantly reduced the amount of final signal obtained (Fig. 4.26).

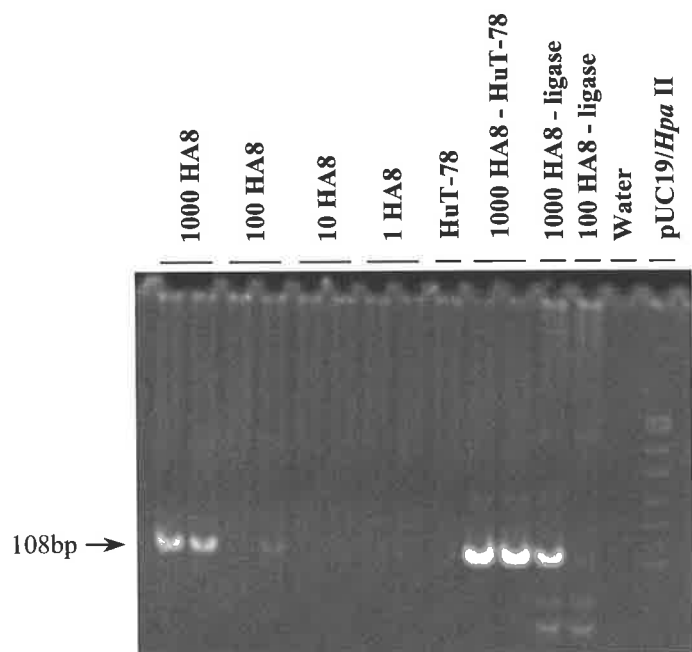
4.8 Selection against Extrachromosomal forms

Since the nested LP-PCR had been shown to successfully amplify integrated HIV DNA in the context of large amounts of background DNA (see sections 4.6 and 4.7), the selectivity of the assay against amplification of extrachromosomal forms needed to be assessed. In order to investigate this, a construct closely mimicking the linear extrachromosomal form (lin) was prepared and precisely equated for HIV copy-number against the HA8 standards using GAG PCR (see section 2.2.4.ii). Known amounts of the linear construct and HA8 standards were then subjected to nested LP-PCR and signal intensities compared. Strong bands were present corresponding to amplification of 10⁴ copies of the linear construct indicating that although efficient amplification of the integrated standards was occurring, substantial amplification of the linear construct was also occurring (Fig. 4.27). Subsequent repetition of this experiment confirmed the non-selective nature of the assay against extrachromosomal forms (data not shown).

Figure 4.24 Optimising 1st-round LP-PCR Cycle-Number

Nested LP-PCR performed on various amounts of HA8 (based on integrated HIV DNA copy-number) in the presence of 1.2µg of HuT-78 chromosomal DNA. First-round LP-PCR was cycled for either 30 (**A.**) or 22 (**B.**) cycles. Nested PCRs were cycled for 22 cycles. Control reactions in which PCRs were performed on HA8 chromosomal DNA prepared in the absence of linker ligation (-ligase), water alone or in the absence of HuT-78 chromosomal DNA (-HuT-78) are shown. DNA mass markers (pUC19/*Hpa* II) are indicated.

A. 30 cycles



B. 22 cycles

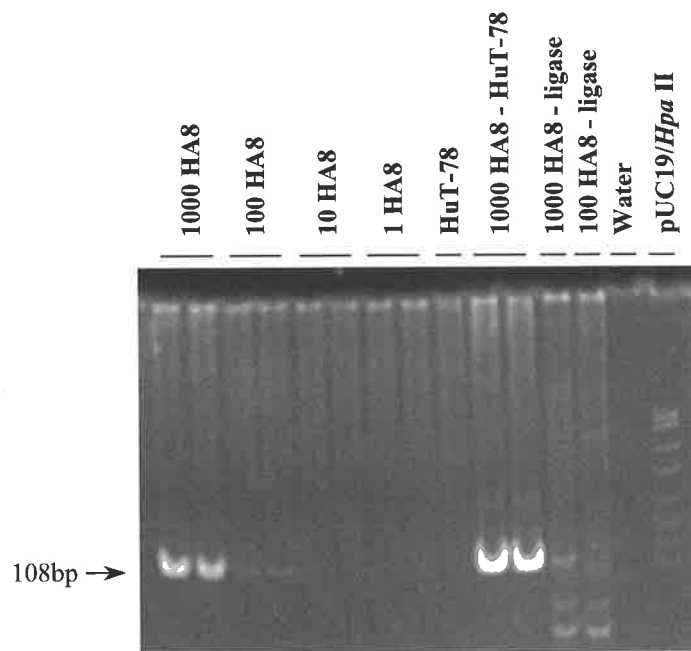
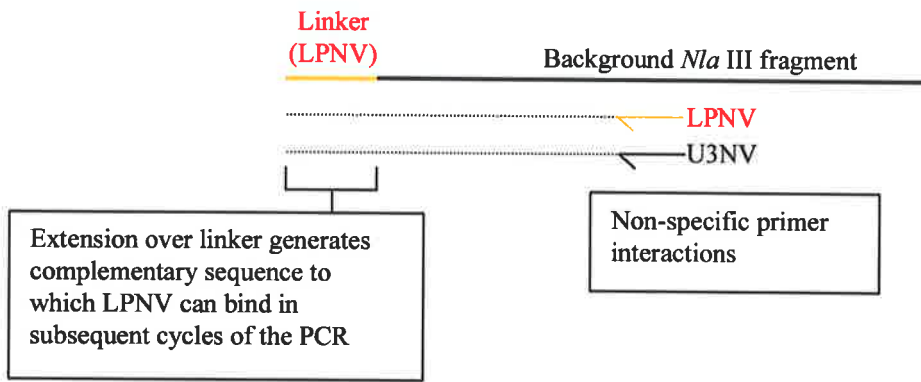


Figure 4.25

Non-specific mechanisms by which background *Nla* III DNA fragments can be amplified. **A.** Non-specific annealing and subsequent extension of either LPNV or U3NV. **B.** Non-specific annealing and subsequent extension of single-stranded *Nla* III DNA fragments.

A.



B.

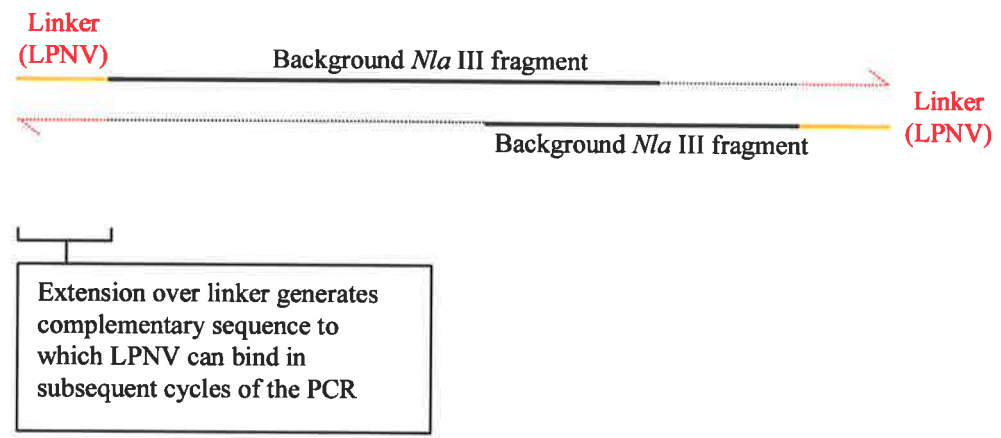


Figure 4.26 Optimising 1st-round LP-PCR Primer-Annealing Temperature

First-round LP-PCR performed (in duplicate) on 10^3 copies of HA8 (based on HIV copy-number) using a primer-annealing temperature of either 63°C or 58°C. Nested PCRs (22 cycles) were performed on 1/100th of 1st-round products as described in section 4.6.4.ii. A control LP-PCR (using a primer-annealing temperature of 58°C) performed on water alone (Water) and DNA mass markers (pUC19/*Hpa* II) are indicated.

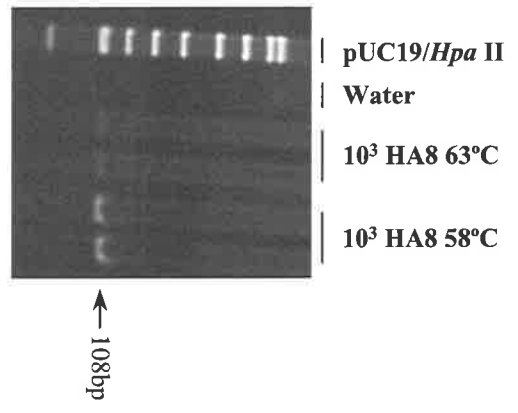
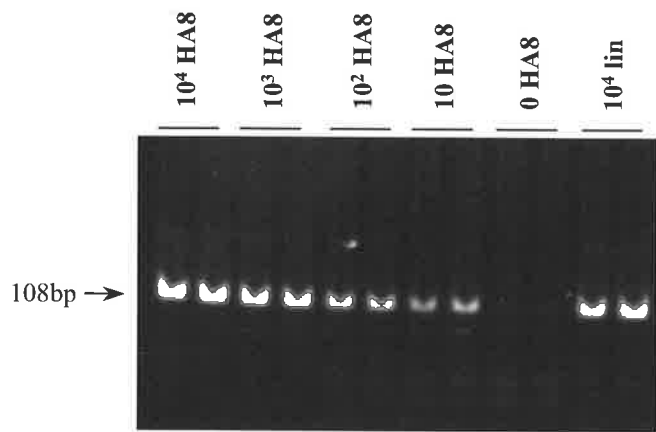


Figure 4.27 LP-PCR Amplification of the lin HIV DNA Construct

Nested LP-PCR amplification of standards mimicking integrated (HA8) and linear extrachromosomal (lin) HIV DNA. The size of the expected nested PCR product is indicated. PCR products were run through an 8% PAGE gel and visualised by ethidium bromide staining.



4.8.1 Addressing the Poor Selectivity

The poor specificity of LP-PCR for integrated HIV DNA indicated that the selection enzyme *Bgl* II was not efficiently stopping LP-PCR-mediated amplification of extrachromosomal forms. Since the *Bgl* II digestion conditions had been extensively optimised (see section 4.2), alternative solutions to why selection against the linear extrachromosomal form was not occurring were proposed. Consequently, it was postulated that re-ligation of *Bgl* II fragments during the linker-ligation step allowed reconstitution of the *Nla* III fragment spanning the 3'LTR/PPT junction sequence (see Fig. 4.3A, B and C) present within all extrachromosomal forms and the 3'LTR/5'LTR junction (see Fig. 4.3C) in the 2-LTR extrachromosomal form. Re-ligation in this fashion would allow LP-mediated PCR amplification over these regions and negate the selective effect of *Bgl* II digestion.

In order to address this problem, two solutions were proposed:

Bgl II digestion after linker addition: If *Bgl* II digestion was performed after the linker ligation step and prior to LP-PCR, selection against the extrachromosomal forms would be restored. However, as this option would involve either the re-optimisation of *Bgl* II digestion (in the presence of ligase buffering conditions) or an additional phenol/chloroform/isoamylalcohol and/or ethanol precipitation step prior to *Bgl* II digestion (to remove the ligation buffer) it was not pursued. In addition, the LP-PCR amplification step following *Bgl* II digestion may have had to be re-optimised considering the significantly different buffering conditions required for direct PCR.

“Filling-in” *Bgl* II 3'-underhangs: The destruction of the complementary nature of sticky ends produced by *Bgl* II digestion to ensure that re-ligation of *Bgl* II termini did not occur was expected to restore the selective nature of LP-PCR for integrated HIV-1 forms. This involved the extension of the 3'-underhang generated by *Bgl* II digestion using Klenow DNA polymerase in the presence of particular nucleotides. Since the Klenow-mediated “filling-in” reaction could potentially be incorporated into the LP-PCR assay without significantly disrupting the established protocol, optimisation of this process was investigated.

4.8.2 Incorporation of the Klenow-Mediated “Fill-In” Step

The process of filling in the *Bgl* II-generated sticky ends is outlined in figure 4.28 (Scenario A). Complete extension over the *Bgl* II site to generate blunt-ends required the presence of the nucleotides G, A, T and C. However, in order to prevent any blunt-end ligation from occurring during the linker ligation reaction after sample incubation with Klenow DNA polymerase, only two nucleotides were “filled-in” (G and A). This was expected to generate non-complementary sticky-ends which would not ligate together in the presence of T4 DNA ligase. Since chain-elongation by Klenow DNA polymerase does not occur in the 3' to 5' direction, the 4bp 3' overhang generated by *Nla* III digestion remained unaffected by this process (Fig. 4.28, Scenario B). However, care was taken to ensure that the enzyme did not have 3'→5' exonuclease activity as this would chew back and destroy the integrity of the 4bp 3'-overhangs generated by *Nla* III digestion which were required for linker ligation. Therefore, for all fill-in reactions, Klenow 3'→5'exo⁻ enzyme (NEB) was used.

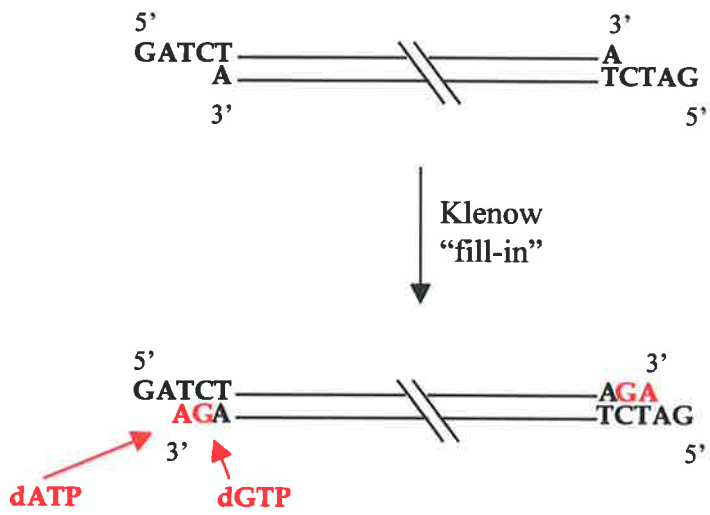
In order to minimise error associated with sample recovery and losses when transferring a sample(s) between tubes, every effort was made to restrict the number of steps in the protocol. Therefore, it was necessary to evaluate the potential for the “fill-in” reaction to be performed either in the restriction digestion buffering conditions, or the restriction digestion buffering conditions supplemented with factors to generate conditions closely matching those of the buffer supplied with the Klenow enzyme.

Consequently, an initial experiment was performed to establish the relative efficiencies of the Klenow 3'→5'exo⁻ enzyme in either; 1) the conditions used for restriction enzyme digestion of target DNA, 2) the conditions used for restriction enzyme digestion of target DNA supplemented with DTT to a final concentration of 7.5mM or, 3) the commercially provided buffering conditions (1×EcoPol Buffer). A diagrammatic representation of the procedure is outlined in figure 4.29. Briefly, HuT-78 DNA was digested with *Nla* III and *Bgl* II as outlined in section 4.2 and subsequently incubated with 5mM dGTP, 5U of Klenow 3'→5'exo⁻ and 50μCi of α-³²P-dATP under the various buffering conditions (see above and section 4.2). Following centrifugation

Figure 4.28

The effect of Klenow (3'→5' exo⁻)-mediated “fill-in” of *Bgl* II-cleaved (Scenario A) and *Nla* III-cleaved (Scenario B) DNA in the presence of dATP and dGTP.

Scenario A: *Bgl* II



Scenario B: *Nla* III

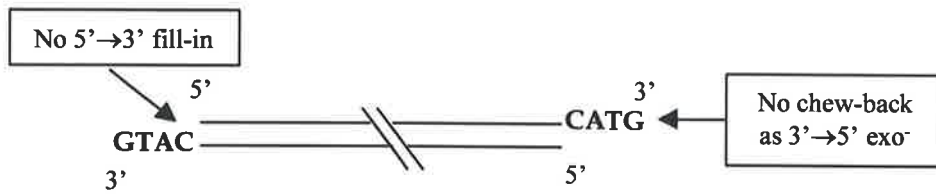
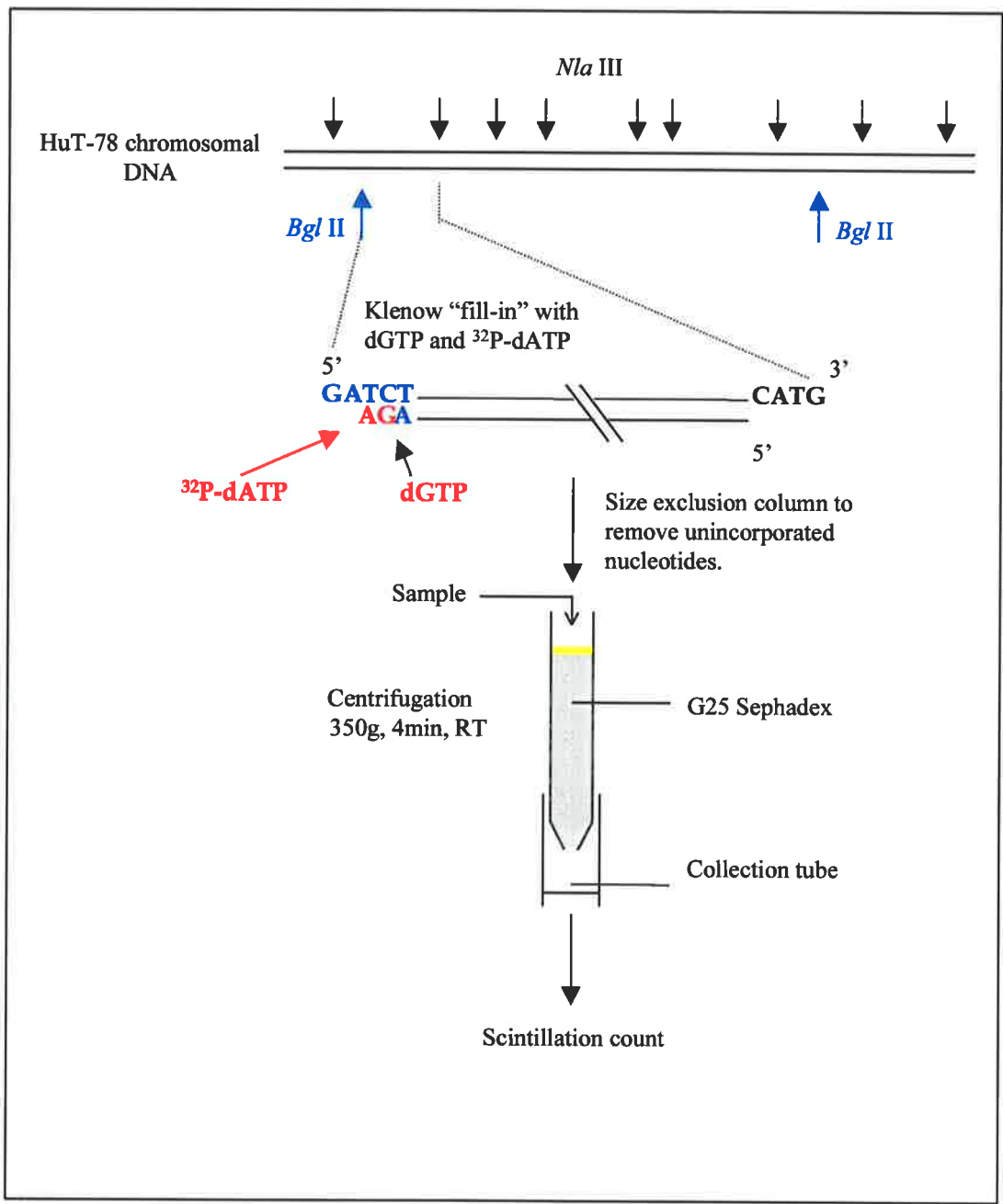


Figure 4.29

Experimental design to assess the relative efficiencies of Klenow (3'→5'exo⁻)-mediated “fill-in” of *Bgl* II-cleaved DNA performed in different buffering conditions. HuT-78 chromosomal DNA is initially cleaved with both *Nla* III and *Bgl* II and then incubated with Klenow (3'→5'exo⁻) in the presence of dGTP and ³²P-labelled dATP (³²P-dATP). The incorporation of radioisotope into large DNA fragments is assessed by scintillation counting after running the sample through a size-exclusion column (G25 Sephadex).



through a size exclusion column, the relative efficiency of the Klenow enzyme in each of the buffering conditions was indirectly measured by the levels of incorporation of ^{32}P -label into the resulting eluate. It is worth noting, that successful addition of dGTP to the 3'-termini is required before addition of (^{32}P -)dATP can occur. Therefore, this experiment was designed to assay the addition of both nucleotides to the 3'-termini of *Bgl* II restricted DNA. Since addition of only one nucleotide (dGTP) is necessary to destroy the potential of *Bgl* II fragments to religate, the efficiencies observed in this experiment were expected to be an underestimation of the functional efficiency ultimately required in the assay.

The results indicated that substantial addition of ^{32}P -dATP to *Bgl* II termini was occurring under all three buffering conditions relative to the negative control lacking the Klenow enzyme (-Klenow) (Table 4.1). However, since the experiment was not performed in duplicate, no standard error could be calculated. Consequently the experimental variation expected using this protocol was unknown. In addition, significant counts were present in the negative control indicating that the size exclusion protocol used to remove unincorporated ^{32}P -dATP from samples was inefficient. Taken together, it was difficult to accurately assess the efficiencies of the Klenow-mediated "fill-in" reaction under the various buffering conditions used in this experiment. However, our data indicated that addition of dATP to target DNA termini under the three conditions tested was occurring.

Table 4.1. Effect of various reaction conditions on Klenow DNA polymerase efficiency

Reaction Conditions	counts/minute
Digest buffer - DTT	290000
Digest Buffer + DTT	300000
EcoPol Buffer	380000
EcoPol Buffer - Klenow	150000

To investigate whether the above strategy did in fact reduce the PCR signal obtained with the linear DNA form, and to further compare buffer conditions for the commercially provided Klenow 3'→5'exo⁻ reaction, an experiment comparing the relative detection efficiencies of known amounts of both the HA8 and lin standards by LP-PCR under the two Klenow reaction conditions was performed. A flow-chart outlining the experimental design is presented in figure 4.30. Duplicate samples of either the HA8 standard or the lin construct were digested as described previously (see section 4.2) and either directly filled-in under digestion conditions adjusted for DTT concentration (Fig. 4.30, Samples 13-18), or ethanol precipitated prior to filling-in under commercial buffering conditions (Fig. 4.30, Samples 1-6). The results indicated that when the buffering conditions used for *Nla* III digestion were modified through the addition of DTT (to a final concentration of 7.5mM), very little LP-PCR-mediated amplification of the lin construct occurred (Fig. 4.31). In comparison, amplification of the lin construct was seen using the commercially provided Klenow buffering conditions. In addition, when *Bgl* II digestion was employed after linker ligation (Fig. 4.30, Samples 7-12), a small amount of lin construct amplification was observed in one of the duplicate samples. The variation in duplicate samples observed in this set of samples (Fig. 4.31, samples 7-12) and samples in which “fill-in” reaction were performed in the commercially available buffering conditions (Fig. 4.31, samples 1-6), may have resulted from DNA recovery variations associated with the additional ethanol precipitation step required in these sample sets.

Taken together, the results presented in table 4.1 and figure 4.31 suggested that Klenow-mediated “fill-in” of *Bgl* II sites under the *Bgl* II/*Nla* III buffering conditions supplemented with DTT prior to linker ligation, was sufficient for use in the LP-PCR assay.

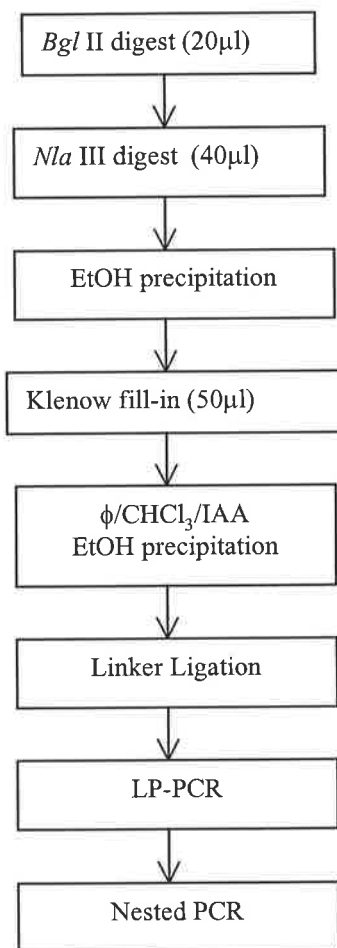
4.9 Finalised LP-PCR Protocol for the Detection of Integrated HIV-1 Forms

The finalised LP-PCR protocol for the detection of integrated HIV DNA is presented in figure 4.32 and figure 4.33. The final evaluation of the ligation efficiency achieved, the

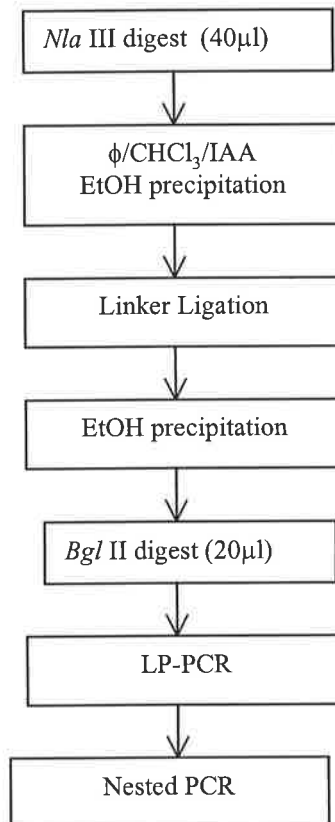
Figure 4.30

Flow charts outlining the three experimental protocols used to avoid re-ligation of *Bgl* II-cleaved DNA during the linker-ligation step of LP-PCR. Samples 1-6 involved the “fill-in” reaction performed in 1×EcoPol Buffer. Samples 7-12 were not subjected to a “fill-in” step but were instead digested with *Bgl* II after linker ligation had been performed. Samples 13-18 involved the “fill-in” reaction performed in the digestion buffering conditions adjusted by adding DTT to a final concentration of 7.5mM. Each of these protocols were performed on known copies of both HA8 and lin DNA. The results are presented in figure 4.31.

Samples 1-6
**1×EcoPol
buffer**



Samples 7-12
***Bgl* II Digestion
after linker ligation**



Samples 13-18
**Digestion Buffer
+DTT**

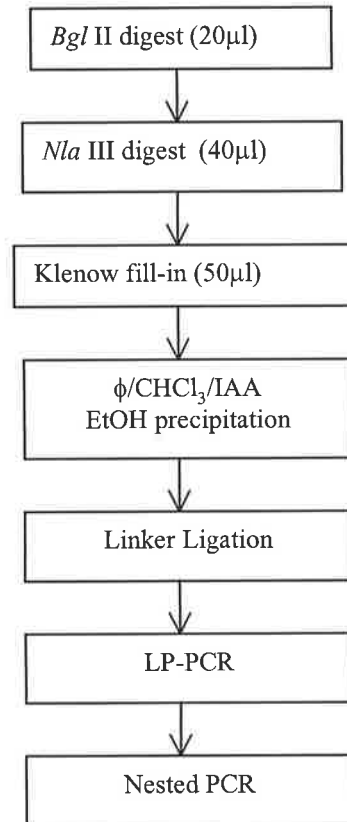


Figure 4.31 Optimising the Selective Nature of LP-PCR for Integrated HIV DNA

Three experimental protocols (see figure 4.30 for details) used to avoid re-ligation of *Bgl* II-cleaved DNA during the linker-ligation step of LP-PCR were compared by amplifying known amounts of standards mimicking both the integrated (HA8) and the linear extrachromosomal (lin) HIV DNA forms. The nested LP-PCR protocol used was the same as that described in section 4.6.4.ii. A control reaction in which the nested PCR was performed on water alone (Water) is indicated. DNA mass markers (pUC19/*Hpa* II) are shown.

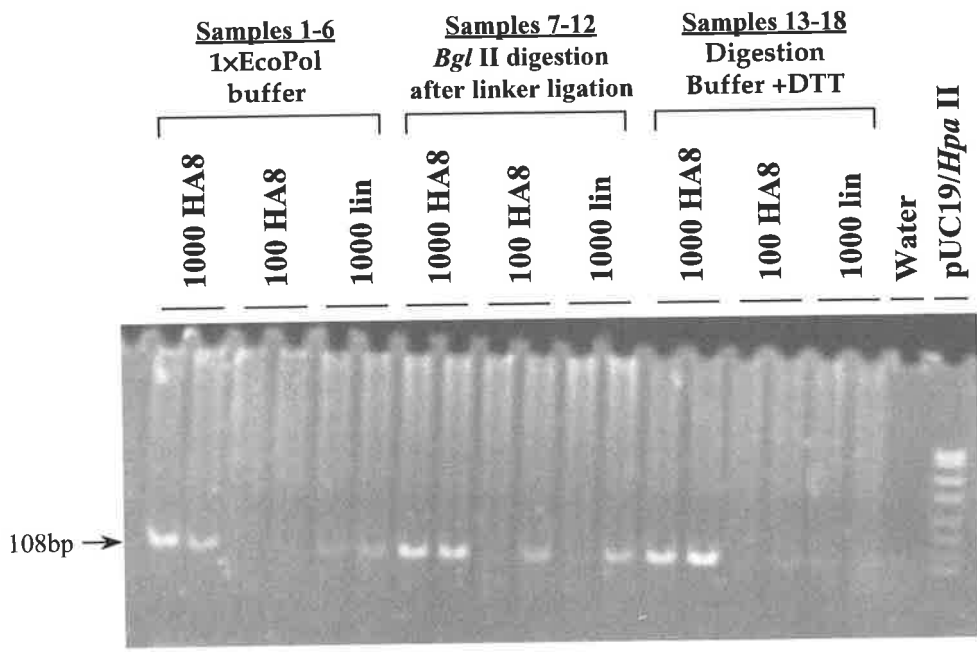


Figure 4.32

Diagram of the finalised nested LP-PCR protocol. The positions of all primers and the probe (U3-106) used to detect the final nested PCR product is shown.

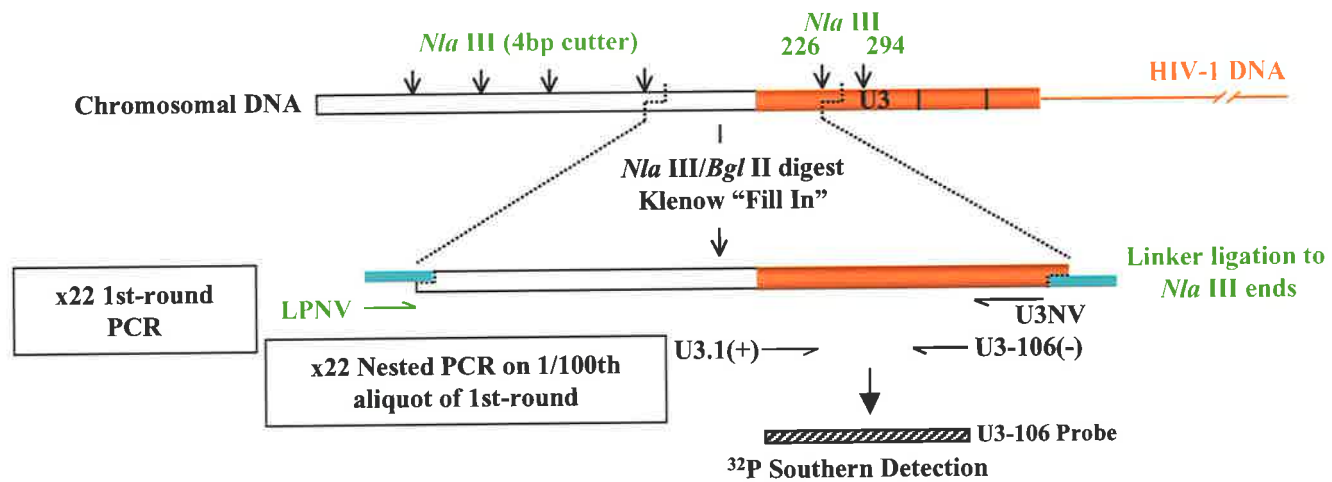
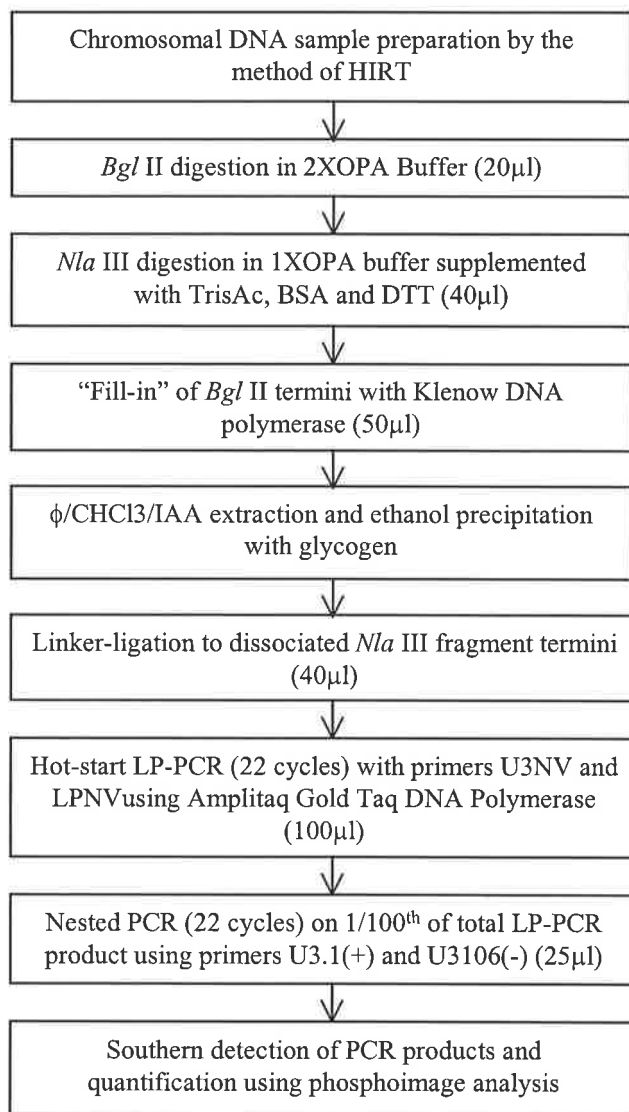


Figure 4.33

Flow-chart showing the key steps of the finalised nested LP-PCR protocol.



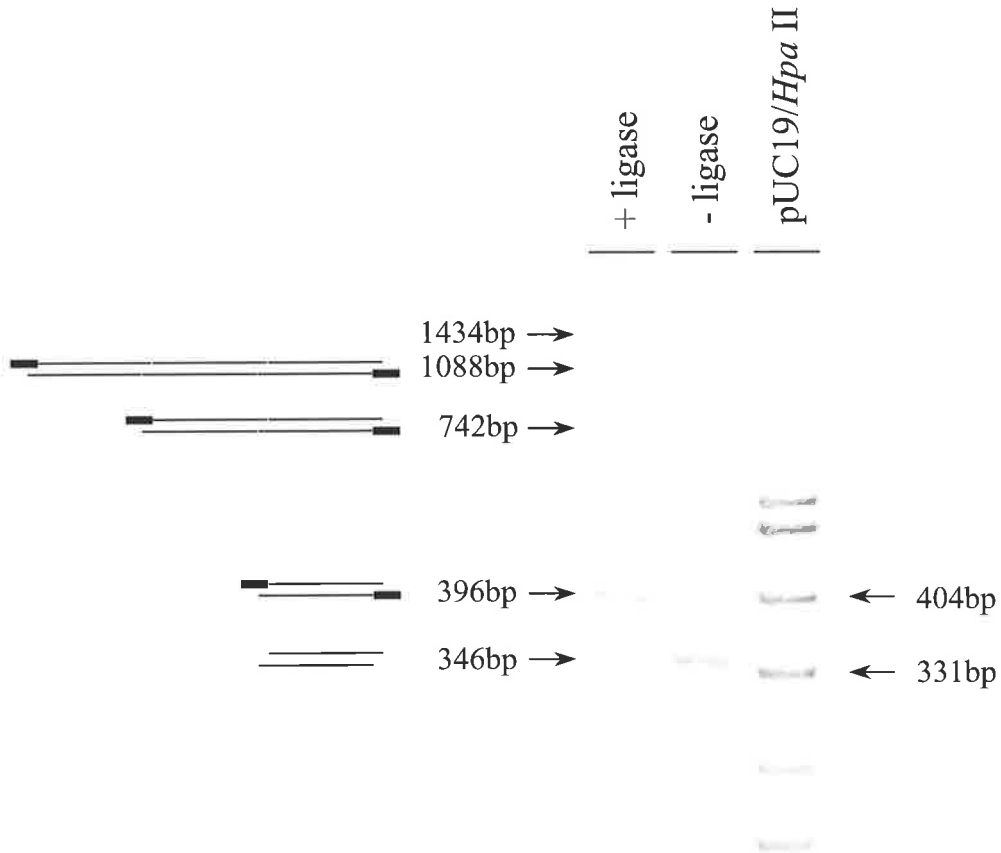
selective nature of the assay for integrated HIV DNA forms and the level of sensitivity achieved in a background of 1.2 μ g (2×10^5 cell-equivalents) of HuT-78 chromosomal DNA is presented below.

4.9.1 Linker Ligation Efficiency

The sensitivity of LP-PCR was expected to heavily depend on the efficiency with which linkers ligate to *Nla* III termini. Although the conditions of linker ligation had been optimised with respect to the type and concentration of the ligase used, and the reaction volume, these experiments had been performed using small amounts of target DNA (10^4 copies of pBSHIV_{HXB2}) in the absence of background HuT-78 DNA. Therefore, to confirm that the efficiency of linker ligation to *Nla* III termini approached 100% in the presence of large amounts of DNA, the addition of linkers to 1.2 μ g of a purified 346bp DNA fragment (generated by the cleavage of pBluescript with *Nla* III) was assessed. Following ligation (under conditions identical to those used in the LP-PCR procedure), the resulting DNA species in each sample was analysed by Southern hybridisation techniques using the T7 oligonucleotide as an internal probe (Table 2.1). In contrast to a control reaction performed without ligase, fragments devoid of linkers could not be detected in the ligated sample even after long exposure times (Fig. 4.34). Following PhosphorImage quantification of all bands (see appendix 4.1), the signal intensity of the band corresponding to the monomeric *Nla* III fragment with attached linkers (396bp product) accounted for greater than 50% of the signal obtained in the unligated control sample (346bp product). The additional bands detected in the ligated sample corresponded to dimers and higher order concatamers of *Nla* III fragments (up to 4 fragments in length, all with linkers ligated to termini) indicating that inter-molecular ligation of *Nla* III fragments was also occurring. As an internal check, the sum of individual band intensities in the ligated sample was closely comparable to the signal intensity of the unligated product (see Appendix 4.1). It is worth noting, that while attempts were made to create conditions that minimised the extent of inter-molecular ligation of *Nla* III fragments, concatamers of up to 8 *Nla* III fragments (assuming an average size of 256bp and an AmpliTaq extension rate of 2kb/min) will still be amplified by the LP-PCR assay when a 1 minute extension time is used.

