INHIBITING HIV-1 USING RNA INTERFERENCE (RNAi) TO TARGET NOVEL HIV DEPENDENCY FACTORS (HDFs)

Mishka D Blondeel

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

DECLARATION

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

........................................

........ day of May, 2010
DEDICATION

To my long-suffering family and Martin
Publications and presentations

ABSTRACT

Three separate recent publications used genome-wide RNA interference (RNAi) to screen for novel host factors that are required for HIV-1 infection and replication. This was achieved using small interfering RNAs (siRNAs) to silence the expression of ~21,000 human genes and determining the effect of each gene's loss of function on HIV-1 replication. Collectively, several hundred genes have now been implicated as novel HIV-1 host factors (termed HIV-1 Dependency Factors, HDFs). However, differences in study design resulted in little overlap and limited interpretive value from the three published datasets. To identify novel HDFs that are potential targets for anti-HIV therapy, five putative HDFs (SPTBN1, TMED2, KIAA1012, PRDM14 and SP110) were chosen for validation. RNAi effecters (both siRNAs and expressed short hairpin RNAs) were used to silence the selected genes. Gene suppression was measured by quantitative RT-PCR assay and two candidate genes were studied further (SPTBN1 and SP110) based on efficient mRNA inhibition (over 90%). As efforts to deliver the RNAi effecters to a T-cell line were unsuccessful, the effect of this knockdown on HIV-1 replication (both early- and late-stage) was assessed in cultured TZM-bl cells, a HeLa-derived cell line that expresses HIV-1 entry receptors and an integrated luciferase reporter for HIV-1 transcriptional activity (also used in the first genome-wide RNAi screen). An initial viral challenge assay with Subtype C-enveloped pseudovirus showed a 60% decrease in TZM-bl luciferase reporter activity in cells with suppressed SPTBN1 function, while knockdown of SP110 showed no effect on reporter activity. The final experiment, using fully-replicating Subtype B virus, showed a 75% decrease in late-stage viral replication when SPTBN1 expression was suppressed. In addition, SP110 suppression was confirmed to have no effect on TZM-bl reporter activity during any stage of HIV-1 replication. In conclusion, SPTBN1, but not SP110, is required for late-stage HIV-1 replication, though these results need to be confirmed in CD4+ T-cells. The absence of several important viral accessory factors from
the virus used in the genome-wide screen may explain these findings and emphasises the need for using physiologically representative viral and cellular models to study the viral/cellular interactome.
ACKNOWLEDGEMENTS

I would like to thank my supervisors, Dr M S Weinberg and Prof P Arbuthnot, for their academic support and guidance.

I would like to acknowledge Dr Papathanasopoulos for providing access to the HIV-1 Pathogenesis Research Laboratory and for generously providing the pseudovirus used in this study and thank Mark Killick for his guidance and assistance in that laboratory.

My personal thanks go to all the members of the Antiviral Gene Therapy Research Unit (AGTRU), especially Victoria Green and Abdullah Ely for their advice and support.

Finally, I am grateful for the financial assistance I received from the German Academic Exchange Service, DAAD (Bonn), the South African Medical Research Council (MRC) and the Poliomyelitis Research Foundation, as well as the Postgraduate Merit Award granted by the University of the Witwatersrand.
# TABLE OF CONTENTS

DECLARATION ........................................................................................................ II
DEDICATION ........................................................................................................... III
PUBLICATIONS AND PRESENTATIONS ................................................................ IV
ABSTRACT .............................................................................................................. V
ACKNOWLEDGEMENTS .......................................................................................... VII
TABLE OF CONTENTS ............................................................................................ VIII
LIST OF FIGURES .................................................................................................. XIII
LIST OF TABLES ..................................................................................................... XV
ABBREVIATIONS .................................................................................................. XVI

CHAPTER 1 : INTRODUCTION ................................................................................ 1

1.0 Chapter summary .............................................................................................. 1

1.1 HIV-1/AIDS .................................................................................................... 2

1.1.1 Global and local burden of the AIDS pandemic ........................................... 2

1.1.2 AIDS pathogenesis ..................................................................................... 3

1.1.3 The HIV-1 replication cycle ...................................................................... 4

1.2 Host factors required for HIV-1 replication .................................................... 5

1.2.1 Entry ........................................................................................................... 6

1.2.2 Uncoating and Reverse Transcription ......................................................... 7

1.2.3 Inward trafficking and nuclear import ....................................................... 7

1.2.4 Integration .................................................................................................. 8
CHAPTER 1: INTRODUCTION AND BACKGROUND

1.2.5 Transcription
1.2.6 RNA processing and export
1.2.7 Translation, processing and trafficking
1.2.8 RNA encapsulation and assembly
1.2.9 Virion budding
1.2.10 Virus transfer

1.3 Current approaches to managing HIV-1/AIDS

1.4 RNA interference (RNAi)

1.4.1 The endogenous RNAi pathway
1.4.2 RNAi as a reverse genetics tool
1.4.3 RNAi as an exciting and powerful new gene therapy

1.5 RNAi approaches targeting HIV-1

1.5.1 Viral versus cellular targets

1.6 Genome-wide siRNA screens for novel HIV-1 host factors

1.6.1 Selection of novel host factors for validation in this study

1.7 Aims and Objectives

CHAPTER 2: METHODS AND MATERIALS

2.1 RNAi effecters

2.1.1 shRNA-encoding sequences
2.1.2 Specificity and structure of shRNA sequences
2.1.3 siRNA SMARTpools

2.2 Previously constructed plasmids used in this study
2.3 Molecular cloning ............................................................................................................34
  2.3.1 Cloning of shRNA expression cassettes .................................................................34
  2.3.2 Cloning of target reporter plasmids .......................................................................36
  2.3.3 Transformation and verification of clones ............................................................38

2.4 Mammalian cell culture ..................................................................................................40

2.5 Transfection of DNA and RNA into cultured cells .......................................................41
  2.5.1 Lipofectamine™ 2000 transfections ........................................................................41
  2.5.2 ESCORT™ V transfections ....................................................................................41
  2.5.3 DharmaFECT®-4 transfections ...............................................................................42

2.6 Dual-Luciferase reporter assays .......................................................................................42

2.7 qRT-PCR assays .............................................................................................................43

2.8 Electroporation ...............................................................................................................45

2.9 Delivery of DNA and RNA to SupT1 cells by cationic lipid-based transfection reagents ..............................................................................................................46

2.10 Cytotoxicity assays .......................................................................................................47

2.11 Virus production ...........................................................................................................48

2.12 Viral titration ................................................................................................................49

2.13 Viral challenge assays ..................................................................................................50
  2.13.1 Env-pseudotyped viral challenge ..........................................................................51
  2.13.2 Replication-competent viral challenge ..................................................................52

2.14 Statistical analyzes ......................................................................................................53
CHAPTER 3 : RESULTS .................................................................................................................. 54

3.1 Criteria-based selection of five genes .................................................................................. 54

3.2 Gene-specific knockdown by Pol III-expressed shRNAs ...................................................... 57
  3.2.1 Design of shRNA expression cassettes ......................................................................... 57
  3.2.2 Cloning of shRNA expression cassettes and reporter targets ...................................... 58
  3.2.3 Knockdown of luciferase reporter targets ..................................................................... 59

3.3 Accurate detection of mRNA by quantitative RT-PCR ......................................................... 62
  3.3.1 Testing of non-targeting controls .................................................................................. 62
  3.3.2 Testing of positive controls .......................................................................................... 64
  3.3.3 Other controls used in optimising qRT-PCR assays .................................................... 66

3.4 shRNA-mediated knockdown of endogenous mRNA targets ............................................. 67

3.5 Knockdown of endogenous mRNA by siRNA SMARTpools ............................................. 70

3.6 Knockdown efficacy of shRNAs derived from pooled siRNAs .......................................... 72

3.7 Delivery to SupT1 cells by electroporation ......................................................................... 74

3.8 Delivery to SupT1 cells by transfection reagents ............................................................... 75

3.9 Cytotoxicity assay ................................................................................................................ 78

3.10 The effect of SPTBN1 and SP110 knockdown on HIV-1 replication ................................. 80
  3.10.1 Single-round challenge of transfected cells with pseudovirus .................................. 80
  3.10.2 Two-part challenge of transfected cells with replication-competent virus .............. 82

CHAPTER 4 : DISCUSSION .......................................................................................................... 84

4.1 The possible roles of SPTBN1 in HIV-1 replication .......................................................... 84

4.2 SP110 and HIV-1 replication ............................................................................................... 85
4.3  Comparison of viruses used in this study and by Brass et al. (2008)........86

4.4  Gene selection .........................................................................................................................88

4.5  Broader perspectives ...............................................................................................................88

   4.5.1  Large-scale lentiviral shRNA transduction of T-cells .....................................................88

   4.5.2  Alternative techniques for characterizing the viral/cellular interactome .......90

4.6  Factors of RNAi experiments highlighted by this study ...............................................91

   4.6.1  Primer design in PCR-based cloning of shRNAs .............................................................91

   4.6.2  Target mRNA accessibility to the RNAi machinery ..........................................................92

   4.6.3  The importance of reliable reference genes and controls in qRT-PCR assays 
          ...........................................................................................................................................93

   4.6.4  Off-target effects and the benefits of Dharmacon’s ON-TARGETplus
          SMARTpool technology ..........................................................................................................94

   4.6.5  siRNA-related cytotoxicity ...............................................................................................95

   4.6.6  Delivery of RNAi effecters to a T-cell line by electroporation and lipid-based
          reagents ..................................................................................................................................98

   4.6.7  Alternative delivery methods ............................................................................................98

CHAPTER 5 : CONCLUSIONS ......................................................................................................102

APPENDIX A: SEQUENCE INFORMATION .................................................................................104

APPENDIX B: EXPERIMENTAL PROCEDURES .........................................................................112

APPENDIX C: SELECTION OF GENES ......................................................................................119

REFERENCES ..............................................................................................................................127
LIST OF FIGURES

Figure 1-1: The number of people living with HIV/AIDS in each country worldwide .......... 3
Figure 1-2: HIV-1 genome structure .................................................................................. 5
Figure 1-3: The mammalian RNAi pathway ....................................................................... 17
Figure 1-4: Model of putative HDF roles in the HIV lifecycle .................................. 27
Figure 2-1: Example of shRNA design approach ............................................................. 30
Figure 2-2: Schematic of two-step PCR approach used in construction of shRNA cassettes .................................................................................................................................................. 35
Figure 2-3: Reporter target cloning strategy ....................................................................... 38
Figure 2-4: Schematic of two-part viral challenge assay ............................................... 51
Figure 3-1: Example of restriction digest and sequence data used to verify correct clones. ................................................................................................................................................................. 60
Figure 3-2: Knockdown efficiency of shRNAs against luciferase reporter targets ........ 61
Figure 3-3: Testing knockdown of SPTBN1, TMED2, KIAA1012 and SP110 by non-specific shRNA controls .................................................................................................................................................. 63
Figure 3-4: Knockdown of positive control HTATSF1 .................................................... 64
Figure 3-5: Knockdown of LTR-driven luciferase by positive controls shLTR and shTat...66
Figure 3-6: Standard curve of plasmid standards for each primer set .............................. 67
Figure 3-7: Knockdown of endogenous SPTBN1 mRNA by shRNAs ............................. 68
Figure 3-8: Knockdown of endogenous TMED2 mRNA by shRNAs ............................. 69
Figure 3-9: Knockdown of endogenous KIAA1012 mRNA by shRNAs ......................... 69
Figure 3-10: Knockdown of endogenous SP110 mRNA by shRNAs .............................. 70
Figure 3-11: Knockdown of endogenous SPTBN1 mRNA by siRNA SMARTpools .......... 71
Figure 3-12: Knockdown of SPTBN1 and SP110 by shRNAs derived from SMARTpool siRNAs ................................................................................................................................................................. 73
Figure 3-13: Delivery of plasmid DNA and siRNA to SupT1 cells ..............................................76
Figure 3-14: Knockdown by siRNA SMARTpools in SupT1 cells ..............................................77
Figure 3-15: Knockdown of SPTBN1 and SP110 by siRNA-derived shRNAs in SupT1 cells. ........................................................................................................................................78
Figure 3-16: MTT assay for RNAi-related cytotoxicity ..................................................................79
Figure 3-17: Changes in pseudovirus activity in cells transfected with shRNA ......................81
Figure 3-18: Changes in pseudovirus activity in cells transfected with siRNAs .......................82
Figure 3-19: Changes in HIV-1p81A-4 activity in cells transfected with siRNAs ......................83
LIST OF TABLES