Figure 4.34 Efficiency of Linker-Ligation to DNA Fragments Produced Following *Nla* III Digestion

A purified *Nla* III fragment (346bp) from pBluescript was used to assess the efficiency of linker ligation under conditions identical to those used in the LP-PCR procedure. Ligation products from reactions performed in the presence or absence of ligase are indicated (arrows), and the proposed structures of the monomeric (396bp), dimeric (742bp) and trimeric (1088bp) *Nla* III fragments with ligated linkers at each end are indicated on the left side. Signals were quantified using Image Quant software (see Appendix 4.1).



4.9.2 Selectivity of LP-PCR for Integrated HIV-1 DNA

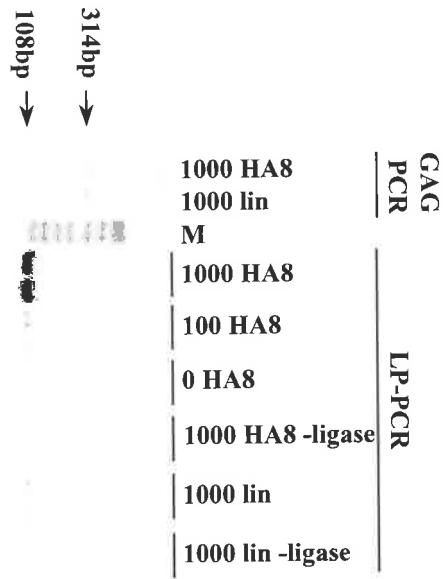
4.9.2.i Selectivity in a cell-free assay

In order to assess the selective nature of LP-PCR for integrated HIV DNA forms it was necessary to generate stocks of both the HA8 standard and the linear construct which contained precisely the same amount of HIV-1 DNA as determined by GAG PCR. Following dilution, both HA8 and “lin” samples were subject to GAG PCR (see section 2.2.5.iii) and the results are presented in figure 4.35 (left side). Clearly, little difference existed with respect to the HIV-1 DNA content between these two samples. Upon PhosphorImage analysis, the HIV-1 DNA signal within the 1000 “lin” sample was found to be marginally higher than that in the 1000 HA8 sample (data not shown). Therefore, the measure of selection efficiency against the “lin” form observed in the subsequent LP-PCR amplification experiment presented below represents a conservative estimate of the selective nature of the assay.

To estimate the sensitivity of LP-PCR for linear HIV DNA, 1000 copies of the engineered HIV-1 linear construct (normalised against the HA8 standard by GAG PCR), and dilutions of the HA8 standard, were subjected to LP-PCR. The signal intensity arising from amplification of the linear construct fell below the signal intensity observed after amplification of 100 copies of the HA8 standard (Fig. 4.35, right side). Phosphorimage analysis was used to quantify band intensities more precisely and to generate a standard curve from the 0, 100 and 1000 HA8 standards (see Appendix 4.2). Using this curve, the signal seen with the 1000 lin sample was found to correlate to a signal intensity of $\approx 7.5\%$ that of the 1000 HA8 signal. This indicated that LP-PCR was between 10 and 20-fold more specific for integrated HIV-1 DNA than the linear HIV-1 DNA form. When coupled with the HIRT chromosomal extraction technique, which removes $>80\%$ of the extrachromosomal forms (114, 248), the sensitivity of the final technique for all extrachromosomal forms was expected to be $<1.5\%$ of the sensitivity for integrated HIV DNA.

Figure 4.35 Selectivity of LP-PCR

HA8 copy-number standards spiked onto 2×10^5 cell-equivalents of HuT-78 chromosomal DNA were used in these experiments. Control reactions performed in the absence of ligase are indicated (-ligase). DNA size markers (pUC19 plasmid DNA restricted with the *Hpa* II restriction enzyme) are indicated (M). LP-PCR analysis on varying dilutions of the HA8 standard and 1000 copies of the construct mimicking the HIV linear extrachromosomal DNA species (lin). The copy number of the linear construct relative to the HA8 standard was confirmed by GAG PCR (first two lanes, left side). Band intensities were quantified using Image Quant software (see Appendix 4.2).



4.9.2.ii Selectivity following infection in cell-culture

To further confirm the above findings using DNA prepared from cells undergoing acute HIV infection, HuT-78 cells were infected with HIV_{HXB2} (0.5 TCID₅₀ units per cell; see section 2.2.1) in the presence or absence of the integration inhibitor L-731,988 (111). The inhibitor of reverse transcription, 3TC (used at a concentration of 10 μ M), served as a negative control lacking any HIV DNA forms. Preparations of chromosomal DNA were made by the method of HIRT (114) and subjected to PCR amplification of the β -globin gene (see section 2.2.5.i) to determine the cell-equivalent DNA content of each sample. Following analyses of 100 cell-equivalents of cellular DNA using the GAG-PCR (see section 2.2.5.iii) and LP-PCR (see section 2.2.5.v) protocols, strong signals corresponding to total and integrated HIV DNA were observed by 26h p.i. in drug-free cultures, respectively (Fig. 4.36A, B and C). As expected, cultures infected in the presence of 3TC were negative for both total and integrated HIV DNA. The accumulation of integrated HIV DNA by 26h p.i. was abolished by L-731,988 (Fig. 4.36B), while the accumulation of extrachromosomal HIV DNA was unaffected (Fig. 4.36A). In addition to confirming the activity of L-731,988 in culture, this result clearly demonstrated the specific nature of LP-PCR for integrated HIV DNA as any extrachromosomal DNA contaminating the preparations of chromosomal DNA did not significantly contribute to the final signal observed.

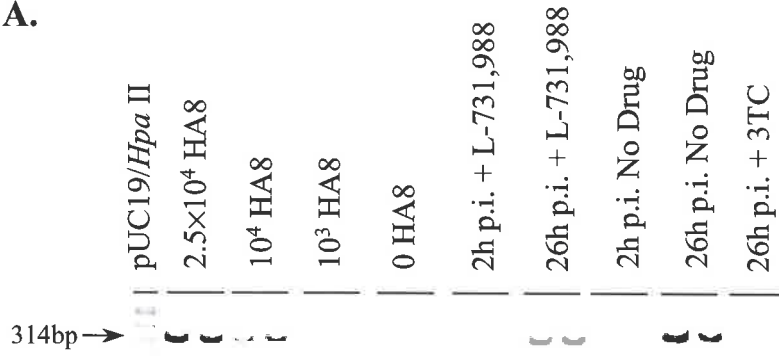
4.9.3 Sensitivity of LP-PCR

In order to evaluate the sensitivity of the LP-PCR procedure, increasing dilutions of the HA8 standard was spiked into 2×10^5 cell-equivalents (1.2 μ g) of HuT-78 chromosomal DNA. When subject to LP-PCR analysis, these preparations routinely exhibited signals arising in samples containing 10 copies of the integrated standard (Fig. 4.37). Due to the intensity of the signals achieved at this dilution and since this was the smallest amount of the HA8 standard tested, it is likely that the sensitivity of the assay under these conditions (see section 2.2.5.v) may be considerably greater. However, since higher levels of sensitivity were not required for either the analysis of HIV-1 integration kinetics (see Chapter 5) or anti-integration drug analyses (see Chapter 6) in cell-culture, a precise measurement of the sensitivity was not undertaken. The high level of

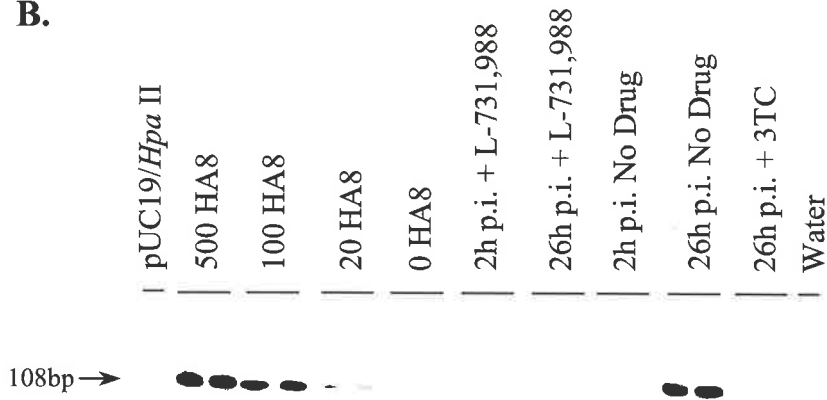
Figure 4.36

Analysis of viral DNA accumulation following cell-free infection performed in the presence or absence of inhibitors. Hut-78 cells were infected using the centrifugal enhancement protocol at 0.5 TCID₅₀ units per cell and cellular DNA prepared at 26h p.i. as described (see Materials and Methods). 3TC and L-731,988 were used as specific inhibitors of reverse transcription and integration, respectively. **A.** Integrated HIV DNA accumulation as measured by LP-PCR performed on HIRT pellet (chromosomal) DNA preparations. **B.** Total reverse transcribed DNA as measured by GAG-PCR performed on combined HIRT supernatant (extrachromosomal) and HIRT pellet (chromosomal) DNA samples. **C.** Graphical representation of the accumulation of integrated DNA. Data was obtained by PhosphorImage analysis of the bands presented in **A.**

A.



B.



C.

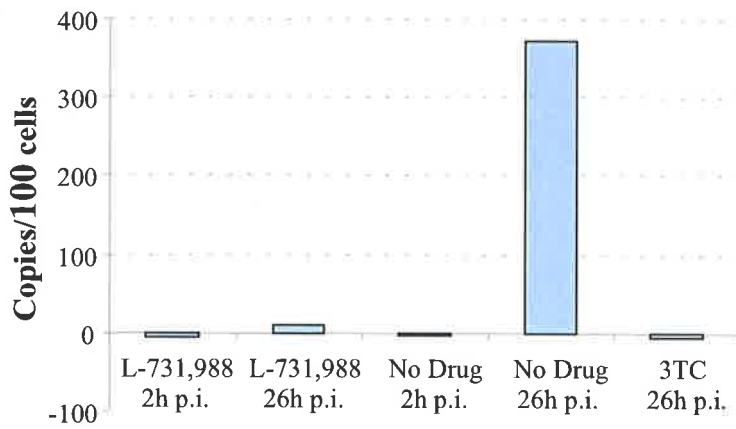
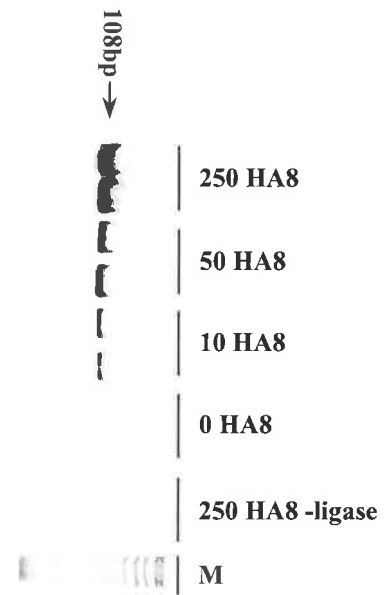


Figure 4.37 Sensitivity of LP-PCR

Sensitivity of LP-PCR as judged by amplification of the HA8 standards.

HA8 copy-number standards spiked onto 2×10^5 cell-equivalents of HuT-78 chromosomal DNA were used in these experiments. Control reactions performed in the absence of ligase are indicated (-ligase). DNA size markers (pUC19 plasmid DNA restricted with the *Hpa* II restriction enzyme) are indicated (M).



sensitivity achieved using the LP-PCR assay suggests that minimal loss of target DNA template had occurred during sample preparation and that each step within the LP-PCR procedure had been sufficiently optimised. Furthermore, since control reactions in which ligase was not present (Fig. 4.37, 250 HA8 -ligase) resulted in only a faint signals, the presence of all bands observed were highly dependent on the successful ligation of LPNV to the termini produced by *Nla* III digestion. In addition, this control indicates that asymmetric amplification of target template sequence at either the 5'LTR or 3'LTR by U3NV during the LP-PCR step was not significantly contributing to the final signal observed following the nested PCR.

4.10 Comparison of LP-PCR with the nested-*Alu* PCR

Protocol

4.10.1 The nested-*Alu* PCR protocol for the Detection of Integrated HIV DNA

The nested-*Alu* PCR technique for detecting integrated HIV DNA was first described by Chun and co-workers (49) and is outlined in section 1.7.3. Briefly, a primer designed to anneal within highly conserved human *Alu* repeat elements (approximately 9×10^5 elements/haploid genome) is used with an HIV-specific primer to mediate amplification of the left-side viral LTR and upstream chromosomal sequence (Fig. 1.14). Amplified products are then diluted and re-amplified using a set of nested primers. However, to avoid primer-annealing within both viral LTR regions and minimise non-specific asymmetric amplification, we chose to perform the 1st-round PCRs using the PBS-659(-) primer (Table 2.1) in exchange for the *Alu*-LTR 3' primer (49, 248). In addition, 1/1000th (instead of 1/400th) of the 1st-round PCR was used in a 20-cycle nested PCR to minimise the potential contribution from input template DNA amplification by second-round PCR alone, to the final signal.

4.10.2 Comparison of LP-PCR and nested-*Alu* PCR: Clonal Cell Line DNA

While the nested-*Alu* PCR protocol had been shown to efficiently detect the integrated DNA within the ACH-2 and U1 cell lines (49), there had been reports that this procedure was unable to detect HIV DNA within the 8E5 cell line (Secondo Sonza, personal communication). To assess the ability of each technique to detect different integration events, chromosomal preparations of the clonal cell lines ACH-2 and 8E5 (each containing 1 copy of integrated provirus) were made. Chromosomal DNA preparations were equated for viral DNA by GAG PCR (see section 2.2.5.iii) quantification (Fig. 4.38A, left side) and then subjected to the LP-PCR and nested *Alu*-PCR procedures to detect integrated HIV DNA. While integrated DNA within the ACH-2 cell line was efficiently detected, the nested *Alu*-PCR method was unable to detect integrated DNA in the preparation of 8E5 chromosomal DNA (see Fig. 4.38A, right side). In contrast, the LP-PCR procedure allowed the efficient detection of integrated DNA present in both cell lines (Fig. 4.38A, middle).

Since the chromosomal sequence immediately adjacent to the integrated provirus in each of these cell lines was known (see section 4.6.5), the location of the first *Alu* repeat element upstream of the integrated provirus was determined (Fig. 4.38B). Using BLAST software, the integrated proviral DNA within the ACH-2 and 8E5 cell lines was mapped to chromosomes 7 and 13, respectively. A search (using the *Alu* 164 primer sequence) in the upstream chromosomal DNA revealed the position of all *Alu* repeat elements immediately upstream of each integrated provirus. The nearest *Alu* element upstream of proviral DNA in the correct orientation for *Alu*-PCR amplification was positioned $\approx 4\text{kb}$ away for ACH-2 cells, but $\approx 12\text{kb}$ away for 8E5 cells. We therefore proposed that the *Alu*-PCR protocol used by us and others (49) was not optimised to permit the amplification of sequences of 12kb or more and would therefore be unable to detect integrated HIV DNA within the 8E5 cell line. Although we believe this to be the most likely explanation, it is also possible that the 5'→3' exonuclease activity associated with the *rTth* DNA polymerase (183) used in this assay may have also contributed to the absence of signal arising from 8E5 chromosomal DNA. A 5'→3' exonuclease activity is present in most thermostable DNA polymerases and directs the 5'→3' hydrolysis of any

Figure 4.38 Comparison of LP-PCR with Nested-Alu PCR

Comparison of PCR detection of integrated HIV DNA by LP-PCR and *Alu*-PCR. **A.** Chromosomal DNA was isolated from the ACH-2 or 8E5 cell lines and equated for total HIV DNA content by GAG-PCR (314bp band on left marked with thin arrow). Sizes of expected bands for LP-PCR (measuring integrated HIV DNA; middle panel) are given on the left side (104bp fragment marked by thick arrow), while the expected size of the product obtained following *Alu*-PCR (also measuring integrated HIV DNA; right panel) is indicated on the right side (351bp fragment). **B.** Position and orientation of *Alu*164 primer sequences in cellular DNA up to 20kb upstream of integrated HIV DNA in the ACH-2 and 8E5 cell lines. Data was obtained by sequencing cellular DNA upstream of the integration sites (Dr. Raman Kumar, unpublished) and using this to search for homologous sequences in the human genome with BLAST software. In both cases (ACH-2 and 8E5), sequences within the human genome displaying 100% homology to the sequences obtained by Dr. Raman Kumar (unpublished) were identified.

ACH-2 Sequence:

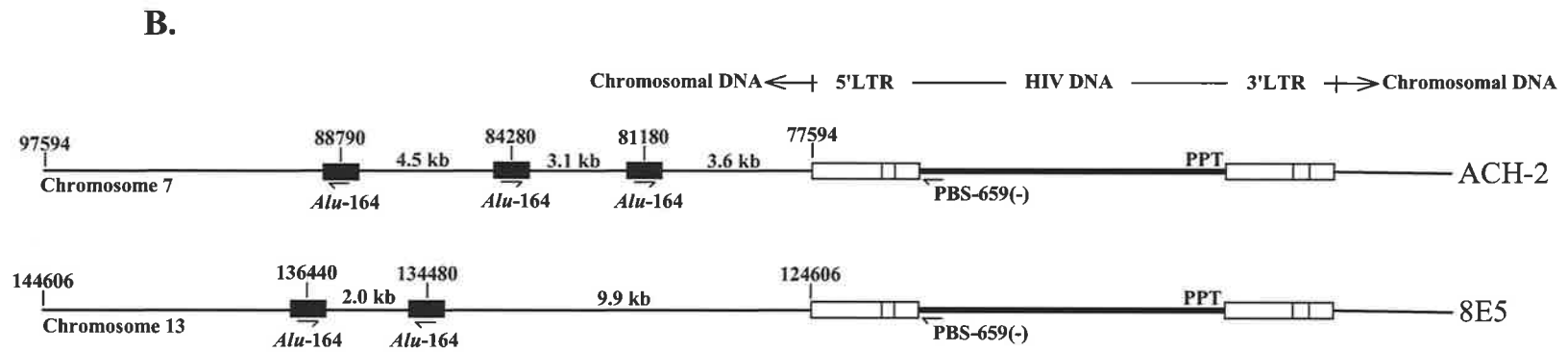
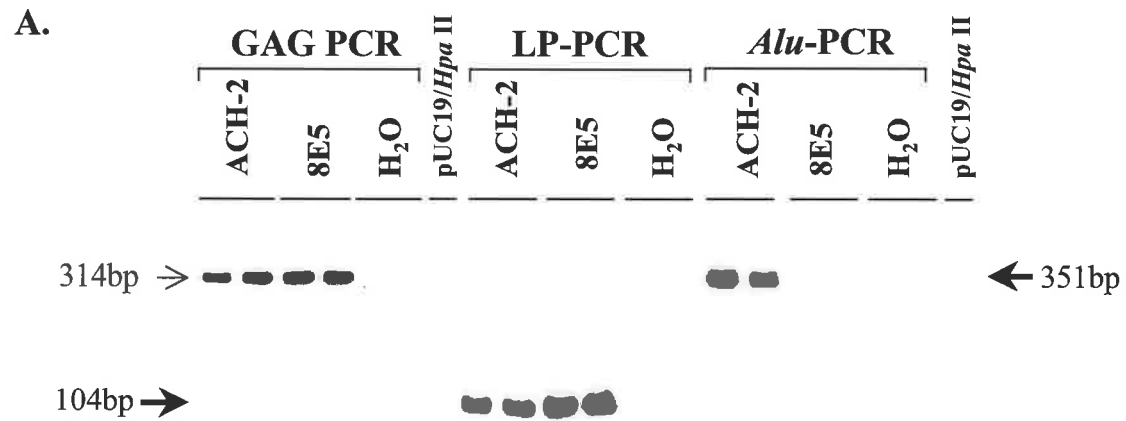
Homo sapiens clone RP11-162O1 on chromosome 7.

Accession number: AC083863

8E5 Sequence:

Homo sapiens clone RP11-631G24 on chromosome 13.

Accession number: AL391374



DNA ahead of a nascent DNA chain. We confirmed the presence of this activity in *rTth* polymerase by performing a PCR, where one of the two flanking primers also acted as an internal primer, and showed that the larger of the two possible PCR products predominated (data not shown). In the context of *Alu*-PCR, this activity would be expected to direct the digestion of annealed *Alu*-164 primers (and extended *Alu*-164 primers) ahead of nascent DNA chains throughout the genome. Thus, the presence of the 5'→3' exonuclease activity may determine the mechanism of proviral DNA amplification by the *Alu*-PCR protocol. For example, in each cycle of a PCR, extension of an *Alu*-164 primer bound at position 84280 in chromosome 7 (see Fig. 4.38B, ACH-2) would ultimately lead to the digestion of primers bound at position 81180 in the chromosome 7 sequence. Therefore, amplification of the 5'-U3 HIV DNA sequence in the ACH-2 cellular DNA is likely to be mediated by the extension of primers bound to the more distal of the two *Alu* repeat elements upstream of the site of integration. It follows therefore, that the presence of multiple *Alu* repeat elements (in the correct orientation for successful *Alu*-PCR amplification) upstream of the site of proviral insertion may compromise the efficiency of amplification, as the thermostable polymerase directing the synthesis of distal nascent DNA chains would be constantly digesting more proximal DNA. In conclusion, successful *Alu*-PCR amplification of randomly integrated HIV DNA may be less dependent on the presence of an *Alu* repeat element in the correct orientation immediately upstream of the site of proviral integration, and more dependent on the location (distance and orientation) of *Alu* repeat elements further upstream. The highly complex combinations of priming and degradation events that will occur in *Alu*-PCR when *rTth* DNA polymerase is used would be expected to limit the number of integration events that can be successfully amplified. It is worth noting, that unless a strand displacement activity is present, long-range thermostable polymerases lacking the 5'→3' exonuclease will allow the amplification of a maximum of 50% of all random integrants. This is because such enzymes will be unable to amplify HIV DNA integrated immediately adjacent to an *Alu* repeat element in the incorrect orientation (see for example, Fig. 4.38B, 8E5).

4.11 Discussion

In order to specifically measure integrated HIV proviral DNA, a novel PCR-based assay termed linker-primer PCR (LP-PCR) was developed that can theoretically detect 94% of all integrated forms while selecting against amplification of all three extrachromosomal forms in infected cells. It was demonstrated that this assay allowed the detection of 10 copies of integrated HIV-1 DNA in a background of 2×10^5 cell-equivalents of human chromosomal DNA while cross-detecting approximately 7.5% of extrachromosomal HIV DNA. When coupled with the HIRT chromosomal extraction technique, which removes >80% of the extrachromosomal forms (114, 248), the combined procedure is highly specific, with an estimated contaminating signal equivalent to <1.5% of all extrachromosomal forms present.

Very few applications measuring integrated provirus loads would be expected to require higher levels of sensitivity. In fact, to assess the kinetics of integrated HIV DNA accumulation in the presence or absence of drugs (see section 4.9.1.ii, Chapter 5 and Chapter 6), the assay was de-sensitised to allow the quantitative assessment of the large amounts of integrated HIV DNA present following HIV-1 infection in cell-culture. Another potential application of the LP-PCR assay is the assessment of the integrated HIV DNA load in populations of cells isolated from patients on HAART to determine sites of viral persistence. All primers used in the linker-primer procedure (U3NV, U3PNV and U3.1(+); see Table 2.1) were designed to anneal within highly conserved regions of the HIV genome. Furthermore, the *Nla* III site present at position 226 in the HIV_{HXB2} genomic sequence (accession number K03455) is perfectly conserved across isolates within clade B (55/55 isolates compared; data not shown) and between the consensus sequences for all known clades. Together with the high levels of sensitivity and specificity for integrated HIV DNA using the LP-PCR procedure, these design features imply that this protocol is well suited for the identification of sites of viral persistence (as judged by the continued presence of integrated HIV DNA) within patients.

Recently, integrated HIV DNA was identified by a nested-*Alu* PCR protocol to persist within highly purified populations of CD4⁺ memory T cells isolated from patients

undergoing successful HAART that had displayed undetectable plasma virus levels for up to 2 years (46, 49). The mean level of integrated HIV DNA within these memory T cells was estimated to be 324 copies/ 10^6 cells using integrated HIV DNA within the ACH-2 cell line as a copy-number standard. However, our findings indicate that although the nested-*Alu* PCR procedure is able to efficiently detect integrated HIV DNA within the ACH-2 cell line, it will be unable to detect a significant proportion of random integration events (such as those occurring following infection). This therefore implies that the levels of integrated HIV DNA within populations of CD4⁺ memory T cells observed by Chun and colleagues is likely to represent an underestimation of the real levels.

To correct for the discrepancy outlined above, copy-number standards that precisely mimic the random nature of HIV DNA integrants following infection (rather than using clonal cell line DNA) can be used to accurately quantify the proviral load as measured by nested-*Alu* PCR. During the preparation of this manuscript, a short article was published in which the accumulation of integrated HIV DNA levels following infection was monitored using a modified version of the nested-*Alu* PCR protocol (32). In this study, the copy-number standards used consisted of chromosomal DNA isolated from cells two weeks after infection with HIV in cell culture. Although the potential for the preferential outgrowth of cells containing particular integrants was not addressed, integrated DNA present within this standard preparation was concluded to be random with respect to cellular sequence. Since extrachromosomal DNA forms were expected to have degraded by this time, the integrated HIV DNA load in this sample was estimated by GAG PCR. Although this approach would allow the accurate quantification of integrated HIV DNA by nested-*Alu* PCR following infection, the efficiency with which the total number of HIV integrants were detected would remain unchanged. In conclusion, the nested-*Alu* PCR procedure is likely to allow the accurate quantification of integrated HIV DNA when the correct copy-number standards are used. However, the ability of the LP-PCR procedure to detect a greater amount of HIV integrants may ultimately make it a more appropriate protocol for the highly sensitive detection of randomly integrated HIV DNA.

Chapter 5

Kinetics of HIV Nucleic Acid Accumulation Following a One-Step Infection of T Cells

5.1 Introduction

5.1.1 Background

Integration of viral DNA into the host cells' genome is an essential part of the HIV life-cycle (76, 151, 218, 238). This process is unique to retroviruses and is primarily mediated by the virally-encoded integrase enzyme (IN). Although the process of reverse transcription generating viral cDNA has been relatively well characterised in cells (131), the efficiency with which the newly synthesised HIV DNA is inserted into the cellular chromosome is poorly defined. Following reverse transcription, one of four main fates awaits the newly synthesised viral DNA: 1) integration in a colinear fashion into the cellular chromosome; 2) circularisation via recombination generating a circular form containing one long terminal repeat (1-LTR circle); 3) circularisation via direct head to tail ligation generating a circular form containing two long terminal repeats (2-LTR circle); 4) remaining as a linear viral DNA molecule (see section 1.2.4 and Fig. 1.6). A number of studies (using both cell-free and cell-to-cell infection models) have indicated that both the linear and 1-LTR DNA species represent the major extrachromosomal HIV forms following infection in cells, while the 2-LTR circular DNA species is present at very low levels (8, 160). The relatively detailed knowledge available about extrachromosomal viral DNA compared to integrated viral DNA is due to the relative ease with which the free forms can be detected. Each of the extrachromosomal viral DNA forms can be readily detected using Southern hybridisation protocols performed on DNA extracted from cells. Moreover, the 1-LTR and 2-LTR viral DNA species can be individually detected by variants of the polymerase chain reaction (PCR) using primers flanking the LTR region and the LTR/LTR-junction region, respectively. However, to date there has been no direct assessment of the accumulation of integrated HIV DNA over time following either cell-

free or cell-to-cell transmission of HIV. This has been primarily due to the absence of a technique that allows the specific detection of integrated HIV DNA.

5.1.2 Aims

As outlined in chapter 4, two PCR-based protocols were developed (LP-PCR and a modified nested-*Alu* PCR) that allow the highly sensitive and specific detection of integrated HIV DNA present within preparations of cellular chromosomal DNA. In this study, each of these assays was used to monitor the accumulation of integrated viral DNA following a near one-step, synchronous cell-free HIV infection of cultured T cells. In particular, we hoped to determine the kinetics of HIV DNA integration into the cellular chromosome, and the proportions of the total reverse transcribed DNA within cells that go on to successfully integrate. To allow a comprehensive profile of viral DNA species following infection to be established, additional PCR procedures were used to assess the accumulation of total viral DNA (GAG-PCR) and 2-LTR viral DNA (2-LTR-PCR) forms.

5.2 Establishing a Synchronous Cell-Free Infection Model

To assess the kinetics of viral DNA accumulation, a high-multiplicity, synchronous, one-step infection model was established in which the majority (if not all) cells were initially infected. This was expected to allow the accumulation of all viral DNA species to levels readily detected by the methods available and to reduce the numbers of uninfected cells available for second-round infection by progeny virus.

5.2.1 Synchronous Entry

A highly synchronous infection is defined as one in which the majority of infection events initiate at the same time and progress through the life cycle at approximately the same rate. When virus is incubated over long periods with susceptible cells, infections initiate at random over the time of viral exposure generating a staggered profile of the nucleic acids kinetics. To overcome this, virus was exposed to cells for a short time-period before extensive washing to remove any particles not bound to the cell surface.

In addition, virion adsorption was performed at a temperature that was non-permissive for viral entry (4°C). The subsequent incubation of cells and virus at 37°C (in the presence of centrifugal enhancement; see section 5.2.3 below) was expected to allowed the coordinated entry of bound virus into cells thereby further synchronising the infection process.

5.2.2 Optimising the Infection Efficiency: DEAE-Dextran Pre-treatment of HuT-78 Cells

In order to optimise the infection efficiency, the use of DEAE-dextran was investigated. Although the pre-treatment of cells with DEAE-dextran polymer has been reported to increase the efficiency with which HIV interacts with the cell surface (147), its' effectiveness remains controversial (108). To examine this in our system, the TCID₅₀ of a freshly isolated virus stock was determined with or without DEAE-dextran pre-incubation. The infection protocol and methods used to generate a TCID₅₀ value for a viral stock are outlined in section 2.1.2. Following infection of HuT-78 cells with serial dilutions of freshly isolated HIV_{HXB2} virus stock, a higher TCID₅₀ value was obtained in the absence of DEAE-dextran pre-treatment (Table 5.1). This indicated that the DEAE-dextran pre-treatment of target cells did not enhance viral infectivity under our infection conditions.

5.2.3 Centrifugal Enhancement of Virus Infection

The use of a high-speed spin to enhance the number of successful infection events following the mixing virus either HIV or herpesvirus type 6 (HHV-6) virus with susceptible cells was first described by Pietroboni and co-workers (197). This method (termed centrifugal enhancement) was shown to increase the infectivity of a virus stock by a factor of approximately 10.

Initially, the efficiency with which an "in-house" centrifugal enhancement protocol (Li, *et al.*, unpublished) enhanced infection of HuT-78 cells (MOI of 1 TCID₅₀ unit per cell) was assessed. This protocol involved a two-step process in which initially half the virus inoculum was centrifuged with cells at 20°C for 25 minutes at 2000×g. Supernatant

Table 5.1 TCID₅₀ Determined With/Without DEAE-dextran

Cells (either pre-treated with DEAE-dextran (upper rows) or not pre-treated (lower rows)) were exposed to dilutions of neat virus stock and scored for the presence (+) or absence (-) of syncytia after 9 days. The TCID₅₀ (the highest dilution causing syncytia in 50% of inoculated cultures) was calculated according to a published method (102).

DEAE-dextran pre-treatment	Dilution of Virus Stock			
	1/1000	1/10000	1/100000	
	+	+	-	
	+	+	-	
	+	-	-	TCID₅₀ = 6.8×10⁴
	+	-	-	
	+	-	-	
	+	-	-	

No DEAE-dextran pre-treatment	Dilution of Virus Stock			
	1/1000	1/10000	1/100000	
	+	+	-	
	+	-	-	
	+	+	-	TCID₅₀ = 2.2×10⁵
	+	+	-	
	+	+	-	
	+	+	-	

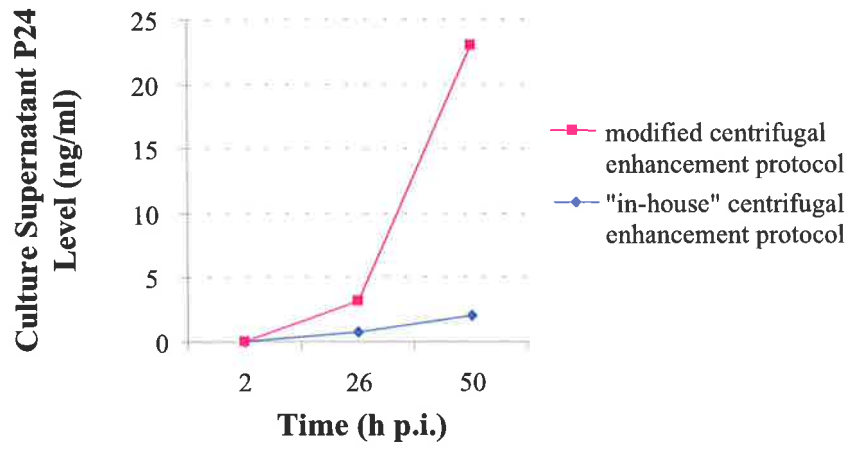
containing non-internalised virus was then removed, and the cells were resuspended in the remaining virus and centrifuged as before. The infection efficiency of this protocol was assessed by evaluating the levels of P24 released (indicating progeny virion release) into the culture supernatant at various times after infection. Using this protocol in duplicate assays, an average P24 level of approximately 0.8ng/ml was obtained at 26h p.i. (Fig 5.1A, "in-house" centrifugal enhancement protocol) which increased two-fold between 26h p.i. and 50h p.i.. In an attempt to increase the infection efficiency, the centrifugal enhancement protocol was modified to more closely resemble that published previously (197) by centrifuging the entire virus inoculum with cells (once only) at 2500×g for 1hr at 37°C. In addition, cells were allowed to recover in fresh medium at 37°C for 15 minutes immediately after centrifugation before the extensive washing steps to remove non-internalised virus were performed. Using this protocol (MOI of 1 TCID₅₀ unit per cell), levels of P24 equivalent to approximately 3ng/ml were observed in the culture supernatant (Fig. 5.1A, modified centrifugal enhancement protocol) by 26h p.i.. Furthermore, a substantial level of ongoing virion release from infected cells into the culture supernatant was evident between 26h p.i. and 50h p.i.. Taken together, these results indicated that the modified centrifugal enhancement protocol directed a more efficient infection compared to the previously published "in-house" protocol. A detailed description of the modified centrifugal enhancement protocol is presented in section 2.2.1.

To establish whether increasing the multiplicity of infection (MOI) translated to a higher degree of cellular infection, the MOI was increased from 1 TCID₅₀ unit per cell to 5 TCID₅₀ units per cell. When cultures that were infected using the revised centrifugal enhancement protocol, and either an MOI of 1 TCID₅₀ unit per cell or 5 TCID₅₀ units per cell, were assessed for supernatant P24 levels at 26h p.i., no significant differences were observed (Fig. 5.1 B). This result suggested that an input virus inoculum equivalent to an MOI of 1 TCID₅₀ unit per cell was sufficient to give maximum infection levels using the revised centrifugal enhancement protocol, and that other factors (*eg.* cellular proteins required for infection) were limiting virus production when an MOI of 5 TCID₅₀ unit per cell was used.

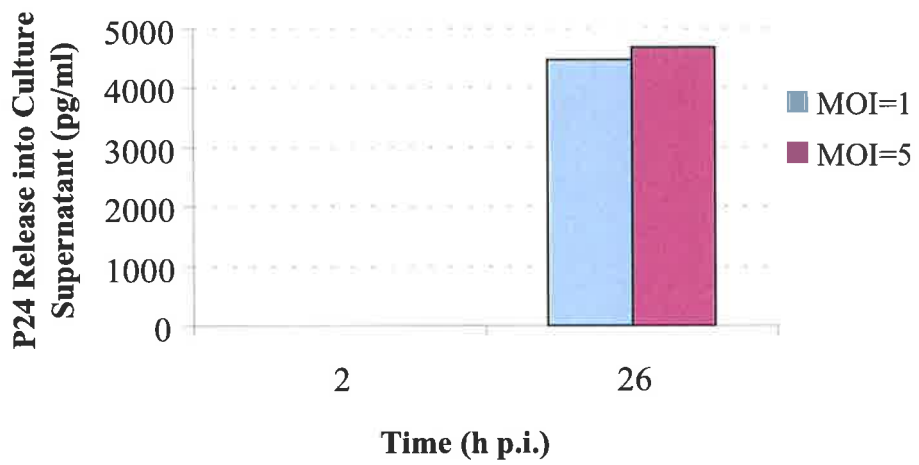
Figure 5.1 Assessing Effects of Centrifugal Enhancement and MOI on HIV Infection of HuT-78 Cells

The efficiency with which HIV was able to productively infect HuT-78 cells was measured by determining P24 levels in culture supernatants at various time points after infection. (A) Graphical representation of P24 data following infection performed using either the modified centrifugal enhancement or the “in-house” centrifugal enhancement protocol (see body text for an overview of these protocols). (B) Graphical representation of P24 data showing the effect increasing the multiplicity of infection (MOI) from 1 to 5 had on progeny virion release following HIV infection of HuT-78 cells using the modified centrifugal enhancement protocol.

A.



B.



5.3 Kinetics of Integrated, Total and 2-LTR Viral DNA Accumulation Following Infection of T Cells

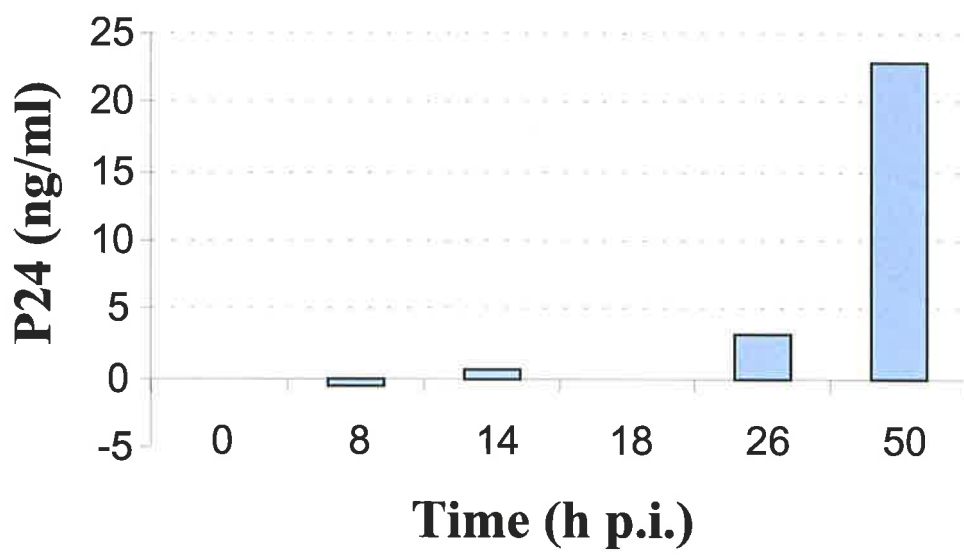
In order to study the kinetics of HIV DNA accumulation, the above one-step infection of HuT-78 cells with cell-free virus was used (see section 5.2). Cellular DNA was extracted by an optimised HIRT protocol (see section 5.3.2.i) at various times after infection and the levels of integrated, total and 2-LTR viral DNA species were determined using PCR-based assays. Since a primary aim of this study was to determine the time following infection when near full-length viral cDNA could be first detected and to correlate this with the time at which integrated HIV DNA could be first detected, extensive sampling early after infection was performed. Hourly harvests until 10h p.i. were collected, followed by sampling at 14, 18, 26 and 50h p.i. to identify the major trends associated with the accumulation of each species later in infection. It is worth noting, that since the centrifugal enhancement protocol required a minimum of 2h to complete, the first time-point ($t = 2\text{h p.i.}$) was the earliest that could possibly be taken. Therefore, the 2h p.i. time-point corresponds to the time at which cells were aliquotted into the wells of the culture trays following infection, and reflects the time elapsed after elevating cultures to 37°C following virus adsorption to cells.

5.3.1 Kinetics of Viral DNA Accumulation: P24 Release

Viral release into the culture supernatant following infection (as measured by P24 release) was evident by 26h p.i., indicating that one round of replication was complete by this stage (Fig. 5.2). Consequently, the end point of all viral DNA analysis (see below) was taken as 26h p.i. to ensure that any contribution to the various viral DNA species from secondary cell-free infection events was minimised.

Figure 5.2 HIV P24 release following infection of HuT-78 cells

P24 levels in culture supernatants were measured at various time points (h p.i.) following centrifugally enhanced infection of HuT-78 cells with HIV_{HXB2} (1 TCID₅₀ unit per cell).



5.3.2 β -globin Analysis and Adjustment of HIRT Pellet (Chromosomal) DNA Fractions

5.3.2.i *Purification of HIRT Pellets*

Cells were removed from culture wells at various times after infection and subjected to HIRT DNA extraction (see (114) and section 2.2.3). Briefly, this method allows the independent purification of the chromosomal DNA and the extrachromosomal DNA fractions. Extrachromosomal DNA predominantly fractionates within the HIRT supernatant, while higher molecular weight DNA precipitates with SDS to form the HIRT pellet. However, chromosomal preparations purified by the published technique routinely failed to allow the PCR amplification of target sequences (data not shown). Since SDS has been shown to be a potent inhibitor of thermostable polymerases (203), the absence of detectable signal was attributed to SDS contamination of the chromosomal pellets. Therefore, to minimise SDS contamination of chromosomal DNA, all ethanol precipitations were performed at room temperature in the presence of sodium chloride (rather than sodium acetate). Furthermore, chromosomal DNA was washed extensively with 70% ethanol at room temperature. This resulted in highly purified preparations of chromosomal DNA from which target DNA sequences could be efficiently amplified (see below). Since it was recognised that the absolute DNA recovery of each preparation may vary, samples were assessed and adjusted for β -globin content prior to PCR analyses.

5.3.2.ii *β -globin Content of Experimental Samples*

Purified chromosomal preparations were subjected to PCR amplification of the human β -globin gene to quantify and normalise the amounts of chromosomal DNA present within experimental samples. Following amplification, the β -globin content of sample DNA preparations were determined by comparison with known cell-equivalents of the HA8 standard (see section 2.2.4.i) amplified in a similar manner. The β -globin PCR protocol used to analyse chromosomal samples was optimised (by sequentially reducing cycle-numbers to 25 cycles) such that amplification of 50 cell-equivalents of the HA8

standard occurred in a highly linear fashion (compare Fig. 5.3A.i. (30 cycles, $R^2 < 0.90$) and 5.3B.ii. (25 cycles, $R^2 > 0.99$)). Under these limited cycling conditions, the DNA content of unknown samples giving signal intensities equivalent to 250 cell-equivalents of HA8 DNA could be estimated with reasonable accuracy ($R^2 > 0.96$; Fig. 5.3B.iii.).

Using the optimised (quantitative) β -globin PCR protocol, approximately 100 cell-equivalents (assuming 100% DNA recovery from 2×10^5 cells) of each of the experimental HIRT chromosomal preparations were assayed in duplicate against known amounts of the HA8 chromosomal standard (Fig 5.4A.). Samples were volume adjusted based on the results of initial β -globin PCR quantification and upon re-analyses of 50 cell-equivalents (single reactions), little variation between samples was observed (Fig. 5.4B.).

5.3.3 Mitochondrial DNA Analysis of HIRT Supernatant (Extrachromosomal) DNA Fractions

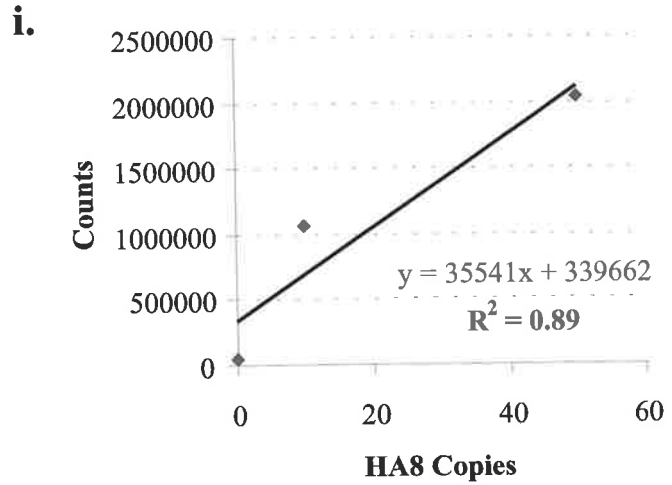
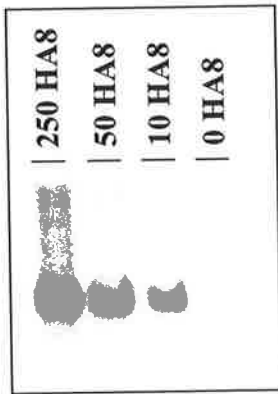
The efficiency with which extrachromosomal DNA was recovered within the HIRT supernatant fractions (see section 2.2.3 for extraction protocol) was assessed by analysing each sample for the presence of mitochondrial DNA. The semi-quantitative nature of the 23-cycle mitochondrial PCR protocol used was established by amplifying serial 5-fold dilutions of the HIRT supernatant sample extracted at 2h p.i.. When the resulting band intensities were quantified and graphed, a plateau in PCR signal intensities was evident indicating that saturation of PCRs had occurred when higher amounts (cell-equivalents) of sample DNA were amplified (Fig. 5.5A and B). However, when the signals obtained following amplification of the three most dilute samples (100 c.e., 20 c.e. and 0 c.e.) were graphed, a linear curve resulted (Fig. 5.5C; $R^2 > 0.9999$) indicating that PCR saturation was not occurring when the amount of DNA amplified was limited to ≤ 100 c.e. of DNA within HIRT supernatant fractions. Therefore, all future analyses determining the mitochondrial content of HIRT supernatant preparations were performed on approximately 100 cell-equivalents of DNA (assuming 100% DNA recovery from 2×10^5 cells) using a 23 cycle PCR (see section 2.2.5.ii).

Figure 5.3 Optimising the β -globin PCR

Optimisation experiments were performed on various dilutions of the HA8 integrated chromosomal standard. PCR products were detected by Southern hybridisation using the Glo probe (see Table 2.2). (A) 30 cycles of the β -globin PCR. (i.) Graphical analysis of band intensities in (A). (B) 25 cycles of the β -globin PCR. (ii.) Graphical analysis of band intensities corresponding to amplification of 50, 10 and 0 copies of the HA8 standard in (B). (iii.) Graphical analysis of band intensities corresponding to amplification of 250, 50, 10 and 0 copies of the HA8 standard in (B).

Counts (y axes of graphs) were obtained using ImageQuant software. The trendline curve equations ($y =$) and regression values (R^2) of each graph are shown.

A. 30 cycles



B. 25 cycles

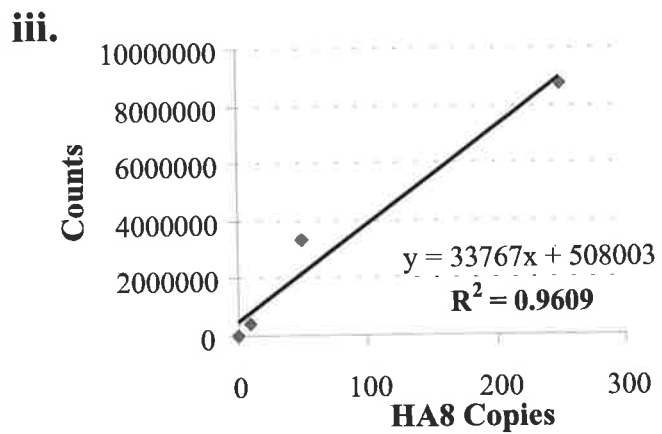
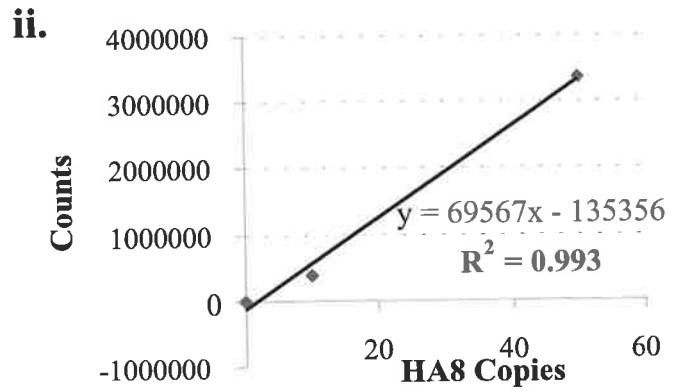
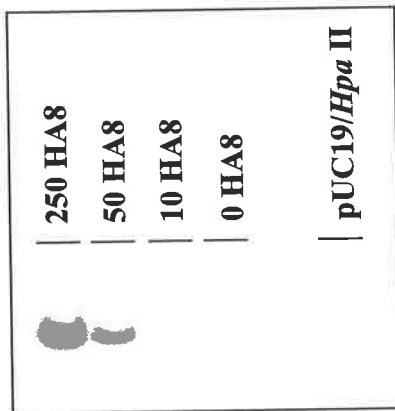
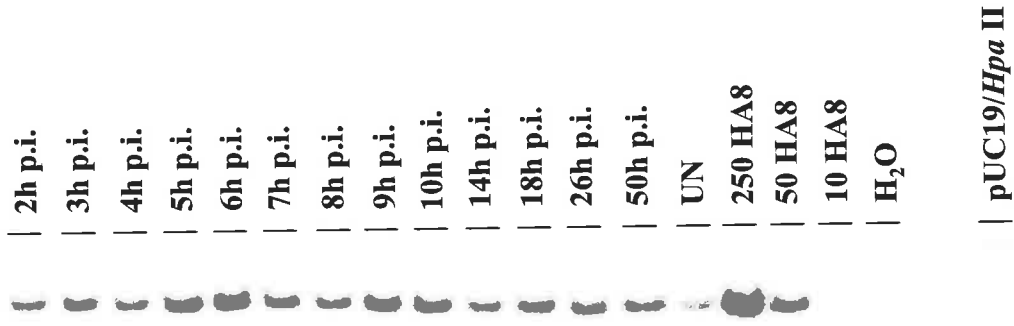


Figure 5.4 β -globin PCR Analysis of Experimental Samples

PCRs were performed on ≈ 50 cell equivalents of chromosomal DNA extracted from HuT-78 cells at various times after infection with HIV. Uninfected control infections (UN), standards (HA8) and DNA markers (pUC19/*Hpa* II) are indicated. PCR products were detected by Southern analysis using the Glo probe (see Table 2.2). **(A)** Initial analysis of β -globin content performed in duplicate. **(B)** Secondary analysis (single PCRs) of experimental samples following volume adjustment based on the results presented in **A**.

B.



A.

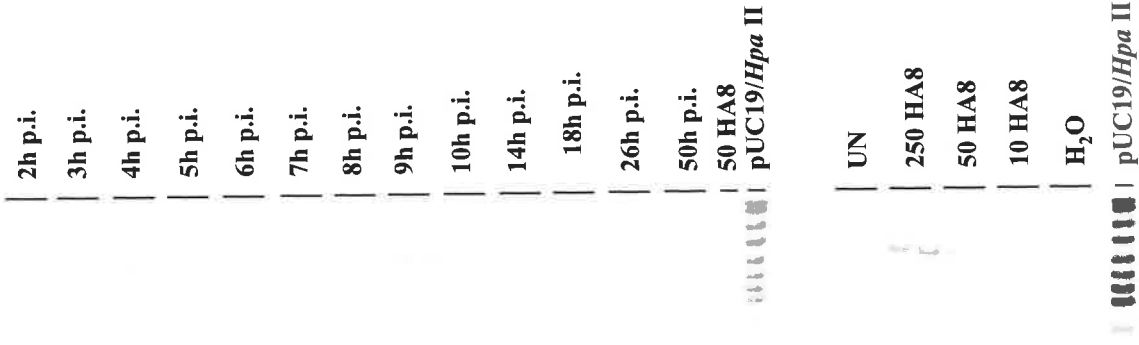
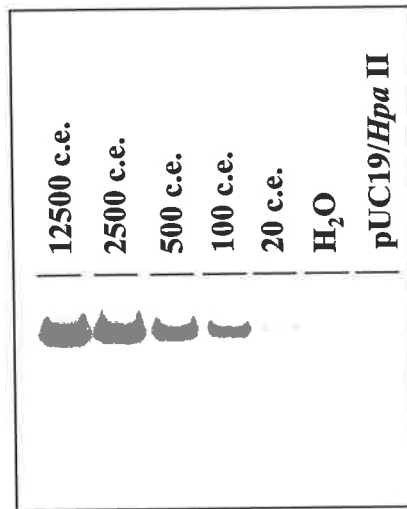


Figure 5.5 Optimisation of the Mitochondrial PCR

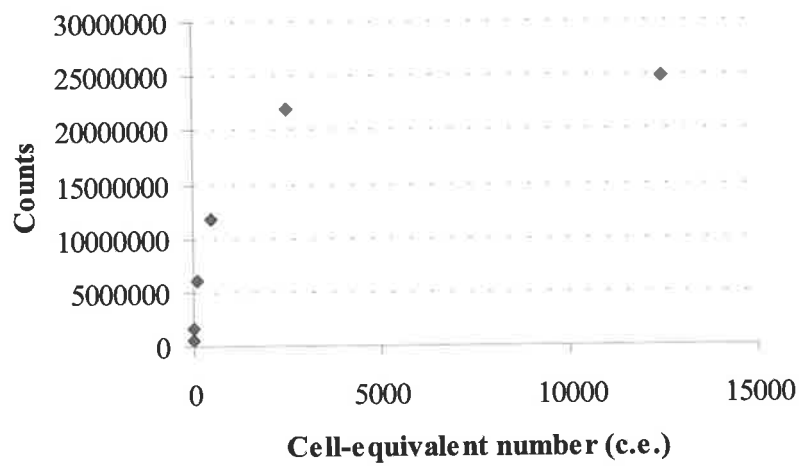
(A) Mitochondrial PCRs (23 cycles) were performed on various dilutions of the HIRT supernatant samples extracted at 2h p.i. to determine the highest amount of initial template DNA that was still amplified in a logarithmic manner. PCR products were detected by Southern analysis using the Mit probe (see Table 2.2). (B) Graph of all signal intensities from A. plotted against the corresponding cell equivalent number (c.e.) used. (C) Graph of the signal intensities obtained following amplification of 0, 20 and 100 cell equivalents of the 2h p.i. sample (see A.)

Counts (y axes of graphs) were obtained using ImageQuant software. The trendline curve equation ($y =$) and regression value (R^2) of the graph in C. is shown.

A.



B.



C.

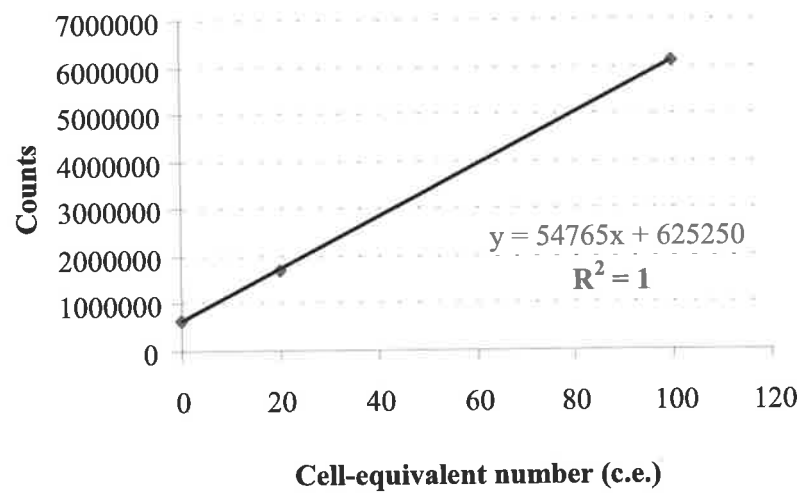
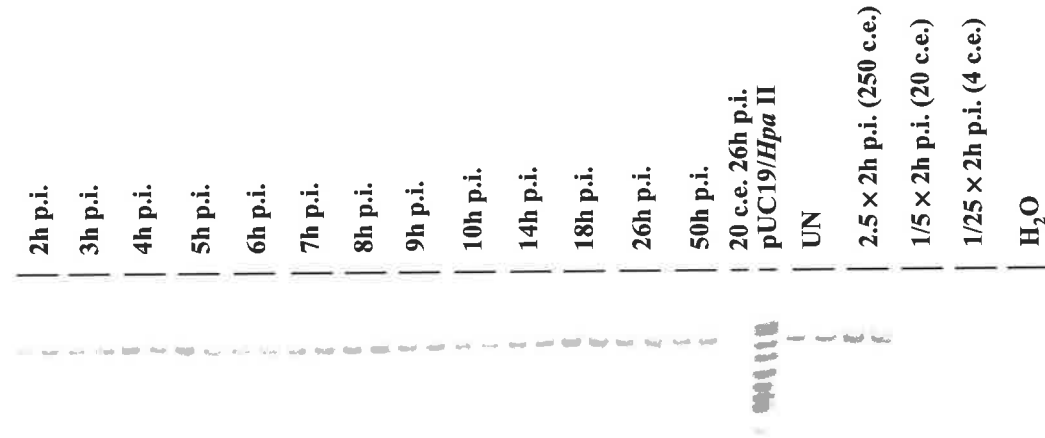


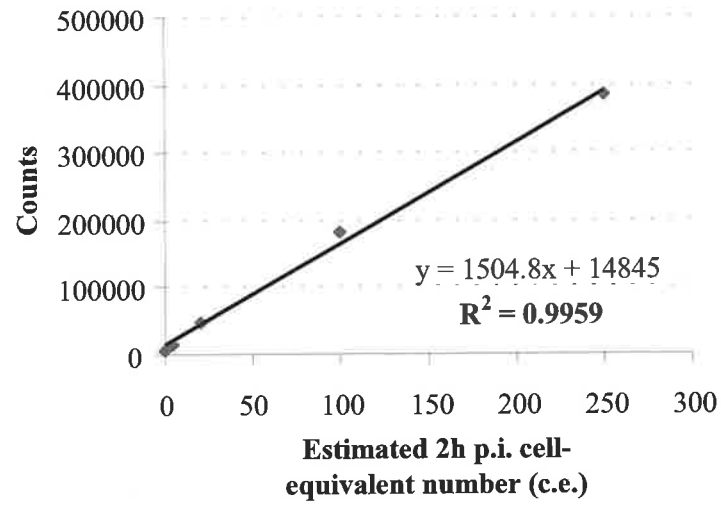
Figure 5.6 Mitochondrial PCR Analysis of Experimental Samples

(A) Duplicate mitochondrial PCRs were performed on ≈ 100 cell-equivalents (c.e.) of HIRT supernatant DNA (based on 100% recovery from 2×10^5 cells) isolated at various times after infection. Uninfected controls (UN) and DNA markers (pUC19/*Hpa* II) are shown. Dilutions of the sample isolated at 2h p.i. were also analysed and used to generate the standard curve shown in (B). Counts (y axes of the graph) were obtained using ImageQuant software. The trendline curve equation ($y =$) and regression value (R^2) of the graph in (B) is shown.

A.



B.



Using the above protocol, the mitochondrial content of each purified experimental HIRT supernatant preparation was determined in duplicate (Fig. 5.6A). Dilutions of DNA extracted at 2h p.i. were again amplified confirming that amplification was occurring in a highly linear fashion ($R^2 > 0.995$; Fig. 5.6B). The average signal intensity of all samples was determined and this value used to represent the signal achieved from 100 cell-equivalents (c.e.) of DNA. All samples were determined to vary from the mean by less than a factor of 1.4-fold (100 ± 39 copies) with the majority (10/13) differing by less than 1.25-fold (data not shown). This degree of variation is commonly observed following the PCR-mediated amplification of duplicate DNA samples and likely reflects errors associated with the PCR procedure and/or the Southern detection method used, and not a significant variation in extraction efficiencies. Consequently, the experimental samples of extrachromosomal DNA were considered to require no further adjustment before assessing levels of both total and 2-LTR viral DNA.

5.3.4 Kinetics of Total Viral DNA Accumulation

5.3.4.i Optimising the GAG PCR

The total viral DNA complement was measured by mixing HIRT supernatant and HIRT pellet DNA fractions from the same time points and analysing the pooled samples by PCR for the presence of GAG DNA sequences. The targeted GAG sequence is synthesised after the 1st template switch during the reverse transcription process (see section 1.2.2). Therefore, amplification of this sequence was expected to reflect the presence of mid- to late-stage reverse transcribed viral cDNA (*ie.* viral cDNA molecules that had successfully undergone the 1st template switch).

The GAG-PCR procedure was initially optimised by amplifying serial dilutions of the HA8 DNA standards (in a background of 1×10^4 c.e. of HuT-78 DNA) in the presence of two different $MgCl_2$ concentrations, for 23 cycles. Although signal intensities did not differ greatly, reactions performed in the presence of 2.5mM $MgCl_2$ allowed the weak detection of 10 HIV copies, whereas reactions performed in the presence of 4mM $MgCl_2$ detected only 100 HIV copies (Fig. 5.7A.). Furthermore, after amplification of HA8 DNA at 2.5mM $MgCl_2$, a linear curve was generated when the signal intensities

were plotted against the input HIV copy-number confirming that amplification was occurring in a logarithmic fashion ($R^2 > 0.98$; Fig 5.7B.).

5.3.4.ii Application of the GAG PCR

Prior to analysing all experimental samples for the presence of reverse transcribed HIV DNA, 200 c.e. of DNA preparations (HIRT supernatant + HIRT pellet) from selected time-points were subjected to GAG-PCR amplification to confirm that the levels of viral DNA present within cells were below that of the uppermost HA8 DNA standard (10^4 HIV copies). The signal intensities obtained from samples isolated at 2h, 4h, 8h, 14h, 18h, 26h and 50h p.i. were all below that obtained following amplification of 1×10^4 HIV copies present within the HA8 standard (Fig. 5.8). However, to increase the linearity of the standard curve over this range, the cycle-number of the GAG-PCR procedure was reduced to 20 in all subsequent analyses. Furthermore, to ensure that samples would not fall out of the standard curve linear range, an additional HA8 standard (5×10^4 copies) was amplified. Taken together, these modifications were expected to allow the accurate quantification of viral DNA present at all time-points within samples equivalent to 500 c.e. of DNA isolated from infected cells.

The signals obtained following GAG PCR-mediated amplification of DNA within experimental samples (Fig. 5.9) indicated that near full-length reverse transcribed DNA was detected as soon as 3h after infection, which is in close agreement with previous studies (8, 138, 160). PhosphorImage analysis of bands (see appendix 5.1 for data) showed that total DNA had reached a level of approximately 30 copies/cell at 14h p.i. and declined to levels of approximately 20 copies per cell by 26h p.i. (Fig. 5.9 and 5.10). The reduction in the total viral DNA complement over this time ($\approx 40\%$; also seen in Fig. 5.8) was attributed to the degradation of extrachromosomal HIV DNA within the cellular environment. This result is consistent with previous observations that significant proportions of reverse transcribed DNA degrade within the intracellular environment following both cell-free and cell-to-cell infection (8, 32). The GAG signal at 50h p.i. was higher than at 26h p.i. implying that some degree of either 2nd-round cell-free or cell-to-cell infection (superinfection) may have occurred during this time. It is worth noting that the intensities of bands corresponding to amplification of 1/10

Figure 5.7 Optimising the GAG-PCR: The Effect of [MgCl₂]

(A) The effect of two different MgCl₂ concentrations (4mM, left side and 2.5mM, right side) on GAG PCR amplification of various dilutions of the HA8 integrated HIV DNA standard. (B) Graph of the band intensities (Counts) obtained following GAG PCR amplification of dilutions of the HA8 integrated HIV DNA. Counts (y axis of graph) were obtained using ImageQuant software from the result presented in A. (2.5mM MgCl₂, right side). The trendline curve equation (y =) and regression value (**R**²) of the graph in C. is shown.

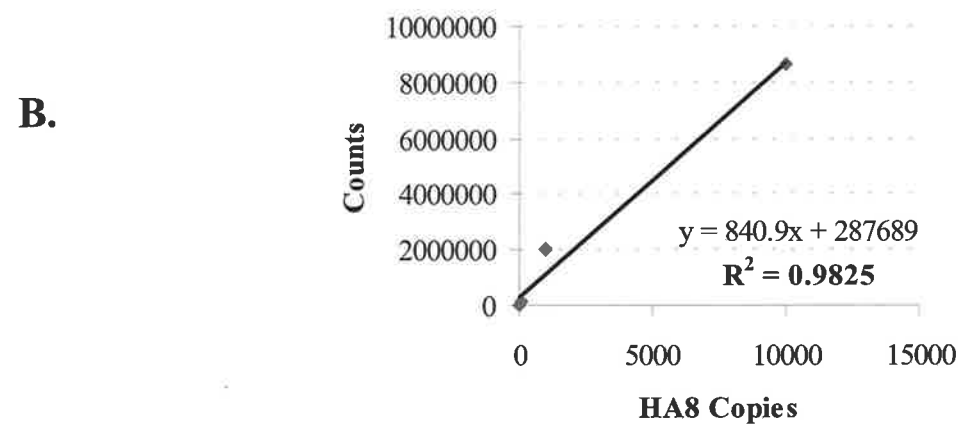
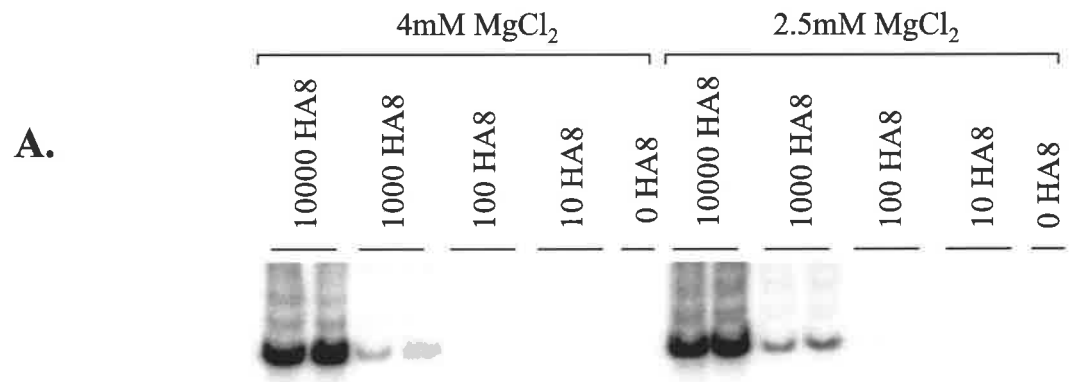


Figure 5.8 GAG-PCR Analysis of Selected Experimental Samples

GAG PCR (23 cycles) was performed in duplicate on 200 cell-equivalents (c.e.) of total DNA extracted at various time-points after infection. DNA markers (pUC19/*Hpa* II) are shown. Purified HuT-78 chromosomal DNA was added to reactions where necessary to bring the total amounts of DNA/reaction to 1×10^4 c.e.. The graph (lower panel) was generated from the signal intensities of the amplified integrated HIV DNA standards (HA8). Counts (y axis of graph) were obtained using ImageQuant software. The trendline curve equation ($y =$) and regression value (R^2) of the graph is shown.

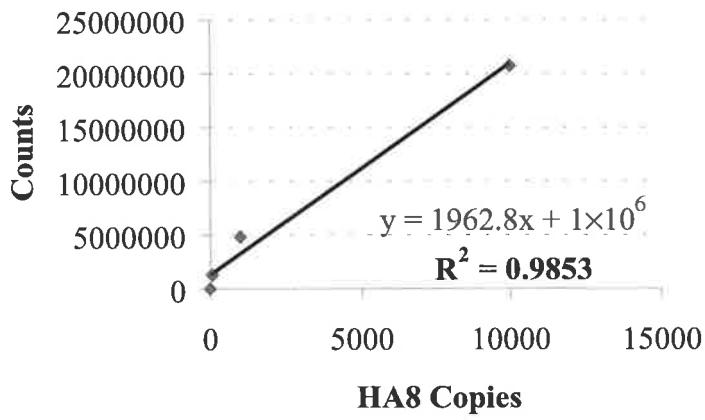
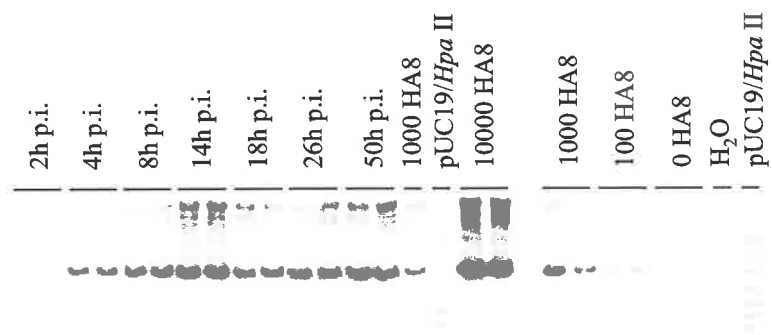
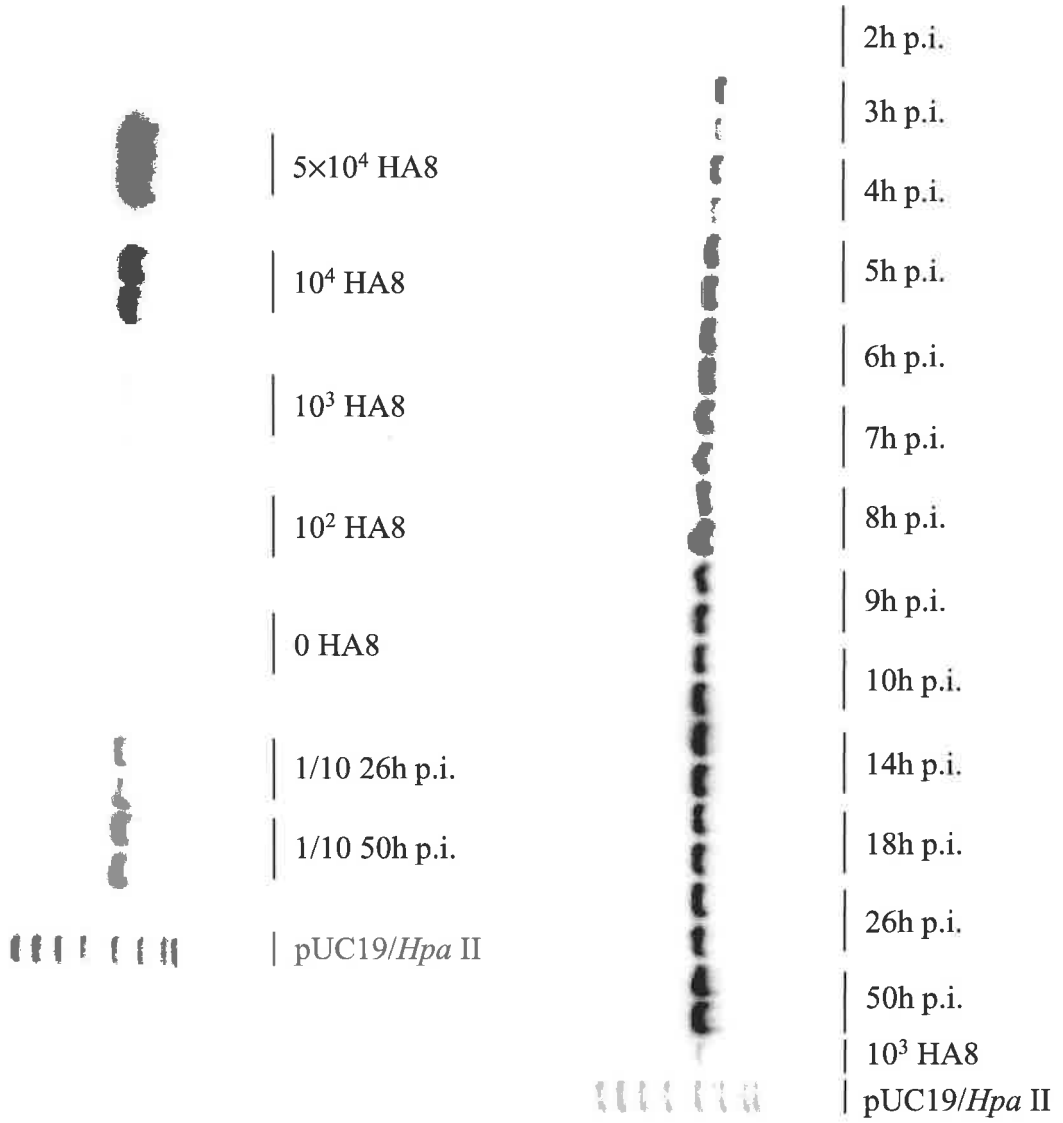


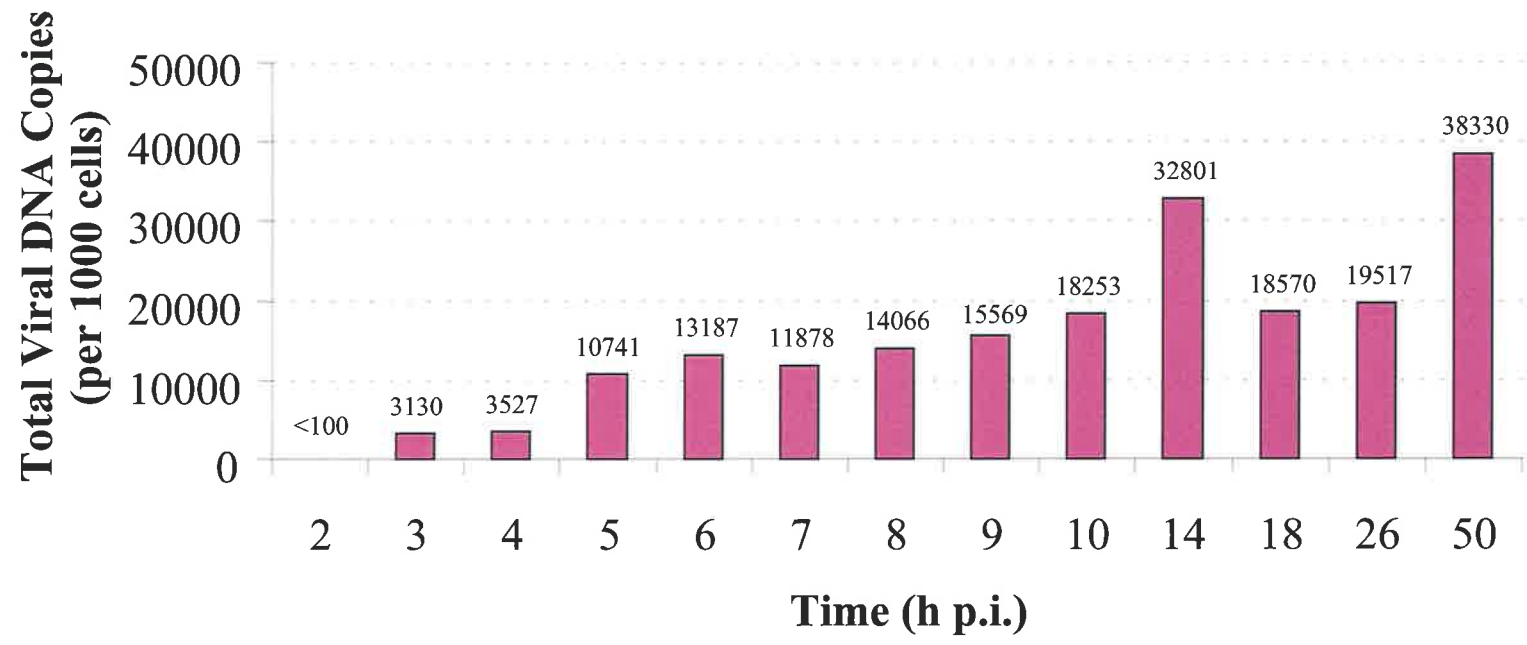
Figure 5.9 GAG-PCR Analysis of Experimental Samples

GAG PCR (in duplicate) was performed on 500 cell-equivalents (c.e.) of total DNA extracted at various time-points after infection. 1/10 dilutions (10 c.e.) of total DNA isolated at 26h p.i. and 50h p.i. were also amplified (lower gel, right side). Purified HuT-78 chromosomal DNA was added to reactions where necessary to bring the total amounts of DNA/reaction to 5×10^4 c.e.. DNA markers (pUC19/*Hpa* II) are indicated.



***Figure 5.10 Total HIV DNA (GAG) Accumulation Following HIV
Infection of HuT-78 Cells***

Graphical representation of the data presented in Figure 5.9 and Appendix 5.1.



dilutions of the 26h and 50h p.i. samples (Fig. 5.9, lower gel, right side) were quantified and confirmed to give comparable results to that observed from the neat samples (data not shown). Since the highest amount of total DNA was observed at 14h p.i., this figure (≈ 30000 copies per 1000 cells) was taken as measuring the total amount of viral DNA (100%) synthesised following infection in this system (see Fig. 5.21).