Table 3-1: The final five putative HDFs selected for validation in this study .................56
Table 3-2: Results of electroporation experiments (average of data in triplicate).............75
Table A-1: Genomic location of shRNA target sequences.............................................104
Table A-2: Genomic location of siRNA-derived shRNA target sequences..........................105
Table A-3: Primer sequences used in two-step PCR-amplification of shRNA expression cassettes for SPTBN1, TMED2 and KIAA1012...............................................................106
Table A-4: Primer sequences used in two-step PCR-amplification of shRNA expression cassettes for PRDM14 and SP110........................................................................107
Table A-5: Primer sequences used in single-round PCR-amplification of siRNA-derived shRNA expression cassettes ..................................................................................108
Table A-6: shRNA Target Oligonucleotide Sequences.....................................................109
Table A-7: qRT-PCR primers .........................................................................................110
Table A-8: Control siRNA/shRNAs and target primers ....................................................111
Table B-1: Preset electroporation protocols ....................................................................117
Table B-2: Electroporation protocols from the literature ................................................118
Table C-1: Genes that decreased HIV-1 activity by 75% or more.....................................120
Table C-2: Genes that scored with two or more siRNAs in Part 1 (left) and Part 2 (right) .................................................................................................................................121
Table C-3: Part 1 genes excluded due to important cellular functions ...............................122
Table C-4: Final Part 1 gene selection .............................................................................123
Table C-5: Part 2 genes excluded due to important cellular functions ...............................124
Table C-6: Part 2 genes with five or fewer homologs .......................................................125
Table C-7: Final Part 2 gene selection .............................................................................126
ABBREVIATIONS

(Only those used more than once are listed)

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>°C</td>
<td>degrees Celsius</td>
</tr>
<tr>
<td>A/T/C/G</td>
<td>Adenosine, Thymine, Cytosine, Guanine-residue (DNA bases)</td>
</tr>
<tr>
<td>ABCE1</td>
<td>ATP-binding cassette, sub-family E (OABP), member 1</td>
</tr>
<tr>
<td>AdV/AAV</td>
<td>Adenovirus/Adeno-associated virus</td>
</tr>
<tr>
<td>AIDS</td>
<td>Acquired immune-deficiency syndrome</td>
</tr>
<tr>
<td>ANOVA</td>
<td>Analysis of variance</td>
</tr>
<tr>
<td>APOBEC</td>
<td>apolipoprotein B mRNA editing enzyme, catalytic polypeptide-like</td>
</tr>
<tr>
<td>ARV</td>
<td>Antiretroviral</td>
</tr>
<tr>
<td>BAF</td>
<td>Barrier to auto-integration factor</td>
</tr>
<tr>
<td>BDH</td>
<td>Burani Design Holding</td>
</tr>
<tr>
<td>BLAST</td>
<td>Basic local alignment search tool</td>
</tr>
<tr>
<td>bp</td>
<td>base-pair</td>
</tr>
<tr>
<td>BSA</td>
<td>Bovine serum albumin</td>
</tr>
<tr>
<td>CA</td>
<td>Capsid</td>
</tr>
<tr>
<td>CCR5</td>
<td>C-C chemokine receptor type 5</td>
</tr>
<tr>
<td>CD4</td>
<td>Cluster of differentiation 4</td>
</tr>
<tr>
<td>CMV</td>
<td>Cytomegalovirus</td>
</tr>
<tr>
<td>CXCR4</td>
<td>C-X-C chemokine receptor 4</td>
</tr>
<tr>
<td>DEAE</td>
<td>Diethylaminoethyl</td>
</tr>
<tr>
<td>DMEM</td>
<td>Dulbecco’s modified Eagle’s medium</td>
</tr>
<tr>
<td>DMSO</td>
<td>Dimethylsulfoxide</td>
</tr>
<tr>
<td>DNA / cDNA</td>
<td>deoxyribonucleic acid / complementary DNA</td>
</tr>
<tr>
<td>EBV</td>
<td>Epstein-Barr virus</td>
</tr>
<tr>
<td>EDTA</td>
<td>Ethylenediaminetetraacetic acid</td>
</tr>
<tr>
<td>Env</td>
<td>Envelope</td>
</tr>
<tr>
<td>ER</td>
<td>endoplasmic reticulum</td>
</tr>
<tr>
<td>ESCRT</td>
<td>Endosomal sorting complex required for transport</td>
</tr>
<tr>
<td>FBS</td>
<td>Fetal bovine serum</td>
</tr>
<tr>
<td>g (500 × g)</td>
<td>Unit of gravitational force</td>
</tr>
<tr>
<td>g</td>
<td>gram</td>
</tr>
<tr>
<td>Gag</td>
<td>Group-specific antigen</td>
</tr>
<tr>
<td>GAPDH</td>
<td>Glyceraldehyde-3-phosphate dehydrogenase</td>
</tr>
<tr>
<td>GFP/eGFP</td>
<td>Green fluorescent protein/enhanced GFP</td>
</tr>
<tr>
<td>HAART</td>
<td>Highly active antiretroviral therapy</td>
</tr>
<tr>
<td>HBV</td>
<td>Hepatitis B virus</td>
</tr>
<tr>
<td>HDAC</td>
<td>Histone deacetylase</td>
</tr>
<tr>
<td>HDF</td>
<td>HIV-1 dependency factor</td>
</tr>
<tr>
<td>HEK293</td>
<td>Human embryonic kidney cell line 293</td>
</tr>
<tr>
<td>HEPES</td>
<td>4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid</td>
</tr>
</tbody>
</table>
HIV-1: Human immunodeficiency virus type 1
hnRNP: Heterogeneous nuclear ribonucleoprotein
HTATSF1: HIV-1 Tat-specific factor 1
IN: Integrase
KIAA1012: NCBI gene name, encodes TRS85 homologue
LB: Luria-Bertani
LEDGF: Lens epithelium-derived growth factor
LTR: Long terminal repeat
Luc/FF Luc: Luciferase / Firefly Luc
m/µ/n/p/f (e.g. mg/µg/ng/pg): milli-/micro-/nano-/pico/femto- (e.g. femtograms)
M: Molar
miRNA: microRNA
MTT: (3-(4,5-Dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide
NARRRP: NIH AIDS research and reference reagent program
NC: Nucleocapsid
NCBI: National Centre for Biotechnology Information
NEB: New England Biolabs
Nef: Negative regulatory factor
NIH: National Institute of Health
ns: not significant
nt: nucleotide (DNA base)
NTP/dNTP: Nucleoside triphosphate / deoxy-NTP
ORF: Open reading frame
PBS: Phosphate-buffered saline
PCR/qRT-PCR: Polymerase chain reaction / quantitative RT-PCR
PIC: Pre-integration complex
Pol: Polymerase
PR: Protease
PRDM14: PR domain-containing 14
PTGS/TGS: Post-transcriptional gene silencing / transcriptional gene silencing
RISC: RNA-induced silencing complex
RNA/mRNA: Ribonucleic acid / messenger RNA
RNAi: RNA interference
RNase: Ribonuclease
RT: Reverse Transcriptase
RTC: Reverse-transcription complex
shRNA/siRNA: Short hairpin RNA / short interfering RNA
SP110: NCBI gene name, encodes nuclear body protein
SPTBN1: Spectrin, beta, non-erythrocytic 1
TAP: Transcytosis-associated protein
Taq: Thermus aquaticus
TAR: Trans-activation response
Tat: transcriptional activator protein
TBE: Tris/borate-EDTA
TCID50: 50% tissue culture infectious dose
TMED2: Trans-membrane emp24 domain-containing protein 2
TRBP: TAR RNA binding protein
TRC: The RNAi Consortium
TRC: The RNAi Consortium
TRIM: Tripartite motif
TSA: Trichostatin A
U: Enzyme unit (amount required to digest 1 µg of DNA in an hour)
UNAIDS: The joint United Nations programme on HIV/AIDS
URL: Universal resource locator (website address)
UTR: Untranslated region
VC/NV: Virus control / no virus
Vif: Viral infectivity factor
Vpr: Viral protein R
Vpu: Viral protein U
CHAPTER 1: INTRODUCTION

1.0 Chapter summary

Acquired Immune-Deficiency Syndrome (AIDS) is an infectious disease caused by the Human Immunodeficiency Virus (HIV) that is now a pandemic affecting 33.4 million people worldwide, with 5.7 million of these living in South Africa. HIV-1 is a retrovirus that makes extensive use of its host cell’s pathways during its replication cycle in cells of the immune system. Antiretroviral (ARV) therapies targeting various steps in the HIV-1 replication cycle can suppress viral activity in infected patients and prevent the onset of AIDS (Palella et al., 1998). However, these therapies are not without adverse side effects and cannot eradicate latent viral reservoirs. Alternative antiviral strategies include gene therapy approaches, the most promising of which uses RNA (ribonucleic acid) interference (RNAi) to specifically and potently silence gene expression and thereby suppress critical viral and/or cellular gene functions that are required for efficient HIV-1 replication (Barichievy et al., 2009). RNAi-based silencing of viral genes has suppressed HIV-1 infection in several cellular and animal models. However, the error-prone Reverse Transcriptase enzyme of HIV-1 leads to the rapid evolution of mutant strains, resulting in resistance developing to both ARVs and RNAi-based therapies in the long term. By targeting HIV-1 host factors the risk of viral mutational escape can be alleviated. However, several concerns exist regarding the possible side-effects of inhibiting a cellular gene function, as well as the existence of functional redundancy in the pathways used by the virus.

Apart from gene therapy applications, RNAi is widely used as a tool to explore gene function [reviewed by (Castanotto and Rossi 2009)]. RNAi has been widely used to reveal the complex interplay that occurs between cellular and viral gene products during
HIV-1 infection and replication. The advent of high-throughput technology has allowed RNAi to be used at a genome-wide scale and has resulted in the identification of hundreds of novel HIV-1 host factors. However, as the cellular and viral models used in these studies do not represent a natural infection validation of these candidate host factors is required.

1.1 HIV-1/AIDS

1.1.1 Global and local burden of the AIDS pandemic

The HIV/AIDS pandemic is placing an increasingly heavy burden on global health, with 33.4 million people living with the disease worldwide (see Figure 1-1). Over two thirds of these people (22.4 million) are living in Sub-Saharan Africa. This region also accounts for 32% of all new HIV infections and 72% of global AIDS-related deaths (1.9 million and 1.4 million people, respectively) (UNAIDS 2009). South Africa’s population has the largest number of infections in the world (5.7 million people) (UNAIDS 2008).
Figure 1-1: The number of people living with HIV/AIDS in each country worldwide

Just over thirty-three million people are living with HIV/AIDS world-wide. Sub-Saharan Africa’s infected population stands at 22.4 million people (UNAIDS 2009).

1.1.2 AIDS pathogenesis

AIDS is a complex and variable syndrome whose pathogenesis is not yet completely understood. However, it is generally characterized by progressive immune deficiency that results in susceptibility to viral, bacterial and fungal opportunistic infections [reviewed by (El-Atrouni et al., 2006; Martinez and Temesgen 2006; Tsigrelis et al., 2006)]. AIDS patients also have increased risk of developing tumours of the gastrointestinal tract, lungs or skin that are mostly caused by co-infections with oncogenic (cancer-causing) viruses like Epstein-Barr Virus, EBV [reviewed by (Angeletti et al., 2008)]. Finally, neurological symptoms (including encephalopathies, meningitis and AIDS dementia complex) often develop in late-stage AIDS patients [reviewed by (Levy et al., 1985; Navia et al., 1986)], as well as non-specific symptoms like weight-loss, diarrhea and fevers.

The course of the disease begins with an acute stage following HIV-1 infection, progresses to a chronic asymptomatic stage that can last for 10 years or more and culminates in the clinical symptoms of AIDS and subsequent death. During the acute stage, high levels of virus target the T-helper lymphocytes that express the CD4 (cluster of differentiation 4) surface marker and cause them to undergo apoptosis (programmed cell death) (Laurent-Crawford et al., 1991; Lu et al., 1994). The immune system is in a chronically activated state throughout the course of the infection, contributing to the virus-mediated depletion of CD4+ T-cell numbers (Hazenberg et al., 2003), especially those bearing the CCR5 (C-C chemokine receptor type 5) marker that are concentrated in the gut mucosal lining, but also newly generated in the thymus and circulating in peripheral blood (Brenchley et al., 2004). As the T-helper cells are critical for an effective immune
response, their loss is directly involved in the progression of the disease. When the CD4+ T-cell count becomes lower than 200 cells per microlitre of blood, signs and symptoms of AIDS generally ensue (Hogg et al., 2001).