5.3.5 Kinetics of Integrated Viral DNA Accumulation: LP-PCR

The LP-PCR procedure was designed to specifically detect integrated HIV proviral forms within the context of large amounts of background DNA in a highly sensitive manner (see chapter 4). A high degree of sensitivity would be required when evaluating patients for their integrated HIV proviral load, particularly in those patients responding well to HAART. However, the sensitivity required to monitor the accumulation of integrated DNA forms over time following infection of cultured T cells, in which large amounts of all viral DNA forms are present, was expected to be considerably lower. Therefore, in the following experiments the number of cycles in the nested PCR reaction within the LP-PCR procedure was dropped from 22 to 20 (see comments in section 1.8.2).

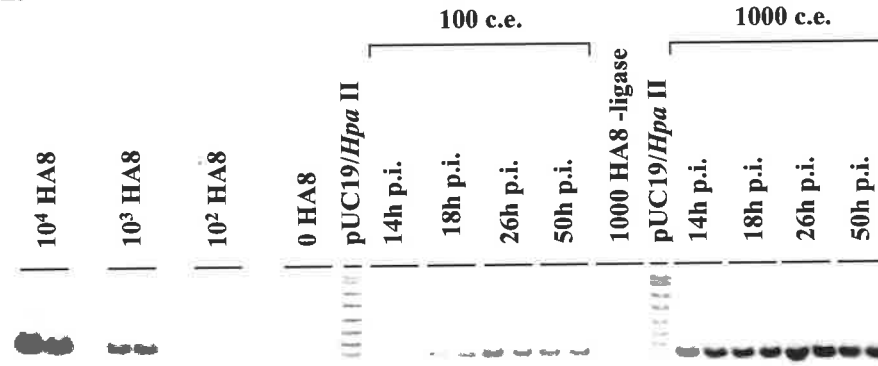
An initial experiment was performed in which either 1000 or 100 cell-equivalents of cellular DNA (in 1×10^4 c.e. of background DNA) extracted at various time points after infection were subjected to LP-PCR. In addition, serial 10-fold dilutions of the HA8 integrated standards were amplified by LP-PCR and the resulting signals quantified. When signal intensities were plotted against integrated HIV copies present in the HA8 standard, amplification in a linear fashion was observed when 1000 copies of the template was initially used (Fig. 5.11A. and C.). However, a plateau in PCR signal was evident when higher amounts of input template DNA were used (Fig. 5.11B.). Furthermore, the LP-PCR-mediated amplification of 1000 c.e. of sample DNA (isolated at 14, 18, 26 and 50h p.i.) resulted in signals that corresponded in intensity to levels outside the linear range of the HA8 standard curve (between 1000 HA8 and 10000 HA8; see Fig. 5.11A.). In contrast, the signal intensities generated following amplification of 100 c.e. of chromosomal DNA isolated at 14, 18, 26 and 50h p.i. corresponded to points within the range of linear amplification (see Fig. 5.11A.).

Figure 5.11 Optimising LP-PCR

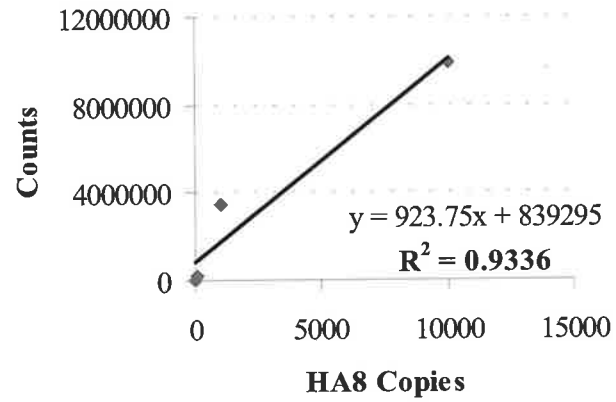
(A) LP-PCR was performed (in duplicate) on 100 cell-equivalents (c.e.) (middle) or 1000 c.e. (right-side) of chromosomal DNA isolated at various times after infection. LP-PCR amplification (in duplicate) of various dilutions of the HA8 integrated HIV DNA standard is shown on the left-side. Purified HuT-78 chromosomal DNA was added to reactions where necessary to bring the total amounts of DNA/reaction to 1×10^4 c.e.. DNA size markers are indicated (pUC19/*Hpa* II). PCR products were detected by Southern hybridisation using the U3-106 probe (see Table 2.2). (B) Graph of all amplified HA8 signal intensities (0 HA8, 10^2 HA8, 10^3 HA8 and 10^4 HA8; Counts) versus sample HA8 copy-number. (C) Graph of amplified HA8 signal intensities (0 HA8, 10^2 HA8, and 10^3 HA8; Counts) versus sample HA8 copy-number.

Counts (y axes of graphs) were obtained using ImageQuant software. The trendline curve equations ($y =$) and regression values (R^2) of each graph is shown.

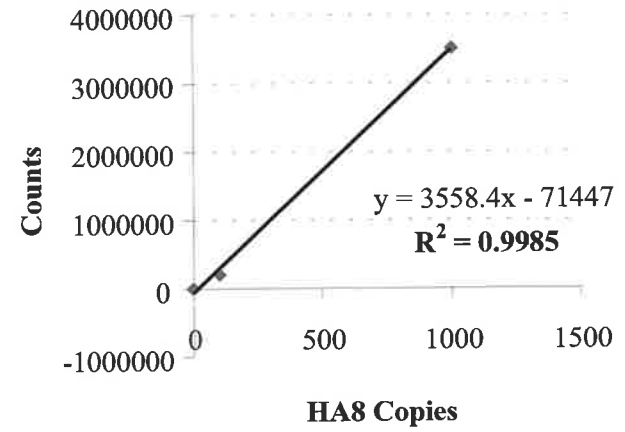
A.



B.



C.



Therefore, 100 c.e. of chromosomal DNA at each time point were used to measure integrated HIV DNA. However, the cycle-number of the nested PCR was returned to 22 to increase the sensitivity ensuring that the time at which integrated viral DNA first appeared following infection was accurately determined. This cycle-number increase was expected to increase the assay sensitivity without elevating the signal intensities observed at 24h and 48h p.i. to levels outside the linear range of the standard curve.

Based on the above, the integrated viral DNA content within chromosomal DNA samples extracted at each of the time-points following infection was determined. Duplicate aliquots containing 100 cell-equivalents of chromosomal DNA (in 500 c.e. of background HuT-78 DNA) were subjected to the LP-PCR protocol (with a 22-cycle nested PCR). In addition, 5-fold serial dilutions of the HA8 integrated DNA standard was amplified in an identical fashion (Fig. 5.12). The signal intensities from the amplification of each duplicate standard were quantified, averaged and plotted as a function of the input copy number to generate a standard curve. The linear nature of the resulting curve confirmed that amplification was occurring at a constant, logarithmic rate throughout the procedure ($R^2 > 0.97$; Appendix 5.2).

Integrated DNA within experimental samples was first detected by LP-PCR at 4h p.i. (that is, 1h after the first appearance of newly synthesised viral DNA; compare Fig. 5.9 with Fig. 5.12) and continued to accumulate throughout the infection, reaching levels equivalent to approximately 2.7 copies/cell by 26h p.i. (Fig. 5.12 and Fig. 5.13). In comparison, maximum levels of newly reverse transcribed DNA (as measured by GAG PCR; see section 5.3.4) within the time-frame of a one-step infection reached 30 copies/cell (14h p.i.). Therefore, integrated HIV DNA (as measured by LP-PCR) was found to account for 8% of the total viral DNA complement at 14h p.i. (see Fig. 5.21). Control experiments involving first-round amplification of the 26h p.i. sample and 500 c.e. of the HA8 standard in the absence of ligation did not give significant signals (Fig. 5.12, lower gel, - ligase lanes).

Figure 5.12 LP-PCR Analysis of Experimental Samples

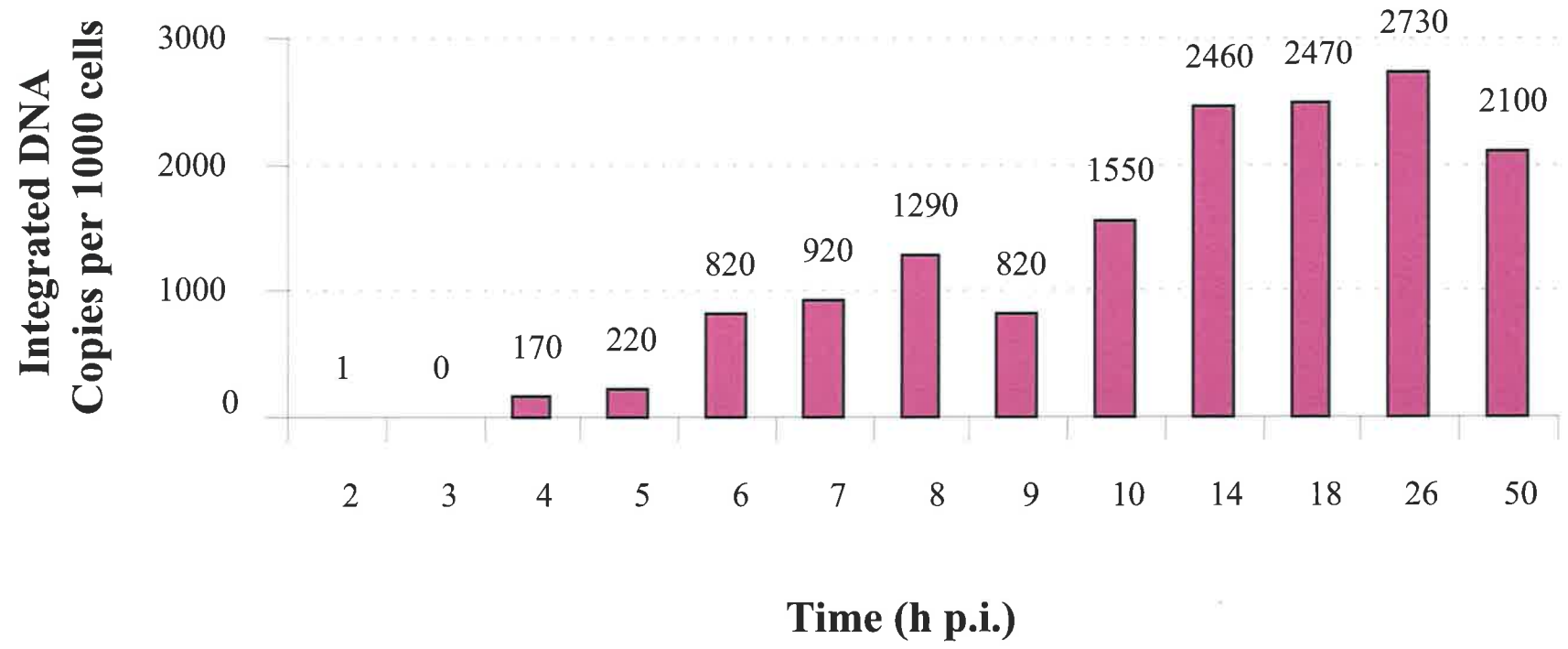
LP-PCR amplification (in duplicate) was performed on 100 cell-equivalents (c.e.) of chromosomal (HIRT pellet) DNA extracted at various time-points after infection.

Purified HuT-78 chromosomal DNA was added to reactions where necessary to bring the total amounts of DNA/reaction to 5×10^2 c.e.. Control reactions in which linker ligation was performed in the absence of T4 DNA ligase are indicated (-ligase). DNA markers (pUC19/*Hpa* II) are indicated.

		2h p.i.
	Uninfected	3h p.i.
		4h p.i.
	500 HA8	5h p.i.
		6h p.i.
	100 HA8	7h p.i.
		8h p.i.
	20 HA8	9h p.i.
		10h p.i.
	0 HA8	14h p.i.
		18h p.i.
	500 HA8 -ligase	26h p.i.
		50h p.i.
	26h p.i. -ligase	20 HA8
		pUC19/ <i>Hpa</i> II
	H ₂ O	
	pUC19/ <i>Hpa</i> II	

Figure 5.13 Integrated HIV DNA Accumulation (as measured by LP-PCR) Following HIV Infection of HuT-78 Cells

Graphical representation of the data presented in Figure 5.12 and Appendix 5.2.



5.3.6 Kinetics of Integrated Viral DNA Accumulation: Nested *Alu*-PCR

A modification of the previously published nested-*Alu* PCR protocol (49) for the detection of integrated HIV DNA was developed in our laboratory to allow a comparison to be made between this protocol and the LP-PCR protocol (see section 4.10). To further compare the two strategies, the nested-*Alu* PCR protocol (in addition to LP-PCR; see section 5.3.5) was used in parallel to assess levels of integrated HIV DNA within each chromosomal DNA preparation. For a detailed explanation of the modified nested-*Alu* PCR protocol used, see sections 2.2.5.iv and 4.10.

Using the modified nested-*Alu* PCR protocol, integrated HIV DNA was not detected until 7h p.i. and accumulated to levels equivalent to 1386 copies per 1000 cells by 26h p.i. (Fig. 5.14 and 5.15). Control reactions involving first-round amplification of the 26h p.i. sample performed in the absence of the *Alu*164 primer (see Table 2.1), produced signals equivalent to approximately 20 copies of integrated HIV per 1000 cells (see Fig. 5.14, right side and Appendix 5.3). This equated to $\approx 1.4\%$ of the signal obtained at 26h p.i. when 1st-round amplification was performed with the *Alu*164 primer and likely resulted from the nested amplification of both integrated HIV DNA and the extrachromosomal HIV DNA contaminating the HIRT pellet. Moreover, since these control signals represented a small percentage of the signal obtained at 26h p.i., the assessment of integrated DNA levels at earlier time-points would not be expected to be influenced by this phenomenon.

It is worth noting that although the lowest HA8 standard amplified contained 100 copies of HIV DNA, the signal obtained at 7h p.i. (using the nested-*Alu* PCR assay) corresponded to a signal intensity equivalent to approximately 5 copies of the HA8 integrated standard (see Appendix 5.3 for data). This sensitivity level is closely comparable to that achieved using LP-PCR; (see section 5.3.5).

5.3.7 Kinetics of 2-LTR Viral DNA Accumulation

In order to assess levels of 2-LTR circular viral DNA following infection, a PCR-based assay that specifically detected these forms was developed. Primers (R7 and U3PNV;

Figure 5.14 Nested-Alu PCR Analysis of Experimental Samples

Nested-*Alu* PCR amplification (in duplicate) was performed on 1000 cell-equivalents (c.e.) of chromosomal (HIRT pellet) DNA extracted at various time-points after infection. Purified HuT-78 chromosomal DNA was added to reactions where necessary to bring the total amounts of DNA/reaction to 1×10^4 c.e.. Control reactions in which the 1st-round PCR was performed in the absence of the *Alu*164 primer (see Table 2.1) are indicated (-*Alu* Control). PCR products were detected by Southern analysis using the U3-106 probe (see Table 2.2) and quantified using ImageQuant software (see Appendix 5.3). DNA size markers are indicated (pUC19/*Hpa* II).

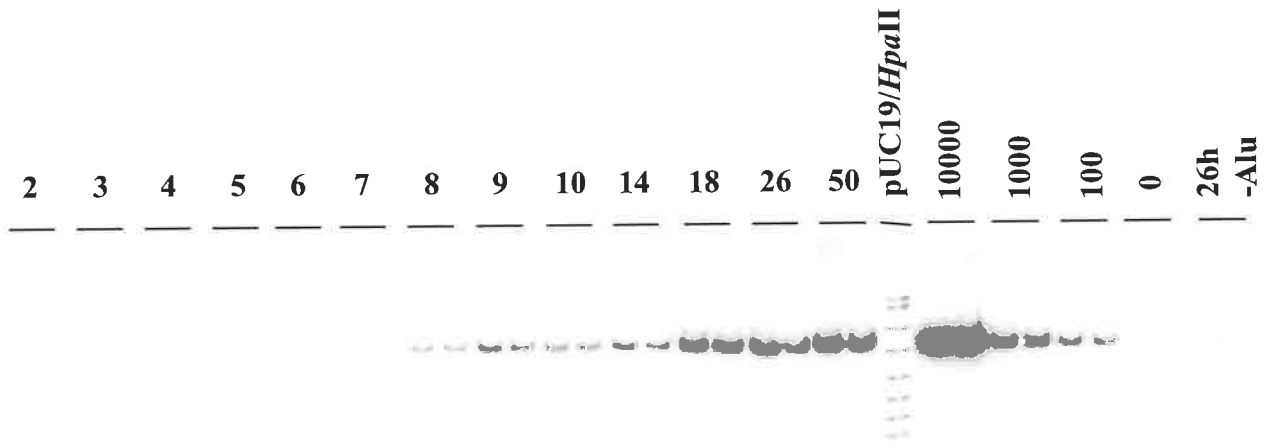
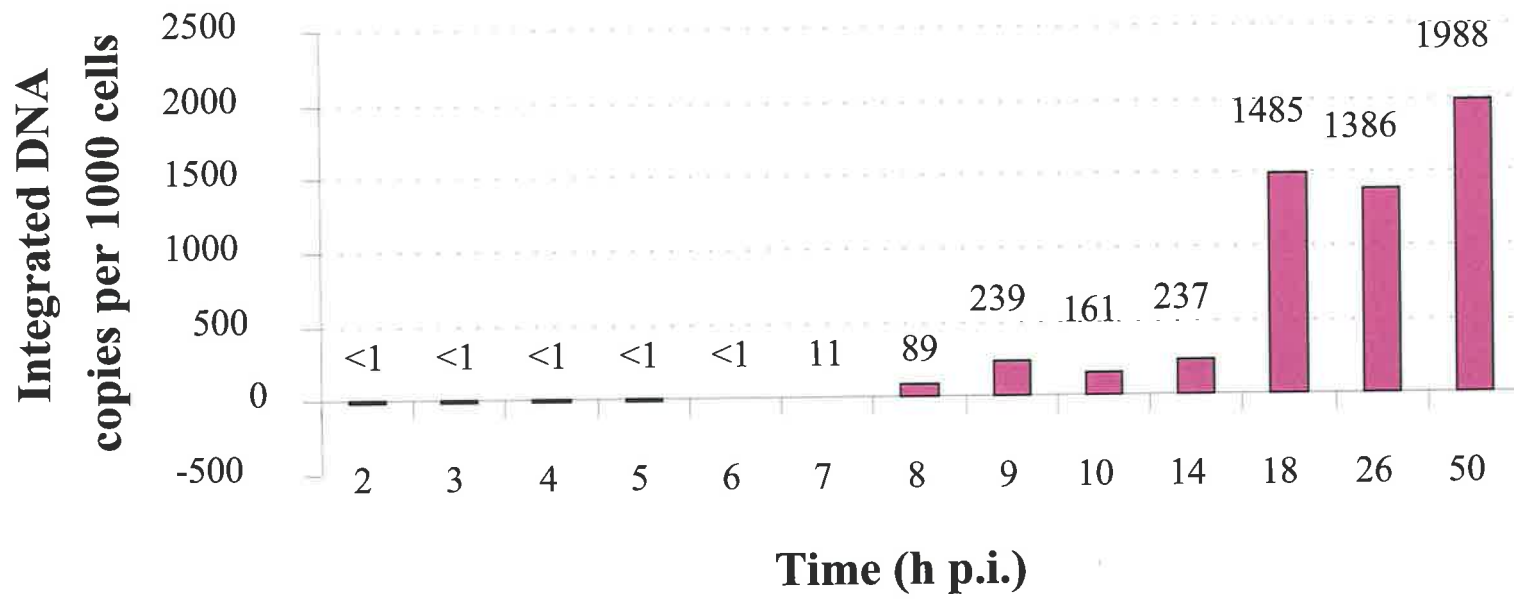


Figure 5.15 Integrated HIV DNA Accumulation (as measured by nested Alu-PCR) Following HIV Infection of HuT-78 Cells

Graphical representation of the data presented in Figure 5.14 and Appendix 5.3.



see Table 2.1) flanking the junction region between the two LTR sequences were designed and assessed for their ability to amplify a control construct mimicking the 2-LTR viral DNA species (see Fig. 5.16). The 2-LTR control construct was generated by PCR amplification of HIRT supernatant samples taken from a cell-free infection of HuT-78 cells at 26h p.i. using primers R7 and U3NV (see Table 2.1). Amplification was performed using *Pfu* polymerase (Stratagene) to generate a 2-LTR junction product which was then ligated into the *Eco* RV site of pBluescript KS(+) vector (Stratagene). The copy number of the 2-LTR control stock solution was initially estimated by spectrophotometer analysis (A_{260} ; see section 2.2.4.iv) and then confirmed by PCR against the HA8 chromosomal mix using primers U3.1(+) and U3PNV in the presence of 1000 c.e. of background HuT-78 DNA (Fig. 5.17 and Appendix 5.4).

Experimental samples (see section 5.3.5) were analysed by a 26 cycle PCR for 2-LTR HIV DNA using 500 c.e. of total cellular DNA (HIRT supernatant + pellet fractions) and primers R7 and U3PNV (see Table 2.1 and Fig. 5.16 for primer sequences and positions, and section 2.2.5.vi for protocol details). A standard of 10^4 copies of the 2-LTR control construct spiked into 500 c.e. of HuT-78 DNA was shown to be amplified in a logarithmic fashion (Fig. 5.18 and Appendix 5.5). Initial results clearly showed continuing 2-LTR accumulation from 7h p.i. throughout the course of infection, which is in close agreement with previous studies (138). However, an unexpectedly low value was obtained for the 26h p.i. sample. Since this time point was to be used for our end-point analysis, we re-analysed this sample. Whereas the re-analysis produced similar values to those previously obtained for the 18h p.i. and 50h p.i. samples, the new 26h p.i. result was now comparable to the adjacent signals (Fig. 5.19 and Appendix 5.6). The initially low figure obtained at 26h p.i. most likely resulted from either poor transfer or poor probing efficiencies of this sample during Southern analysis. A graphical representation of the accumulation of 2-LTR viral DNA following infection, using the re-analysed data from the 18h p.i., 26h p.i. and 50h p.i. samples, is presented in figure 5.20. In conclusion, levels of viral 2-LTR DNA accumulated to levels of approximately 0.4 copies per cell by 26h p.i. reflecting approximately 1% of the total viral DNA synthesised following infection (Fig. 5.21).

Figure 5.16 Construction of the 2-LTR Control Construct

Sequence spanning the 2-LTR junction was amplified by PCR using primers R7 and U3NV and cloned into the *Eco* RV site of pBluescript (as outlined in section 2.2.4.iii). The positions of the R7 and U3NV primers, and the primers used to establish the copy-number of this construct (U3.1(+) and U3PNV; see also Figure 5.17) are shown.

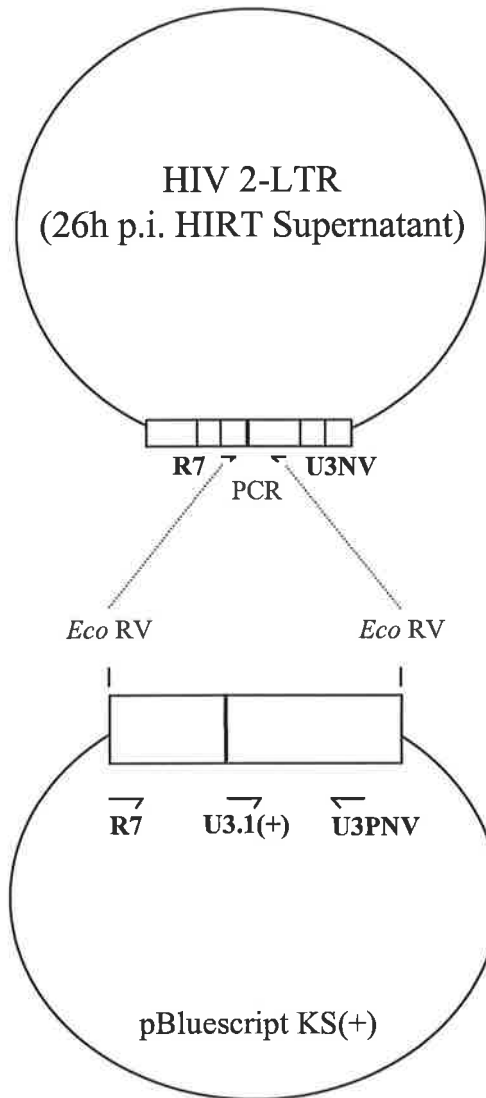


Figure 5.17 Establishing Copy-Numbers of the 2-LTR Control Construct

Copy numbers were established by PCR using the U3.1(+) and U3PNV primers (see Table 2.1 and section 2.2.4.iii) and known amounts of the HA8 control standard. PCR products were detected by Southern hybridisation using the U3-106 probe (see Table 2.2) and quantified using ImageQuant software (see Appendix 5.4). DNA size markers are indicated (pUC19/*Hpa* II).

1 | pUC19/*Hpa* II
2 | 2-LTR (\approx 50 copies)
3 | 2-LTR (\approx 250 copies)
4 | 0 HA8
5 | 40 HA8
6 | 200 HA8
7 | 1000 HA8

Figure 5.18 2-LTR PCR Analysis of Experimental Samples

2-LTR PCR amplification (in duplicate) was performed on 500 cell-equivalents (c.e.) of extrachromosomal (HIRT supernatant) DNA extracted at various time-points after infection. PCR products were detected by Southern analysis using the U3-106 probe (see Table 2.2) and quantified using ImageQuant software (see Appendix 5.5). DNA size markers are indicated (pUC19/*Hpa* II).

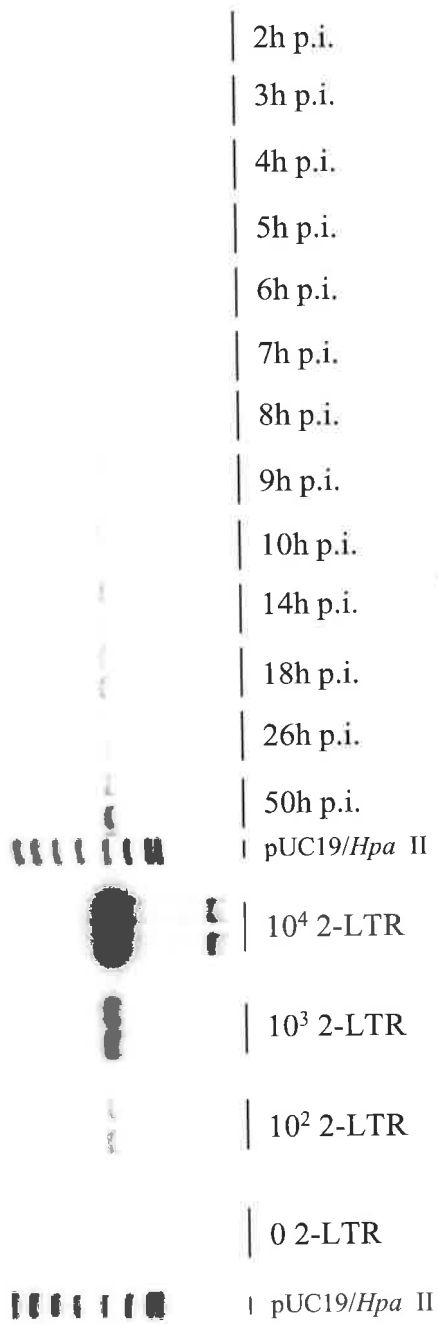


Figure 5.19 2-LTR PCR Analysis of Experimental Samples: Re-Analysis of Selected Time-Points

2-LTR PCR amplification (in duplicate) was performed on 1000 cell-equivalents (c.e.) of extrachromosomal (HIRT supernatant) DNA extracted at various time-points after infection. Single amplification reactions of known amounts of the 2-LTR control construct were performed to allow quantification of signal intensities (see Appendix 5.4). PCR products were detected by Southern analysis using the U3-106 probe (see Table 2.2). DNA size markers are indicated (pUC19/*Hpa* II).

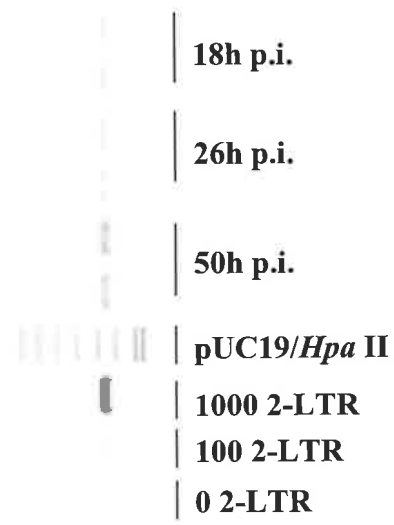
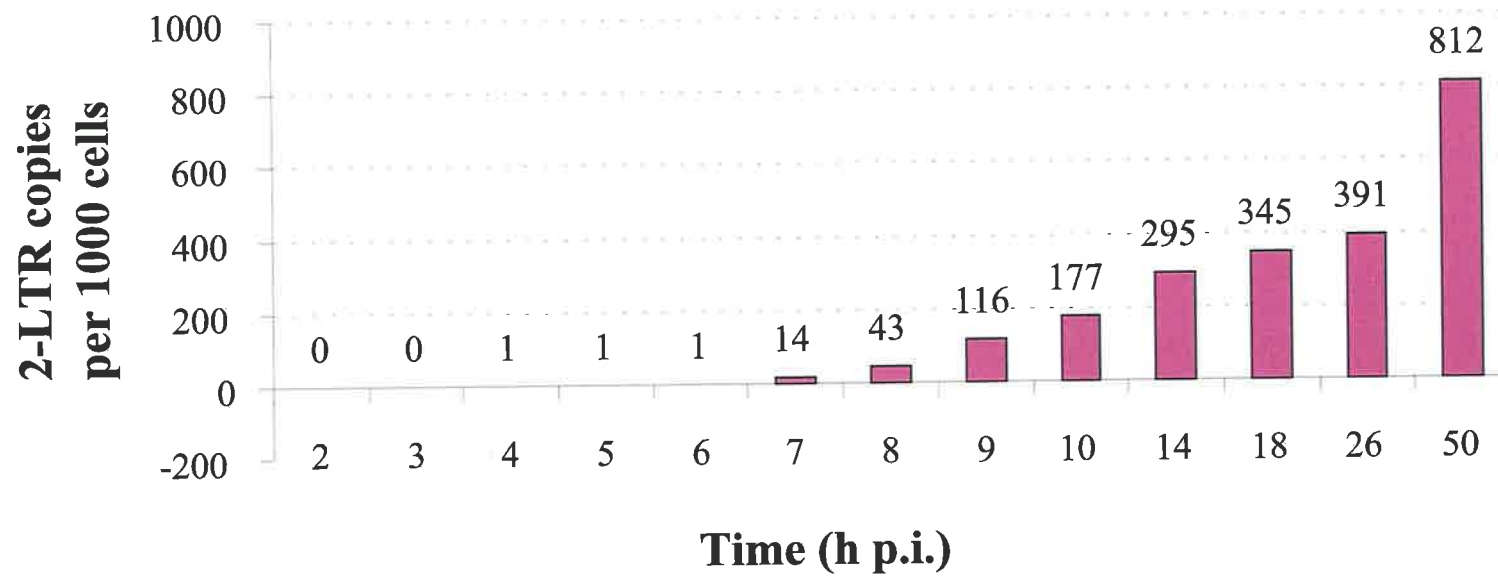


Figure 5.20 2-LTR Circular HIV DNA Accumulation Following HIV Infection of HuT-78 Cells

Graphical representation of the data presented in Figure 5.18 and Appendix 5.5. The figures obtained for the later time-points (18, 26 and 50h p.i.) were taken from the data presented in Figure 5.19 and Appendix 5.6.



5.4 Discussion

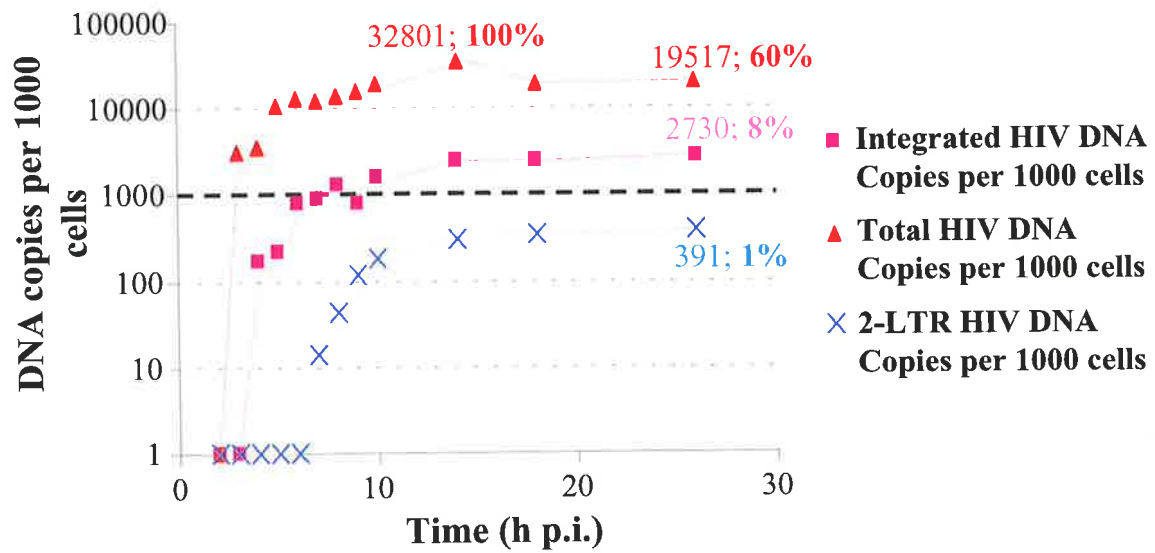
This study monitored the kinetics of accumulation of the total, 2-LTR and integrated HIV DNA forms following acute infection of a T cell line in cell culture. In order to specifically measure integrated proviral DNA at each time point, a novel PCR-based assay termed linker-primer PCR (LP-PCR) was developed that was able to theoretically detect 94% of all integrated forms while selecting against the amplification of extrachromosomal HIV DNA (see chapter 4). Such an assay was required as, without rigorous purification, chromosomal preparations invariably contain significant amounts of extrachromosomal DNA.

In our infection model, near full-length DNA species were first detected at 3h p.i., with the appearance of integrated and 2-LTR forms at the later time points of 4h p.i. and 7h p.i., respectively. The total viral DNA synthesised following infection peaked at approximately 14h p.i., after which levels decreased significantly (by $\approx 40\%$) until 26h p.i.. Similar decreases in the total viral DNA complement following infection have been observed by several groups (8, 32, 160), and have been attributed to degradation of viral by cellular nucleases. Integrated HIV DNA, as measured by LP-PCR was found to comprise approximately 8% of the total viral DNA complement present at 14h p.i. (Fig. 5.21). While we believe this figure to represent an efficient process, our figure is lower than that observed following cell-to-cell HIV infection by Barbosa *et.al.* (8) who indirectly estimated that approximately one in three (34%) newly synthesised viral DNA molecules go on to successfully integrate. However, many differences exist between the two HIV infection studies with regard to the type of viral transmission tested and the viral strain and cell types used. It is quite plausible that integration efficiencies may be affected either by the strain of HIV used (for example, as a consequence of alternative PIC constitution and/or assembly) or the cell types infected (for example, due to differential expression patterns of particular host cell proteins which may be involved in the early events of HIV replication). In addition, the high multiplicity of our T cell infection model gave rise to large numbers of viral DNA molecules, which might be expected to compete for viral and cellular factors involved in a variety of early events in HIV replication including integration. It is possible that in the infected host, where the effective MOI would be expected to be lower resulting in

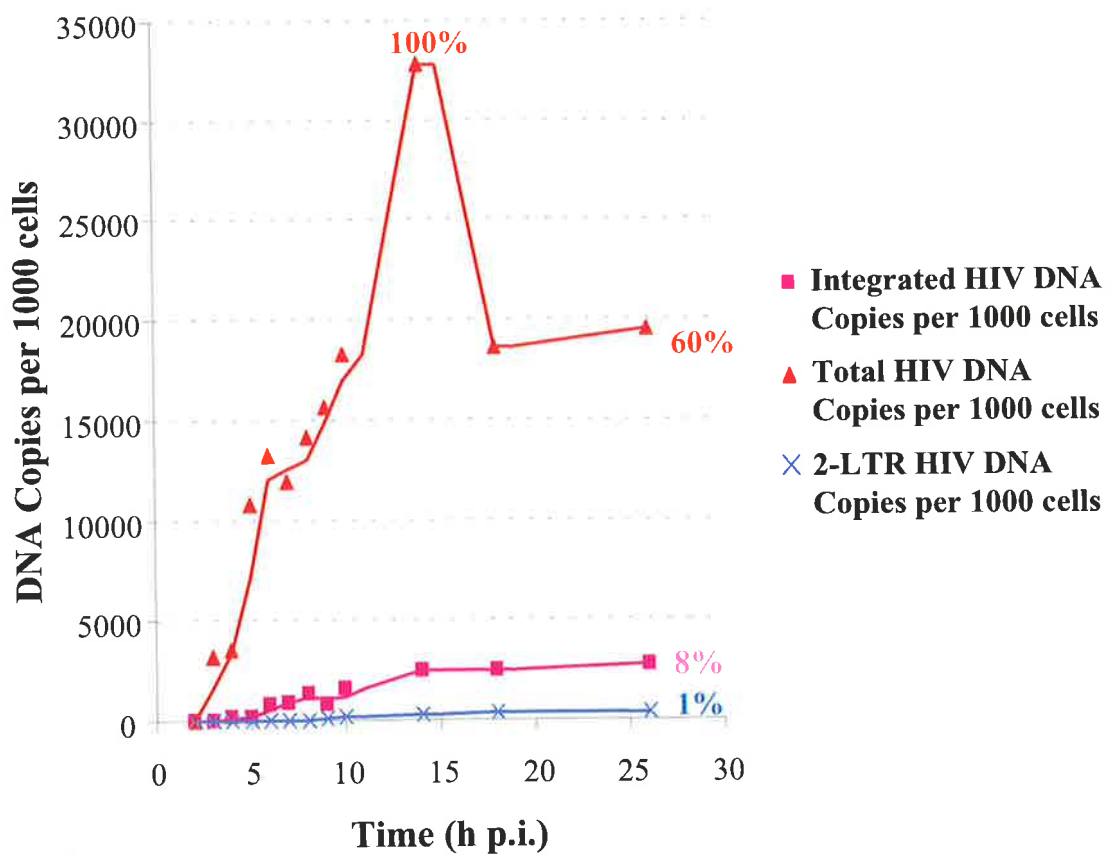
***Figure 5.21 Total, Integrated and 2-LTR Viral DNA Accumulation
Following HIV Infection of HuT-78 Cells***

Graphs were generated from the data (2-26h p.i.) presented in Appendices 5.1 (Total HIV DNA), 5.2 (Integrated HIV DNA), 5.5 (2-LTR HIV DNA) and 5.6 (2-LTR HIV DNA). The total HIV DNA synthesised following infection (indicated as 32801; **100%**) is derived from the 14h p.i. time-point. Other percentages shown indicate the proportion of each viral DNA form at 26h p.i. (expressed as a percentage of the total amount of HIV DNA synthesised following infection). **(A)** DNA copies per 1000 cells (y axis) graphed on a logarithmic scale. **(B)** DNA copies per 1000 cells (y axis) graphed on a linear scale.

A.



B.



less competition for such factors, the efficiency of HIV integration may be higher. Consequently, care should be taken when using the integration efficiencies observed in high MOI cell-free infection systems for predicting the integration efficiencies *in vivo*. Furthermore, it should be emphasised that our infection model involves the use of actively growing T cells. Major differences exist *in vivo* with respect to not only the various activation states of T cells but also the cell type(s) infected. Cells of the macrophage/monocyte lineage are generally considered to be non-dividing cells (259), and the early events in infection of these cells differ markedly to those observed in proliferating T cells (117, 231, 260). Thus, the kinetics of integration within the monocyte/macrophage cell lineage should be considered as a separate study.

In this study, a significant difference was observed between the times at which integrated HIV DNA was observed when either LP-PCR or the modified nested-*Alu* PCR protocol was used. Evidence obtained from earlier experiments suggested that the nested-*Alu* PCR would be unable to amplify all randomly integrated HIV DNA (see section 4.10.2). Although this protocol could detect integrated HIV DNA within the HA8 integrated standard with high sensitivity (only 4 independent integration events), the reduced detection sensitivity of randomly integrated DNA may explain why levels of integrated DNA were below levels of detection at earlier time-points (4-6h p.i.) when the nested-*Alu* PCR protocol was used.

Since 2-LTR circles only occur in the nucleus, they have been used as markers of nuclear entry (22, 138). Consequently, we expected to observe the appearance of 2-LTR circles at the same time as integrated viral DNA forms. However, 2-LTR DNA forms were first observed 3h after the initial detection of integrated forms. This may be attributed to the lower degree of sensitivity achieved using the 2-LTR PCR procedure compared to LP-PCR (compare Figs. 5.18 and 5.12) and the substantially lower amounts of these forms in the infected cell, rather than a complete absence of the 2-LTR viral DNA species at these earlier time points. The 2-LTR viral DNA species was present in smaller amounts than the integrated viral DNA form at all time points tested. This apparent preference for integration rather than 2-LTR circle formation presumably reflects the more specific, enzyme-mediated process of the former. Alternatively, it may be related to the amounts of cellular enzymes required for DNA circularisation within

the nucleus, or a requirement for a change in arrangement and/or composition of the PIC in order for these enzymes to gain access to the DNA termini.

Degradation, integration and 2-LTR circle formation was shown to account for approximately 49% of the total viral DNA synthesised following infection by 26h p.i. (see Fig. 5.21). The majority of the remaining viral DNA (51%) was assumed to exist in either a 1-LTR or linear viral DNA form. Since the 1-LTR DNA levels were not quantified following infection, the relative proportions of these two viral DNA species could not be estimated. However, previous work by Barbosa and co-workers suggests, that although linear viral DNA predominated early after infection, approximately equimolar amounts of the linear and 1-LTR form are present after a single round of cell-to-cell infection (8). It is worth noting that a minor percentage of viral DNA may also exist in alternative forms, such as auto-integrated species (81), within the cell.

Within the time-frame (26h p.i.) of the kinetics study, the newly synthesised viral DNA observed likely resulted from reverse transcription of incoming virion RNA rather than newly synthesised full-length HIV-1 RNA (213). Syncytia were observed at 26h p.i. indicating that some cell-to-cell infection events may have initiated during the 18h p.i. to 26h p.i. time period. However, since a significant increase in the amount of total DNA was not observed between the 18h and 26h p.i. time points (see Fig. 5.9 and 5.10), it is unlikely that reverse transcription of newly synthesised full-length RNA was contributing significantly to our results at 26h p.i..

In conclusion, we have established a system in which the amounts of integration can be measured over the course of an infection with HIV-1. We have also demonstrated that HIV-1 integration is a rapid and relatively efficient process under one-step infection conditions and defined the levels of total, integrated and 2-LTR HIV DNA forms during the course of our infection. During the period after this chapter was written (and a manuscript submitted for publication (247)), a short report was published (32) that further supported the conclusions of this work. A detailed comparison of the findings presented by these researchers with the results obtained in our study is discussed in chapter 7 (General Discussion).

Chapter 6

Specific Inhibition of HIV-1 Integration in Cell Culture: Evaluating Putative Inhibitors of HIV-1 Integrase

6.1 Background

The process of retroviral integration, in which newly reverse transcribed viral DNA is inserted into the host cell chromosome, is essential for a productive infection (76, 106, 151, 218, 238). Consequently, this process has been identified as a prime candidate for antiviral therapy. Since HIV integration is mediated by the virally-encoded integrase protein, the majority of studies have focussed on the identification of compounds that block the action of this enzyme. To date, the high throughput screening of potential integrase inhibitors has primarily been performed in cell-free systems using either purified integrase alone or within the context of a partially purified pre-integration complex (26, 82, 84, 109, 110, 118, 163). Since these assays can be designed to test for inhibition of either the formation of the initial stable complex, 3'-end processing, strand transfer or disintegration (the reverse of strand-transfer) (see section 1.6.3.i), they can not only rapidly identify potential inhibitors, but can also provide preliminary evidence as to their mode of action. However, inhibitors targeting the integrase protein and/or preintegration complexes identified in this manner are frequently cytotoxic or do not exhibit antiviral activities when tested in cell culture (200).

Recently, a number of compounds identified in cell-free assays have been shown to inhibit viral replication in cell culture without displaying significant cytotoxicity (79, 111, 140, 188, 208, 215, 261, 262). AR177 (a G-quartet-containing oligonucleotide that forms highly stable intermolecular tetrad structures; see section 1.6.4)), members of the bisaroyl hydrazine family of integrase inhibitors (*eg.* L17) and flavones (*eg.* quercetin dihydrate) have been shown to inhibit cell-free integration reactions in the nM to low μ M range (41, 84, 85, 173). Of these compounds, AR177 was further characterised and

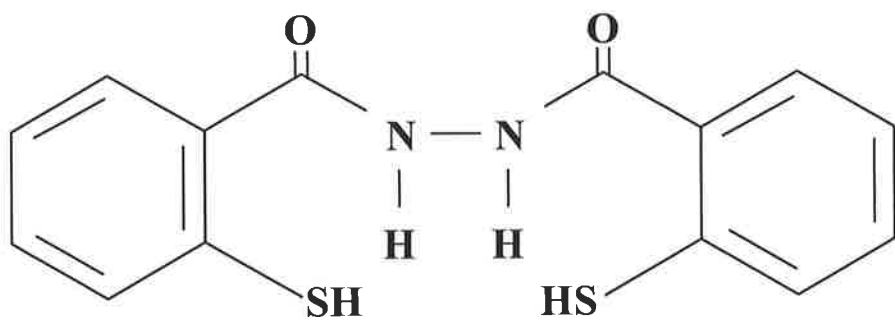
shown to inhibit syncytia formation and productive infection in cell culture, albeit at higher concentrations than those observed for integrase inhibition in cell-free assays (79, 188). In addition, a new class of integration inhibitors containing a diketo acid moiety has been described (77, 111) (see section 1.6.3). Acute infections performed in the presence of such compounds (L-731,988 and L-708,906) did not yield infectious progeny virus and also resulted in the accumulation of large amounts of circular DNA forms incapable of integration. In addition, mutations conferring resistance to these drugs in cell culture consistently mapped to defined regions within the integrase protein. Although these results strongly suggested that the antiviral effect observed was due to a specific block of the integration process in infected cells, a direct evaluation of whether the drugs inhibited the accumulation of integrated HIV-1 DNA was not performed.

In addition to those drugs outlined above, new compounds that have been identified in inhibitor screens to block integration in cell-free assays are continually being reported. In this study, a new drug called L17 is examined for its ability to block integration in cell culture. L17 is a member of the bisaroyl hydrazine family of integrase inhibitors initially described by Zhao and co-workers (279) and consists of two sulfhydrylated aromatic ring structures that are spaced by an N-N linkage (Fig. 6.1). This compound was recently shown to inhibit integration with a 50% inhibitory concentration (IC_{50}) value of ≈ 20 nM in cell-free assays, and to prevent the productive infection of T cells with an IC_{50} of ≈ 5 μ M (Dr. Terrence Burke Jr., personal communication). Although its ability to inhibit purified integrase *in vitro* has been shown, the exact mechanism of cytoprotection in cell culture remains to be elucidated.

This chapter describes an approach that can be used to evaluate potential inhibitors identified in cell-free systems for their ability to specifically inhibit the accumulation of integrated HIV-1 DNA following acute infection in cell culture. Using a modified nested *Alu*-polymerase chain reaction (*Alu*-PCR) to quantify HIV provirus in cells (see section 4.10.1), a total of seven compounds (including control drugs) were tested for their ability to block integration of newly synthesised HIV-1 DNA into T cell genomic DNA. The accumulation of extrachromosomal HIV DNA was also monitored to establish whether inhibition of viral infection resulted from the specific inhibition of

Figure 6.1

Structure of L17



viral integration or inhibition of events at, or prior to, reverse transcription of the viral genome.

6.2 The Cell-free Infection Model

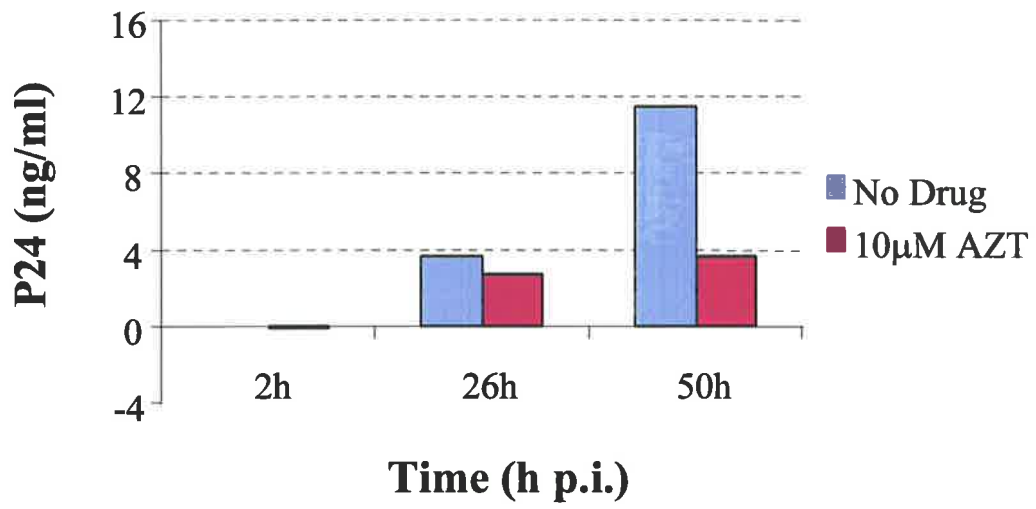
To assess HIV integration following infection of T cells in this study, an infection model was established that was essentially the same as that used in section 5.2, except that the MOI used was slightly lower (0.5 TCID₅₀ units per cell). At this MOI, the levels of P24 in the culture supernatant at 50h p.i. following infection of HuT-78 cells with HIV was approximately half that observed when an MOI of 1 TCID₅₀ unit per cell was used (compare Fig. 6.2 to Fig. 5.2). Infections performed in the presence of 10 μ M AZT (an inhibitor of reverse transcription) exhibited low levels of P24 in the culture supernatant over the course of the infection (Fig. 6.2). A slight increase in the levels of P24 present at 26h and 50h p.i. compared to 2h p.i. was routinely observed in the presence of AZT and also other inhibitors of viral replication such as 3TC. This was attributed to detachment (after washing) of virus that had bound to cell surfaces, but failed to enter cells, during the infection process.

6.3 The Effect of L17 and AR177 on HIV Integration

To evaluate the effects L17 and AR177 on HIV replication in cell culture (in particular integration of viral DNA into the host cell chromosome), cells were infected with HIV in the presence or absence of drugs. DNA was extracted from pooled triplicate infections by the method of HIRT (see section 2.2.3) at 2h and 26h p.i., and assessed for the presence of integrated and extrachromosomal viral DNA. AZT (an inhibitor of reverse transcription) was used in this study (at a concentration of 10 μ M) as a positive control for the inhibition of extrachromosomal HIV DNA synthesis prior to integration. The concentration of AR177 used in this study (10 μ M) was based on previous reports of cytoprotection in cell culture (79, 188), while L17 was used at 30 μ M, a 25% lower concentration than that shown to cause significant cytotoxicity under our culture conditions (see section 6.3.1 below).

Figure 6.2 HIV P24 release following infection of HuT-78 cells

P24 levels in culture supernatants were measured at various time points (h p.i.) following centrifugally enhanced infection of HuT-78 cells with HIV_{HXB2} (0.5 TCID₅₀ unit per cell) in the presence (10µM AZT) or absence (No Drug) of AZT.



6.3.1 Cytotoxicity

Prior to evaluation for effects on the accumulation of viral nucleic acids following infection in cell culture, the cytotoxicity of L17 and AR177 was assessed. Previous reports had indicated that AR177 was non-cytotoxic to a variety of cell lines, even at concentrations of drug above 100 μ M (79, 188). Cytotoxicity was therefore examined under our cell culture conditions using a concentration of 100 μ M. In contrast, the cytotoxicity profile of L17 was unknown and a more rigorous evaluation of this compounds' cytotoxicity under our culture conditions was therefore performed.

HuT-78 cells were incubated in the presence of serial 2-fold dilutions of L17 (ranging from final concentrations of 10-320 μ M in culture media) and assessed for cytotoxicity by trypan blue exclusion 48h after plating (see section 2.2.12). Significantly lower cell counts due to slower cell growth and/or cell death (>5%) were observed at concentrations greater than 40 μ M (see Table 6.1). In contrast, the compounds AR177 and AZT (at concentrations of 100 μ M and 10 μ M, respectively), did not slow cell growth or exhibit cytotoxic effects over this time-period compared to drug-free controls.

6.3.2 Infection in the Presence of L17 and AR177

In order to allow entry into cells, cells were pre-incubated with L17 (30 μ M) and AR177 (10 μ M) for 30min prior to the exposure to virus. In contrast, AZT (10 μ M) was pre-incubated for 16h (overnight) with cells to allow entry and conversion to the active form by phosphorylation within the intracellular environment. All viral stocks and media (including that used for washes) contained relevant drugs throughout the infection procedure.

In the absence of drug, infected cultures displayed extensive syncytia formation by 26h p.i. indicating that a productive infection had occurred. This was inhibited in the presence of L17, AR177 and AZT (data not shown). Supernatant P24 levels at 26h p.i. were not assessed since progeny virion release by this time was expected to be minimal (compare No Drug and AZT at 26h p.i., Fig. 6.2). However, at 50h p.i., the supernatant P24 levels in cultures infected in the presence of both L17 and AR177 were

Table 6.1 HuT-78 Cytotoxicity of L17, AR177 and AZT

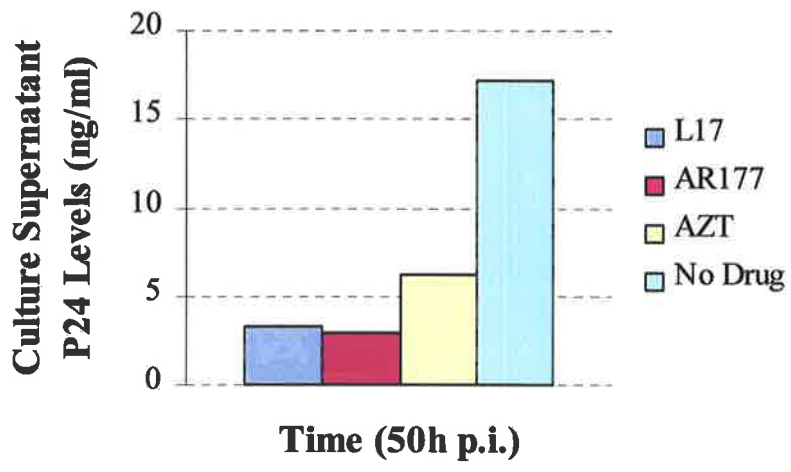
HuT-78 cells were incubated with known concentrations of drugs (μM) and assessed 48h later for cell viability by trypan blue exclusion. Aliquots from triplicate wells were counted and averaged (Average Cell Counts), and then expressed as a percentage of the average count obtained 48h following culture in the absence of drug (% of No Drug).

n/a = Not applicable

Drug	uM	Average Cell Counts (cells/ml @ 48h)	% of No Drug
L17	320	340000	14
	160	541000	23
	80	963000	41
	40	2240000	95
	20	2330000	99
	10	2340000	99
AR177	100	2380000	101
AZT	10	2340000	99
No Drug	n/a	2360000	100

Figure 6.3 P24 Release Following HIV Infection of HuT-78 Cells in the Presence/Absence of L17 and AR177

P24 levels in culture supernatants were measured at 50h p.i. following centrifugally enhanced infection of HuT-78 cells with HIV_{HXB2} (0.5 TCID₅₀ unit per cell) in the presence of either L17 (30µM), AR177 (10µM) or AZT 10µM. The P24 levels in culture supernatant following a control infection performed in the absence of drug is also indicated (No Drug).



significantly lower than that seen with AZT (Fig. 6.3) indicating that a block in the viral replication cycle had occurred.

6.3.3 Viral Nucleic Acid Accumulation

6.3.3.i Integrated Viral DNA Accumulation

Since a near one-step infection of cells (allowing the substantial accumulation of all HIV DNA forms) was expected to have occurred in drug-free cultures by 26h p.i., this time-point was used to assess the ability of L17 and AR177 to inhibit HIV provirus formation. Chromosomal (HIRT pellet) DNA preparations extracted at 26h p.i. were analysed for the presence of integrated HIV DNA using a modified nested *Alu* PCR protocol (see sections 2.2.5.iv). Prior to performing the nested *Alu*-PCR procedure, purified chromosomal DNA preparations extracted at each time-point were matched for β -globin content against known amounts of the HA8 copy-number standards. When a 25 cycle PCR detecting the single-copy β -globin gene was performed on ≈ 50 cell equivalents (c.e.) of sample chromosomal DNA (as determined by cell-numbers), little difference was observed between all samples with respect to β -globin content (Fig 6.4).

Nested *Alu*-PCR analysis for integrated HIV DNA was performed in triplicate on 1000 c.e. of chromosomal sample DNA. As expected, a large amount of integrated DNA was detected in the chromosomal DNA isolated at 26h p.i. from drug-free cultures (Fig. 6.5). In contrast, integrated HIV DNA was not detected in cultures infected in the presence of L17, AR177 or AZT. This result clearly demonstrated that the accumulation of integrated HIV was blocked in the presence of both L17 and AR177. However, to establish whether the inhibition of proviral DNA accumulation observed resulted from a specific block at the step of viral integration, or a block at events prior to integration (eg. reverse transcription or entry), levels of total viral DNA were assessed in cells at all time points.

Figure 6.4 β -globin PCR Analysis of Experimental Samples: L17 and ARI77

PCRs were performed on ≈ 50 cell equivalents of chromosomal DNA extracted from HuT-78 cells at various times after infection with HIV. Standards (HA8) and DNA markers (pUC19/*Hpa* II) are indicated. PCR products were detected by Southern analysis using the Glo probe (see Table 2.2).

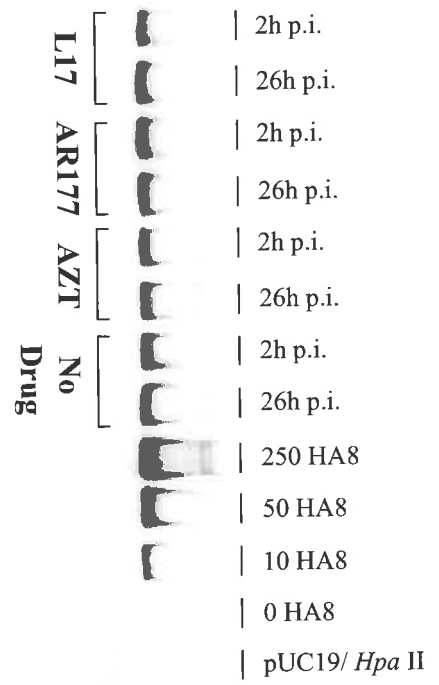
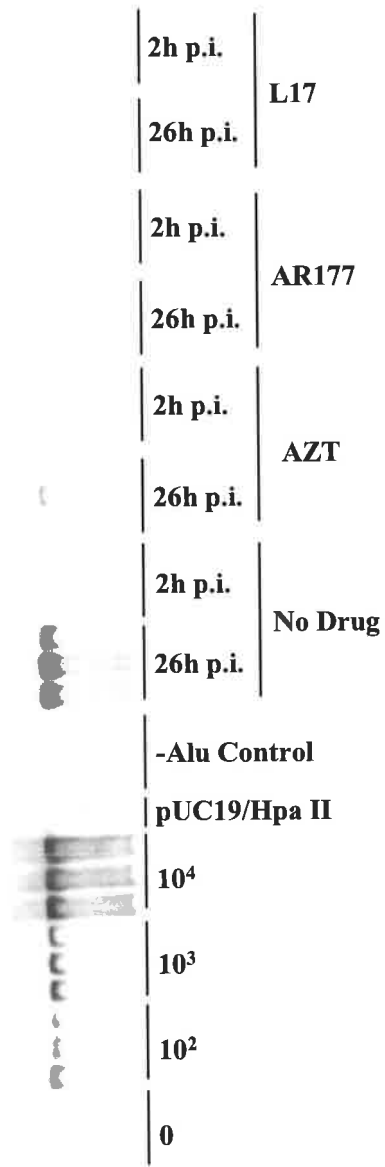


Figure 6.5 Nested-Alu PCR Analysis of Experimental Samples: L17 and ARI77

Nested-*Alu* PCR amplification (in duplicate) was performed on 1000 cell-equivalents (c.e.) of chromosomal (HIRT pellet) DNA extracted at various time-points after infection. Purified HuT-78 chromosomal DNA was added to reactions where necessary to bring the total amounts of DNA/reaction to 1×10^4 c.e.. Control reactions in which the 1st-round PCR was performed in the absence of the *Alu*164 primer (see Table 2.1) are indicated (-*Alu*). PCR products were detected by Southern analysis using the U3-106 probe (see Table 2.2). DNA size markers are indicated (pUC19/*Hpa* II).

Integrated DNA →



6.3.3.ii Total Viral DNA Accumulation

The total viral DNA complement was assessed by GAG-PCR as outlined previously (see section 2.2.5.iii). Concurrently, the relative efficiencies of extrachromosomal DNA recovery in each sample was determined by amplification of mitochondrial DNA within the HIRT supernatant fractions (see section 2.2.5.ii). As expected, substantial levels of reverse transcribed product were observed in total DNA preparations (chromosomal DNA + extrachromosomal DNA) isolated at 26h p.i. in drug-free cultures (Fig. 6.6). Small amounts of viral DNA were also detected when infections were performed in the presence of 10 μ M AZT. This has also been observed by others (32) and likely represents the incomplete block of reverse transcription by AZT. In contrast, reverse transcribed DNA was not detected when infection was performed in the presence of either L17 or AR177 indicating a block at, or prior to reverse transcription in the HIV replication cycle. This result demonstrated that although these two drugs were active in blocking integration reactions as assayed *in vitro*, the primary target of both L17 and AR177 was not the process of viral integration *in vivo*.

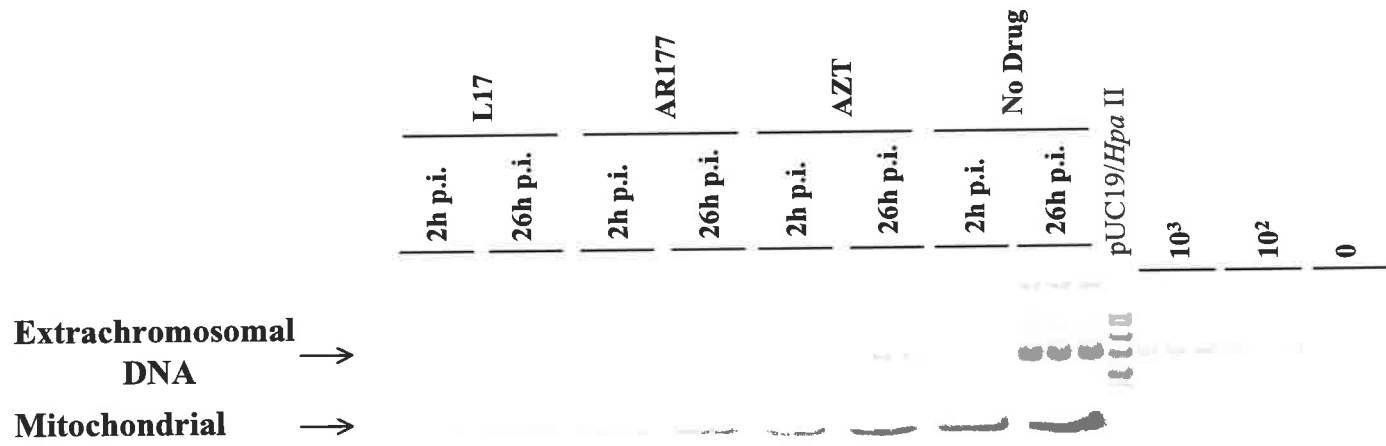
It is worth noting, that while the extrachromosomal DNA extraction efficiency was lower in cultures containing L17 and AR177 than those with AZT, these differences (a maximum of 2.4-fold; Fig. 6.6, compare mitochondrial 26h p.i. L17 to 26h p.i. AZT) cannot account for the large difference (15-fold) in signal intensities observed following GAG-PCR analysis of the same samples (compare GAG 26h p.i. L17 to 26h p.i. AZT). This therefore suggests, that although these drugs do not inhibit the process of integration, they are more potent inhibitors of HIV replication than AZT (at 10 μ M) under our infection conditions.

6.4 The Effects of Quercetin Dihydrate and Diketo Acid Inhibitors on HIV Integration

Following the analysis of both L17 and AR177, samples of 5,8 dihydroxynaphthoquinone, quercetin dihydrate (a flavone) and two recently characterised compounds containing a diketo acid moiety (L-731,988 and L-708,906) became available. Each of these compounds had been shown to inhibit purified

Figure 6.6 GAG PCR Analysis of Experimental Samples: L17 and ARI77

GAG PCR amplification (in triplicate) was performed on 1000 cell-equivalents (c.e.) of extrachromosomal (HIRT supernatant) DNA extracted at various time-points after single infections. Purified HuT-78 chromosomal DNA was added to reactions where necessary to bring the total amounts of DNA/reaction to 1×10^4 c.e.. PCR products were detected by Southern analysis using the U3-106 probe (see Table 2.2). DNA size markers are indicated (pUC19/*Hpa* II). The mitochondrial content of each sample (as measured by a single PCR; see section 2.2.5.ii) is shown below.



integrase *in vitro* (see appendix 6.1) although only the diketo acid-containing inhibitors had been shown to confer protection against HIV infection in cell culture (111). Since AZT (used at 10 μ M) was shown in previous experiments to incompletely inhibit reverse transcription (see section 6.3.3), the reverse transcriptase inhibitor 3TC was used (at a concentration of 10 μ M) as an additional positive control for the inhibition of extrachromosomal HIV DNA accumulation prior to integration.

6.4.1 Cytotoxicity

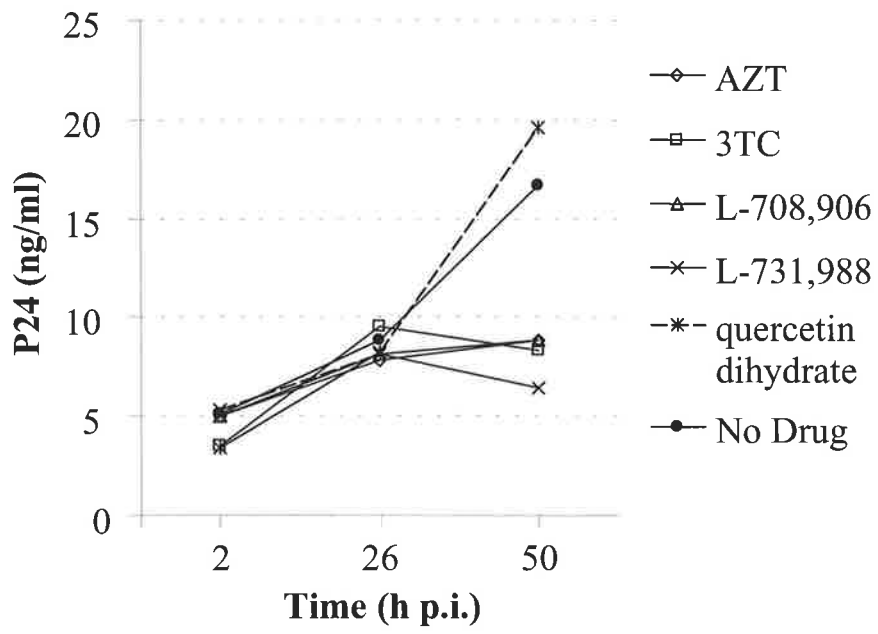
Prior to evaluation for their effects on viral replication, the cytotoxicity of each compound was assessed. Cells were incubated in triplicate in the presence of drugs at concentrations 5-fold above and below the chosen experimental concentration and assessed for cytotoxicity by trypan blue exclusion 48h after plating (see section 2.2.12). All drugs except 5,8 dihydroxynaphthoquinone were non-cytotoxic under infection conditions, even at concentrations 5-fold higher than that used in the assay (see appendix 6.1). The compound 5,8 dihydroxynaphthoquinone (IC₅₀ of 2.5 μ M for strand transfer) was shown to be highly cytotoxic when used at concentrations above 1 μ M and was therefore not analysed further.

6.4.2 Infection in the Presence of Quercetin Dihydrate and Diketo Acid Inhibitors

Duplicate infections were performed in the presence of either 10 μ M L-708,906 or 10 μ M L-731,988 (both containing a diketo acid moiety), or 50 μ M quercetin dihydrate (a flavone). The inhibitors of reverse transcription, AZT and 3TC (used at 10 μ M), served as positive controls for inhibition of extrachromosomal HIV DNA synthesis. In the absence of drug, infected cultures displayed extensive syncytia formation by 26h p.i. (data not shown) and high levels of supernatant P24 by 50h p.i. (Fig. 6.7), indicating that a productive infection had occurred. In cultures with or without drug, levels of P24 rose slightly from 2h p.i. to 26h p.i.. Since levels failed to increase by 50h p.i. in cultures containing AZT and 3TC (indicating that a near complete block to viral replication had been achieved using these drugs), the initial increase in P24 levels in each culture was largely attributed due to detachment of bound virus from the surface of

Figure 6.7

Effect of five compounds on the levels of P24 released into culture supernatants at 2h, 26h and 50h following infection of HuT-78 cells with HIV_{HXB2}.



cells during the infection procedure. With the exception of quercetin dihydrate, all drugs inhibited syncytia formation (data not shown) and P24 release into the culture supernatant at 50h p.i. (Fig. 6.7).

6.4.3 Viral Nucleic Acid Accumulation in the Presence of Quercetin Dihydrate and Diketo Acid Inhibitors

6.4.3.i Integrated Viral DNA Accumulation

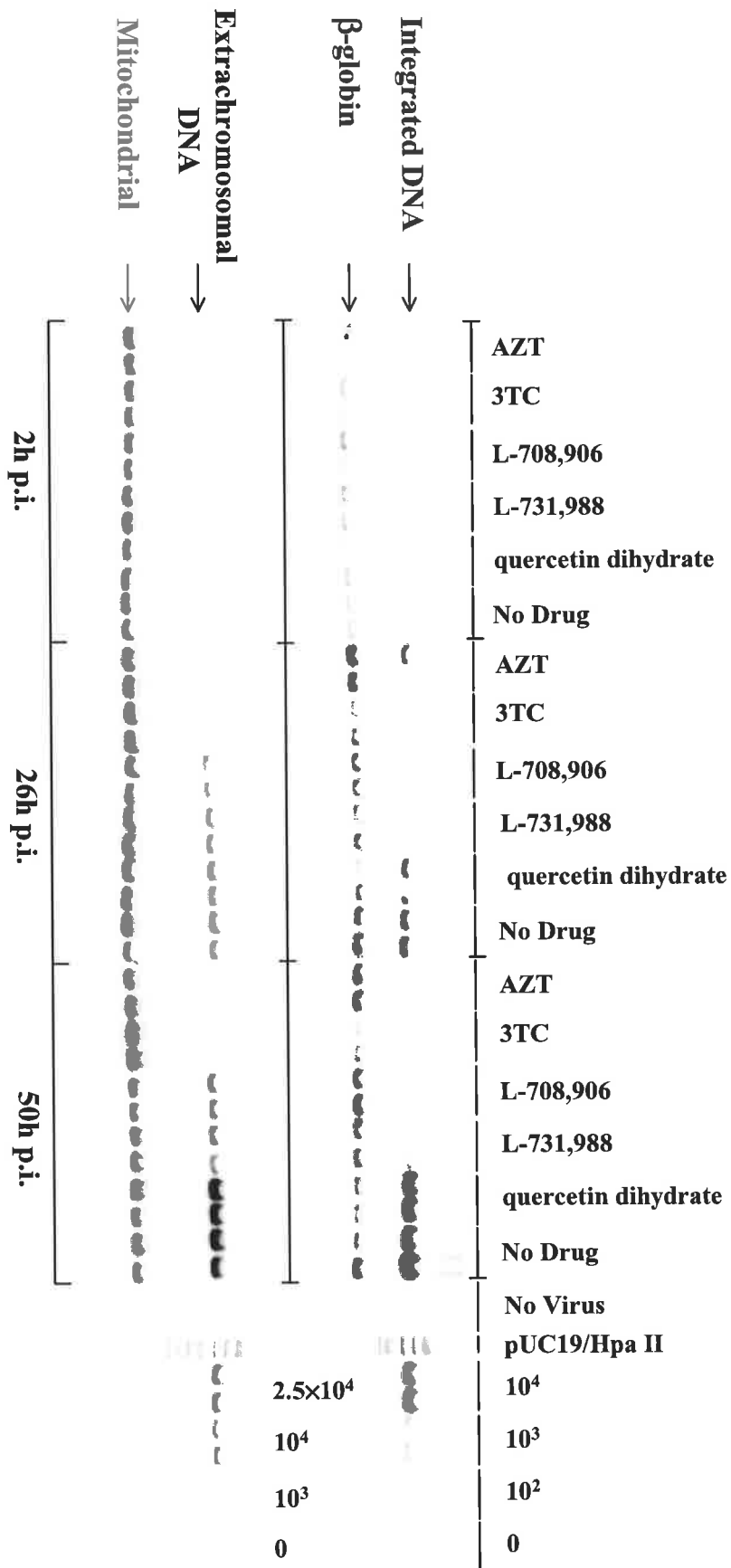
To examine the accumulation of integrated HIV DNA in the presence of each drug, HIRT chromosomal DNA preparations (see section 2.2.3) made from infected cells at 2h, 26h and 50h p.i. were subjected to a modified nested *Alu*-PCR (see section 2.2.5.iv) that specifically detects integrated HIV DNA forms. As expected, integrated HIV DNA was not detected in cultures treated with AZT and 3TC (Fig. 6.8, Integrated DNA and Fig. 6.9A). Similarly, integrated DNA accumulation was not detected in the presence of either L-708,906 or L-731,988. However, consistent with the P24 results, levels of integrated DNA observed in the presence of quercetin dihydrate at both 26h p.i. and 50h p.i. were comparable to those observed in the absence of drug. As a control, first-round PCR without the *Alu*164 primer was performed on the 50h p.i., drug-free chromosomal sample. The absence of a detectable signal confirmed that the signals observed at 50h p.i. in the drug-free samples (Fig. 6.8, No Drug, Integrated DNA) were derived from the first-round PCR amplification of integrated HIV sequences and not the nested PCR amplification of contaminating extrachromosomal forms present in the chromosomal DNA preparations (data not shown).

6.4.3.ii Extrachromosomal Viral DNA Accumulation

Since an absence of integrated HIV DNA might reflect either a specific inhibition of HIV DNA integration or a block in the HIV-1 replication cycle prior to integration, HIRT supernatant fractions from all samples were assayed using a GAG-PCR protocol that detects extrachromosomal HIV DNA, to establish whether reverse transcription was proceeding to completion. As expected, drug-free cultures and those infections performed in the presence of quercetin dihydrate contained significant amounts of

Figure 6.8

Effect of the potential integration inhibitors L-708,906, L-731,988 and quercetin dihydrate on levels of integrated and extrachromosomal HIV DNA following infection of HuT-78 cells with HIV_{HXB2}. PCRs were performed on 1000 cell-equivalents of HIRT pellet and supernatant fractions from duplicate cultures using the PCR protocols for the quantification of integrated and extrachromosomal DNA, respectively (see materials and methods). DNA levels in the presence of each potential integration inhibitor were compared with those detected in a control culture (No Drug) or after treatment with either AZT or 3TC that block DNA synthesis prior to integration. Amplification of the single-copy β -globin gene and mitochondrial DNA were used to control for the cell-equivalent amounts of chromosomal (HIRT pellet) and extrachromosomal (HIRT supernatant) DNA, respectively.



reverse transcribed products at 26h p.i., whereas those performed in with AZT and 3TC displayed negligible levels (Fig. 6.8, Extrachromosomal DNA and Fig. 6.9B). Both L-708,906 and L-731,988 also allowed the accumulation of extrachromosomal DNA by 26h p.i., although at marginally lower amounts than that observed for drug-free cultures. Extrachromosomal DNA then increased substantially from 26h p.i. to 50h p.i. in both drug-free cultures and cultures with quercetin dihydrate, while little further increase was seen in cultures containing L-708,906 and L-731,988 (Fig. 6.9B). Since infected cultures incubated in the absence of drug or the presence of quercetin dihydrate exhibited high levels of P24 by 50h p.i. (see Fig. 6.7) and extensive syncytia by 26h p.i. (data not shown), the increases in extrachromosomal DNA observed after 26h p.i. are likely to reflect *de novo* reverse transcription resulting from secondary infection of HuT-78 cells by progeny virus released from infected cells.