1.1.3 The HIV-1 replication cycle

HIV-1 is a retrovirus (i.e. a virus with an RNA genome that is converted to DNA (deoxyribonucleic acid) by its Reverse Transcriptase (RT) enzyme during replication) that was discovered over 25 years ago (Barre-Sinoussi et al., 1983) and associated with causing AIDS not long after (Gallo et al., 1984). A brief summary of the HIV-1 replication cycle follows: Firstly, the enveloped virus recognizes and binds to entry receptors on its target cells, which include T-helper lymphocytes, macrophages, dendritic cells and microglial cells of the neural system (Klatzmann et al., 1984; Macatonia et al., 1990; Watkins et al., 1990; Kalter et al., 1991). This leads to fusion of the viral envelope with the cell membrane and subsequent entry (Stein et al., 1987). Once inside the cell, the virus sheds its coat (made of capsid, CA, proteins) and forms a reverse transcription complex (RTC) to convert the two copies of its single-stranded RNA genome into double-stranded DNA by viral RT (Bukrinsky et al., 1993). The resulting pre-integration complex (PIC) is then trafficked towards the nucleus where import of the PIC occurs for subsequent integration (mediated by viral Integrase, IN) of the viral DNA into the host genome (Farnet and Haseltine 1991; Bukrinsky et al., 1992; Bukrinskaya et al., 1996).

Viral gene expression requires the initial transcription and translation of the Tat regulatory protein, which then causes transactivation (via the transactivation response, TAR, loop) for high-level long terminal repeat (LTR)-driven transcription of the viral genes encoding the accessory factors Vif (viral infectivity factor), Vpu (viral protein U), Vpr (viral protein R) and Nef (negative regulatory factor), the Env (envelope) and Gag (group-specific antigen) structural proteins and the Pol-encoded RT / ribonuclease (RNase) H
integrase (IN) and protease (PR) enzymes (see Figure 1-2) (Sodroski et al., 1985; Fisher et al., 1986). The Rev regulatory protein plays an important role in allowing partially spliced and unspliced viral transcripts to exit the nucleus for subsequent translation and processing of viral proteins (Malim et al., 1989; Kjems et al., 1991). The Gag and Env proteins are transported to the membrane where the RNA genome and other viral factors are encapsulated. As the assembled viral components bud outward from the host membrane (which has viral Env proteins embedded in it) a new virion is formed (Bugelski et al., 1995). During budding, virion maturation occurs when the PR enzyme cleaves the Pol and Gag precursor proteins into their subunits, readying the virus for infection of a new cell (Peng et al., 1989).

Figure 1-2: HIV-1 genome structure.

Depiction of HIV-1’s nine open reading frames (ORFs) — Gag, Env, Pol, Tat, Rev, Vpu, Vif, Vpr and Nef — flanked by long terminal repeats (LTRs) in the HIV-1 genome. Group-specific antigen (Gag) polyprotein is cleaved to form MA (p17 matrix protein), CA (p24 capsid protein), NC (p7 nucleocapsid protein) and p6. Pol encodes the viral enzymes protease (PR), reverse transcriptase (RT)/RNase H and integrase (IN). The envelope protein (Env) has surface (SU, gp120) and transmembrane (TM, gp41) subunits; Vif, Vpr, Vpu and Nef are viral accessory proteins while Tat and Rev are viral regulatory proteins (Robinson 2002).

1.2 Host factors required for HIV-1 replication

Apart from the viral gene functions described above, several host factors have been shown to be involved in each stage of HIV-1 infection and replication (Goff 2007). An overview of these factors is given below as a background to the novel host factors that were validated in this study. While small RNA species (microRNAs, miRNAs) of viral and
cellular origin, as well as dozens of cellular and humoral immune system factors, have been shown to be involved in regulating viral replication (Kaur and Mehra 2009; Mothe et al., 2009; Ouellet et al., 2009), these factors are not included in this overview, as they are not the focus of this study.

1.2.1 Entry

During cellular entry of HIV-1, the viral Env glycoprotein, gp120, binds to host receptor CD4 and a chemokine co-receptor, mostly CCR5 or CXCR4 (C-X-C chemokine receptor 4) (Doranz et al., 1997), which are brought into close proximity by F-actin (Iyengar et al., 1998), whose movement is regulated by moesin of the exrin-radixin-moesin protein family (Barrero-Villar et al., 2009), within cholesterol-rich lipid microdomains (Manes et al., 2000; Liao et al., 2001). Actin and myosin drives the movement of CD4-bound virus particles along filopodia from the cell periphery to the cell body where fusion occurs between the lipid bilayers of the viral envelope and the plasma membrane (Lehmann et al., 2005). The viral fusion protein (gp41) recruits a guanine nucleotide exchange factor (GEF) for the Rho GTPase, which together with CXCR4-activated cofilin (actin depolymerizing factor), acts to remodel the cortical actin barrier beneath the plasma membrane at the fusion site (Zhang et al., 1999; Yoder et al., 2008). The viral accessory protein, Nef, also appears to enable the entering viral capsid to cross this actin barrier (Campbell et al., 2004). In addition, histone deacetylase (HDAC) 6 is activated that stabilizes microtubules, possibly for anchoring the fusion pore (Westermann and Weber 2003; Valenzuela-Fernandez et al., 2005).
1.2.2 Uncoating and Reverse Transcription

Uncoating of the viral capsid after entry is mediated by cyclophilin A (CypA) and Tripartite motif-containing (TRIM) 5α. These host factors recognize and bind viral CA protein and promotes its premature disassembly ((Luban et al., 1993; Braaten et al., 1996; Sokolskaja et al., 2004; Chatterji et al., 2006; Stremlau et al., 2006). While human TRIM5α weakly affects HIV-1, other species’ orthologues have more potent effects (Stremlau et al., 2005). It appears that activation of the actin polymerization nucleator, Arp2/3 (Komano et al., 2004) is required for subsequent short-range actin-driven transport of the viral capsid to microtubule tracks where the RTC is formed (Bukrinskaya et al., 1998). Subsequently, viral nucleocapsid (NC) protein (which coats the RNA) recruits the cellular tRNA\textsuperscript{Lys} primer to the viral primer binding site for initiation of cDNA synthesis (Chan and Musier-Forsyth 1997).

Reverse transcription is downregulated by APOBEC3G/F (apolipoprotein B mRNA-editing enzyme, catalytic polypeptide-like 3G and F). This cytidine deaminase converts cytosines to uracils during reverse transcription, resulting in DNA that is either degraded by cellular uracil-DNA glycosidases and repair enzymes (Mangeat et al., 2003) or has G-A hypermutations (Lecossier et al., 2003). In addition, APOBEC3G is incorporated into virions where it hinders subsequent rounds of infection. However Vif (viral infectivity factor) counteracts this action by binding it, inducing its degradation and preventing incorporation into virions (Sheehy et al., 2002; Stopak et al., 2003).

1.2.3 Inward trafficking and nuclear import

Trafficking of the PIC to the nuclear pore appears to occur by retrograde microtubule transport, probably using dynein motors (McDonald et al., 2002). Subsequently, the PIC is actively transported across the nuclear envelope, an aspect that
enables HIV-1 (and other lentiviruses) to infect non-dividing cells where the nuclear
envelope is not fragmented during mitosis. Several host factors have been suggested as
being required for nuclear import, including importin 7, a nuclear import receptor for
ribosomal RNA and histones (Fassati et al., 2003), Nup98, a component of the nuclear
pore complex (Ebina et al., 2004), as well as transfer RNAs (tRNAs) and their precursors
(Zaitseva et al., 2006), which may act as transport carriers.

1.2.4 Integration

Integration comprises processing of the 3’ ends of the viral DNA just after reverse
transcription, followed by strand transfer in the nucleus and subsequent repair of the
recombination intermediate. Integration is mainly controlled by viral IN via its interaction
with host LEDGF (lens epithelium-derived growth factor)/p75 and BAF (barrier to auto-
integration factor). Other IN-interacting host factors include INI1 (integrase interactor 1)
(Kalpana et al., 1994) and hRAD18 (Mulder et al., 2002).

LEDGF is a nuclear transcriptional co-activator that acts as a chromatin-
associated receptor for the PIC (Emiliani et al., 2005; Llano et al., 2006b). It interacts with
IN (thereby preventing its degradation by the proteasome – (Llano et al., 2004) and directs
proviral integration to active transcription units (Cherepanov et al., 2003; Ciuffi et al.,
2005). RNAi-mediated knockdown studies, as well as knockout mutant and
overexpression studies, show an essential role for LEDGF in HIV-1 infection (Llano et al.,
2006a; Vandekerckhove et al., 2006; Shun et al., 2007).

BAF prevents intramolecular integration of the viral DNA by compacting the cDNA
and forming DNA-protein structures on the HIV-1 long terminal repeat (LTR) ends for
insertion of the provirus into the host chromosome (Lee and Craigie 1994; Chen and
Engelman 1998). LAP2α, a laminin-associated component of nuclear membrane, binds to
BAF. Both are components of the PIC and are required for productive integration
In addition, emerin (a component of the inner nuclear membrane that interacts with BAF) seems to mediate association of the PIC with chromatin thus promoting integration (Lee et al., 2001; Bengtsson and Wilson 2006).

1.2.5 Transcription

A vast network of 183 nuclear proteins that interact with the HIV-1 Tat transcriptional activator protein in vitro (i.e. in cultured cells) has recently been described (Gautier et al., 2009). This Tat interactome could be subdivided into functional modules including transcriptional activators, transcriptional repressors, replication and chromosome organization factors and nuclear structure components. Several transcription factor binding sites occur in the integrated viral DNA that modify the surrounding chromatin to control transcriptional activity by recruiting HATs (histone acetyltransferases) and HDACs. Some of the most well-known factors are members of the Sp1 (Harrich et al., 1989) and NF-κB signaling pathways (Kawakami et al., 1988).

PCAF (p300/CBP-associated factor) (Deng et al., 2000) and TRBP (TAR RNA binding protein) (Gatignol et al., 1991) recruit Tat and bind the HIV-1 TAR loop, respectively. Cyclin T1 and cyclin-dependant kinase 9, CDK9 (subunits of the positive transcriptional elongation factor b complex, P-TEFb) (Bieniasz et al., 1998) are recruited by Tat to relieve RNA Polymerase II stalling for the production of full-length viral transcripts (Yankulov and Bentley 1998). Recently, integrase interactor 1 (INI1), the key subunit of the SWI/SNF (switch/sucrose non-fermentable) chromatin remodeling complex, was found to downregulate basal LTR promoter activity and may induce one of two states post-integration: either latency or high-level, Tat-dependant gene expression (Boese et al., 2009).
1.2.6 RNA processing and export

The transcribed 9 kb viral RNA encoding the Gag and Gag-Pol precursor protein can be spliced to a 4 kb RNA encoding Vif, Vpr, Vpu and Env or further spliced to a 2 kb RNA encoding Tat, Rev and Nef (Schwartz et al., 1990; Purcell and Martin 1993). HIV-1 RNA splicing is controlled by viral exonic splicing silencers (ESS) and an intronic splicing silencer (ISS) that are bound by the cellular hnRNP (heterogeneous nuclear ribonucleoprotein) A/B family (Bilodeau et al., 2001) or by hnRNP H. In addition, viral exonic splicing enhancers (ESE) are bound by the host SR (serine-arginine) proteins (Stoltzfus and Madsen 2006). HIV-1 Tat-specific factor 1 (HTATSF1), previously defined as a Tat regulatory factor (Zhou and Sharp 1996) has recently been shown not to function in stimulation of transcription but rather in regulating the relative levels of spliced and unspliced viral transcripts (Miller et al., 2009).

The fully-spliced 2 kb RNA encoding Rev is exported to the cytoplasm for translation, after which Rev shuttles back to the nucleus and binds the Rev-responsive element (RRE) in the introns of the 9 kb and 4 kb RNAs. The nuclear export signal of Rev then allows the bound RNA to be exported via the Ran-CRM1 (exportin 1) pathway (Fridell et al., 1996), involving host factors Rev-interacting protein, hRIP (Bogerd et al., 1995), Ran binding protein 1, RanBP1 (Zolotukhin and Felber 1997) and Sam68 (Reddy et al., 1999). A Rev-independent RNA export pathway using the constitutive transport element (CTE) can also be used (Bray et al., 1994) that relies on host factors Sam68, transcytosis-associated protein (TAP) and RNA helicase A (Tang et al., 1997; Li et al., 1999; Reddy et al., 2000).
1.2.7 Translation, processing and trafficking

HIV-1 can use cap-independent translation by means of an internal ribosomal entry site (IRES) in its 5' untranslated region (UTR) (Brasey et al., 2003) and in the Gag ORF (Buck et al., 2001). This appears to be the favoured mode of translation as the viral PR inhibits a crucial host factor for cap-dependant translation, eukaryotic translation initiation factor 4GI, eIF4GI (Ventoso et al., 2001). Finally, for Gag-Pol translation to occur, an RNA pseudoknot forms that allows for read-through of the Gag stop codon by inducing a translational frameshift (Kang 1998).