6.5 Discussion

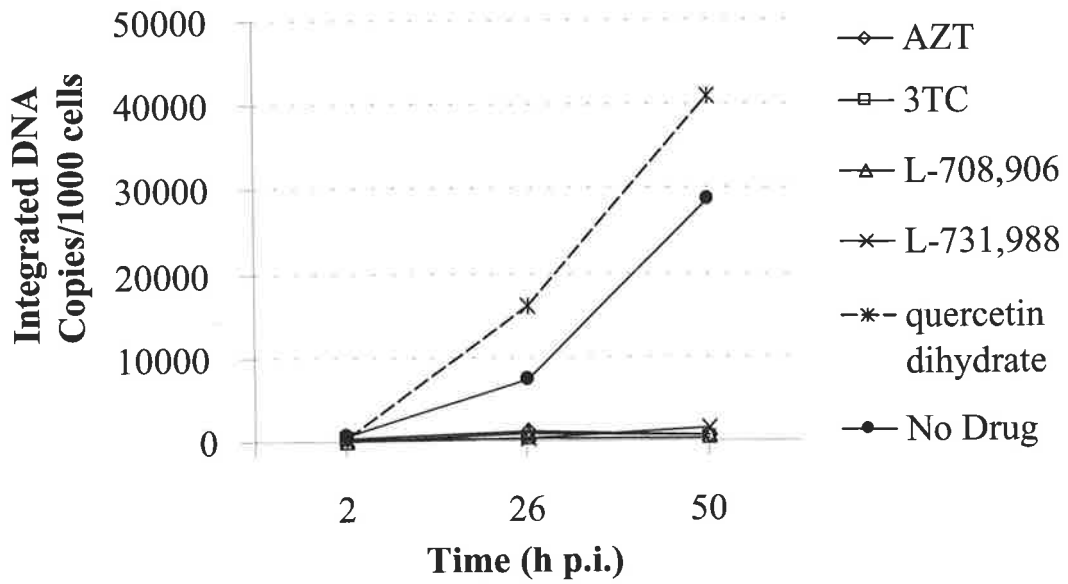
In this study, two diketo acid compounds (L-708,906 and L-731,988) inhibited the accumulation of integrated HIV-1 DNA without altering the synthesis of extrachromosomal HIV cDNA in a single-round of viral replication. Although this result is consistent with inhibition of the viral integrase protein, the drugs could also be blocking transport of newly synthesised viral DNA to the nucleus. This however is unlikely since high levels of circular viral DNA (used as a marker of viral entry into the nucleus) have been observed following infection in the presence of these drugs (111). Our results are in close agreement with previous reports indicating that viral reverse transcription is unaffected by these diketo compounds (111). It has also been shown that PICs isolated from cells infected in the presence of L-731,988 were unable to facilitate the integration of HIV DNA into a ϕ X174 DNA target substrate in a cell-free system (111). Taken together, these results indicate that L-708,906 and L-731,988 selectively block the HIV-1 integration reaction in cell culture.

Although shown in biochemical assays to inhibit the 3'-processing and strand transfer reactions at 20 μ M and 12 μ M, respectively (206), quercetin dihydrate (a weak DNA intercalator and topoisomerase 2 inhibitor) had no antiviral activity (at 50 μ M) in our experiments, based on P24 release into the culture supernatant, syncytia formation and

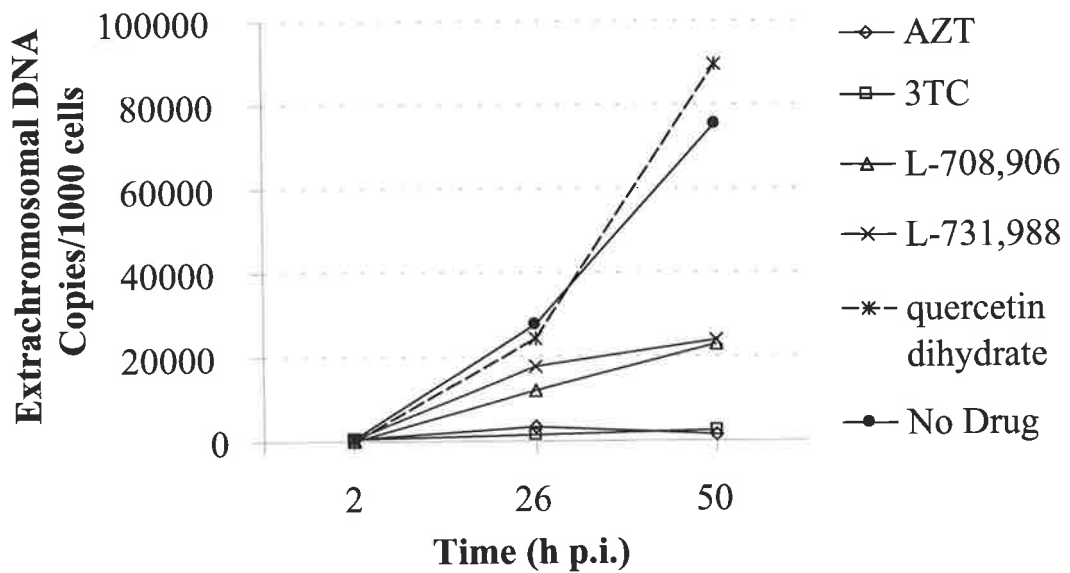
Figure 6.9

Graphical representation of data presented in Figure 6.8. Graphed values are averages of duplicate samples. **A.** Integrated DNA levels at 2h, 26h and 50h p.i.: values obtained by PhosphorImage analysis of Southern blots were adjusted based on β -globin content. **B.** Extrachromosomal DNA accumulation at 2h, 26h and 50h p.i. after adjustment for mitochondrial DNA content.

A.



B.



the accumulation of both integrated and extrachromosomal viral DNA. This finding further confirms previous observations that compounds identified in cell-free assays do not necessarily inhibit integration in cell culture. Such compounds may be denied access or inefficiently transported to their primary site(s) of action within cells. Alternatively, interactions with unrelated components within the cell might degrade or sequester these compounds making them unavailable to exert their effect.

Like AZT and 3TC, AR177 inhibited the accumulation of both integrated HIV DNA forms and extrachromosomal DNA, indicating a block in viral replication at, or prior to, reverse transcription. Our finding is consistent with recent studies showing that the primary target of AR177 is the viral gp120 protein (79) and underscores the importance of performing cell-based assays to define the precise targets of drugs within cells.

AR177 has been shown to interfere with the binding of a monoclonal antibody raised against the V3 loop of gp120, and mutations that confer viral resistance to AR177 in cell culture map to residues within the loop regions of gp120 protein (79). Along with our findings, these data suggest that the primary target of AR177 is the process of viral entry. However, it is worth noting that blocks in the production of viral DNA could potentially result from an inhibition of viral entry, an inhibition of replication complex (or PIC) formation, or a direct effect on the viral reverse transcription process. Like AR177, L17 was shown to inhibit not only the accumulation of integrated HIV DNA, but also newly reverse transcribed DNA. Although this finding suggests that the primary viral target of this drug in cell culture is unlikely to be the process of integration, the precise target of L17 cannot be elucidated without further analysis. Furthermore, until mutations conferring viral resistance to this drug are mapped, the possibility that this drug inhibits viral replication both at, or prior to, reverse transcription, as well as integration, cannot be eliminated.

In this study, we have described an efficient assay for monitoring the accumulation of integrated HIV DNA over time following infection of cells with HIV-1. When coupled with the quantitative detection of viral extrachromosomal DNA (both linear and circular forms), this assay can rapidly evaluate potential anti-integration drugs, identified in cell-free screening systems, for their ability to specifically block the HIV-1 integration process in cell culture. Similar experiments using peripheral blood mononuclear cells

(PBMCs) isolated from HIV seronegative patients will provide further data on drug efficacy in cell culture.

Chapter 7

General Discussion

7.1 Summary and Discussion

Integration is required by retroviruses such as HIV both for replication and to allow persistence within chromosomal DNA for the life-span of the infected cell. It is this latter feature that has provided a major obstacle for virus eradication *in vivo*, even when viral replication has been largely controlled (as measure by plasma viral load) by HAART (44, 47, 49, 86). Although acknowledged as a mandatory step in HIV infection, until now there was little known about the efficiency with which integration is accomplished in cells. The lack of detailed knowledge on this important step in the HIV life-cycle is in sharp contrast to other aspects of HIV virology.

The results presented in this thesis can essentially be divided into two sections. Firstly, the development of a PCR-based assay designed to specifically detect integrated HIV-1 DNA (see Chapters 3 and 4). Secondly, the application of this assay to assess both the kinetics of HIV integration and the effects various compounds have on integration frequencies, following infection of T cells (see Chapters 5 and 6). Of the two novel PCR-based assays presented, only the LP-PCR procedure allowed the detection of randomly integrated HIV DNA in both a highly sensitive and specific manner. Although the DP-PCR procedure was able to amplify individual HIV integration events, the extensive interactions of the semi-degenerate primers with cellular DNA (presented in Chapter 3) ultimately compromised the efficiency of amplification of specific integrants. Similarly, amplification of background DNA was problematic in the LP-PCR procedure. However, although not abolished completely, methods to minimise non-specific priming events were developed thereby allowing amplification of randomly integrated HIV DNA in a quantitative manner. The optimisation and validation of the LP-PCR protocol is presented in Chapter 4. This procedure was shown to specifically detect integrated HIV DNA (present within the HA8 integrated standard) in a highly sensitivity manner (10 integrated HIV copies/ 2×10^5 cells). This level of

sensitivity is comparable to that achieved by others for the detection of integrated HIV DNA within the ACH-2 clonal cell line using an alternative assay (the nested-*Alu* PCR) (49). However, in contrast to the nested-*Alu* PCR procedure, LP-PCR was shown to allow the amplification of integrated HIV DNA within the 8E5 cell line. Furthermore, LP-PCR directed the amplification of integrated HIV-1 DNA earlier after infection than the nested-*Alu* PCR protocol (see sections 5.3.5 and 5.3.6). Taken together, these results suggested that the LP-PCR protocol amplified a significantly higher proportion of randomly integrated HIV DNA than the nested-*Alu* PCR protocol. The inefficiency with which the nested-*Alu* PCR protocol was able to amplify randomly integrated HIV DNA was attributed to the long distance of *Alu* repeat sequences from the location of many HIV proviral insertion points, and to the presence of a 5'→3' exonuclease activity in the thermostable DNA polymerase used. These findings have significant ramifications for those research groups using the nested-*Alu* PCR procedure to quantify integrated HIV DNA, either in cell culture or within infected individuals. Unless copy-number standards consisting of truly randomly integrated HIV DNA are used, an accurate assessment of the proviral content cannot be made with this technique. Conversely, when a clonal cell line such as the ACH-2 cell line (1 integrant/cell; 100% of integrants theoretically detected) is used as a copy number standard, a substantial underestimation of the integrated DNA content within a sample would be expected. To address this issue, Butler and co-workers recently reported the construction of copy-number standards containing randomly integrated HIV DNA (32). Briefly, cells were infected with HIV and subsequently passaged until extrachromosomal DNA forms had degraded. As extrachromosomal HIV DNA was not present, the integrated HIV DNA content could therefore be estimated by PCR amplification of a sequence within the GAG coding region. Although a reasonable diversity of integrants would be expected to be present in such DNA preparations, the potential for outgrowth of cells containing particular integrants was not addressed. Furthermore, although the use of truly randomly integrated DNA standards would be expected to allow the accurate quantification of integrated HIV DNA within an unknown sample, the sensitivity of the nested-*Alu* PCR protocol would remain compromised due to the incomplete detection of all random integrants. This is exemplified in the publication by Butler and coworkers (32), in which the minimum copy-number standard detected was $\approx 10^3$ integrated HIV copies (in a background of 5×10^4 c.e. of cellular DNA). I believe therefore, that the LP-PCR

procedure developed in this study is a better assay than the previously published nested-*Alu* PCR protocol for the highly sensitive detection of integrated HIV DNA in both cell culture and in cells isolated from HIV-positive patients.

Using the newly developed LP-PCR protocol, the kinetics of HIV DNA integration following a near one-step infection of T cells was assessed. In addition, the accumulation of all reverse transcribed viral DNA, and 2-LTR viral DNA, was monitored using different PCR-based assays to allow an extensive profile of viral DNA metabolism following infection to be made (Chapter 5). Integrated HIV-1 DNA was first detected 4h after infection, approximately 1h after the first appearance of mid to late-stage reverse transcribed viral DNA, and accounted for approximately 8% of all reverse transcribed DNA synthesised following infection. These results closely paralleled those of Butler and co-workers (32), who found using the nested-*Alu* PCR protocol (and appropriate standards) that the majority (approximately 95%) of reverse transcribed DNA did not successfully integrate into the cellular chromosome. However, Butler and co-workers were unable to detect integrated HIV DNA before 12h p.i.. This is surprising since nuclear transport of newly synthesised viral DNA was observed (as judged by 2-LTR accumulation) in this same study by ≈ 4 h p.i.. We therefore believe that absence of integrated HIV DNA detected at earlier time-points is likely to reflect the lower level of detection sensitivity achieved by these researchers (approximately 900 integrated copies/ $\approx 5 \times 10^4$ cells) and not a delay in the integration process upon nuclear entry of viral DNA. Furthermore, the 2-LTR extrachromosomal HIV DNA species was confirmed in our study to be a very minor species ($\approx 1\%$) following infection. This finding is in close agreement with the observations made by Butler and colleagues (32) that $\approx 0.7\%$ of total viral DNA made following HIV infection was converted to the 2-LTR viral DNA form. Since circular DNA forms appeared to be relatively stable within the intracellular environment, the large amount of viral DNA degradation observed by us and others (8, 32, 138) to be occurring following infection, is likely to reflect decay of linear DNA forms.

One of the first studies investigating the kinetics of HIV-1 DNA accumulation following infection of T cells was performed by Kim and co-workers (138). While integrated HIV DNA was not assessed in this study, under their infection conditions and

using Southern hybridisation techniques, linear viral DNA and circular forms were observed at 4h p.i. and 8h p.i., respectively. Furthermore, linear DNA was found within the nucleus by 5h p.i.. Although in close agreement with our findings (3h and 7h p.i. for linear and 2-LTR, respectively), the slight delay in the appearance of these forms (≈ 1 h) observed in this study may reflect the lower detection sensitivity of the Southern hybridisation techniques used compared to the PCR assays used in our study. Moreover, the lower levels (approximately 10-fold) of infection achieved in cultures under their infection conditions may have also contributed to the absence of these viral DNA forms at earlier time points. In addition to analysing the accumulation of viral DNA, Kim and co-workers also assessed levels of viral RNA synthesis by Northern analysis following infection. The earliest RNA species observed were the 2-kb RNAs (encoding regulatory proteins such as *tat*, *rev* and *nef*), which were apparent from 12h p.i.. The 4.3-kb and full-length RNA species (encoding *env* and *gag/pol* proteins, respectively) were apparent from 16h p.i.. However, similar analyses performed using similar infection conditions but a more sensitive RT-PCR procedure, showed that viral RNA species could be detected between 8-12h p.i. (144). Although differing slightly with respect to the time at which viral RNA first appeared, these results suggested that there was a significant lag between the time at which viral DNA was present in the nucleus (and presumably able to integrate) and the time at which substantial levels of viral RNA synthesis were achieved. This is also implied from the results obtained in our study, where P24 release was not observed until 26h p.i., even though integrated HIV DNA was present by 4h p.i.. Thus, approximately 20h (or $\approx 80\%$ of the viral life-cycle assuming one-round completed by 26h p.i.) was required after integration before viral proteins were able to accumulate to sufficient levels to mediate progeny virion formation. Taken together, these studies suggest that the rate at which a full-round of viral replication is able to occur within cells is likely to be more heavily dependent on events associated with transcription and translation of viral genes than on the completion of events up to, and including, viral DNA integration. The reasons for this delay are at present unknown.

In addition to assessing the kinetics of viral DNA accumulation following infection of T cells, an assay was established that was able to accurately assess the effects of various compounds on HIV DNA integration in cell culture. The accumulation of total HIV

DNA was also assessed to determine whether any inhibition of integration observed reflected an inhibition of events before or after reverse transcription. The nested-*Alu* PCR procedure was used here instead of the LP-PCR protocol, as it was substantially easier to perform and the highly accurate quantification of integrated HIV DNA levels following infection was not required. Of the seven compounds tested (including two reverse transcriptase inhibitors), only two compounds (L-731,988 and L-708,906) were shown to inhibit the integration of HIV DNA without significantly affecting the reverse transcription process. The finding that L-731,988 (previously reported to inhibit HIV replication in cell culture (111)) exerted its effect at the level of viral DNA integration, was further confirmed using the LP-PCR protocol (see section 4.9.1.ii). Furthermore, Butler and colleagues recently demonstrated (using their nested-*Alu* PCR protocol) that integration was abolished during infections performed in the presence of L-731,988 (32). To our knowledge, these are the first experiments clearly demonstrating the direct effect of compounds on the HIV integration process. The identification of lead compounds followed by the ability to accurately assess whether such compounds specifically affect HIV integration, provides a solid basis for future rational drug design targeting the integration process. Furthermore, future experiments assessing dilutions of compounds for effects on integration may allow a measurement of inhibition efficiency to be made. However, since this would require the accurate (and sensitive) assessment of integration levels following infection, the LP-PCR assay should be used instead of nested-*Alu* protocols for such experiments.

Currently, antiviral drugs are available that target either the processes of reverse transcription or processing of the Gag-Pol polyprotein (see section 1.6.1). Although the frequency with which escape mutants were generated was dramatically reduced when protease inhibitors were included with reverse transcriptase inhibitors in clinical treatment regimens (103, 105), viruses exhibiting drug-resistance have been isolated from patients on such therapies (278). Furthermore, a number of these drugs have been reported to mediate substantial side-effects in some patients (see section 1.6.1). Therefore, drugs specifically targeting the integration of HIV DNA will provide an important adjunct to current treatment strategies. The presence of a drug targeting a third vital process in the HIV replication cycle should further reduce the chance that viral strains resistant to therapy will be generated. Furthermore, since processes analogous to integration have not been identified within human cells, the side-effects

associated with the use of such anti-integration drugs (although currently unknown) may be minimal.

7.2 Future Work

When compared to most other PCR procedures (including the nested-*Alu* PCR procedure), the LP-PCR assay developed in this study is a relatively labour-intensive protocol and consequently not well suited to the simultaneous processing of large numbers of samples. To simplify the technique, and avoid chance error associated with Southern transfer and hybridisation techniques, the nested PCR could be performed in a real-time PCR machine. Although not affecting sample preparation time prior to LP-PCR, conversion to a real-time PCR format (already achieved for nested-*Alu* PCR protocol) would considerably reduce the workload associated with this assay by allowing the levels of amplified product within each reaction to be immediately assessed upon completion of PCR cycling.

In this study, the kinetics of HIV DNA accumulation following infection of T cells with HIV_{HXB2} was assessed. To assess whether the strain of virus or cell type affects the rate of viral DNA synthesis and integration, similar studies could be performed using alternative T cell tropic viruses (*eg.* NL4.3) and/or target cell-types (*eg.* Jurkat, Molt-4). Cells of the macrophage/monocyte lineage are generally considered to be non-dividing cells (231, 260) and therefore might be expected to exhibit different rates of viral DNA synthesis and integration following infection (with appropriate viral strains) in culture. Moreover, in contrast to T cells and macrophages, cells such as astrocytes (present in the CNS) appear not to support full rounds of viral replication. Astrocytes within brain sections of infected individuals have been shown to contain early markers of HIV-1 infection (*eg.* viral DNA and more rarely, gene products such as Nef), indicating that virus is able to enter and initiate reverse transcription in these cells (186, 207, 217, 241, 242). However, later markers of HIV-1 infection (such as structural gene transcripts and products) were present in small minority (<10%) of these brain-derived astrocytes indicating that a block in the viral replication cycle may be occurring in these cells *in vivo*. The possibility that astrocytes may constitute an additional HIV reservoir consisting of low-producing or non-producing infected cells has been supported by a

number of experiments on astrocyte cell lines in cell culture (for review see (16)). Although some transfection studies have indicated that blocks may exist at the level of both viral transcription and/or translation (99, 100), the precise site(s) at which viral replication is potentially blocked in astrocytes following infection remains unknown. Furthermore, since only events after integration of viral DNA can be assessed in these transfection experiments, blocks at earlier stages in the HIV life-cycle could not be addressed. Therefore, determining whether the efficiency with which integration occurs in these cells is altered will have important implications for understanding astrocyte infection *in vivo*. The PCR-based assays presented in this thesis provide a means of dissecting viral replication blocks following infection in a variety of cells and virus strains.

It is also worth emphasising that the principle of the LP-PCR procedure presented in this study can be applied to other retroviruses (and indeed for the detection of other known elements) present within chromosomal DNA. For example, researchers in our laboratory have demonstrated that cell-to-cell HTLV-I infection is blocked prior to virion release in cell culture (as judged by release of the viral antigen P19 into the culture supernatant). Moreover, viral transcription appeared not to be occurring following infection (10). However, mid to late-stage viral DNA products were detected following infection indicating that virus was able to enter cells and reverse transcribe the genomic RNA. Preliminary evidence (obtained using a modified nested-*Alu* PCR protocol) has suggested that integration can occur following cell-to-cell HTLV-I infection although the efficiency with which viral DNA integrates was lower than that observed following HIV infection (Purins *et al.*, unpublished). These findings could be further confirmed and more precisely quantified using the LP-PCR modified to detect integrated HTLV-1 DNA.

Specific assays for integrated HIV DNA have been used previously to establish sites of viral persistence in HIV-positive patients on HAART (45, 47, 49, 86). Both CD4⁺ memory T cells and monocytes isolated from such patients have been shown to contain significant amounts of integrated DNA able to direct the production of infectious virus in cell culture (49, 231). However, to date few other cell types within infected individuals on HAART have been investigated for the presence of integrated HIV DNA. Recently, HIV DNA was detected by a *gag*-PCR protocol in purified populations of

CD4⁻CD8⁻ (DN) T cells isolated from patients on HAART (171). Furthermore, in a separate study, a nested-*Alu* PCR protocol was used to demonstrate detectable integrated HIV DNA within highly purified populations of DN T cells in 8/9 HIV-infected patients (149). Since nested-*Alu* PCR protocols have been shown in this thesis to substantially underestimate levels of integrated HIV DNA within cells, LP-PCR may be a more appropriate tool for assessing the extent of integrated HIV DNA in different cell types. Such experiments could be combined with coculture experiments to assess culturable virus in cells harbouring integrated HIV DNA, and provide valuable information regarding viral persistence in patients receiving HAART.

To further characterise inhibitors that are found to specifically block HIV-1 integration in T cells, and provide additional evidence to support their development as a viable antiviral compound, similar experiments conducted using alternative cell types should be performed. For example, PBMCs isolated from uninfected individuals contain various subsets of T cells, macrophages and monocytes. The precise mechanism of HIV DNA integration may differ between these cell types due to variation in the amounts of cellular proteins and/or other intracellular components expressed within these cells. Furthermore, subtle differences in the amino acid composition between the integrase proteins of different viral strains may influence the interactions of various compounds with the protein. Therefore, the analysis of integration inhibitors using genetically distinct viral strains and diverse target cell types may allow a more accurate prediction of drug-efficacy *in vivo*.

Bibliography

1. **Gregg, M. B.** 1981. Pneumocystis pneumonia--Los Angeles. *Morb Mortal Wkly Rep.* **30**:250-2.
2. **Acel, A., B. E. Udashkin, M. A. Wainberg, and E. A. Faust.** 1998. Efficient gap repair catalyzed in vitro by an intrinsic DNA polymerase activity of human immunodeficiency virus type 1 integrase. *J Virol.* **72**:2062-71.
3. **Ahmed, R., and D. Gray.** 1996. Immunological memory and protective immunity: understanding their relation. *Science.* **272**:54-60.
4. **Aiyar, A., P. Hindmarsh, A. M. Skalka, and J. Leis.** 1996. Concerted integration of linear retroviral DNA by the avian sarcoma virus integrase in vitro: dependence on both long terminal repeat termini. *J Virol.* **70**:3571-80.
5. **Allawi, H. T., and J. SantaLucia, Jr.** 1998. Nearest neighbor thermodynamic parameters for internal G.A mismatches in DNA. *Biochemistry.* **37**:2170-9.
6. **Allawi, H. T., and J. SantaLucia, Jr.** 1998. Nearest-neighbor thermodynamics of internal A.C mismatches in DNA: sequence dependence and pH effects. *Biochemistry.* **37**:9435-44.
7. **Asante-Appiah, E., S. H. Seeholzer, and A. M. Skalka.** 1998. Structural determinants of metal-induced conformational changes in HIV-1 integrase. *J Biol Chem.* **273**:35078-87.
8. **Barbosa, P., P. Charneau, N. Dumey, and F. Clavel.** 1994. Kinetic analysis of HIV-1 early replicative steps in a coculture system. *AIDS Res Hum Retroviruses.* **10**:53-9.
9. **Barre-Sinoussi, F., J. C. Chermann, F. Rey, M. T. Nugeyre, S. Chamaret, J. Gruest, C. Dautet, C. Axler-Blin, F. Vezinet-Brun, C. Rouzioux, W. Rozenbaum, and L. Montagnier.** 1983. Isolation of a T-lymphotropic retrovirus from a patient at risk for acquired immune deficiency syndrome (AIDS). *Science.* **220**:868-71.
10. **Benovic, S., T. Kok, A. Stephenson, J. McInnes, C. Burrell, and P. Li.** 1998. De novo reverse transcription of HTLV-1 following cell-to-cell transmission of infection. *Virology.* **244**:294-301.

11. **Bhat, T. N., E. T. Baldwin, B. Liu, Y. S. Cheng, and J. W. Erickson.** 1994. Crystal structure of a tethered dimer of HIV-1 proteinase complexed with an inhibitor. *Nat Struct Biol.* **1**:552-6.
12. **Blankson, J. N., D. Persaud, and R. F. Siliciano.** 2002. The challenge of viral reservoirs in hiv-1 infection. *Annu Rev Med.* **53**:557-93.
13. **Bleul, C. C., M. Farzan, H. Choe, C. Parolin, I. Clark-Lewis, J. Sodroski, and T. A. Springer.** 1996. The lymphocyte chemoattractant SDF-1 is a ligand for LESTR/fusin and blocks HIV-1 entry. *Nature.* **382**:829-33.
14. **Bor, Y. C., M. D. Miller, F. D. Bushman, and L. E. Orgel.** 1996. Target-sequence preferences of HIV-1 integration complexes in vitro. *Virology.* **222**:283-8.
15. **Bowerman, B., P. O. Brown, J. M. Bishop, and H. E. Varmus.** 1989. A nucleoprotein complex mediates the integration of retroviral DNA. *Genes Dev.* **3**:469-78.
16. **Brack-Werner, R.** 1999. Astrocytes: HIV cellular reservoirs and important participants in neuropathogenesis. *AIDS.* **13**:1-22.
17. **Brinkman, K., J. A. Smeitink, J. A. Romijn, and P. Reiss.** 1999. Mitochondrial toxicity induced by nucleoside-analogue reverse-transcriptase inhibitors is a key factor in the pathogenesis of antiretroviral-therapy-related lipodystrophy. *Lancet.* **354**:1112-5.
18. **Brodin, P., M. Gottikh, C. Auclair, and J. F. Mouscadet.** 1999. Inhibition of HIV-1 integration by mono- & bi-functionalized triple helix forming oligonucleotides. *Nucleosides Nucleotides.* **18**:1717-8.
19. **Brodin, P., J. S. Sun, J. F. Mouscadet, and C. Auclair.** 1999. Optimization of alternate-strand triple helix formation at the 5"-TpA-3" and 5"-ApT-3" junctions. *Nucleic Acids Res.* **27**:3029-34.
20. **Brown, P. O., B. Bowerman, H. E. Varmus, and J. M. Bishop.** 1989. Retroviral integration: structure of the initial covalent product and its precursor, and a role for the viral IN protein. *Proc Natl Acad Sci U S A.* **86**:2525-9.
21. **Bujacz, G., M. Jaskolski, J. Alexandratos, A. Wlodawer, G. Merkel, R. A. Katz, and A. M. Skalka.** 1996. The catalytic domain of avian sarcoma virus integrase: conformation of the active-site residues in the presence of divalent cations. *Structure.* **4**:89-96.

22. **Bukrinsky, M., N. Sharova, and M. Stevenson.** 1993. Human immunodeficiency virus type 1 2-LTR circles reside in a nucleoprotein complex which is different from the preintegration complex. *J Virol.* **67**:6863-5.
23. **Bukrinsky, M. I., N. Sharova, M. P. Dempsey, T. L. Stanwick, A. G. Bukrinskaya, S. Haggerty, and M. Stevenson.** 1992. Active nuclear import of human immunodeficiency virus type 1 preintegration complexes. *Proc Natl Acad Sci U S A.* **89**:6580-4.
24. **Bukrinsky, M. I., N. Sharova, T. L. McDonald, T. Pushkarskaya, W. G. Tarpley, and M. Stevenson.** 1993. Association of integrase, matrix, and reverse transcriptase antigens of human immunodeficiency virus type 1 with viral nucleic acids following acute infection. *Proc Natl Acad Sci U S A.* **90**:6125-9.
25. **Burke, C. J., G. Sanyal, M. W. Bruner, J. A. Ryan, R. L. LaFemina, H. L. Robbins, A. S. Zeft, C. R. Middaugh, and M. G. Cordingley.** 1992. Structural implications of spectroscopic characterization of a putative zinc finger peptide from HIV-1 integrase. *J Biol Chem.* **267**:9639-44.
26. **Burke, T. R., Jr., M. R. Fesen, A. Mazumder, J. Wang, A. M. Carothers, D. Grunberger, J. Driscoll, K. Kohn, and Y. Pommier.** 1995. Hydroxylated aromatic inhibitors of HIV-1 integrase. *J Med Chem.* **38**:4171-8.
27. **Bushman, F. D.** 1999. Host proteins in retroviral cDNA integration. *Adv Virus Res.* **52**:301-17.
28. **Bushman, F. D.** 1994. Tethering human immunodeficiency virus 1 integrase to a DNA site directs integration to nearby sequences. *Proc Natl Acad Sci U S A.* **91**:9233-7.
29. **Bushman, F. D., and R. Craigie.** 1991. Activities of human immunodeficiency virus (HIV) integration protein in vitro: specific cleavage and integration of HIV DNA. *Proc Natl Acad Sci U S A.* **88**:1339-43.
30. **Bushman, F. D., T. Fujiwara, and R. Craigie.** 1990. Retroviral DNA integration directed by HIV integration protein in vitro. *Science.* **249**:1555-8.
31. **Bushman, F. D., and M. D. Miller.** 1997. Tethering human immunodeficiency virus type 1 preintegration complexes to target DNA promotes integration at nearby sites. *J Virol.* **71**:458-64.
32. **Butler, S. L., M. S. Hansen, and F. D. Bushman.** 2001. A quantitative assay for HIV DNA integration in vivo. *Nature Med.* **7**:631-4.

33. **Cai, M., R. Zheng, M. Caffrey, R. Craigie, G. M. Clore, and A. M. Gronenborn.** 1997. Solution structure of the N-terminal zinc binding domain of HIV-1 integrase. *Nat Struct Biol.* **4**:567-77.
34. **Cara, A., A. Cereseto, F. Lori, and M. S. Reitz, Jr.** 1996. HIV-1 protein expression from synthetic circles of DNA mimicking the extrachromosomal forms of viral DNA. *J Biol Chem.* **271**:5393-7.
35. **Carr, A., K. Samaras, D. J. Chisholm, and D. A. Cooper.** 1998. Pathogenesis of HIV-1-protease inhibitor-associated peripheral lipodystrophy, hyperlipidaemia, and insulin resistance. *Lancet.* **351**:1881-3.
36. **Carteau, S., C. Hoffmann, and F. Bushman.** 1998. Chromosome structure and human immunodeficiency virus type 1 cDNA integration: centromeric alphoid repeats are a disfavored target. *J Virol.* **72**:4005-14.
37. **Chan, P. P., and P. M. Glazer.** 1997. Triplex DNA: fundamentals, advances, and potential applications for gene therapy. *J Mol Med.* **75**:267-82.
38. **Chang, D. D., and P. A. Sharp.** 1989. Regulation by HIV Rev depends upon recognition of splice sites. *Cell.* **59**:789-95.
39. **Chen, H., and A. Engelman.** 1998. The barrier-to-autointegration protein is a host factor for HIV type 1 integration. *Proc Natl Acad Sci U S A.* **95**:15270-4.
40. **Chen, J. C., J. Krucinski, L. J. Miercke, J. S. Finer-Moore, A. H. Tang, A. D. Leavitt, and R. M. Stroud.** 2000. Crystal structure of the HIV-1 integrase catalytic core and C-terminal domains: a model for viral DNA binding. *Proc Natl Acad Sci U S A.* **97**:8233-8.
41. **Cherepanov, P., J. A. Este, R. F. Rando, J. O. Ojwang, G. Reekmans, R. Steinfeld, G. David, E. De Clercq, and Z. Debyser.** 1997. Mode of interaction of G-quartets with the integrase of human immunodeficiency virus type 1. *Mol Pharmacol.* **52**:771-80.
42. **Choe, H., M. Farzan, Y. Sun, N. Sullivan, B. Rollins, P. D. Ponath, L. Wu, C. R. Mackay, G. LaRosa, W. Newman, N. Gerard, C. Gerard, and J. Sodroski.** 1996. The beta-chemokine receptors CCR3 and CCR5 facilitate infection by primary HIV-1 isolates. *Cell.* **85**:1135-48.
43. **Chow, S. A., K. A. Vincent, V. Ellison, and P. O. Brown.** 1992. Reversal of integration and DNA splicing mediated by integrase of human immunodeficiency virus. *Science.* **255**:723-6.

44. **Chun, T. W., L. Carruth, D. Finzi, X. Shen, J. A. DiGiuseppe, H. Taylor, M. Hermankova, K. Chadwick, J. Margolick, T. C. Quinn, Y. H. Kuo, R. Brookmeyer, M. A. Zeiger, P. Barditch-Crovo, and R. F. Siliciano.** 1997. Quantification of latent tissue reservoirs and total body viral load in HIV-1 infection. *Nature*. **387**:183-8.
45. **Chun, T. W., R. T. Davey, Jr., M. Ostrowski, J. Shawn Justement, D. Engel, J. I. Mullins, and A. S. Fauci.** 2000. Relationship between pre-existing viral reservoirs and the re-emergence of plasma viremia after discontinuation of highly active anti-retroviral therapy. *Nature Med*. **6**:757-61.
46. **Chun, T. W., D. Engel, M. M. Berrey, T. Shea, L. Corey, and A. S. Fauci.** 1998. Early establishment of a pool of latently infected, resting CD4(+) T cells during primary HIV-1 infection. *Proc Natl Acad Sci U S A*. **95**:8869-73.
47. **Chun, T. W., and A. S. Fauci.** 1999. Latent reservoirs of HIV: obstacles to the eradication of virus. *Proc Natl Acad Sci U S A*. **96**:10958-61.
48. **Chun, T. W., D. Finzi, J. Margolick, K. Chadwick, D. Schwartz, and R. F. Siliciano.** 1995. In vivo fate of HIV-1-infected T cells: quantitative analysis of the transition to stable latency. *Nature Med*. **1**:1284-90.
49. **Chun, T. W., L. Stuyver, S. B. Mizell, L. A. Ehler, J. A. Mican, M. Baseler, A. L. Lloyd, M. A. Nowak, and A. S. Fauci.** 1997. Presence of an inducible HIV-1 latent reservoir during highly active antiretroviral therapy. *Proc Natl Acad Sci U S A*. **94**:13193-7.
50. **Clapham, P.** 1997. HIV and chemokines: Ligands sharing cell-surface receptors. *Trends in Cell Biology*. **7**:264-268.
51. **Clapham, P. R., and R. A. Weiss.** 1997. Immunodeficiency viruses. Spoilt for choice of co-receptors. *Nature*. **388**:230-1.
52. **Clark, S. J., M. S. Saag, W. D. Decker, S. Campbell-Hill, J. L. Roberson, P. J. Veldkamp, J. C. Kappes, B. H. Hahn, and G. M. Shaw.** 1991. High titers of cytopathic virus in plasma of patients with symptomatic primary HIV-1 infection. *N Engl J Med*. **324**:954-60.
53. **Clouse, K. A., D. Powell, I. Washington, G. Poli, K. Strebel, W. Farrar, P. Barstad, J. Kovacs, A. S. Fauci, and T. M. Folks.** 1989. Monokine regulation of human immunodeficiency virus-1 expression in a chronically infected human T cell clone. *J Immunology*. **142**:431-8.