The viral structural protein precursors, Gag and GagPol are translated in the cytoplasm on soluble polysomes and then directly transported to the membrane where multimerization (promoted by host ATP-binding cassette, sub-family E (OABP), member 1 - HP68/ABCE1 (Zimmerman et al., 2002)) causes a conformational change that increases its affinity for the membrane (Lindwasser and Resh 2001; Lingappa et al., 2006). Membrane trafficking of Gag is inhibited by TRIM22 that is induced by the Type I Interferon response (Barr et al., 2008). In addition, Gag appears to bud into intracellular vesicles for release from the cell by budding exosomes (Gould et al., 2003; Perlman and Resh 2006). This pathway is especially important in macrophages (Nguyen et al., 2003, Pelchen-Matthews et al., 2003) and requires the adaptor protein AP-3 for sorting into endosomes (Dong et al., 2005). Recently, the role of dynein motor complexes and late endosomes in trafficking the viral RNA genome and Gag for virion production has been shown (Lehmann et al., 2009).

Although Env and Gag trafficking to the membrane can occur independently, when co-expressed they affect each other's localization (Lodge et al., 1997; Bhattacharya et al., 2006). The Env protein is trafficked via the Golgi body after translation by ribosomes bound to the endoplasmic reticulum (ER). During transit, it forms a trimer, is processed
into its subunits (gp120 and gp41) by cellular endoproteases and is modified by
glycosylation (Earl et al., 1991).

1.2.8 RNA encapsulation and assembly

Virion assembly and release is driven by the viral NC protein that stabilizes and
encapsulates the RNA genome during virion formation. Incorporation of the Env proteins
and assembly at the membrane occurs on cholesterol-enriched microdomains or lipid rafts
(Ono and Freed 2001) and requires the 47-kDa tail interacting protein, TIP47 (Blot et al.,
2003). As the entry receptor CD4 inhibits Env incorporation, it is down-modulated by viral
Nef, Vpu and Env soon after entry (Chen et al., 1996). Several host proteins are also
incorporated into the virion: These include cyclophilin A, CypA (Franke et al., 1994),
APOBEC3G (Mariani et al., 2003), tumour susceptibility gene 101, TSG101 (Garrus et al.,
2001) and vacuolar protein sorting 28, Vps28 (Hammarstedt and Garoff 2004), immune
molecules HLA-DR (human leukocyte antigen DR), β2-microglobulin and MHC (major
histocompatibility complex) class I proteins (Orentas and Hildreth 1993; Tremblay et al.,
1998), as well as heat shock protein 70, Hsp70 (Gurer et al., 2002), INI1/SNF5 (Yung et
al., 2001) and lysyl-tRNA synthetase (Kleiman et al., 2004). These factors are either
passively incorporated (by being embedded in the part of the plasma membrane that buds
off) or included due to functional interactions with various viral products.

1.2.9 Virion budding

Budding is mediated by the endosomal sorting complex required for transport
(ESCRT) components I, II and III (Usami et al., 2009). Briefly, this machinery is recruited
to the plasma membrane by Gag, which is ubiquitylated at low frequency (Gottwein and
Krausslich 2005), via its associations with AIP1 (atrophin-interacting protein 1) (Strack et
and TSG101, a component of the Class E vacuolar protein sorting (VPS) machinery (Garrus et al., 2001; VerPlank et al., 2001). Several other VPS proteins, as well as the charged multivesicular body (MVB) proteins, CHMPs, that form the ESCRT III complex are involved in mediating efficient viral budding (Zamborlini et al., 2006).

Other host factors involved in the process include Gag-interacting annexin II (Ryzhova et al., 2006) and endophilins that interact with the ESCRT machinery and induce membrane curvature for virion budding (Farsad et al., 2001). Tetherin (BST-2) has recently been shown to inhibit the spreading and release of budding viral particles. One of the membrane-spanning domains of this protein is embedded in the viral envelope while the other is anchored in the host cell membrane, effectively “tethering” the virions to the host cell. This action is countered by viral Vpu that causes proteasomal degradation of tetherin, as well as CD4 (Bour et al., 1999; Levesque et al., 2003), via the β-TrCP2-dependent pathway (Neil et al., 2008; Douglas et al., 2009; Mangeat et al., 2009).

1.2.10 Virus transfer

Mature virus passes between cells mainly by the formation of virological synapses (Jolly and Sattentau 2004) that are induced by Env-CD4-coreceptor interactions and Env-induced activation of lymphocyte function-associated antigen 1, LFA1 (Arthos et al., 2008). Other requirements include actin and tubulin cytoskeletal rearrangements, polarization of the microtubule organizing centre (MTOC) and junction stabilization by cellular adhesion molecules (Chen et al., 2007; Jolly et al., 2007). These can form between T-cells (Jolly et al., 2004) and T-cells and dendritic cells (Stoll et al., 2002; McDonald et al., 2003).

Viral transfer between HIV-1 CD4+ T-cells is promoted by trafficking of Env and Gag to tetraspanin-enriched membrane domains for assembly and budding (Jolly and Sattentau 2007). Dendritic cells can take up virus particles by the C-type lectin, DC-SIGN
(dendritic cell-specific ICAM3 (intercellular adhesion molecule 3)-grabbing non-integrin) (Geijtenbeek et al., 2000) that internalizes the virus into endocytic vesicles without becoming infected. When contact is made with a CD4+ T-cell, the virus is passed to the T-cell across a virological synapse mimicking the process of antigen presentation across an immunological synapse (McDonald et al., 2003). Polysynapses have also been described whereby virus is passed from one infected T-cell to multiple target cells (Rudnicka et al., 2009). Other modes of virus transport across cell-to-cell contacts use tunneling nanotubes (Sherer and Mothes 2008; Sowinski et al., 2008) and filopodial bridges (Sherer et al., 2007) as physical cellular extensions for virus transfer.

1.3 Current approaches to managing HIV-1/AIDS

The last 25 years have seen the development of several classes of ARV therapies that target various steps of the HIV-1 lifecycle. These include eight nucleoside/nucleotide analogues that inhibit viral RT activity [reviewed by (Cihlar and Ray 2010)], four non-nucleoside RT inhibitors [reviewed by (de Bethune 2010)] and nine approved PR inhibitors [reviewed by (Wensing et al., 2010)]. In addition, inhibitors of viral entry and fusion (by maraviroc and enfuvirtide, respectively) [reviewed by (Tilton and Doms 2010)] and one targeting integration (raltegravir) have recently been approved, while another IN strand-transfer inhibitor (called elvitegravir) [reviewed by (McColl and Chen 2010)] is currently undergoing clinical trials. ARVs are the only available treatment options at present. These drugs effectively suppress HIV-1 replication when their regimen is adhered to consistently. However, adverse reactions in some patients, as well as poor access to the drugs in developing countries, often results in non-compliance to treatment (Deeks 2006).

HIV-1’s relatively short lifecycle (1.5 days) and low-fidelity RT enzyme contribute to an uncommonly high viral mutation rate of $\sim 2 \times 10^{-5}$ mutations per base per replication
cycle (Huang and Wooley 2005). This leads to rapid evolution of viral escape mutations during single-drug therapy resistance to all known ARVs. In response, a highly active antiretroviral therapy (HAART) treatment programme was developed, in which two or more drugs (usually from different classes) are combined. HAART has been very successful, particularly in developed countries, at reducing the morbidity and mortality of HIV/AIDS (Palella et al., 1998). The extensive development of ARVs has provided both first-line therapies (for previously untreated, ARV-naïve, patients) and second-line options for those that develop resistance or do not respond to initial therapy, while new drugs continue to be tested for efficacy from each ARV class (McColl and Chen 2010). However, as HAART cannot prevent viral resistance developing (particularly in non-compliant patients) nor eradicate latent reservoirs of transcriptionally silent integrated provirus (Pierson et al., 2000) novel approaches are constantly being sought.

While the United States Food and Drug Administration (US FDA) have yet to approve a single vaccine for clinical use (AIDSinfo 2008), there are currently around 30 vaccine candidates being tested for clinical efficacy (IAVI 2010). In November 2007, the field was dealt a major blow when advanced clinical trials (Human Phase 2b) of the Merck vaccine (based on Ad5 – Adenovirus Type 5) were halted. The vaccine could neither prevent infection nor lower viral load in infected patients, while potentially increasing the risk of infection (Watkins et al., 2008). As a result, the focus has shifted to basic research of alternative strategies, including gene therapy. While catalytic ribozymes, antisense molecules, aptamers and RNA decoys have all been used to good effect to target HIV-1 [reviewed by (Sherer et al., 2007; Reyes-Darias et al., 2008)], the most exciting advances have been in using the molecular tool RNA interference (RNAi) for the specific silencing (termed knockdown) of any gene of choice.
1.4 RNA interference (RNAi)

1.4.1 The endogenous RNAi pathway

The Nobel prize-winning discovery of gene silencing in nematode worms mediated by double-stranded RNA, termed RNAi (Fire et al., 1998), was followed by the finding that 21-base pair (bp) small RNA duplexes (termed small interfering RNAs, siRNAs) could direct sequence-specific inhibition of mammalian genes with complementary sequences (Elbashir et al., 2001). Subsequently, this inhibition was found to occur via the same pathway that generates 21-23 nt microRNAs (miRNAs), of which more than 700 have now been cloned, according to miRBase (Griffiths-Jones et al., 2008). These sequences have diverse functions in development and cellular physiology [reviewed by (Bala et al., 2009; Coolen and Bally-Cuif 2009; Iorio and Croce 2009; Liang and Qin 2009; Yang et al., 2009)].

Both siRNAs and miRNAs are processed by RNase III-like endonucleolytic complexes in the multi-step RNAi pathway (Figure 1-3). Genome-encoded long RNA Pol II-derived primary miRNA (pri-miRNA) transcripts fold into an imperfect hairpin structure that is cleaved by the Micro-processor complex comprising of the Drosha nuclease and its cofactor, DGCR8 (DiGeorge critical region-8) (Lee et al., 2003; Gregory et al., 2004; Han et al., 2004; Lee et al., 2004). The mechanism by which Drosha “measures” the duplex before cleavage is still under debate (Han et al., 2006; Zeng and Cullen 2006), but this process has recently been shown to occur during transcription in association with 5'-3' and 3'-5' exonucleases that degrade intronic sequences before splicing occurs (Morlando et al., 2008). In addition, mirtrons (a subclass of miRNAs within intronic sequences) are not processed by Drosha at all but are rather generated during the splicing reaction (Okamura et al., 2007; Ruby et al., 2007).
Figure 1-3: The mammalian RNAi pathway.

Long primary microRNA (pri-miRNA) transcripts form hairpin structures that are recognised and cleaved by the nuclease Drosha and its cofactor DGCR8. The resulting pre-miRNAs are exported from the nucleus and subsequently processed by Dicer/TRBP (HIV TAR RNA binding protein) into ~22-nt mature miRNAs with 5' phosphates and 2 nucleotide overhangs on the 3' end. The strand with the thermodynamically most unstable 5' end is incorporated into an RNA-induced silencing complex (RISC) as a guide strand, leading the complex to target mRNAs of complementary sequence for cleavage by Ago2 (in the case of perfect complementarity) or translational suppression (imperfect binding).
The ~70 nucleotide (nt) pre-miRNAs generated by Drosha processing are exported from the nucleus by G protein-coupled Exportin 5 (Yi et al., 2003; Lund et al., 2004). In the cytoplasm, these pre-miRNAs are further processed by Dicer and its cofactors TRBP, a factor that provides an interesting link between HIV-1 and RNAi, which may be used by the virus to subvert the RNAi machinery and prevent downregulation of viral gene expression (Bennasser et al., 2006), and the protein, PACT (Kok et al., 2007). Dicer “measures” 21-22-nt from the Drosha cleavage site to create mature miRNAs with 5’ phosphates and 3’ 2-nt overhanging ends (Bernstein et al., 2001; Hutvagner et al., 2001). Double-stranded RNA is similarly processed into 21-25 nt siRNAs. Subsequently, the duplex is unwound (by an unknown helicase) and one strand (the mature miRNA or guide strand) is incorporated into the RNA-induced silencing complex, RISC (Gregory et al., 2005), selected based on the relative thermodynamic stability of the duplex ends (Khvorova et al., 2003; Schwarz et al., 2003).

Perfect complementarity with the guide sequence results in target mRNA cleavage by the Argonaute (Ago) 2 component of the RISC (Chendrimada et al., 2005), while imperfect binding mediates translational suppression and mRNA destabilization [reviewed by (Ma et al., 2009)]. The most important sequence in the latter mechanism is the “seed” or “nucleus”, a seven base-pair region at the 5’ end of the guide strand extending from nucleotides 2 to 8 (Doench and Sharp 2004). miRNA seed region matching allows a single miRNA to potentially target hundreds of genes, usually in their 3’ untranslated regions (Grimson et al., 2007). When these genes are regulators of transcription or alternative splicing themselves, the compound effect of the miRNA on gene regulation can become complex [reviewed by (Makeyev and Maniatis 2008)]. The exact mechanism of miRNA translational suppression is incompletely understood but seems to include deadenylation of poly-A tails and decapping (or inhibiting 5’ cap recognition) of mRNA that leads to destabilisation and exonuclease degradation, respectively, as well as inhibition of the translation initiation machinery [reviewed by (Ma et al., 2009)].
The mechanisms described above mediate post-transcriptional gene silencing by the RNAi pathway. Recently, a novel transcriptional gene silencing process in the nucleus has been described, whereby promoter-targeted siRNAs induce chromatin modifications associated with transcriptional silencing, including histone 3 lysine 9 (H3K9) dimethylation and H3K27 trimethylation (Mette et al., 2000; Morris et al., 2004; Strunnikova et al., 2005; Weinberg et al., 2006). Although the exact mechanisms are still unclear, Ago1, DNA methyltransferase 3a (DNMT3a) and HDAC1 appear to initiate siRNA-mediated transcriptional gene silencing, TGS, while DNMT1 is necessary for maintaining the silencing effect (Hawkins et al., 2009). In addition, Ago1 and the Polycomb group component, EZH2, were associated with miRNA-mediated TGS of the polymerase (RNA) III (DNA directed (POLR3D) cell cycle gene promoter [reviewed by (Kim et al., 2008)]. There is some controversy as to whether TGS targets the DNA or associated RNA of the promoter (Verdel et al., 2004; Han et al., 2007). Finally, as the silencing mediated by TGS is stabler than PTGS and also heritable, the prospect of its application in gene therapies is great. Recently, prolonged TGS of a sequence within the HIV-1 LTR promoter has been reported in a T-cell line model using a retrovirally delivered shRNA (Turner et al., 2009).

1.4.2 RNAi as a reverse genetics tool

The RNAi pathway can be harnessed for use as a molecular tool by exogenously introducing synthetic siRNAs or vectors that express short hairpin RNAs (shRNAs) that enter the miRNA pathway post- and pre-Dicer processing, respectively. These shRNAs/siRNAs can be engineered for targeting specific genes for knockdown, which has rapidly become the most popular reverse genetic approach to studying gene function. Some of the organisms in which this approach has recently been used include: mammals (in vitro) (Chen et al., 2010; Shahbazian et al., 2010; Zalckvar et al., 2010), invertebrates (Posnien et al., 2009; Rehm et al., 2009; Yang et al., 2010), yeast (Hoffmann et al., 2009;
Bolz et al., 2010), plants (Petsch et al., 2009; Zhang et al., 2009), protoplasts (Zhai et al., 2009), parasites (Stephens et al., 2007; de la Fuente et al., 2010; Verner et al., 2010) and viruses (Heikkila et al., 2010; Hirsch 2010; Karlas et al., 2010). The advantage of RNAi over previous approaches like mutagenesis screens and mouse knock-out models is that it is far simpler, quicker and less costly to implement (even at a genome-wide scale) while providing potent gene inhibition to study loss-of-function effects.

1.4.3 RNAi as an exciting and powerful new gene therapy

Apart from functional studies, novel gene therapy strategies based on RNAi have been developed and even patented for a wide range of diseases [reviewed by (Grinberg 2008)]. These include neurodegenerative diseases [reviewed by (Farah 2007)], cancers [reviewed by (Ashihara et al., 2009)], respiratory disorders like asthma [reviewed by (Huang and Chiang 2009)], cardiovascular disease (Hu et al., 2009; Qiu et al., 2009) and infectious diseases (Yokota et al., 2007; Gao et al., 2008; Lambeth et al., 2009). Because of the ease with which RNAi effectors can be designed (using bioinformatic approaches) to target any gene, it is now possible to develop therapeutics for targets that have before now been declared “undruggable”, as well as for diseases that are rare or newly emerging. Moreover, diseases that manifest early or are heritable can be addressed by the advent of delivery of RNAi to human embryonic stem cells (HESCs) [reviewed by (Rassouli and Matin 2009)]. These cells are not only pluripotent (i.e. differentiate into every cell type during development) but also self-renewable (Thomson et al., 1998), making RNAi-modified HESCs an attractive prospect for treating degenerative disorders.
1.5 RNAi approaches targeting HIV-1

There have been numerous studies using RNAi to silence HIV-1 gene expression, including the *gag*, *pol*, *vif*, *tat*, *rev*, *env* and *nef* genes and the LTR promoter region (Coburn and Cullen 2002; Hu *et al.*, 2002; Jacque *et al.*, 2002; Park *et al.*, 2003; Boden *et al.*, 2004; Das *et al.*, 2004). However, the high viral mutation rate makes it difficult to develop antiviral RNAi therapies with long-lasting efficacy. As with antiretroviral drug resistance, escape mutants arise rapidly to evade RNAi-mediated targeting, via single base changes within the target region or by evolving secondary structures that render the target region inaccessible (Boden *et al.*, 2003; Das *et al.*, 2004; Westerhout *et al.*, 2005). In response, researchers have combined shRNAs targeting several highly conserved viral regions simultaneously (ter Brake *et al.*, 2006), while multiple siRNAs targeting *gag*, *pol*, *int* and *vpu* conserved sequences have been shown to effectively inhibit various strains of HIV-1, as well as viral replication in a chronically infected cell line and in primary human peripheral blood lymphocytes (Chang *et al.*, 2005). Finally, a combination of three potent shRNA sequences that target highly conserved HIV-1 sequences was recently shown to inhibit replication of a HIV-1 molecular clone in shRNA vector-transduced T-cells without viral escape for more than three months (von Eije *et al.*, 2009).

The relative advantages and disadvantages of using siRNAs versus shRNA expression vectors in this setting are well reviewed (Subramanya *et al.*, 2010). The main issues raised are that siRNAs mediate transient knockdown (a few days), are not effectively delivered to T-cell or macrophage primary cells and cannot be incorporated into combinatorial therapies, while shRNAs are more time-, cost- and labour-intensive to produce and when stably expressed from lentiviral (HIV-derived) vectors pose safety concerns.
1.5.1 Viral versus cellular targets

An alternative strategy to prevent viral escape from RNAi is to target cellular genes that HIV-1 requires for efficient replication. While HIV-1 can evade RNAi effectors targeting its own genome with relative ease, targeting essential HIV-1 host factors would require the virus to evolve a novel or alternate method of exploiting the host cellular pathways. Another advantage to this approach is that the problem of eradicating latent integrated provirus may be overcome by targeting host factors involved in pre-integration stages of the replication cycle. Disadvantages of targeting host factors include HIV-1’s ability to exploit multiple cellular pathways in parallel. This redundancy that has evolved between host and pathogen creates networks of interactions that are difficult to define and interrupt via therapeutic means. Targeting host genes may also have detrimental side-effects on normal cellular function: while knockdown of both CD4 and CXCR4 have been shown to inhibit HIV-1 replication (Martinez et al., 2002; Novina et al., 2002; Park et al., 2002), CD4 is essential to immune function while CXCR4 is important for hematopoietic stem cell function and differentiation, thus limiting these factors’ value as good therapeutic targets (Moore 1997; Lapidot and Kollet 2002). In contrast, the CCR5 entry co-receptor poses an attractive target for several reasons: 1) it is not essential for immune function, 2) individuals with a 32 base-pair genetic deletion (CCR5Δ32) are healthy and 3) these individuals either resist HIV-1 infection (if homozygous for the deletion) or progress much slower to AIDS (heterozygous carriers) (Samson et al., 1996; Eugen-Olsen et al., 1997; Garred et al., 1997). Indeed, a recent report shows that transplantation of CCR5Δ32 stem cells to an HIV-1 infected patient inhibited the virus to undetectable levels for 20 months after the transplant, without HAART (Hutter et al., 2009). The only apparent drawback of targeting this factor is that HIV-1 switches tropism during infection from using CCR5 as coreceptor to CXCR4 (Arien et al., 2006). In addition, a recent study in South Africa has shown a concerning increase in primary HIV-1 strains that use CXCR4 as entry co-receptor (Connell et al., 2008). However, by targeting conserved viral sequences in
addition to such cellular factors a more potent and durable therapy is made possible. As more host factors are identified and characterised, a greater diversity of such therapies can be developed.

1.6 Genome-wide siRNA screens for novel HIV-1 host factors

Before the development of genome-wide libraries of RNAi effectors approximately 50 host factors were known to be involved in regulating HIV-1 replication [reviewed by (Goff 2007)]. Since then, high-throughput technology and robotic systems have developed that have been used in several genome-wide siRNA screens in the last two years. These screens have collectively implicated several hundred novel host factors that appear to be involved in HIV-1 replication, effectively leap-frogging our understanding of this process.

The first published study (Brass et al., 2008) used a genomic siRNA screen to determine which of ~21,000 human genes HIV-1 needs to replicate. The authors discovered 273 host genes that could be classified by their criteria as HIV-1 dependency factors (HDFs), of which 237 were novel (see Figure 1-4, pg 27). Two independent research teams (Konig et al., 2008; Zhou et al., 2008) closely followed with reported discoveries of 213 and 267 novel HDFs respectively.

A meta-analysis comparing these three datasets, as well as several other studies’ datasets (Chertova et al., 2006; Hao et al., 2008; Krishnan et al., 2008; Studamire and Goff 2008) showed limited overlaps from individual pairwise comparisons (Bushman et al., 2009). However, network analysis of the novel genes recovered formed several densely connected cellular-viral gene clusters representing cellular subsystems involved in HIV-1 replication. These include the proteasome, various subunits and factors of RNA Polymerase II, the Mediator complex, various transcription factors required for Tat activation and elongation of transcription, RNA binding and splicing proteins, protein
chaperones involved in protein folding and sorting, Tat activation and budding and finally factors involved in tRNA synthase functioning and transport (Bushman et al., 2009).

The small overlap between these datasets mostly reflects fundamental differences in study design. All three groups used cell lines engineered to allow artificial infection by HIV-1. The pattern of gene expression in these cells differs greatly from CD4+ T-cells and macrophages (HIV-1 natural host cell types). This implies that gene expression products in these cell lines may allow interactions with viral proteins that do not represent true host-virus interactions occurring during a natural infection. In addition, viruses used were either lab-adapted strains or molecular clones missing several important viral accessory proteins (including Nef, Vpu and Vpr) or pseudotyped with the envelope protein of Vesicular Stomatitis Virus (VSV-G) that uses a different mode of entry to HIV-1. Screening for host factors using such virus models could thus mask many important viral/host interactions.

1.6.1 Selection of novel host factors for validation in this study

To address the need to validate these candidate factors, five of the novel HDFs implicated by Brass et al. (2008) as being required for HIV-1 infection were selected: SPTBN1 (spectrin beta non-erythrocytic 1), TMED2 (transmembrane emp24 domain trafficking protein 2), KIAA1012 (a gene of relatively unknown function), PRDM14 (PR domain containing 14) and SP110 (SP110 nuclear body protein), using unbiased criteria as described in Section 3.1. According to Brass et al., SPTBN1, TMED2 and KIAA1012 are involved in the early phase (pre-Gag translation) of the HIV-1 lifecycle, while PRDM14 and SP110 are required for later stages involving assembly and budding of virions (see Figure 1-4, pg 27).
**1.6.1.1 SPTBN1**

SPTBN1 is one member of a family of beta-spectrin genes. Spectrin is an actin-crosslinking and molecular scaffold protein and functions in determining cell shape, arranging transmembrane proteins and organelle organization (Taylor-Harris et al., 2005). In addition, it appears to play a role in regulating T-lymphocyte function via ankyrin, which binds directly to CD45, sequestering it from being recruited to the cell surface and abrogating Jurkat T-cell activation (Pradhan and Morrow 2002). Finally, this spectrin is directly cleaved by HIV-1 protease, along with other focal adhesion plaque proteins. Brass et al. hypothesize that SPTBN1 may be involved in inward trafficking of the virus via its interaction with the actin network.