54. **Coates, J. A., N. Cammack, H. J. Jenkinson, A. J. Jowett, M. I. Jowett, B. A. Pearson, C. R. Penn, P. L. Rouse, K. C. Viner, and J. M. Cameron.** 1992. (-)-2'-deoxy-3'-thiacytidine is a potent, highly selective inhibitor of human immunodeficiency virus type 1 and type 2 replication in vitro. *Antimicrob Agents Chemother.* **36**:733-9.
55. **Coffin, J. M., S. H. Hughes, and H. E. Varmus.** 1997. *Retroviruses.* Cold Spring Harbor Laboratory Press.
56. **Craigie, R., T. Fujiwara, and F. Bushman.** 1990. The IN protein of Moloney murine leukemia virus processes the viral DNA ends and accomplishes their integration in vitro. *Cell.* **62**:829-37.
57. **Cullen, B. R.** 1986. Trans-activation of human immunodeficiency virus occurs via a bimodal mechanism. *Cell.* **46**:973-82.
58. **Daniel, R., R. A. Katz, and A. M. Skalka.** 1999. A role for DNA-PK in retroviral DNA integration. *Science.* **284**:644-7.
59. **Davies, J. F., 2nd, Z. Hostomska, Z. Hostomsky, S. R. Jordan, and D. A. Matthews.** 1991. Crystal structure of the ribonuclease H domain of HIV-1 reverse transcriptase. *Science.* **252**:88-95.
60. **Davis, A. J., P. Li, and C. J. Burrell.** 1997. Kinetics of viral RNA synthesis following cell-to-cell transmission of human immunodeficiency virus type 1. *J Gen Virol.* **78**:1897-906.
61. **De Clercq, E.** 2001. New Developments in Anti-HIV Chemotherapy. *Curr Med Chem.* **8**:1529-1558.
62. **Depienne, C., A. Mousnier, H. Leh, E. Le Rouzic, D. Dormont, S. Benichou, and C. Dargemont.** 2001. Characterization of the nuclear import pathway for HIV-1 integrase. *J Biol Chem.*
63. **Depienne, C., P. Roques, C. Creminon, L. Fritsch, R. Casseron, D. Dormont, C. Dargemont, and S. Benichou.** 2000. Cellular distribution and karyophilic properties of matrix, integrase, and Vpr proteins from the human and simian immunodeficiency viruses. *Exp Cell Res.* **260**:387-95.
64. **Dornadula, G., H. Zhang, O. Bagasra, and R. J. Pomerantz.** 1997. Natural endogenous reverse transcription of simian immunodeficiency virus. *Virology.* **227**:260-7.
65. **Dornadula, G., H. Zhang, B. VanUitert, J. Stern, L. Livornese, Jr., M. J. Ingerman, J. Witek, R. J. Kedanis, J. Natkin, J. DeSimone, and R. J.**

- Pomerantz.** 1999. Residual HIV-1 RNA in blood plasma of patients taking suppressive highly active antiretroviral therapy. *Jama.* **282**:1627-32.
66. **Drelich, M., R. Wilhelm, and J. Mous.** 1992. Identification of amino acid residues critical for endonuclease and integration activities of HIV-1 IN protein in vitro. *Virology.* **188**:459-68.
67. **Eckert, D. M., and P. S. Kim.** 2001. Design of potent inhibitors of HIV-1 entry from the gp41 N-peptide region. *Proc Natl Acad Sci U S A.* **98**:11187-92.
68. **Ellison, V., H. Abrams, T. Roe, J. Lifson, and P. Brown.** 1990. Human immunodeficiency virus integration in a cell-free system. *J Virol.* **64**:2711-5.
69. **Ellison, V., and P. O. Brown.** 1994. A stable complex between integrase and viral DNA ends mediates human immunodeficiency virus integration in vitro. *Proc Natl Acad Sci U S A.* **91**:7316-20.
70. **Engelman, A.** 1999. In vivo analysis of retroviral integrase structure and function. *Adv Virus Res.* **52**:411-26.
71. **Engelman, A., F. D. Bushman, and R. Craigie.** 1993. Identification of discrete functional domains of HIV-1 integrase and their organization within an active multimeric complex. *EMBO J.* **12**:3269-75.
72. **Engelman, A., and R. Craigie.** 1995. Efficient magnesium-dependent human immunodeficiency virus type 1 integrase activity. *J Virol.* **69**:5908-11.
73. **Engelman, A., and R. Craigie.** 1992. Identification of conserved amino acid residues critical for human immunodeficiency virus type 1 integrase function in vitro. *J Virol.* **66**:6361-9.
74. **Engelman, A., A. B. Hickman, and R. Craigie.** 1994. The core and carboxyl-terminal domains of the integrase protein of human immunodeficiency virus type 1 each contribute to nonspecific DNA binding. *J Virol.* **68**:5911-7.
75. **Engelman, A., K. Mizuuchi, and R. Craigie.** 1991. HIV-1 DNA integration: mechanism of viral DNA cleavage and DNA strand transfer. *Cell.* **67**:1211-21.
76. **Englund, G., T. S. Theodore, E. O. Freed, A. Engleman, and M. A. Martin.** 1995. Integration is required for productive infection of monocyte-derived macrophages by human immunodeficiency virus type 1. *J Virol.* **69**:3216-9.
77. **Espeseth, A. S., P. Felock, A. Wolfe, M. Witmer, J. Grobler, N. Anthony, M. Egbertson, J. Y. Melamed, S. Young, T. Hamill, J. L. Cole, and D. J. Hazuda.** 2000. HIV-1 integrase inhibitors that compete with the target DNA

- substrate define a unique strand transfer conformation for integrase [In Process Citation]. *Proc Natl Acad Sci U S A.* **97**:11244-9.
78. **Esposito, D., and R. Craigie.** 1999. HIV integrase structure and function. *Adv Virus Res.* **52**:319-33.
79. **Este, J. A., C. Cabrera, D. Schols, P. Cherepanov, A. Gutierrez, M. Witvrouw, C. Pannecouque, Z. Debyser, R. F. Rando, B. Clotet, J. Desmyter, and E. De Clercq.** 1998. Human immunodeficiency virus glycoprotein gp120 as the primary target for the antiviral action of AR177 (Zintevir). *Mol Pharmacol.* **53**:340-5.
80. **Farnet, C. M., and F. D. Bushman.** 1997. HIV-1 cDNA integration: requirement of HMG I(Y) protein for function of preintegration complexes in vitro. *Cell.* **88**:483-92.
81. **Farnet, C. M., and W. A. Haseltine.** 1991. Circularization of human immunodeficiency virus type 1 DNA in vitro. *J Virol.* **65**:6942-52.
82. **Farnet, C. M., B. Wang, J. R. Lipford, and F. D. Bushman.** 1996. Differential inhibition of HIV-1 preintegration complexes and purified integrase protein by small molecules. *Proc Natl Acad Sci U S A.* **93**:9742-7.
83. **Fenyo, E. M., J. Albert, and B. Asjo.** 1989. Replicative capacity, cytopathic effect and cell tropism of HIV. *AIDS.* **3**:S5-12.
84. **Fesen, M. R., K. W. Kohn, F. Leteurtre, and Y. Pommier.** 1993. Inhibitors of human immunodeficiency virus integrase. *Proc Natl Acad Sci U S A.* **90**:2399-403.
85. **Fesen, M. R., Y. Pommier, F. Leteurtre, S. Hiroguchi, J. Yung, and K. W. Kohn.** 1994. Inhibition of HIV-1 integrase by flavones, caffeic acid phenethyl ester (CAPE) and related compounds. *Biochem Pharmacol.* **48**:595-608.
86. **Finzi, D., J. Blankson, J. D. Siliciano, J. B. Margolick, K. Chadwick, T. Pierson, K. Smith, J. Lisziewicz, F. Lori, C. Flexner, T. C. Quinn, R. E. Chaisson, E. Rosenberg, B. Walker, S. Gange, J. Gallant, and R. F. Siliciano.** 1999. Latent infection of CD4⁺ T cells provides a mechanism for lifelong persistence of HIV-1, even in patients on effective combination therapy. *Nature Med.* **5**:512-7.
87. **Fischer, U., J. Huber, W. C. Boelens, I. W. Mattaj, and R. Luhrmann.** 1995. The HIV-1 Rev activation domain is a nuclear export signal that accesses an export pathway used by specific cellular RNAs. *Cell.* **82**:475-83.

88. **Fitzgerald, M. L., and D. P. Grandgenett.** 1994. Retroviral integration: in vitro host site selection by avian integrase. *J Virol.* **68**:4314-21.
89. **Folks, T. M., D. Powell, M. Lightfoote, S. Koenig, A. S. Fauci, S. Benn, A. Rabson, D. Daugherty, H. E. Gendelman, and M. D. Hoggan.** 1986. Biological and biochemical characterization of a cloned Leu-3- cell surviving infection with the acquired immune deficiency syndrome retrovirus. *J Exp Med.* **164**:280-90.
90. **Fouchier, R. A., and M. H. Malim.** 1999. Nuclear import of human immunodeficiency virus type-1 preintegration complexes. *Adv Virus Res.* **52**:275-99.
91. **Fouchier, R. A., B. E. Meyer, J. H. Simon, U. Fischer, and M. H. Malim.** 1997. HIV-1 infection of non-dividing cells: evidence that the amino-terminal basic region of the viral matrix protein is important for Gag processing but not for post-entry nuclear import. *EMBO J.* **16**:4531-9.
92. **Fu, D. X., A. Jinno, N. Shimizu, Y. Haraguchi, and H. Hoshino.** 1999. Isolation and characterization of a monoclonal antibody that inhibits HIV-1 infection. *Microbes Infect.* **1**:677-84.
93. **Fujiwara, T., and K. Mizuuchi.** 1988. Retroviral DNA integration: structure of an integration intermediate. *Cell.* **54**:497-504.
94. **Furtado, M. R., D. S. Callaway, J. P. Phair, K. J. Kunstman, J. L. Stanton, C. A. Macken, A. S. Perelson, and S. M. Wolinsky.** 1999. Persistence of HIV-1 transcription in peripheral-blood mononuclear cells in patients receiving potent antiretroviral therapy. *N Engl J Med.* **340**:1614-22.
95. **Gallay, P., T. Hope, D. Chin, and D. Trono.** 1997. HIV-1 infection of nondividing cells through the recognition of integrase by the importin/karyopherin pathway. *Proc Natl Acad Sci U S A.* **94**:9825-30.
96. **Gallo, R. C., S. Z. Salahuddin, M. Popovic, G. M. Shearer, M. Kaplan, B. F. Haynes, T. J. Palker, R. Redfield, J. Oleske, B. Safai, and et al.** 1984. Frequent detection and isolation of cytopathic retroviruses (HTLV-III) from patients with AIDS and at risk for AIDS. *Science.* **224**:500-3.
97. **Gao, F., E. Bailes, D. L. Robertson, Y. Chen, C. M. Rodenburg, S. F. Michael, L. B. Cummins, L. O. Arthur, M. Peeters, G. M. Shaw, P. M. Sharp, and B. H. Hahn.** 1999. Origin of HIV-1 in the chimpanzee *Pan troglodytes troglodytes*. *Nature.* **397**:436-41.

98. **Gazdar, A. F., D. N. Carney, P. A. Bunn, E. K. Russell, E. S. Jaffe, G. P. Schechter, and J. G. Guccion.** 1980. Mitogen requirements for the in vitro propagation of cutaneous T-cell lymphomas. *Blood*. **55**:409-17.
99. **Gorry, P., D. Purcell, J. Howard, and D. McPhee.** 1998. Restricted HIV-1 infection of human astrocytes: potential role of nef in the regulation of virus replication. *J Neurovirol*. **4**:377-86.
100. **Gorry, P. R., J. L. Howard, M. J. Churchill, J. L. Anderson, A. Cunningham, D. Adrian, D. A. McPhee, and D. F. Purcell.** 1999. Diminished production of human immunodeficiency virus type 1 in astrocytes results from inefficient translation of gag, env, and nef mRNAs despite efficient expression of Tat and Rev. *J Virol*. **73**:352-61.
101. **Goulaouic, H., and S. A. Chow.** 1996. Directed integration of viral DNA mediated by fusion proteins consisting of human immunodeficiency virus type 1 integrase and *Escherichia coli* LexA protein. *J Virol*. **70**:37-46.
102. **Grist, N. R., C. A. Ross, and E. J. Bell.** 1974. *Diagnostic Methods in Clinical Virology*, Second Edition ed. Blackwell Scientific Publications, Oxford.
103. **Gulick, R. M., J. W. Mellors, D. Havlir, J. J. Eron, C. Gonzalez, D. McMahon, D. D. Richman, F. T. Valentine, L. Jonas, A. Meibohm, E. A. Emini, and J. A. Chodakewitz.** 1997. Treatment with indinavir, zidovudine, and lamivudine in adults with human immunodeficiency virus infection and prior antiretroviral therapy. *N Engl J Med*. **337**:734-9.
104. **Gupta, P., R. Balachandran, M. Ho, A. Enrico, and C. Rinaldo.** 1989. Cell-to-cell transmission of human immunodeficiency virus type 1 in the presence of azidothymidine and neutralizing antibody. *J Virol*. **63**:2361-5.
105. **Hammer, S. M., K. E. Squires, M. D. Hughes, J. M. Grimes, L. M. Demeter, J. S. Currier, J. J. Eron, Jr., J. E. Feinberg, H. H. Balfour, Jr., L. R. Deyton, J. A. Chodakewitz, and M. A. Fischl.** 1997. A controlled trial of two nucleoside analogues plus indinavir in persons with human immunodeficiency virus infection and CD4 cell counts of 200 per cubic millimeter or less. AIDS Clinical Trials Group 320 Study Team. *N Engl J Med*. **337**:725-33.
106. **Hansen, M. S., S. Carreau, C. Hoffmann, L. Li, and F. Bushman.** 1998. Retroviral cDNA integration: mechanism, applications and inhibition. *Genet Eng (N Y)*. **20**:41-61.

107. **Hansen, M. S., G. J. Smith, 3rd, T. Kafri, V. Molteni, J. S. Siegel, and F. D. Bushman.** 1999. Integration complexes derived from HIV vectors for rapid assays in vitro. *Nat Biotechnol.* **17**:578-82.
108. **Haraguchi, Y., Y. Takeuchi, and H. Hoshino.** 1997. Inhibition of plating of human T cell leukemia virus type I and syncytium-inducing types of human immunodeficiency virus type 1 by polycations. *AIDS Res Hum Retroviruses.* **13**:1517-23.
109. **Hazuda, D., C. U. Blau, P. Felock, J. Hastings, B. Pramanik, A. Wolfe, F. Bushman, C. Farnet, M. Goetz, M. Williams, K. Silverman, R. Lingham, and S. Singh.** 1999. Isolation and characterization of novel human immunodeficiency virus integrase inhibitors from fungal metabolites. *Antivir Chem Chemother.* **10**:63-70.
110. **Hazuda, D., P. Felock, J. Hastings, B. Pramanik, A. Wolfe, G. Goodarzi, A. Vora, K. Brackmann, and D. Grandgenett.** 1997. Equivalent inhibition of half-site and full-site retroviral strand transfer reactions by structurally diverse compounds. *J Virol.* **71**:807-11.
111. **Hazuda, D. J., P. Felock, M. Witmer, A. Wolfe, K. Stillmock, J. A. Grobler, A. Espeseth, L. Gabryelski, W. Schleif, C. Blau, and M. D. Miller.** 2000. Inhibitors of strand transfer that prevent integration and inhibit HIV-1 replication in cells. *Science.* **287**:646-50.
112. **Hermida-Matsumoto, L., and M. D. Resh.** 2000. Localization of human immunodeficiency virus type 1 Gag and Env at the plasma membrane by confocal imaging. *J Virol.* **74**:8670-9.
113. **Heuer, T. S., and P. O. Brown.** 1997. 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.* **36**:10655-65.
114. **Hirt, B.** 1967. Selective extraction of polyoma DNA from infected mouse cell cultures. *J Mol Biol.* **26**:365-9.
115. **Ho, D. D., A. U. Neumann, A. S. Perelson, W. Chen, J. M. Leonard, and M. Markowitz.** 1995. Rapid turnover of plasma virions and CD4 lymphocytes in HIV-1 infection. *Nature.* **373**:123-6.
116. **Ho, W. Z., R. Cherukuri, S. D. Ge, J. R. Cutilli, L. Song, S. Whitko, and S. D. Douglas.** 1993. Centrifugal enhancement of human immunodeficiency virus

- type 1 infection and human cytomegalovirus gene expression in human primary monocyte/macrophages in vitro. *J Leukoc Biol.* **53**:208-12.
117. **Huang, Z. B., M. J. Potash, M. Simm, M. Shahabuddin, W. Chao, H. E. Gendelman, E. Eden, and D. J. Volsky.** 1993. Infection of macrophages with lymphotropic human immunodeficiency virus type 1 can be arrested after viral DNA synthesis. *J Virol.* **67**:6893-6.
118. **Hwang, Y., D. Rhodes, and F. Bushman.** 2000. Rapid microtiter assays for poxvirus topoisomerase, mammalian type IB topoisomerase and HIV-1 integrase: application to inhibitor isolation. *Nucleic Acids Res.* **28**:4884-92.
119. **Ibanez, A., T. Puig, J. Elias, B. Clotet, L. Ruiz, and M. A. Martinez.** 1999. Quantification of integrated and total HIV-1 DNA after long-term highly active antiretroviral therapy in HIV-1-infected patients. *AIDS.* **13**:1045-9.
120. **Innis, M.** 1990. *PCR Protocols - A Guide to Methods and Applications.* Academic Press, New York.
121. **Jacobo-Molina, A., J. Ding, R. G. Nanni, A. D. Clark, Jr., X. Lu, C. Tantillo, R. L. Williams, G. Kamer, A. L. Ferris, P. Clark, and et al.** 1993. Crystal structure of human immunodeficiency virus type 1 reverse transcriptase complexed with double-stranded DNA at 3.0 Å resolution shows bent DNA. *Proc Natl Acad Sci U S A.* **90**:6320-4.
122. **James, W., R. A. Weiss, and J. H. Simon.** 1996. The receptor for HIV: dissection of CD4 and studies on putative accessory factors. *Curr Top Microbiol Immunol.* **205**:137-58.
123. **Jenkins, T. M., A. Engelman, R. Ghirlando, and R. Craigie.** 1996. A soluble active mutant of HIV-1 integrase: involvement of both the core and carboxyl-terminal domains in multimerization. *J Biol Chem.* **271**:7712-8.
124. **Jing, N., C. Marchand, J. Liu, R. Mitra, M. E. Hogan, and Y. Pommier.** 2000. Mechanism of inhibition of HIV-1 integrase by G-tetrad-forming oligonucleotides in Vitro. *J Biol Chem.* **275**:21460-7.
125. **Johnson, M. S., M. A. McClure, D. F. Feng, J. Gray, and R. F. Doolittle.** 1986. Computer analysis of retroviral pol genes: assignment of enzymatic functions to specific sequences and homologies with nonviral enzymes. *Proc Natl Acad Sci U S A.* **83**:7648-52.

126. **Jones, K. S., J. Coleman, G. W. Merkel, T. M. Laue, and A. M. Skalka.** 1992. Retroviral integrase functions as a multimer and can turn over catalytically. *J Biol Chem.* **267**:16037-40.
127. **Jordan, A., P. Defechereux, and E. Verdin.** 2001. The site of HIV-1 integration in the human genome determines basal transcriptional activity and response to Tat transactivation. *EMBO J.* **20**:1726-38.
128. **Ju, G., and A. M. Skalka.** 1980. Nucleotide sequence analysis of the long terminal repeat (LTR) of avian retroviruses: structural similarities with transposable elements. *Cell.* **22**:379-86.
129. **Jurka, J., and T. Smith.** 1988. A fundamental division in the Alu family of repeated sequences. *Proc Natl Acad Sci U S A.* **85**:4775-8.
130. **Karageorgos, L., P. Li, and C. Burrell.** 1993. Characterization of HIV replication complexes early after cell-to-cell infection. *AIDS Res Hum Retroviruses.* **9**:817-23.
131. **Karageorgos, L., P. Li, and C. J. Burrell.** 1995. Stepwise analysis of reverse transcription in a cell-to-cell human immunodeficiency virus infection model: kinetics and implications. *J Gen Virol.* **76**:1675-86.
132. **Katz, R. A., G. Merkel, J. Kulkosky, J. Leis, and A. M. Skalka.** 1990. The avian retroviral IN protein is both necessary and sufficient for integrative recombination in vitro. *Cell.* **63**:87-95.
133. **Katz, R. A., G. Merkel, and A. M. Skalka.** 1996. Targeting of retroviral integrase by fusion to a heterologous DNA binding domain: in vitro activities and incorporation of a fusion protein into viral particles. *Virology.* **217**:178-90.
134. **Katzman, M., and R. A. Katz.** 1999. Substrate recognition by retroviral integrases. *Adv Virus Res.* **52**:371-95.
135. **Katzman, M., and M. Sudol.** 1996. Influence of subterminal viral DNA nucleotides on differential susceptibility to cleavage by human immunodeficiency virus type 1 and visna virus integrases. *J Virol.* **70**:9069-73.
136. **Katzman, M., and M. Sudol.** 1998. Mapping viral DNA specificity to the central region of integrase by using functional human immunodeficiency virus type 1/visna virus chimeric proteins. *J Virol.* **72**:1744-53.
137. **Khan, E., J. P. Mack, R. A. Katz, J. Kulkosky, and A. M. Skalka.** 1991. Retroviral integrase domains: DNA binding and the recognition of LTR sequences. *Nucleic Acids Res.* **19**:851-60.

138. **Kim, S. Y., R. Byrn, J. Groopman, and D. Baltimore.** 1989. Temporal aspects of DNA and RNA synthesis during human immunodeficiency virus infection: evidence for differential gene expression. *J Virol.* **63**:3708-13.
139. **Kimpton, J., and M. Emerman.** 1992. Detection of replication-competent and pseudotyped human immunodeficiency virus with a sensitive cell line on the basis of activation of an integrated beta-galactosidase gene. *J Virol.* **66**:2232-9.
140. **King, P. J., G. Ma, W. Miao, Q. Jia, B. R. McDougall, M. G. Reinecke, C. Cornell, J. Kuan, T. R. Kim, and W. E. Robinson, Jr.** 1999. Structure-activity relationships: analogues of the dicaffeoylquinic and dicaffeoyltartaric acids as potent inhibitors of human immunodeficiency virus type 1 integrase and replication. *J Med Chem.* **42**:497-509.
141. **King, P. J., and W. E. Robinson, Jr.** 1998. Resistance to the anti-human immunodeficiency virus type 1 compound L-chicoric acid results from a single mutation at amino acid 140 of integrase. *J Virol.* **72**:8420-4.
142. **Kjems, J., A. D. Frankel, and P. A. Sharp.** 1991. Specific regulation of mRNA splicing in vitro by a peptide from HIV-1 Rev. *Cell.* **67**:169-78.
143. **Kjems, J., and P. A. Sharp.** 1993. The basic domain of Rev from human immunodeficiency virus type 1 specifically blocks the entry of U4/U6.U5 small nuclear ribonucleoprotein in spliceosome assembly. *J Virol.* **67**:4769-76.
144. **Klotman, M. E., S. Kim, A. Buchbinder, A. DeRossi, D. Baltimore, and F. Wong-Staal.** 1991. Kinetics of expression of multiply spliced RNA in early human immunodeficiency virus type 1 infection of lymphocytes and monocytes. *Proc Natl Acad Sci U S A.* **88**:5011-5.
145. **Kohlstaedt, L. A., J. Wang, J. M. Friedman, P. A. Rice, and T. A. Steitz.** 1992. Crystal structure at 3.5 Å resolution of HIV-1 reverse transcriptase complexed with an inhibitor. *Science.* **256**:1783-90.
146. **Kok, T., P. Li, and C. J. Burrell.** 1993. HIV DNA integration during cell-to-cell transmission of infection: evidence for partially integrated DNA structures in acutely infected cells. *Arch Virol.* **146**:1963-78.
147. **Konopka, K., L. Stamatatos, C. E. Larsen, B. R. Davis, and N. Duzgunes.** 1991. Enhancement of human immunodeficiency virus type 1 infection by cationic liposomes: the role of CD4, serum and liposome-cell interactions. *J Gen Virol.* **72**:2685-96.

148. **Kulkosky, J., K. S. Jones, R. A. Katz, J. P. Mack, and A. M. Skalka.** 1992. Residues critical for retroviral integrative recombination in a region that is highly conserved among retroviral/retrotransposon integrases and bacterial insertion sequence transposases. *Mol Cell Biol.* **12**:2331-8.
149. **Kumar, R., K. Cheney, L. Mundy, N. Vandegraaff, W. Ferguson, D. Shaw, C. Burrell, and P. Li.** 2001. CD4-CD8- double-negative T cells as a reservoir of HIV-1 in patients on antiretroviral therapy. submitted.
150. **LaFemina, R. L., P. L. Callahan, and M. G. Cordingley.** 1991. Substrate specificity of recombinant human immunodeficiency virus integrase protein. *J Virol.* **65**:5624-30.
151. **LaFemina, R. L., C. L. Schneider, H. L. Robbins, P. L. Callahan, K. LeGrow, E. Roth, W. A. Schleif, and E. A. Emini.** 1992. Requirement of active human immunodeficiency virus type 1 integrase enzyme for productive infection of human T-lymphoid cells. *J Virol.* **66**:7414-9.
152. **Lambert, D. M., S. Barney, A. L. Lambert, K. Guthrie, R. Medinas, D. E. Davis, T. Bucy, J. Erickson, G. Merutka, and S. R. Petteway, Jr.** 1996. Peptides from conserved regions of paramyxovirus fusion (F) proteins are potent inhibitors of viral fusion. *Proc Natl Acad Sci U S A.* **93**:2186-91.
153. **Laspia, M. F., A. P. Rice, and M. B. Mathews.** 1989. HIV-1 Tat protein increases transcriptional initiation and stabilizes elongation. *Cell.* **59**:283-92.
154. **Lee, M. S., and R. Craigie.** 1998. A previously unidentified host protein protects retroviral DNA from autointegration. *Proc Natl Acad Sci U S A.* **95**:1528-33.
155. **Lee, M. S., and R. Craigie.** 1994. Protection of retroviral DNA from autointegration: involvement of a cellular factor. *Proc Natl Acad Sci U S A.* **91**:9823-7.
156. **Lee, S. P., J. Xiao, J. R. Knutson, M. S. Lewis, and M. K. Han.** 1997. Zn²⁺ promotes the self-association of human immunodeficiency virus type-1 integrase in vitro. *Biochemistry.* **36**:173-80.
157. **Li, G., M. Simm, M. J. Potash, and D. J. Volsky.** 1993. Human immunodeficiency virus type 1 DNA synthesis, integration, and efficient viral replication in growth-arrested T cells. *J Virol.* **67**:3969-77.

158. **Li, G., M. Simm, M. J. Potash, and D. J. Volsky.** 1993. Human immunodeficiency virus type 1 DNA synthesis, integration, and efficient viral replication in growth-arrested T cells. *J Virol.* **67**:3969-77.
159. **Li, L., K. Yoder, M. S. Hansen, J. Olvera, M. D. Miller, and F. D. Bushman.** 2000. Retroviral cDNA Integration: Stimulation by HMG I Family Proteins. *J Virol.* **74**:10965-10974.
160. **Li, P., and C. J. Burrell.** 1992. Synthesis of human immunodeficiency virus DNA in a cell-to-cell transmission model. *AIDS Res Hum Retroviruses.* **8**:253-9.
161. **Li, P., L. J. Kuiper, A. J. Stephenson, and C. J. Burrell.** 1992. De novo reverse transcription is a crucial event in cell-to-cell transmission of human immunodeficiency virus. *J Gen Virol.* **73**:955-9.
162. **Li, P., A. J. Stephenson, P. A. Brennan, L. Karageorgos, T. Kok, L. J. Kuiper, J. Swift, and C. J. Burrell.** 1994. Initiation of reverse transcription during cell-to-cell transmission of human immunodeficiency virus infection uses pre-existing reverse transcriptase. *J Gen Virol.* **75**:1917-26.
163. **Lin, Z., N. Neamati, H. Zhao, Y. Kiryu, J. A. Turpin, C. Aberham, K. Strebel, K. Kohn, M. Witvrouw, C. Pannecouque, Z. Debyser, E. De Clercq, W. G. Rice, Y. Pommier, and T. R. Burke, Jr.** 1999. Chicoric acid analogues as HIV-1 integrase inhibitors. *J Med Chem.* **42**:1401-14.
164. **Lodi, P. J., J. A. Ernst, J. Kuszewski, A. B. Hickman, A. Engelman, R. Craigie, G. M. Clore, and A. M. Gronenborn.** 1995. Solution structure of the DNA binding domain of HIV-1 integrase. *Biochemistry.* **34**:9826-33.
165. **Luciw, P.** 1996. Human immunodeficiency viruses and their replication., p. 1881-1952. *In* B. N. Fields and D. M. Knipe and P. M. Howley (ed.), *Fields Virology*. Lippincott-Raven Publishers, New York.
166. **Luo, T., J. L. Foster, and J. V. Garcia.** 1997. Molecular Determinants of Nef Function. *J Biomed Sci.* **4**:132-138.
167. **Lutzke, R. A., C. Vink, and R. H. Plasterk.** 1994. Characterization of the minimal DNA-binding domain of the HIV integrase protein. *Nucleic Acids Res.* **22**:4125-31.
168. **Mangasarian, A., and D. Trono.** 1997. The multifaceted role of HIV Nef. *Res Virol.* **148**:30-3.

169. **Mansky, L. M., and H. M. Temin.** 1995. Lower in vivo mutation rate of human immunodeficiency virus type 1 than that predicted from the fidelity of purified reverse transcriptase. *J Virol.* **69**:5087-94.
170. **Marciniak, R. A., and P. A. Sharp.** 1991. HIV-1 Tat protein promotes formation of more-processive elongation complexes. *EMBO J.* **10**:4189-96.
171. **Marodon, G., D. Warren, M. C. Filomio, and D. N. Posnett.** 1999. Productive infection of double-negative T cells with HIV in vivo. *Proc Natl Acad Sci U S A.* **96**:11958-63.
172. **Mazumder, A., D. Cooney, R. Agbaria, M. Gupta, and Y. Pommier.** 1994. Inhibition of human immunodeficiency virus type 1 integrase by 3'-azido-3'-deoxythymidylate. *Proc Natl Acad Sci U S A.* **91**:5771-5.
173. **Mazumder, A., N. Neamati, J. O. Ojwang, S. Sunder, R. F. Rando, and Y. Pommier.** 1996. Inhibition of the human immunodeficiency virus type 1 integrase by guanosine quartet structures. *Biochemistry.* **35**:13762-71.
174. **Mazumder, A., and Y. Pommier.** 1995. Processing of deoxyuridine mismatches and abasic sites by human immunodeficiency virus type-1 integrase. *Nucleic Acids Res.* **23**:2865-71.
175. **McDougal, J. S., P. J. Maddon, A. G. Dalgleish, P. R. Clapham, D. R. Littman, M. Godfrey, D. E. Maddon, L. Chess, R. A. Weiss, and R. Axel.** 1986. The T4 glycoprotein is a cell-surface receptor for the AIDS virus. *Cold Spring Harb Symp Quant Biol.* **51 Pt 2**:703-11.
176. **Meyer, B. E., and M. H. Malim.** 1994. The HIV-1 Rev trans-activator shuttles between the nucleus and the cytoplasm. *Genes Dev.* **8**:1538-47.
177. **Michie, C. A., A. McLean, C. Alcock, and P. C. Beverley.** 1992. Lifespan of human lymphocyte subsets defined by CD45 isoforms. *Nature.* **360**:264-5.
178. **Miller, M. D., C. M. Farnet, and F. D. Bushman.** 1997. Human immunodeficiency virus type 1 preintegration complexes: studies of organization and composition. *J Virol.* **71**:5382-90.
179. **Morikawa, Y., D. J. Hockley, M. V. Nermut, and I. M. Jones.** 2000. Roles of matrix, p2, and N-terminal myristoylation in human immunodeficiency virus type 1 Gag assembly. *J Virol.* **74**:16-23.
180. **Morrison, C., and F. Gannon.** 1994. The impact of the PCR plateau phase on quantitative PCR. *Biochim Biophys Acta.* **1219**:493-8.

181. **Mullis, K., F. Faloona, S. Scharf, R. Saiki, G. Horn, and H. Erlich.** 1986. Specific enzymatic amplification of DNA in vitro: the polymerase chain reaction. *Cold Spring Harb Symp Quant Biol.* **51**:263-73.
182. **Mullis, K., F. Faloona, S. Scharf, R. Saiki, G. Horn, and H. Erlich.** 1992. Specific enzymatic amplification of DNA in vitro: the polymerase chain reaction. 1986. *Biotechnology.* **24**:17-27.
183. **Myers, T. W., and D. H. Gelfand.** 1991. Reverse transcription and DNA amplification by a *Thermus thermophilus* DNA polymerase. *Biochemistry.* **30**:7661-6.
184. **Neamati, N., S. Sunder, and Y. Pommier.** 1997. Design and Discovery of HIV-1 integrase inhibitors. *Drug Discovery Today.* **2**:487-498.
185. **Nicholson, W. J., A. J. Shepherd, and D. W. Aw.** 1996. Detection of unintegrated HIV type 1 DNA in cell culture and clinical peripheral blood mononuclear cell samples: correlation to disease stage. *AIDS Res Hum Retroviruses.* **12**:315-23.
186. **Nuovo, G. J., F. Gallery, P. MacConnell, and A. Braun.** 1994. In situ detection of polymerase chain reaction-amplified HIV-1 nucleic acids and tumor necrosis factor-alpha RNA in the central nervous system. *Am J Pathol.* **144**:659-66.
187. **Oberlin, E., A. Amara, F. Bachelier, C. Bessia, J. L. Virelizier, F. Arenzana-Seisdedos, O. Schwartz, J. M. Heard, I. Clark-Lewis, D. F. Legler, M. Loetscher, M. Baggiolini, and B. Moser.** 1996. The CXC chemokine SDF-1 is the ligand for LESTR/fusin and prevents infection by T-cell-line-adapted HIV-1. *Nature.* **382**:833-5.
188. **Ojwang, J. O., R. W. Buckheit, Y. Pommier, A. Mazumder, K. De Vreese, J. A. Este, D. Reymen, L. A. Pallansch, C. Lackman-Smith, T. L. Wallace, and et al.** 1995. T30177, an oligonucleotide stabilized by an intramolecular guanosine octet, is a potent inhibitor of laboratory strains and clinical isolates of human immunodeficiency virus type 1. *Antimicrob Agents Chemother.* **39**:2426-35.
189. **Organisation, G. W. H.** 1999. Removing obstacles to healthy development. .
190. **Patel, P. H., and B. D. Preston.** 1994. Marked infidelity of human immunodeficiency virus type 1 reverse transcriptase at RNA and DNA template ends. *Proc Natl Acad Sci U S A.* **91**:549-53.

191. **Pauza, C. D.** 1990. Two bases are deleted from the termini of HIV-1 linear DNA during integrative recombination. *Virology*. **179**:886-9.
192. **Pauza, C. D., and J. Galindo.** 1989. Persistent human immunodeficiency virus type 1 infection of monoblastoid cells leads to accumulation of self-integrated viral DNA and to production of defective virions. *J Virol*. **63**:3700-7.
193. **Perelson, A. S., P. Essunger, Y. Cao, M. Vesanen, A. Hurley, K. Saksela, M. Markowitz, and D. D. Ho.** 1997. Decay characteristics of HIV-1-infected compartments during combination therapy. *Nature*. **387**:188-91.
194. **Petit, C., O. Schwartz, and F. Mammano.** 2000. The karyophilic properties of human immunodeficiency virus type 1 integrase are not required for nuclear import of proviral DNA. *J Virol*. **74**:7119-26.
195. **Peyret, N., P. A. Seneviratne, H. T. Allawi, and J. SantaLucia, Jr.** 1999. Nearest-neighbor thermodynamics and NMR of DNA sequences with internal A.A, C.C, G.G, and T.T mismatches. *Biochemistry*. **38**:3468-77.
196. **Phillips, D. M., and A. S. Bourinbaier.** 1992. Mechanism of HIV spread from lymphocytes to epithelia. *Virology*. **186**:261-73.
197. **Pietroboni, G. R., G. B. Harnett, and M. R. Bucens.** 1989. Centrifugal enhancement of human immunodeficiency virus (HIV) and human herpesvirus type 6 (HHV-6) infection in vitro. *J Virol Methods*. **24**:85-90.
198. **Plumyers, W., P. Cherepanov, D. Schols, E. De Clercq, and Z. Debyser.** 1999. Nuclear localization of human immunodeficiency virus type 1 integrase expressed as a fusion protein with green fluorescent protein. *Virology*. **258**:327-32.
199. **Plumyers, W., N. Neamati, C. Pannecouque, V. Fikkert, C. Marchand, T. R. Burke, Jr., Y. Pommier, D. Schols, E. De Clercq, Z. Debyser, and M. Witvrouw.** 2000. Viral entry as the primary target for the anti-HIV activity of chicoric acid and its tetra-acetyl esters. *Mol Pharmacol*. **58**:641-8.
200. **Pommier, Y., and N. Neamati.** 1999. Inhibitors of human immunodeficiency virus integrase. *Adv Virus Res*. **52**:427-58.
201. **Popovic, M., M. G. Sarngadharan, E. Read, and R. C. Gallo.** 1984. Detection, isolation, and continuous production of cytopathic retroviruses (HTLV-III) from patients with AIDS and pre-AIDS. *Science*. **224**:497-500.
202. **Pryciak, P. M., A. Sil, and H. E. Varmus.** 1992. Retroviral integration into minichromosomes in vitro. *EMBO J*. **11**:291-303.

203. **Qiagen.** Critical Factors for Successful PCR.
204. **Qiagen.** 1995. Qiagen genomic DNA handbook.
205. **Raeymaekers, L.** 2000. Basic principles of quantitative PCR. *Mol Biotechnol.* **15**:115-22.
206. **Raghavan, K., J. K. Buolamwini, M. R. Fesen, Y. Pommier, K. W. Kohn, and J. N. Weinstein.** 1995. Three-dimensional quantitative structure-activity relationship (QSAR) of HIV integrase inhibitors: a comparative molecular field analysis (CoMFA) study. *J Med Chem.* **38**:890-7.
207. **Ranki, A., M. Nyberg, V. Ovod, M. Haltia, I. Elovaara, R. Raininko, H. Haapasalo, and K. Krohn.** 1995. Abundant expression of HIV Nef and Rev proteins in brain astrocytes in vivo is associated with dementia. *AIDS.* **9**:1001-8.
208. **Reddy, M. V., M. R. Rao, D. Rhodes, M. S. Hansen, K. Rubins, F. D. Bushman, Y. Venkateswarlu, and D. J. Faulkner.** 1999. Lamellarin alpha 20-sulfate, an inhibitor of HIV-1 integrase active against HIV-1 virus in cell culture. *J Med Chem.* **42**:1901-7.
209. **Reischl, U., and B. Kochanowski.** 1995. Quantitative PCR. A survey of the present technology. *Mol Biotechnol.* **3**:55-71.
210. **Ren, J., R. Esnouf, E. Garman, D. Somers, C. Ross, I. Kirby, J. Keeling, G. Darby, Y. Jones, D. Stuart, and et al.** 1995. High resolution structures of HIV-1 RT from four RT-inhibitor complexes. *Nat Struct Biol.* **2**:293-302.
211. **Reyes, R. A., and G. L. Cockerell.** 1996. Unintegrated bovine leukemia virus DNA: association with viral expression and disease. *J Virol.* **70**:4961-5.
212. **Richman, D. D.** 2001. HIV chemotherapy. *Nature.* **410**:995-1001.
213. **Robinson, H. L., and D. M. Zinkus.** 1990. Accumulation of human immunodeficiency virus type 1 DNA in T cells: results of multiple infection events. *J Virol.* **64**:4836-41.
214. **Robinson, W. E., Jr., M. Cordeiro, S. Abdel-Malek, Q. Jia, S. A. Chow, M. G. Reinecke, and W. M. Mitchell.** 1996. Dicafeoylquinic acid inhibitors of human immunodeficiency virus integrase: inhibition of the core catalytic domain of human immunodeficiency virus integrase. *Mol Pharmacol.* **50**:846-55.
215. **Robinson, W. E., Jr., M. G. Reinecke, S. Abdel-Malek, Q. Jia, and S. A. Chow.** 1996. Inhibitors of HIV-1 replication that inhibit HIV integrase. *Proc Natl Acad Sci U S A.* **93**:6326-31.