**1.6.1.2 TMED2**

TMED2 (p24A) is part of the p24 cargo receptor family. This family has been shown to form complexes with GRASP55 and -65, peripheral membrane components of the Golgi matrix, which sequesters the p24 cargo receptors in the Golgi apparatus. In addition, p24A interacts with both ARF1 and ARFGAP (Majoul et al., 2001), with studies implicating it as an integral receptor for the coatomer protein 1 (COP1) vesicle coat, cycling continuously between the intermediate compartments of the Golgi and the cis-Golgi network (Blum et al., 1999). Recently, p24A has had a novel function described as a regulator of signal-dependent trafficking during the life cycle of PAR-2 (Protease-activated receptor 2) (Luo et al., 2007). Thus this gene may play a role in retrograde Golgi-to-ER trafficking of virus.
1.6.1.3 **KIAA1012**

KIAA1012 is highly expressed in immune cells, both B- and T-lymphocytes, as well as NK (natural killer) cells. With three transcripts expressing five proteins, its function is still uncharacterised. Its homologue, TRS85, is a component of the transport protein particle complex (TRAPP) (Sacher *et al.*, 2000) and is involved in selective autophagy via the Cvt (cytoplasm to vacuole targeting) pathway, for the biogenesis of double membrane-layered Cvt vesicles (Meiling-Wesse *et al.*, 2005). A putative involvement in viral transport via these vesicles has still to be explored.

1.6.1.4 **PRDM14**

PRDM14, apart from containing a PR-domain, a subtype of the SET domain that catalyzes histone methylation, also has several zinc finger conserved domains. Recently, upregulation of this gene was found in two-thirds of all breast cancers, while an effect of reducing cancer cell sensitivity to chemotherapeutics was also described (Nishikawa *et al.*, 2007). Interestingly, PRDM14 is specifically expressed in undifferentiated human ES (embryonic stem) cells (Tsuneyoshi *et al.*, 2008). Taken together, these findings indicate a potential role in transcriptional regulation of viral gene expression.

1.6.1.5 **SP110**

SP110 acts as a nuclear hormone receptor transcriptional co-activator and is associated with the nuclear body complex. This gene is induced by the Interferon response (Regad and Chelbi-Alix 2001), offering immunoprotection against pathogens (Everett and Chelbi-Alix 2007). In contrast, it has also been found to be recruited by several pathogens, specifically Epstein Barr Virus (Nicewonger *et al.*, 2004) and *Anaplasma phagocytophilium* (de la Fuente *et al.*, 2007) where it modulates retinoid signalling and stabilises pathogenic gene expression. Thus a potential role in regulation of viral gene expression and/or the Interferon response is intriguing.
The functions and molecular interactions of these genes, as well as the findings published by Brass et al. (2008), support their putative role as novel HDFs. However, a robust, individual approach is required for careful validation of these genes’ role in HIV-1 infection and replication. Validating each of the several hundred candidate HDFs that have been implicated recently will vastly accelerate our understanding of the way in which complex retroviruses like HIV-1 adapt to use their cellular environment for efficient replication. In addition, such HDFs will potentially provide a large array of novel targets for conventional drug-targeting and newer gene therapy approaches, thereby making an effective anti-HIV treatment a more feasible goal.

Figure 1-4: Model of putative HDF roles in the HIV lifecycle.

The function and sub-cellular location of HDFs was determined using multiple databases (Table S4 of Brass et al., 2008). Proteins in multiple locations represent more than one possible role in the HIV lifecycle. The HDFs selected for this study are circled in red. Newly identified HDFs (red or blue, the latter late-acting); previously implicated HDFs detected (green), or not detected but with a relevant interaction (grey); HIV protein (black): matrix, MA, reverse transcriptase, RT, integrase, IN, envelope (gp41, gp120), ENV. Unfolded protein response, UPR. (Brass et al., 2008)
1.7 Aims and Objectives

The overall aim of this project was to validate the role that *SPTBN1*, *TMED2*, *KIAA1012*, *PRDM14* and *SP110* play in HIV-1 infection and replication. The specific objectives were to:

1) Use expressed shRNAs and siRNAs to silence the expression of these genes by the RNA interference (RNAi) pathway.

2) Quantify the ability of these RNAi effecters to suppress their target sequences in (i) a minimal target/reporter system and (ii) at the endogenous mRNA level using a dual luciferase reporter assay and quantitative RT-polymerase chain reaction (qRT-PCR) assays, respectively. The latter would require optimisation and validation to ensure that the results are accurate and reproducible.

3) Determine if gene loss-of-function, as well as the RNAi process or cellular delivery method, had any detrimental effects on cell viability using a colorimetric assay for metabolic enzyme activity.

4) Optimise delivery of DNA- and RNA-based fluorescent reporters to a T-cell line using electroporation and lipid-based transfection reagents.

5) Determine the effect of each gene's knockdown on HIV-1 replication kinetics in (i) a single infection cycle pseudoviral challenge and (ii) a two-part challenge with fully-replicating virus.
CHAPTER 2: METHODS AND MATERIALS

2.1 RNAi effecters

2.1.1 shRNA-encoding sequences

The original set of shRNAs was designed based on the sequences available in the Broad Institute/The RNAi Consortium (TRC) MISSION™ shRNA library (Moffat et al., 2006; TRC 2008). This library consists of shRNA sequences designed to specifically target the transcripts of a large number of NCBI (National Centre for Biotechnology Information) human and mouse genes and can be accessed online at the following URL: http://www.broadinstitute.org/rnai/trc/lib. The 21-mer candidate shRNA sequences (typically five per gene) span the length of the coding region (25 bp after the initiation of transcription until 150 bp before the end of the transcript) and usually include one targeting the 3’ untranslated region. They are sorted according to a score that is based on siRNA sequence criteria (Elbashir et al., 2001; Reynolds et al., 2004). These include strong base-pairing at the 5’ end, weak base-pairing at the 3’ end, no internal loop formations, a preferred GC-content of 60-70%, no single nucleotide polymorphisms (SNPs) or ambiguous bases, relatively AT-rich sequences from positions 6-11 and avoidance of seven consecutive GC bases or four of the same bases in a row. After sorting the candidates are further selected to avoid overlaps with each other and matches with other human gene sequences.

The TRC 21-nucleotide sense and antisense sequences were joined with a nine-nucleotide loop sequence (5’-CCT GAC CCA-3’) (see Figure 2-1 for an example). A guanine residue (G) at the 5’ end (the same first nucleotide as in the native U6 small nuclear RNA) and six T-residues at the 3’ end ensured efficient initiation and termination
of transcription respectively. Where three or more T-residues occurred in the sense (anti-guide) strand, mismatches were introduced to avoid premature termination of transcription. If three or more T-residues occurred in the antisense (guide) sequence that is important for target sequence recognition, a different shRNA sequence was designed using the HPCDispatcher siRNA site selector application (Heale et al., 2005). This is an online application available at the following universal resource locator (URL): http://www1.infosci.coh.org/hpcdispatcher/siRNA.aspx (HPCDispatcher) that calculates the difference in thermodynamic stability between the ends of a siRNA duplex. Duplexes with the greatest thermodynamic difference were chosen while incorporating the same design rules and modifications described above.

**Figure 2-1: Example of shRNA design approach.**

The top panel shows the duplex sequence from the online TRC MISSION shRNA library. The bottom panel shows 21 base-pairs of the duplex connected by a 9-nucleotide loop, including single-stranded end modifications and mismatches induced (bold) to remove triple uridine residues in the sense strand (anti-guide).

### 2.1.2 Specificity and structure of shRNA sequences

The specificity of all shRNA sequences was assessed using NCBI's online BLAST (basic local alignment and search tool) (Altschul et al., 1990; BLASTn 1990) to detect off-
target matches with other gene sequences. This search tool can be found online at the following URL: http://blast.ncbi.nlm.nih.gov/Blast. The shRNA sequence was compared with the human genomic and transcript database using the megablast program of BLASTn (for nucleotide sequences) that optimises the search for highly similar sequences. In addition, the secondary structure of the shRNA sequence was assessed using the RNA mfold v 3.2 server (Zuker 1995-2010; Mathews et al., 1999; Zuker 2003) using the default parameter settings. This program is available at the following URL: http://mfold.bioinfo.rpi.edu/cgi-bin/rna-form1.cgi.

2.1.3 siRNA SMARTpools

siRNA ON-TARGETplus™ SMARTpools were ordered from Dharmacon® (Thermo Fisher Scientific, Lafayette, CO, USA) targeting SP110 (target transcript NM_004510, catalogue number L-011875-00-0005) and SPTBN1 (target transcript NM_178313, catalogue number L-018149-01-0005). As a non-targeting control, a siRNA targeting the HBx sequence of Hepatitis B Virus, HBV (designed by Prof Piet Herdewijn of the Rega Institute for Medical Research Laboratory of Medicinal Chemistry / Katholieke Universiteit Leuven in Leuwen, Belgium and Prof Joachim Engels of Wolfgang Goethe University in Frankfurt, Germany) was used that comprised a Cy3 fluorescent dye-labelled antisense strand and a sense strand that were annealed together before use (see Table A-8 of Appendix A). All siRNAs were resuspended in buffer (60 mM KCl (Burani Design Holding, BDH, Chemicals, Poole, England), 6 mM HEPES (pH7.5), 0.2 mM MgCl₂ (Saarchem Merck Chemicals, Gauteng, South Africa)) and stored as 20 µM aliquots at -70°C.

Subsequently, new shRNAs were designed based on the individual siRNA target sequences of each SMARTpool. The same nine-nucleotide loop sequence was used to join the 19 bp sense/antisense (antiguide/guide) duplex, with six terminating T-residues at
the 3’ end and a G-residue as the first nucleotide to be transcribed. Three or more T-residues were again avoided by introducing mismatches (if occurring in the sense strand). Where these occurred within the antisense (guide) strand the sequence was synthesized so that ~50% of the oligonucleotides would have a T-residue at that point and ~50% would have a C-residue (as both these nucleotides have pyrimidine bases, they have similar ability to bind the sense strand). Thus a single clone produced from these oligonucleotides would have either a T- or a C-residue. The genomic locations targeted by the shRNAs and siRNAs described above are listed in Tables A-1 and A-2, respectively, of Appendix A.

### 2.2 Previously constructed plasmids used in this study

Several plasmids were used routinely in cloning or in transfections that were not created during this study: pTZ57R (provided with the PureExtreme® InstAclone™ PCR cloning kit from MBI Fermentas, MD, USA) was used in T/A cloning - this method exploits the fact that Taq polymerase adds an A-residue to the 3’ end of amplicons (Clark 1988; Hu 1993). PCR products can easily be inserted into the linearised vector, pTZ57R/T, by ligating to the flanking T-residue of the T/A cloning site (Holton and Graham 1991; Marchuk et al., 1991; Mead et al., 1991).

psiCHECK-2.2 was created by Dr M S Weinberg of our laboratory: an oligonucleotide encoding several restriction sites was inserted into psiCHECK-2™ (Promega, WI, USA), thereby creating a new multiple cloning site. This vector encodes both Renilla luciferase and Firefly luciferase and is useful for rapidly assessing RNAi knockdown efficacy.

pCI-EGFP is a useful reporter plasmid expressing enhanced Green Fluorescent Protein (eGFP) from the Cytomegalovirus (CMV) immediate early enhancer/promoter. It
was constructed by cloning the EGFP ORF into the mammalian expression vector pCI Neo (Promega, WI, USA) (Passman et al., 2000).

phRL-CMV (Promega, WI, USA) is a reporter plasmid that expresses Renilla luciferase from a CMV promoter. This plasmid was used to control for transfection efficiency and variation in cell density during experiments in TZM-bl cells that stably express integrated Firefly luciferase (Section 2.4).

All shRNAs used and created in this study are expressed from the pTZU6+1 vector (kindly provided by Dr David Engelke of the University of Michigan). To control for background activity from the U6 promoter, experiments always included this empty vector as a Mock control. To control for non-specific knockdown effects, several plasmids expressing shRNAs targeted to other sequences were tested for use in knockdown experiments. Two shRNAs targeted to HBV, U6 shRNA 5 and U6 shRNA 10 (Carmona et al., 2006), renamed shHBx-5 and shHBx-10 here, as well as shLacZ (constructed in our laboratory to target the lacZ gene of E. coli) were tested for their ability to knockdown each of the genes targeted in this study.