216. **Saha, K., J. Zhang, A. Gupta, R. Dave, M. Yimen, and B. Zerhouni.** 2001. Isolation of primary HIV-1 that target CD8+ T lymphocytes using CD8 as a receptor. *Nature Med.* **7**:65-72.
217. **Saito, Y., L. R. Sharer, L. G. Epstein, J. Michaels, M. Mintz, M. Louder, K. Golding, T. A. Cvetkovich, and B. M. Blumberg.** 1994. Overexpression of nef as a marker for restricted HIV-1 infection of astrocytes in postmortem pediatric central nervous tissues. *Neurology.* **44**:474-81.
218. **Sakai, H., M. Kawamura, J. Sakuragi, S. Sakuragi, R. Shibata, A. Ishimoto, N. Ono, S. Ueda, and A. Adachi.** 1993. Integration is essential for efficient gene expression of human immunodeficiency virus type 1. *J Virol.* **67**:1169-74.
219. **Sambrook, J., E. F. Fritsch, and T. Maniatis.** 1989. *Molecular Cloning: A Laboratory Manual*, Second ed. Cold Spring Harbor Laboratory Press.
220. **Sanchez, G., X. Xu, J. C. Chermann, and I. Hirsch.** 1997. Accumulation of defective viral genomes in peripheral blood mononuclear cells of human immunodeficiency virus type 1-infected individuals. *J Virol.* **71**:2233-40.
221. **Schmid, C. W., and W. R. Jelinek.** 1982. The Alu family of dispersed repetitive sequences. *Science.* **216**:1065-70.
222. **Schnell, S., and C. Mendoza.** 1997. Enzymological considerations for a theoretical description of the quantitative competitive polymerase chain reaction (QC-PCR). *J Theor Biol.* **184**:433-40.
223. **Schnell, S., and C. Mendoza.** 1997. Theoretical description of the polymerase chain reaction. *J Theor Biol.* **188**:313-8.
224. **Schutz, E., and N. von Ahsen.** 1999. Spreadsheet software for thermodynamic melting point prediction of oligonucleotide hybridization with and without mismatches. *Biotechniques.* **27**:1218-22, 1224.
225. **Sharer, L. R.** 1992. Pathology of HIV-1 infection of the central nervous system. A review. *J Neuropathol Exp Neurol.* **51**:3-11.
226. **Sharkey, M. E., I. Teo, T. Greenough, N. Sharova, K. Luzuriaga, J. L. Sullivan, R. P. Bucy, L. G. Kostrikis, A. Haase, C. Veryard, R. E. Davaro, S. H. Cheeseman, J. S. Daly, C. Bova, R. T. Ellison, 3rd, B. Mady, K. K. Lai, G. Moyle, M. Nelson, B. Gazzard, S. Shaunak, and M. Stevenson.** 2000. Persistence of episomal HIV-1 infection intermediates in patients on highly active anti-retroviral therapy. *Nature Med.* **6**:76-81.

227. **Sherman, P. A., M. L. Dickson, and J. A. Fyfe.** 1992. Human immunodeficiency virus type 1 integration protein: DNA sequence requirements for cleaving and joining reactions. *J Virol.* **66**:3593-601.
228. **Shimotohno, K., S. Mizutani, and H. M. Temin.** 1980. Sequence of retrovirus provirus resembles that of bacterial transposable elements. *Nature.* **285**:550-4.
229. **Shiraishi, T., S. Misumi, M. Takama, I. Takahashi, and S. Shoji.** 2001. Myristoylation of human immunodeficiency virus type 1 gag protein is required for efficient env protein transportation to the surface of cells. *Biochem Biophys Res Commun.* **282**:1201-5.
230. **Simmons, G., J. D. Reeves, S. Hibbitts, J. T. Stine, P. W. Gray, A. E. Proudfoot, and P. R. Clapham.** 2000. Co-receptor use by HIV and inhibition of HIV infection by chemokine receptor ligands. *Immunol Rev.* **177**:112-26.
231. **Sonza, S., A. Maerz, N. Deacon, J. Meanger, J. Mills, and S. Crowe.** 1996. Human immunodeficiency virus type 1 replication is blocked prior to reverse transcription and integration in freshly isolated peripheral blood monocytes. *J Virol.* **70**:3863-9.
232. **Southgate, C. D., and M. R. Green.** 1991. The HIV-1 Tat protein activates transcription from an upstream DNA-binding site: implications for Tat function. *Genes Dev.* **5**:2496-507.
233. **Spina, C. A., J. C. Guatelli, and D. D. Richman.** 1995. Establishment of a stable, inducible form of human immunodeficiency virus type 1 DNA in quiescent CD4 lymphocytes in vitro. *J Virol.* **69**:2977-88.
234. **Stammers, D. K., D. O. Somers, C. K. Ross, I. Kirby, P. H. Ray, J. E. Wilson, M. Norman, J. S. Ren, R. M. Esnouf, E. F. Garman, and et al.** 1994. Crystals of HIV-1 reverse transcriptase diffracting to 2.2 Å resolution. *J Mol Biol.* **242**:586-8.
235. **Stevens, S. W., and J. D. Griffith.** 1994. Human immunodeficiency virus type 1 may preferentially integrate into chromatin occupied by L1Hs repetitive elements. *Proc Natl Acad Sci U S A.* **91**:5557-61.
236. **Stevens, S. W., and J. D. Griffith.** 1996. Sequence analysis of the human DNA flanking sites of human immunodeficiency virus type 1 integration. *J Virol.* **70**:6459-62.

237. **Stevenson, M., S. Haggerty, C. A. Lamonica, C. M. Meier, S. K. Welch, and A. J. Wasiak.** 1990. Integration is not necessary for expression of human immunodeficiency virus type 1 protein products. *J Virol.* **64**:2421-5.
238. **Stevenson, M., T. L. Stanwick, M. P. Dempsey, and C. A. Lamonica.** 1990. HIV-1 replication is controlled at the level of T cell activation and proviral integration. *EMBO J.* **9**:1551-60.
239. **Stutz, F., and M. Rosbash.** 1994. A functional interaction between Rev and yeast pre-mRNA is related to splicing complex formation. *EMBO J.* **13**:4096-104.
240. **Sun, J. S., T. Garestier, and C. Helene.** 1996. Oligonucleotide directed triple helix formation. *Curr Opin Struct Biol.* **6**:327-33.
241. **Takahashi, K., S. L. Wesselingh, D. E. Griffin, J. C. McArthur, R. T. Johnson, and J. D. Glass.** 1996. Localization of HIV-1 in human brain using polymerase chain reaction/in situ hybridization and immunocytochemistry. *Ann Neurol.* **39**:705-11.
242. **Tornatore, C., R. Chandra, J. R. Berger, and E. O. Major.** 1994. HIV-1 infection of subcortical astrocytes in the pediatric central nervous system. *Neurology.* **44**:481-7.
243. **van den Ent, F. M., C. Vink, and R. H. Plasterk.** 1994. DNA substrate requirements for different activities of the human immunodeficiency virus type 1 integrase protein. *J Virol.* **68**:7825-32.
244. **van Gent, D. C., Y. Elgersma, M. W. Bolk, C. Vink, and R. H. Plasterk.** 1991. DNA binding properties of the integrase proteins of human immunodeficiency viruses types 1 and 2. *Nucleic Acids Res.* **19**:3821-7.
245. **van Gent, D. C., C. Vink, A. A. Groeneger, and R. H. Plasterk.** 1993. Complementation between HIV integrase proteins mutated in different domains. *EMBO J.* **12**:3261-7.
246. **Van Lint, C.** 2000. Role of chromatin in HIV-1 transcriptional regulation. *Adv Pharmacol.* **48**:121-60.
247. **Vandegraaff, N., R. Kumar, C. J. Burrell, and P. Li.** 2001. Kinetics of human immunodeficiency virus type 1 (HIV) DNA integration in acutely infected cells as determined using a novel assay for detection of integrated HIV DNA. *J. Virol.* **75**:11253-60.

248. **Vandegraaff, N., R. Kumar, L. Mundy, C. Burrell, and P. Li.** 2002. Evaluation of PCR-based methods for the quantification of integrated HIV DNA. *submitted*.
249. **Vincent, K. A., V. Ellison, S. A. Chow, and P. O. Brown.** 1993. Characterization of human immunodeficiency virus type 1 integrase expressed in *Escherichia coli* and analysis of variants with amino-terminal mutations. *J Virol.* **67**:425-37.
250. **Vink, C., M. Groenink, Y. Elgersma, R. A. Fouchier, M. Tersmette, and R. H. Plasterk.** 1990. Analysis of the junctions between human immunodeficiency virus type 1 proviral DNA and human DNA. *J Virol.* **64**:5626-7.
251. **Vink, C., A. M. Oude Groeneger, and R. H. Plasterk.** 1993. Identification of the catalytic and DNA-binding region of the human immunodeficiency virus type I integrase protein. *Nucleic Acids Res.* **21**:1419-25.
252. **Vink, C., D. C. van Gent, Y. Elgersma, and R. H. Plasterk.** 1991. Human immunodeficiency virus integrase protein requires a subterminal position of its viral DNA recognition sequence for efficient cleavage. *J Virol.* **65**:4636-44.
253. **Vink, C., E. Yeheskiely, G. A. van der Marel, J. H. van Boom, and R. H. Plasterk.** 1991. Site-specific hydrolysis and alcoholysis of human immunodeficiency virus DNA termini mediated by the viral integrase protein. *Nucleic Acids Res.* **19**:6691-8.
254. **Vodicka, M. A., D. M. Koepp, P. A. Silver, and M. Emerman.** 1998. HIV-1 Vpr interacts with the nuclear transport pathway to promote macrophage infection. *Genes Dev.* **12**:175-85.
255. **von Schwedler, U. K., T. L. Stemmler, V. Y. Klishko, S. Li, K. H. Albertine, D. R. Davis, and W. I. Sundquist.** 1998. Proteolytic refolding of the HIV-1 capsid protein amino-terminus facilitates viral core assembly. *EMBO J.* **17**:1555-68.
256. **Wang, J. Y., H. Ling, W. Yang, and R. Craigie.** 2001. Structure of a two-domain fragment of HIV-1 integrase: implications for domain organization in the intact protein. *EMBO J.* **20**:7333-43.
257. **Wattel, E., J. P. Vartanian, C. Pannetier, and S. Wain-Hobson.** 1995. Clonal expansion of human T-cell leukemia virus type I-infected cells in asymptomatic and symptomatic carriers without malignancy. *J Virol.* **69**:2863-8.

258. **Wei, X., S. K. Ghosh, M. E. Taylor, V. A. Johnson, E. A. Emini, P. Deutsch, J. D. Lifson, S. Bonhoeffer, M. A. Nowak, B. H. Hahn, and et al.** 1995. Viral dynamics in human immunodeficiency virus type 1 infection. *Nature*. **373**:117-22.
259. **Weinberg, J. B., M. M. Hobbs, and M. A. Misukonis.** 1984. Recombinant human gamma-interferon induces human monocyte polykaryon formation. *Proc Natl Acad Sci U S A*. **81**:4554-7.
260. **Weinberg, J. B., T. J. Matthews, B. R. Cullen, and M. H. Malim.** 1991. Productive human immunodeficiency virus type 1 (HIV-1) infection of nonproliferating human monocytes. *J Exp Med*. **174**:1477-82.
261. **Weinberg, J. B., D. L. Sauls, M. A. Misukonis, and D. C. Shugars.** 1995. Inhibition of productive human immunodeficiency virus-1 infection by cobalamins. *Blood*. **86**:1281-7.
262. **Weinberg, J. B., D. C. Shugars, P. A. Sherman, D. L. Sauls, and J. A. Fyfe.** 1998. Cobalamin inhibition of HIV-1 integrase and integration of HIV-1 DNA into cellular DNA. *Biochem Biophys Res Commun*. **246**:393-7.
263. **Wetmur, J. G.** 1991. DNA probes: applications of the principles of nucleic acid hybridization. *Crit Rev Biochem Mol Biol*. **26**:227-59.
264. **White, B. (ed.).** 1993. PCR Protocols: Current Methods and Applications. Humana Press Inc., Totowa, NJ.
265. **Wieggers, K., G. Rutter, H. Kottler, U. Tessmer, H. Hohenberg, and H. G. Krausslich.** 1998. Sequential steps in human immunodeficiency virus particle maturation revealed by alterations of individual Gag polyprotein cleavage sites. *J Virol*. **72**:2846-54.
266. **Wilderspin, A. F., and R. J. Sugrue.** 1994. Alternative native flap conformation revealed by 2.3 Å resolution structure of SIV proteinase. *J Mol Biol*. **239**:97-103.
267. **Wlodawer, A.** 1999. Crystal structures of catalytic core domains of retroviral integrases and role of divalent cations in enzymatic activity. *Adv Virus Res*. **52**:335-50.
268. **Wolfe, A. L., P. J. Felock, J. C. Hastings, C. U. Blau, and D. J. Hazuda.** 1996. The role of manganese in promoting multimerization and assembly of human immunodeficiency virus type 1 integrase as a catalytically active complex on immobilized long terminal repeat substrates. *J Virol*. **70**:1424-32.

269. **Wong, J. K., M. Hezareh, H. F. Gunthard, D. V. Havlir, C. C. Ignacio, C. A. Spina, and D. D. Richman.** 1997. Recovery of replication-competent HIV despite prolonged suppression of plasma viremia. *Science*. **278**:1291-5.
270. **Wu, Y., and J. W. Marsh.** 2001. Selective transcription and modulation of resting T cell activity by preintegrated HIV DNA. *Science*. **293**:1503-6.
271. **Yerly, S., L. Kaiser, T. V. Perneger, R. W. Cone, M. Opravil, J. P. Chave, H. Furrer, B. Hirschel, and L. Perrin.** 2000. Time of initiation of antiretroviral therapy: impact on HIV-1 viraemia. The Swiss HIV Cohort Study. *AIDS*. **14**:243-9.
272. **Zack, J. A., S. J. Arrigo, S. R. Weitsman, A. S. Go, A. Haislip, and I. S. Chen.** 1990. HIV-1 entry into quiescent primary lymphocytes: molecular analysis reveals a labile, latent viral structure. *Cell*. **61**:213-22.
273. **Zack, J. A., A. M. Haislip, P. Krogstad, and I. S. Chen.** 1992. Incompletely reverse-transcribed human immunodeficiency virus type 1 genomes in quiescent cells can function as intermediates in the retroviral life cycle. *J Virol*. **66**:1717-25.
274. **Zanussi, S., M. T. Bortolin, M. Giacca, and P. De Paoli.** 2000. Quantitative assessment of integrated and episomal HIV DNA. *AIDS*. **16**:931-3.
275. **Zennou, V., C. Petit, D. Guetard, U. Nerhbass, L. Montagnier, and P. Charneau.** 2000. HIV-1 genome nuclear import is mediated by a central DNA flap. *Cell*. **101**:173-85.
276. **Zhang, H., G. Dornadula, and R. J. Pomerantz.** 1996. Endogenous reverse transcription of human immunodeficiency virus type 1 in physiological microenvironments: an important stage for viral infection of nondividing cells. *J Virol*. **70**:2809-24.
277. **Zhang, H., G. Dornadula, and R. J. Pomerantz.** 1998. Natural endogenous reverse transcription of HIV type 1. *AIDS Res Hum Retroviruses*. **14 Suppl 1**:S93-5.
278. **Zhang, L., B. Ramratnam, K. Tenner-Racz, Y. He, M. Vesanen, S. Lewin, A. Talal, P. Racz, A. S. Perelson, B. T. Korber, M. Markowitz, and D. D. Ho.** 1999. Quantifying residual HIV-1 replication in patients receiving combination antiretroviral therapy. *N Engl J Med*. **340**:1605-13.

279. **Zhao, H., N. Neamati, S. Sunder, H. Hong, S. Wang, G. W. Milne, Y. Pommier, and T. R. Burke, Jr.** 1997. Hydrazide-containing inhibitors of HIV-1 integrase. *J Med Chem.* **40**:937-41.
280. **Zheng, R., T. M. Jenkins, and R. Craigie.** 1996. Zinc folds the N-terminal domain of HIV-1 integrase, promotes multimerization, and enhances catalytic activity. *Proc Natl Acad Sci U S A.* **93**:13659-64.
281. **Zhu, K., M. L. Cordeiro, J. Atienza, W. E. Robinson, Jr., and S. A. Chow.** 1999. Irreversible inhibition of human immunodeficiency virus type 1 integrase by dicaffeoylquinic acids. *J Virol.* **73**:3309-16.

Appendix 1.1 Source Data: Modelling PCR Amplification

Modelling the amplification of target DNA at a PCR efficiency of 100% (left side; **1**) and 80% (right side; **0.8**). The cycle numbers of the PCR are given in the far left column. Initial starting amounts of template DNA (10, 50 and 250) are indicated at cycle 0. This data was used to generate the graphs presented in Figures 1.16 and 1.17.

	1	Efficiency		0.8	Efficiency	
Cycle #	DNA Product	DNA Product	DNA Product	DNA Product	DNA Product	DNA Product
0	10	50	250	10	50	250
1	20	100	500	18	90	450
2	40	200	1000	32	162	810
3	80	400	2000	58	292	1458
4	160	800	4000	105	525	2624
5	320	1600	8000	189	945	4724
6	640	3200	16000	340	1701	8503
7	1280	6400	32000	612	3061	15306
8	2560	12800	64000	1102	5510	27550
9	5120	25600	128000	1984	9918	49590
10	10240	51200	256000	3570	17852	89262
11	20480	102400	512000	6427	32134	160671
12	40960	204800	1024000	11568	57842	289208
13	81920	409600	2048000	20823	104115	520574
14	163840	819200	4096000	37481	187407	937033
15	327680	1638400	8192000	67466	337332	1686660
16	655360	3276800	16384000	121440	607198	3035988
17	1310720	6553600	32768000	218591	1092956	5464779
18	2621440	13107200	65536000	393464	1967320	9836602
19	5242880	26214400	131072000	708235	3541177	17705884
20	10485760	52428800	262144000	1274824	6374118	31870591

Appendix 1.2 Source Data: Modelling the Effect of Saturation on PCR Amplification

The cycle-number of the PCR is indicated in the far left column. Initial input template DNA copy number is indicated at cycle 0. Saturation was modelled by introducing a “saturation factor” (marked as a %) at the cycle below which 1×10^7 copies of the amplified product are present (indicated by boldface script). The “saturation factor” essentially reflects the % decrease in PCR amplification efficiency (from 100%) that is expected to occur due to competition for components within the reaction mix. Data is otherwise the same as for Appendix 1.1 (left side; 1) and assumes a PCR efficiency of 100% prior to the point at which the saturation factor is introduced.

Efficiency		1			
Cycle #	DNA Product	DNA Product	DNA Product	DNA Product	
0	10	50		250	
1	20	100		500	
2	40	200		1000	
3	80	400		2000	
4	160	800		4000	
5	320	1600		8000	
6	640	3200		16000	
7	1280	6400		32000	
8	2560	12800		64000	
9	5120	25600		128000	
10	10240	51200		256000	
11	20480	102400		512000	
12	40960	204800		1024000	
13	81920	409600		2048000	
14	163840	819200		4096000	
15	327680	1638400		8192000	0%
16	655360	3276800		15974400	5%
17	1310720	6553600	0%	30351360	10%
18	2621440	12779520	5%	54632448	20%
19	5242880	24281088	10%	87411917	40%
20	10223616	43705958	20%	104894300	80%

Appendix 3.1

Source data obtained following ImageQuant analysis of the dot-blot presented in Figure 3.26. Duplicate counts for the HA8, LHS, RHS and 2-LTR standards were averaged (Average) and adjusted for background signal (Average-Background). The LTR copies obtained for the LHS, RHS and 2-LTR constructs were obtained using the equation shown in Figure 3.26B ($y=$).

HA8	Counts	Average	Average - Background	LTR Copies
20000	2760328	2987064	2969317	
20000	3213800			
2000	501163	473863	456116	
2000	446563			
200	80533	68625	50878	
200	56717			
0	19568	17748	0	
0	15927			
LHS	231357	204205	186457	876
LHS	177052			
RHS	109727	126075	108328	341
RHS	142423			
2-LTR	225458	243246	225499	1143
2-LTR	261034			

Appendix 3.2

Source data for the results presented in Figure 3.30. The melting temperatures (T_m) were obtained using MeltCalc software (Schutz and von Ahsen, 1999) that can be downloaded at <http://www.meltcalc.com>. T_m temperatures were calculated based on a fixed oligonucleotide concentration (Oligoconc. [μM]) of $0.1\mu\text{M}$ in the absence of DMSO and a MgCl_2 concentration of either 2.5mM (Na eq. [mM] = 300) or 4.5mM (Na eq. [mM] = 300). The DP sequence is shown in the Sequence (perfect match) row and the integrant sequence is shown in the Sequence (mismatch) row. The ΔT_m value reflects the difference in T_m temperatures between a perfectly matched oligonucleotide and the mismatched oligonucleotide.

Integrand #1 @ 2.5mM MgCl₂

Oligoconc. [μ M]	0.1	DMSO [%]	0
Na eq. [mM]	300		
Length [bp]	18	Tm	Δ Tm
Sequence (perfect match)	catcatctcttggaaggg	55.6	23.4
Sequence (mismatch)	catgatctattggaaggg	32.3	

Integrand #1 @ 4.5mM MgCl₂

Oligoconc. [μ M]	0.1	DMSO [%]	0
Na eq. [mM]	630		
Length [bp]	18	Tm	Δ Tm
Sequence (perfect match)	catcatctcttggaaggg	59.4	22.3
Sequence (mismatch)	catgatctattggaaggg	37.1	

Integrand #2

Oligoconc. [μ M]	0.1	DMSO [%]	0
Na eq. [mM]	630		
Length [bp]	18	Tm	Δ Tm
Sequence (perfect match)	catcatctcttggaaggg	59.2	0.0
Sequence (mismatch)	catcatctcttggaaggg	59.4	

Integrand #3

Oligoconc. [μ M]	0.1	DMSO [%]	0
Na eq. [mM]	630		
Length [bp]	18	Tm	Δ Tm
Sequence (perfect match)	catcatctcttggaaggg	58.4	12.3
Sequence (mismatch)	gggcatctcttggaaggg	47.1	

Integrand #4

Oligoconc. [μ M]	0.1	DMSO [%]	0
Na eq. [mM]	630		
Length [bp]	18	Tm	Δ Tm
Sequence (perfect match)	catcatctcttggaaggg	59.4	17.0
Sequence (mismatch)	catcatcgactggaaggg	42.4	

Integrand #5

Oligoconc. [μ M]	0.1	DMSO [%]	0
Na eq. [mM]	630		
Length [bp]	18	Tm	Δ Tm
Sequence (perfect match)	ccccccccgctggaaggg	75.0	0.0
Sequence (mismatch)	ccccccccgctggaaggg	75.0	

Integrand #6

Oligoconc. [μ M]	0.1	DMSO [%]	0
Na eq. [mM]	630		
Length [bp]	18	Tm	Δ Tm
Sequence (perfect match)	aaaaaaaaatatggaaggg	52.7	0.0
Sequence (mismatch)	aaaaaaaaatatggaaggg	52.7	

Appendix 4.1 Linker Ligation Efficiency

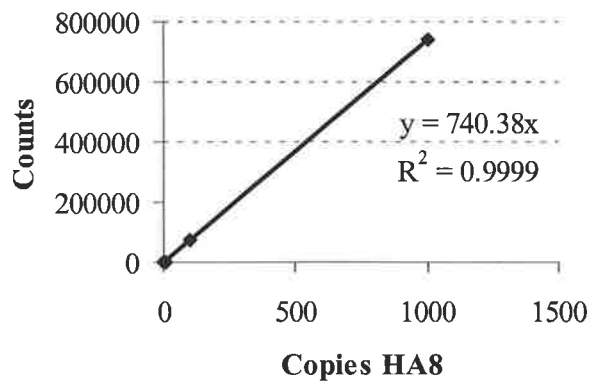
Quantification of the band intensities observed in Figure 4.34 using ImageQuant software (PhosphorImage analysis). The signal intensity of bands is shown as Counts. The counts resulting from the –ligase lane (Unligated) and each band in the +ligase lane (indicated by sizes in bp) are expressed as a percentage of the sum total of counts obtained for each lane (Percentage). The sum of all the counts obtained in the +ligase lane is indicated (SUM *Nla* III ligated).

<u>Band</u>	<u>Counts</u>	<u>Percentage</u>
Unligated	133296	100
396bp	68771	52
742bp	51775	39
1088bp	41778	31
1434bp	3958	3
SUM <i>Nla</i> III ligated	166292	125

Appendix 4.2 Selectivity of LP-PCR

Quantification of the band intensities observed in Figure 4.35 using ImageQuant software (PhosphorImage analysis). The signal intensity of bands is shown as Counts. Counts obtained following duplicate PCR amplification of experimental samples were averaged (Average Counts) and adjusted for background counts (Av. Counts - Background). Band intensities resulting from amplification of known copies of the HA8 integrated standard were quantified and used to generate the standard curve (pictured bottom). The regression value (R^2) of the standard curve is shown. The HA8 copy-number to which the signal intensity following amplification of 1000 copies of the lin construct corresponds is indicated (Equivalent HA8 Copies). This signal intensity is also expressed as a percentage of that obtained following amplification of 1000 copies of the HA8 integrated standard (% of 1000 HA8).

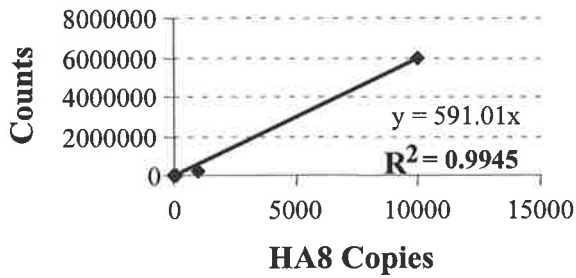
HA8	Counts	Average Counts	Av. Counts - Background	Equivalent HA8 Copies	% of 1000 HA8
1000	737585	747298	740322		
1000	757011				
100	98935	82233	75257		
100	65530				
10	9909	8376	1400		
10	6842				
0	6682	6976	0		
0	7270				
1000lin	59414	61439	54463	74	7.4
1000lin	63463				



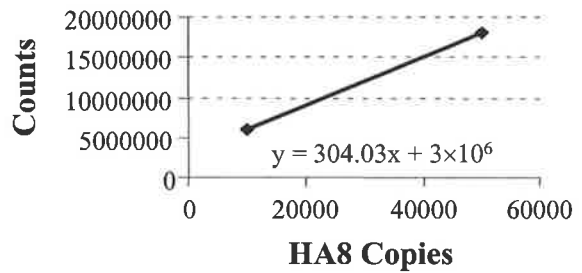
Appendix 5.1 Accumulation of Total HIV DNA Following Infection of HuT-78 Cells: GAG-PCR

Data obtained by quantifying bands (ImageQuant analysis) resulting from GAG-PCR amplification of experimental samples (see Figure 5.9). Bands not quantified are indicated as "Bad Probe". These results were used to generate the graph presented in Figure 5.10. Band intensities resulting from amplification of the HA8 standards were quantified (data not shown) and used to generate the standard curves (pictured top). Counts from duplicate samples were averaged (Average Counts) and adjusted for background counts (Av. Counts - Background). Copies per 500 cells for 2h-10h p.i. samples were calculated using the equation from Standard Curve 0-10000 (top and left). Copies per 500 cells for 14h-50h p.i. samples were calculated using the equation from Standard Curve 10000-50000 (top and right).

Standard Curve 0-10000



Standard Curve 10000-50000

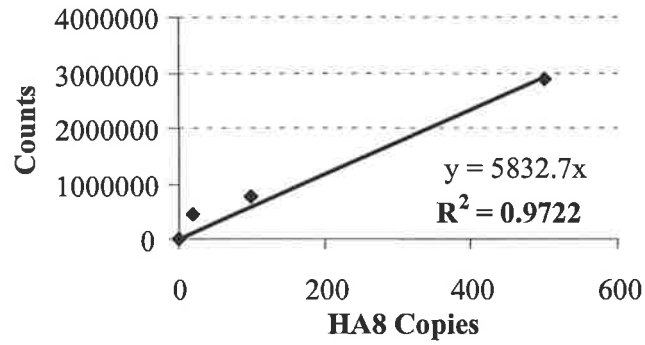


h p.i.	Counts	Average Counts	Av. Counts - Background	Copies per 500 cells	Copies per 1000 cells
2	46132	53105	0	0	0
2	60078				
3	1177423	957711	904606	1531	3061
3	737999				
4	1288409	1074847	1021742	1729	3458
4	861284				
5	2802224	3206657	3153552	5336	10672
5	3611090				
6	3312808	3929542	3876437	6559	13118
6	4546275				
7	3542757	3542757	3489652	5905	11809
7	Bad probe				
8	Bad probe	4189370	4136265	6999	13997
8	4189370				
9	4490159	4633499	4580394	7750	15500
9	4776838				
10	4472454	5426515	5373410	9092	18184
10	6380576				
14	8464344	8018935	7965830	16333	32667
14	7573525				
18	5370757	5520080	5466975	9250	18500
18	5669402				
26	6467427	5799984	5746879	9724	19448
26	5132540				
50	9049517	8859406	8806301	19098	38196
50	8669295				

Appendix 5.2 Accumulation of Integrated HIV DNA Following Infection of HuT-78 Cells: LP-PCR

Data obtained by quantifying bands (ImageQuant analysis) resulting from LP-PCR amplification of experimental samples (see Figure 5.12). Bands not quantified are indicated as “bad probe”. These results were used to generate the graph presented in Figure 5.13. Band intensities resulting from amplification of the HA8 standards were quantified (data not shown) and used to generate the standard curve (pictured top). The regression value (R^2) of the standard curve is shown. Counts from duplicate samples were averaged (Average Counts) and adjusted for background counts (Average Counts - Background). Copies per 100 cells were calculated using the equation ($y=$) from the standard curve (top).

Standard Curve 0-500

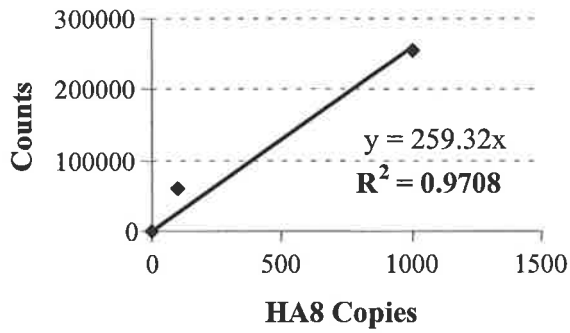


h p.i.	Counts	Average Counts	Average Counts - Background	Copies per 100 cells	Copies per 1000 cells
2	16645	17504	4787	1	10
2	18363				
3	13035	14188	1471	0	0
3	15341				
4	98187	110254	97536	17	170
4	122320				
5	133905	143264	130546	22	220
5	152622				
6	482608	493094	480377	82	820
6	503580				
7	559880	551554	538837	92	920
7	543228				
8	754446	763586	750869	129	1290
8	772726				
9	532280	490176	477458	82	820
9	448071				
10	956679	918352	905635	155	1550
10	880025				
14	1383433	1449656	1436939	246	2460
14	1515879				
18	1591642	1455174	1442457	247	2470
18	1318706				
26	1486546	1604499	1591781	273	2730
26	1722451				
50	1421039	1239791	1227074	210	2100
50	1058543				

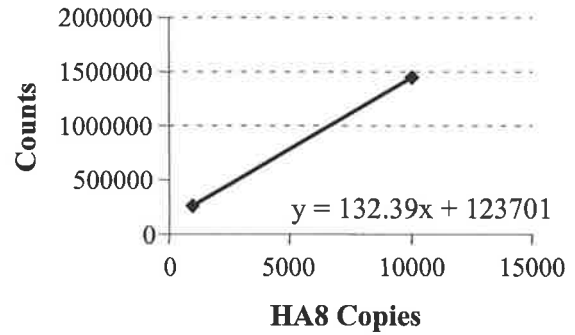
Appendix 5.3 Accumulation of Integrated HIV DNA Following Infection of HuT-78 Cells: Nested-Alu PCR

Data obtained by quantifying bands (ImageQuant analysis) resulting from nested-*Alu* PCR amplification of experimental samples (see Figure 5.14). These results were used to generate the graph presented in Figure 5.15. Band intensities resulting from amplification of the HA8 standards were quantified (data not shown) and used to generate the standard curves (pictured top). Counts from duplicate samples were averaged (Average Counts) and adjusted for background counts (Av. Counts - Background). Integrated DNA copies per 1000 cells for 2-14h p.i. were calculated using the equation from the Standard Curve 0-1000 (top and left). Integrated DNA copies per 1000 cells for the remaining samples were calculated using the equation from the Standard Curve 1000-10000 (top and right). The 26h p.i. samples amplified in the absence of the *Alu*164 primer are indicated (26-*Alu*).

Standard Curve 0-1000



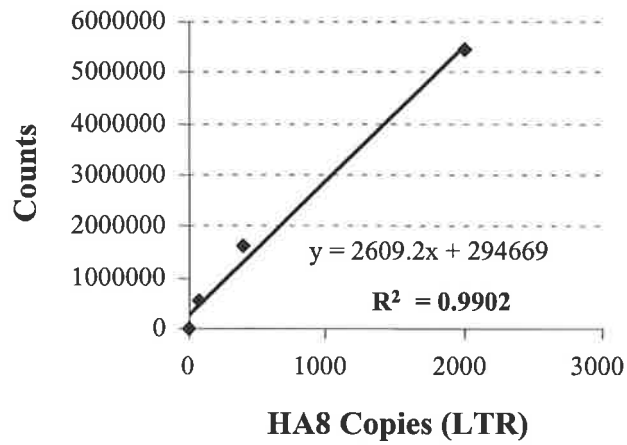
Standard Curve 1000-10000



h p.i.	Counts	Average Counts	Av. Counts - Background	Integrated DNA Copies per 1000 cells
2	1178	1550	-792	-3
2	1921			
3	1576	1625	-717	-3
3	1673			
4	1999	1715	-627	-2
4	1430			
5	1578	1633	-709	-3
5	1687			
6	1690	2145	-196	-1
6	2600			
7	3516	5260	2919	11
7	7004			
8	24495	25315	22974	89
8	26135			
9	75386	64366	62025	239
9	53345			
10	45467	44112	41771	161
10	42756			
14	65555	63809	61468	237
14	62063			
18	337069	322595	320254	1485
18	308121			
26	309469	309469	307128	1386
26	bad probe			
50	430932	389270	386929	1988
50	347607			
26 -Alu	8830	7441	5100	20
26 -Alu	6052			

Appendix 5.4 Equating the 2-LTR Construct

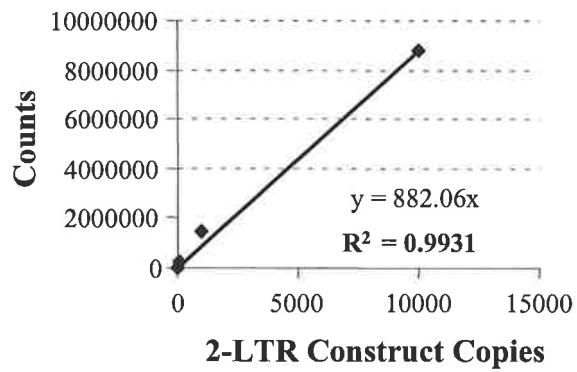
The standard curve (top) was generated from data obtained by quantifying bands (ImageQuant analysis) resulting from amplification of known copies of the HA8 DNA standards (see Fig. 5.17). Counts from duplicate samples were averaged (Average Counts) and adjusted for background counts (Average Counts - background). The regression value (R^2) of the standard curve is shown. The signals obtained following amplification of ≈ 250 copies of the 2-LTR control construct and the equation of the standard curve ($y=$) were used to determine the normalised copy-number (Adjusted Copies).



HA8 (LTR)	Counts	Average Counts	Average Counts -Background	
2000	5237632	5518198	5452874	
2000	5798763			
400	2381262	1690943	1625619	
400	1000623			
80	625541	636276	570953	
80	647011			
0	58750	65324	0	
0	71897			
LTR Construct	Counts	Average Counts	Average Counts -Background	Adjusted Copies
≈250	908136	1048861	983537	264
≈250	1189585			

Appendix 5.5 Accumulation of 2-LTR HIV DNA Following Infection of HuT-78 Cells

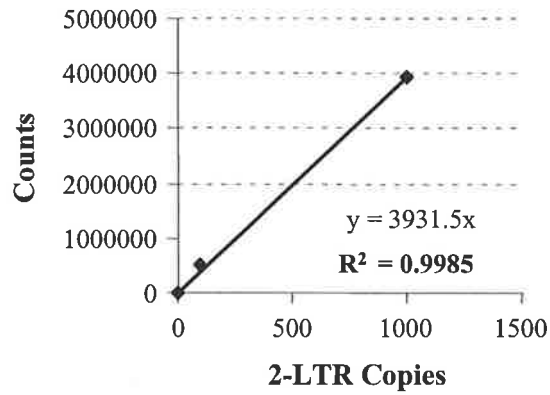
Data obtained by quantifying bands (ImageQuant analysis) resulting from 2-LTR PCR amplification of experimental samples (see Figure 5.18). These results (2h-14h p.i.) were used to generate the graph presented in Figure 5.20. Band intensities resulting from amplification of known copies of the 2-LTR control construct were quantified (data not shown) and used to generate the standard curve (pictured top). Counts from duplicate samples were averaged (Average Counts) and adjusted for background counts (Av. Counts - Background). 2-LTR DNA copies per 500 cells were calculated using the equation ($y=$) from the standard curve.



h p.i.	Counts	Average Counts	Av. Counts - Background	2-LTR Copies per 500 cells	2-LTR Copies per 1000 cells
2	65	128	0	0	0
2	191				
3	90	85	-43	0	0
3	80				
4	109	592	464	1	1
4	1075				
5	530	360	232	0	1
5	190				
6	730	459	331	0	1
6	187				
7	2498	6256	6128	7	14
7	10013				
8	15610	18957	18829	21	43
8	22303				
9	71062	51311	51183	58	116
9	31559				
10	87492	78174	78046	88	177
10	68856				
14	162634	130183	130055	147	295
14	97731				
18	174544	185759	185631	210	421
18	196973				
26	53505	67136	67008	76	152
26	80767				
50	219051	405637	405509	460	919
50	592222				

Appendix 5.6 Accumulation of 2-LTR HIV DNA Following Infection of HuT-78 Cells: Re-Analysis of Selected Time-Points

Data obtained by quantifying bands (ImageQuant analysis) resulting from 2-LTR PCR amplification (re-analysis) of the 18h, 26h and 50h p.i. experimental samples (see Figure 5.19). These results were used to generate the graph presented in Figure 5.20 (later time-points). Band intensities resulting from amplification of known copies of the 2-LTR control construct were quantified and used to generate the standard curve (pictured top). The regression value (R^2) of the standard curve is shown. Counts obtained following duplicate PCR amplification of experimental samples were averaged (Average Counts) and adjusted for background counts (Av. Counts - Background). 2-LTR DNA copies per 1000 cells were calculated using the equation ($y=$) from the standard curve.



2-LTR	Counts	Counts - Background			
1000	3920167	3919721			
100	511518	511072			
0	446	0			
h p.i.	Counts	Average Counts	Av. Counts - Background	2-LTR Copies per 500 cells	2-LTR Copies per 1000 cells
16	626482	678591	678145	172	345
16	730700				
24	782145	769273	768827	196	391
24	756400				
48	1770608	1597021	1596575	406	812
48	1423434				

Appendix 6.1

In vitro activity against HIV-1 IN, cytotoxicity and concentration of drugs used in this study

Drug	Anti-IN Activity(IC₅₀)¹	Cytotoxicity (% cell death)	Cell Culture Concentration Used
5,8 dihydroxynapthoquinone	2.5μM	100% @ 50μM	N/A
Quercetin dihydrate	12μM	<1% @100μM	50μM
AR177	0.05μM	<1% @ 100μM	10μM
L17	0.02μM	<5% @ 40μM	30μM
L-731,988	0.1μM	<2% @ 50μM	10μM
L-708,906	0.1μM	<4% @ 50μM	10μM
3TC	RT Inhibitor	<1% @ 50μM	10μM
AZT	RT Inhibitor	<1% @ 50μM	10μM