Two shRNAs that have been shown previously to knock down HIV-1 were included as positive controls in virus knockdown experiments: shLTR was constructed in our laboratory by S Barichievy and targets the HIV-1 long terminal repeat (LTR) promoter. U6+1 tat/rev shRNA (Scherer et al., 2004) was renamed shTat in this study and targets the viral tat/rev ORF. Two shRNAs targeting confirmed HIV-1 host factors were used: shPL3 (renamed shLEDGF in this study) was created by V Green of our laboratory to target LEDGF (lens epithelium-derived growth factor) (Section 1.2.4); shHTATSF-2 (also created by V Green) targets HTATSF1 (Section 1.2.6). The sequences of these shRNA controls are given in Table A-8 of Appendix A.
2.3 Molecular cloning

2.3.1 Cloning of shRNA expression cassettes

Sequences encoding the first set of shRNAs were designed as two overlapping oligonucleotides. They were synthesized in reverse primer orientation to allow for two-step PCR-amplification of the U6 promoter template, similar to the approach used by Castanotto et al. (2002). Figure 2-2 shows a schematic of the PCR/cloning approach and Tables A-3 and A-4 (Appendix A) list the primer sequences used.
Two-step PCR amplification of the U6 promoter using a universal U6 forward primer and two reverse primers successively resulted in a final product comprising the U6 Pol III promoter, beginning with BamHI and EcoRI restriction sites and ending in a Guanine residue (important for initiation of transcription), followed by the sense, loop and antisense sequences of the shRNA and ending with a poly-T termination signal of 6 Thymine residues. This PCR product (with 3’ adenine residues added by Taq polymerase) could be cloned into T/A vector pTZ57R/T.

First-round PCR amplification of 10 pmol of U6 promoter template was performed using 10 pmol each of forward and reverse primer (R1) and 1 x PCR Master Mix (Promega, WI, USA) (25 U/ml Taq polymerase in a proprietary buffer at pH 8.5; 200 µM of each dNTP and 1.5 mM MgCl₂) in a 20 µl reaction volume. Thermal cycling parameters were as follows: initial denaturation at 95°C for 5 minutes, followed by 30 cycles of denaturation (95°C for 30 seconds), primer annealing (58°C for 45 seconds) and elongation (72°C for 30 seconds) and a 15 minute extension step (at 72°C). Approximately 10 ng of the product was used as template for the second round of PCR, using the same conditions as above with R2 reverse primer and the final step extended to 20 minutes at 72°C.

The PCR-amplified shRNA cassettes were then ligated with pTZ57R/T using the InstAclone PCR cloning kit. After PCR, 5 µl of the reaction volume was analyzed by gel electrophoresis (2% agarose, stained with ethidium bromide (Sigma-Aldrich Corp., MO, USA) (see Appendix B for safety precautions) and run at 5 V/cm in 1 x Tris-borate-EDTA (TBE) buffer (45 mM Tris-borate, 1mM EDTA (ethylenediaminetetraacetic acid), pH 8.3) (Saarchem Merck). When a single clear band was visible under UV light, 540 ng of the ~300 bp PCR product was used directly in the ligation reaction with 150 ng of pTZ57R/T (corresponding to a 3:1 ratio of insert:vector), 5 U T4 DNA Ligase and 1 x ligation buffer (40 mM Tris-HCl, 10 mM MgCl₂, 10 mM dithiothreitol (DTT), 0.5 mM ATP, pH 7.8) made up to a 20 µl reaction volume. When non-specific bands were visible, gel extraction of the correct-sized band was done (using the KOMA EzWay Gel Extraction Kit (KOMA Biotech,
Seoul, Korea) according to the instructions of the supplier) before ligation. Quantification of the DNA used in the ligation was done using the NanoDrop™ 1000 spectrophotometer (Thermo Scientific) at 260 nm relative to 280 nm.

Cloning of the second set of shRNAs, shSPTBN-A, shSPTBN-B, shSPTBN-C, shSPTBN-D, shSP110-A, shSP110-B, shSP110-C and shSP110-D, was done using the pooled siRNA sequences (see Section 2.1.3) and used a slightly different approach (also described by Castanotto et al., 2002): a single round of PCR was performed using one long reverse primer (encoding a portion complementary to the 3’ end of the U6 promoter, followed by the entire shRNA sequence) and a U6 forward primer with four upstream restriction sites: BamHI, XhoI, NotI and NheI. The sequences of these primers are given in Table A-5 (Appendix A).

2.3.2 Cloning of target reporter plasmids

Target reporter plasmids (corresponding to the first set of shRNAs) were constructed so that the three shRNA target sites for each gene were cloned in tandem downstream of the Renilla luciferase ORF (see Figure 2-3, pg 38), expressed from the dual luciferase reporter plasmid, psiCHECK-2.2. A set of four oligonucleotides was designed to encode the various target sequences (Table A-6 of Appendix A and Figure 2-3). The F1 and R1 oligonucleotides were designed to anneal as double stranded DNA sequences encoding the first half of the three target sites and an overhang containing XhoI and EcoRV restriction sites at the 5’ end. Similarly, the F2 and R2 oligonucleotides were designed to anneal as double stranded DNA sequences encoding the second half of the three target sites and a NotI overhang at the 3’ end. The annealed F1/R1 and F2/R2 sequences would also have overhangs at their 3’ and 5’ ends respectively, that would allow them to be ligated together. psiCHECK-2.2 backbone was prepared for directional cloning by digesting 2 µg of plasmid with 2 U XhoI and 1 U NotI restriction enzymes using
Buffer O (MBI Fermentas, Ontario, Canada) (50 mM Tris-HCl, 10 mM MgCl₂, 100 mM NaCl, 0.1 mg/ml bovine serum albumin (BSA), pH 7.5) in a 30 µl reaction volume. This was followed by agarose gel electrophoresis (0.8% agarose, stained with ethidium bromide and run at 5 V/cm in 1 × TBE buffer) and extraction of the digested product with the KOMA EzWay Gel Extraction Kit (KOMA Biotech, Seoul, Korea). The four oligonucleotides for each target sequence were joined to form one double-stranded oligonucleotide by the following steps: First, the 5’ ends of the F2 and R1 oligonucleotides were phosphorylated - 200 pmol of oligonucleotide was incubated with 2 µl of 10 mM ATP (Sigma-Aldrich) and 5 U of T4 Polynucleotide Kinase (PNK) (New England Biolabs, NEB) in 1 × T4 Polynucleotide Kinase Reaction Buffer (NEB) (70 mM Tris-HCl, 10 mM MgCl₂, 5 mM DTT, pH 7.6) in a 20 µl reaction volume at 37°C for 1 hour. Next, the F1/R1 and F2/R2 pairs were annealed to form two double-stranded oligonucleotides: 100 pmol each of the oligonucleotides were incubated at 80°C for 5 minutes, followed by slow cooling to ~25°C. Finally, these were three-way ligated into the prepared backbone: 60 fmol of vector backbone was added to 180 fmol each of insert and incubated with 5 U of T4 DNA Ligase and 1 × ligation buffer (composition described above) in a 20 µl reaction volume for 1 hour at 16°C.
Figure 2-3: Reporter target cloning strategy.

The oligonucleotides encoding the shRNA target sequences for TMED2. After ligating the oligonucleotide pairs, they were positionally cloned into psiCHECK2.2 digested with XhoI and NotI.

2.3.3 Transformation and verification of clones

Chemically competent DH5α *Escherichia coli* (*E. coli*) cells (see Appendix B) were transformed with the ligation mixes by incubating 100 μl cells with 10 μl of the ligation reaction for 30 minutes on ice, heat-shocking at 42°C for 90 seconds, followed by further incubation on ice for 5 minutes. Transformed cells were plated onto sterile Agar (see Appendix B) containing ampicillin (100 μg/ml) and spread with 40 μl X-Gal (20 mg/ml,
5-bromo-4-chloro-3-indolyl b-D-galactopyranoside) and 8 µl IPTG (100 mg/ml, Isopropyl β-D-1-thiogalactopyranoside) and grown overnight at 37°C. In recombinant clones, the lacZ gene is interrupted by the ligation reaction and thus white colonies are formed (Sambrook and Russell 2001). White colonies were selected and used to inoculate 4 ml Luria-Bertani broth supplemented with ampicillin for overnight culture (see Appendix B). Subsequent plasmid preparation was done using the High Pure Plasmid Isolation kit (Roche Applied Science, Mannheim, Germany) as described in Appendix B. Target clones were screened by EcoRV (MBI Fermentas) digestion (10 U in Buffer R (MBI Fermentas) (10 mM Tris-HCl, 10 mM MgCl₂, 100 mM KCl, 0.1 mg/ml BSA, pH 8.5) for 1 hour at 37°C) and shRNA clones with XbaI/HindIII double digestion (2 U of HindIII, 1 U of XbaI and 1 × Tango buffer (MBI Fermentas) (33 mM Tris-acetate, 10 mM Mg-acetate, 66 mM K-acetate, 0.1 mg/ml BSA, pH 7.5) in a 30 µl reaction volume for 1 hour at 37°C), followed by agarose (stained with ethidium bromide) gel electrophoresis and visualization under UV light.

Positive clones were verified using automated sequencing with the BigDye Terminator v3.1 Cycle Sequencing Kit (Applied Biosystems, Foster City, CA, USA). The cycle sequencing reaction was set up with 100 ng of template plasmid, 2 µl Ready Reaction Mix (Applied Biosystems), 1 × BigDye Terminator v3.1 Sequencing Buffer (Applied Biosystems) and 10 pmol primer, made up to 10 µl with distilled water (dH₂O). Sequencing of pTZU6-shRNA clones was done with M13 reverse primer (5'-CAG GAA ACA GCT ATG AC-3') or the U6 forward primer; target primer (5'-CGA CGA TCT GCC TAA GAT GT-3') was used for psiCHECK-target clones. Thermal cycling conditions were as follows: 96°C for 1 minute, followed by 25 cycles of: 96°C for 10 seconds, 50°C for 5 seconds and 60°C for 4 minutes. The reaction product was purified by ethanol/EDTA precipitation (see Appendix B) and the pellet resuspended in Hi-Di™ Formamide (Applied Biosystems) for subsequent electrophoresis performed by Inqaba Biotec (Pretoria, SA) or internally on the ABI Prism 3100-Avant Genetic Analyzer (Applied Biosystems). Sequence data analysis was performed using the Sequencing Analysis V3.3 programs (Applied
Biosystems) and FinchTV v 1.4.0 (Geospiza, Seattle, WA, USA). Bulk preparations were made of all sequence-verified plasmids for use in molecular cloning and transfections. These preparations were made using the QIAGEN® Plasmid Maxi Kit (QIAGEN, Hilden, Germany) as detailed in Appendix B.

2.4 Mammalian cell culture

116 cells (human embryonic kidney cells expressing the T-antigen of SV40 virus and Cre Recombinase) were cultured in Dulbecco’s Modified Eagle’s Medium (DMEM) supplemented with 4.5 g/L Glucose and 2 mM L-Glutamine (BioWhittaker®, Lonza, Walkersville, Maryland, USA) and 10% heat-inactivated fetal bovine serum, FBS (Gibco®, Invitrogen, CA, USA), referred to as complete medium. Cells were incubated at 37°C with 5% CO₂ and passaged at 80% confluency as described in Appendix B. TZM-bl cells (HeLa-derived cells expressing CD4, CCR5 and CXCR4 HIV receptors, as well as luciferase and lacZ reporter genes under control of the HIV Tat-inducible LTR promoter) were obtained through the NIH AIDS Research and Reference Reagent Program (NARRRP), Division of AIDS, NIAID, NIH (Catalogue number 8129) from Dr. John C. Kappes, Dr. Xiaoyun Wu and Tranzyme Inc. TZM-bl cells were also cultured in complete medium (10% FBS) and passaged with 0.25%Trypsin / 1 mM EDTA (Gibco) (see Appendix B). SupT1 cells (NARRRP Catalogue Reagent 100, obtained from Dr. James Hoxie (Smith et al., 1984)), were grown in suspension in complete medium (10% FBS). Their passaging is also described in Appendix B.
2.5 Transfection of DNA and RNA into cultured cells

For all transfections cells seeded the previous day were transfected at 50-80% confluency. The following protocols are for 24-well culture plates (2 cm² culture surface area). Transfections in 96-well culture plates (0.4 cm² surface area) were carried out in the same way but scaled down by a factor of 0.2.

2.5.1 Lipofectamine™ 2000 transfections

A maximum amount of 1 µg DNA was diluted with OptiMEM® Medium to a 50 µl volume. The required volume of Lipofectamine™ 2000 (Invitrogen, CA, USA) (using a ratio of 1 µl Lipofectamine™ 2000 to 1 ug DNA) was similarly diluted into 50 µl OptiMEM® Medium (Invitrogen). After five minutes incubation at room temperature, the dilutions were mixed and allowed to incubate for 20 minutes at room temperature to allow lipoplexes (DNA-lipid complexes) to form. Thereafter the total volume (100 µl) was added drop-wise to the cell culture medium (500 µl). The culture plate was gently rocked to distribute the complexes evenly and incubated for the required length of time at 37°C.

2.5.2 ESCORT™ V transfections

For this reagent, the culture medium was replaced with fresh medium (without antibiotics) from 30 minutes to two hours before the transfection. A maximum amount of 700 ng DNA was diluted with 60 µl ESCORT V Transfection Buffer (Sigma-Aldrich). As a ratio of 3 µl ESCORT V Transfection Reagent (Sigma-Aldrich) to 1 µg DNA was used, 2.1 µl of Transfection Reagent was diluted with 60 µl Transfection Buffer. The dilutions were mixed gently, combined and mixed again before incubating at room temperature for 20 minutes. The rest of the procedure is the same as described in Section 2.5.1.
2.5.3 DharmaFECT®-4 transfections

Transfections with this reagent were carried out using the same procedures described in Section 2.5.1, except that a 96-well format was used. siRNAs were used instead of DNA and precautions were therefore taken to avoid RNase contamination of the samples and the siRNAs were stored on ice when not being handled. Moreover, dilutions of the siRNA stocks to required final concentrations were done using the buffer described in Section 2.1.3. Transfections were optimised as explained in Section 2.9 and thereafter 0.4 µl of the transfection reagent was used per 100 nM siRNA, unless otherwise specified.

2.6 Dual-Luciferase reporter assays

To evaluate the ability of the shRNA expression cassettes to knock down their reporter targets, a dual luciferase reporter assay was performed. Because each shRNA’s target sequence was cloned into the 3’ UTR of the Renilla (Renilla reniformis) luciferase reporter gene, the level of knockdown could be determined by measuring Renilla Luciferase activity relative to that of control Firefly (Photinus pyralis) Luciferase (expressed separately on the same plasmid). The 116 cells were seeded 24 hours prior to transfection at 1.2 × 10⁵ cells per well in 24-well culture plates (Nunclon™ Δ Surface, Thermo Fisher Scientific, Rochester, NY, USA). Cells were then transfected at 10:1 and 5:1 ratios of shRNA:target using Lipofectamine™ 2000 (see Section 2.5.1). In both cases a total of 1 µg of DNA was transfected, comprising: 800 ng shRNA with 80 ng target (for 10:1 ratio) and 750 ng shRNA with 150 ng target (5:1), as well as 100 ng of pCI-EGFP reporter plasmid (to monitor transfection efficiency). Mock (pTZU6+1) and non-specific controls (U6 shRNA 5, renamed shHBx-5) were included in both experiments.
Forty-eight hours post-transfection, a Dual-Luciferase® Reporter Assay (Promega, WI, USA) was performed. After aspirating the culture medium, cells were lysed in the recommended volume of 1 × passive lysis buffer (PLB) (Promega, proprietary formulation) and incubated at 37°C for 15-20 minutes. After resuspending the lysates to remove clumps, 12 µl of each sample were transferred to the wells of a white-walled 96-well plate (Corning® Costar®, Corning Inc., MA, USA) and loaded into the Veritas dual-injection luminometer (Turner Biosystems, CA, USA). Sufficient volumes (i.e. 50 µl per sample plus 500 µl for injector priming) of 1 × Luciferase Assay Reagent II (LAR II) and Stop & Glo® Reagent (Promega, proprietary formulations) were made up according to the instructions of the supplier. Luciferase activity readings were then taken (in relative light units, RLU): first Firefly luciferase, after addition of LAR II, followed by Renilla luciferase after Stop & Glo Reagent addition. The protocol used for measurements was adapted for 50 µl from the Promega protocol for 100 µl of reagent per sample.

2.7 qRT-PCR assays

Quantitative real-time reverse-transcription PCR (qRT-PCR) was performed in order to verify the knockdown efficacy of the shRNA constructs against their endogenous mRNA targets. qRT-PCR primers (see Table A-7 of Appendix A) were designed to amplify each of the 5 genes, as well as housekeeping genes glyceraldehyde-3-phosphate dehydrogenase (GAPDH) and β-actin. TZM-bl, HEK293T and 116 cells were seeded in 12-well (2.5 × 10^5 cells per well) or 24-well (1.2 × 10^5 cells per well) tissue culture plates (Nunclon™ Δ Surface, Thermo Fisher Scientific). Twenty-four hours later, the maximum amount of DNA was transfected for each well format (i.e. 24-well plates: 950 ng of shRNA-expressing plasmid or control plasmid and 50 ng pCI-EGFP; 12-well plates: 1.9 µg shRNA or control plasmid with 100 ng pCI-EGFP). Lipofectamine™ 2000 was used for
116 and HEK293T cells (as described in Section 2.5.1), while ESCORT™ V transfection reagent (Sigma-Aldrich) was used to transflect TZM-bl cells (see Section 2.5.2).

RNA was extracted 48 hours post-transfection using TRI-Reagent® (Sigma-Aldrich). Culture medium was aspirated and the cells lysed using the recommended amount of TRI-Reagent (1 ml per 10 cm² of culture dish area). After resuspension and 5 minutes incubation at room temperature, chloroform (200 µl per ml TRI-Reagent) was added to the homogenate, vortexed until thoroughly mixed and the phases allowed to separate before centrifugation at 12 000 × g for 15 minutes at 4°C. The upper aqueous phase (containing the RNA) was then transferred to a clean microfuge tube and isopropanol added (500 µl per ml TRI-Reagent) before storing at -70°C overnight. The following day (or when needed) the RNA was pelleted by centrifugation at 12 000 × g for 30 minutes at 4°C. The supernatant was discarded, 75% ethanol added (1 ml per ml TRI-Reagent) and the RNA washed by brief vortexing. After centrifugation at 7 500 × g for 5 minutes, the pellet was allowed to air-dry for ~5 minutes before resuspension in diethyl pyrocarbonate (DEPC)-treated water (Sigma-Aldrich) and incubation at 55°C for 10 minutes. Thereafter the RNA was stored on ice and used directly in reverse transcription before long-term storage at -70°C.

The QuantiTect® Reverse Transcription kit (QIAGEN, Hilden, Germany) was used for reverse transcription of RNA to cDNA. The reaction was scaled down so that 500 ng or 1 µg of RNA was incubated with 1 µl gDNA Wipeout buffer in a 7 µl reaction volume at 42°C for two minutes and then placed on ice. Thereafter 0.5 µl Quantiscript® Reverse Transcriptase (a proprietary mixture of the QIAGEN® products Omniscript® Reverse Transcriptase and SensiScript® Reverse Transcriptase that also contains RNase inhibitor), 0.5 µl RT Primer mix (a blend of oligo-dT and random hexamer primers including Mg²⁺ ions and dNTPs) and 1 × Quantiscript RT buffer were added to the first reaction to a final volume of 10 µl and incubated at 42°C for 20 minutes before heat-
inactivation at 95°C for three minutes and placing on ice. The resulting cDNA (4 µl) was used in subsequent qPCR with 10 pmol of the designed primers (Table A-7, Appendix A) and 1 × LightCycler® FastStart DNA MasterPLUS SYBR Green I (Roche Applied Science, Mannheim, Germany) (proprietary formulation comprising FastStart Taq DNA Polymerase, reaction buffer, MgCl₂, SYBR Green I dye and dNTP mix) in a 20 µl reaction volume. The Roche LightCycler® 2.0 instrument was used for the PCR using the following parameters: a hotstart at 95°C for 10 minutes was followed by 35-50 cycles of denaturation (95°C for 10 seconds), annealing (60°C for 10 seconds) and elongation (72°C for 10 seconds). Quantification occurred during the elongation stage, while melting curve analysis (95°C for 30 seconds, 65°C for 30 seconds, followed by slow (ramp of 0.2) heating to 95°C with continuous quantification) ensured PCR product specificity. After PCR, the Relative Quantification (Monocolour) module of the LightCycler® LCS4 v 4.0.0.23 software (Roche Applied Science, Mannheim, Germany) was used for data analysis.

2.8 Electroporation

Electroporation of SupT1 cells with pCI-EGFP, was done on the CE Module of a GenePulser Xcell apparatus (Bio-Rad Laboratories Inc., Hercules, CA, USA) using variations of several protocols as outlined in Tables B-1 and B-2 of Appendix B. Buffers used included DMEM, with or without FBS (10%) and DMSO (1.25%) (Invitrogen), OptiMEM (Invitrogen), Gene Pulser® electroporation buffer (Bio-Rad) (patented formulation, International Publication Number WO 2008/134 200 A1, World Intellectual Property Organization), phosphate-buffered (PB) sucrose (272 mM sucrose (Saarchem Merck), 7 mM K₃PO₄ (pH 7.4) (Sigma-Aldrich), 1 mM MgCl₂) and HEPES-buffered (HB) dextrose (20 mM HEPES (pH 7.3) (BDH Chemicals), 137 mM NaCl (Saarchem Merck),
0.7 mM Na₂HPO₄ (Saarchem Merck) and 6 mM dextrose (Saarchem Merck)) (Chu et al., 1987; Baum et al., 1994). Transfection efficacy (i.e. the proportion of cells that fluoresce green) was assessed by fluorescence microscopy (Axiovert 100 M, Carl Zeiss AG, Germany) after 24 hours of outgrowth post-electroporation. The total number of cells and the number of fluorescing cells were counted in a field of view approximately 2 cm × 2 cm under 200 × magnification. The viability of the cells was also assessed using 0.4% Trypan-blue stain (Biowhittaker). While dead or dying cells are permeable to the blue dye and thus stain blue, viable cells exclude it and appear colourless. To stain the cells, they were resuspended in an equal volume of the stain and loaded onto a haemocytometer slide for counting (see Appendix B).

2.9 Delivery of DNA and RNA to SupT1 cells by cationic lipid-based transfection reagents

Two siRNA pools and a recommended reagent for siRNA transfection, DharmaFECT® 4, were purchased from Thermo Scientific Dharmacon® (Thermo Fisher Scientific Inc., Lafayette, CO, USA). Delivery of siRNA to SupT1 cells using DharmaFECT® 4 transfection reagent was optimised as follows: SupT1 cells were seeded in 100 µl of medium at 5 × 10³, 1 × 10⁴ and 2.5 ×10⁴ cells per well of a 96-well tissue culture plate (Techno Plastic Products AG, TPP®, Switzerland). Twenty-four hours later, 100 nM siRNA was transfected in duplicate with 0.05, 0.1, 0.2, 0.3, 0.4 or 0.5 µl DharmaFECT®-4 (20 µl volume transfection mix). Forty-eight hours later, cells were examined using a fluorescence microscope (Axiovert 100 M, Carl Zeiss AG, Germany) for detection of the siRNAs (labelled with the sulfoindocyanine fluorescent probe, Cy3) within cells. Delivery of the siRNAs to TZM-bl cells using DharmaFECT®-4 was assessed similarly: cells were seeded in 100 µl of medium at 10 000 and 25 000 cells per well of a
96-well tissue culture plate (TPP®). Twenty-four hours later, 100 nM Cy3-labelled siRNA was transfected in duplicate with 0.2, 0.4 or 0.6 µl DharmaFECT®-4 (20 µl volume transfection mixes) or 0.4 µl Lipofectamine™ 2000. Forty-eight hours later, cells were examined by fluorescence microscopy for Cy3 fluorescence within cells. shRNAs were transfected into SupT1 cells and TZM-bl cells using ESCORT™ V transfection reagent (Sigma-Aldrich) as explained in Section 2.5.2.

2.10 Cytotoxicity assays

Various assays exist for determining the metabolic and physical condition of cells. A colorimetric assay was used here that measures the ability of mitochondrial Reductase enzymes to convert the yellow MTT (3-(4,5-Dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide, a tetrazole) substrate to purple Formazan (Mosmann 1983). The activity of these enzymes is reduced in cells that are experiencing metabolic dysfunction - for example, as a result of treatment with a cytotoxic agent like Trichostatin A, TSA. This organic compound selectively inhibits the class I and II mammalian HDAC enzyme families (Yoshida et al., 1990) and activates expression of apoptotic genes (Henderson et al., 2003).

For the MTT assay, 5 × 10³ TZM-bl cells per well were seeded in DMEM containing 10% FBS in a 96-well tissue culture plate (TPP®). Eighteen hours later, cells were treated with 0.4 µl ESCORT™ V for shRNA transfections (100 ng of shRNA or control plasmid and 10 ng pCI-EGFP) or 0.4 µl DharmaFECT® 4 for 100 nM siRNA transfections. Controls wells (in triplicate) included untransfected and untreated cells, as well as cells treated with three concentrations (1000 nM, 500 nM and 100 nM) of TSA (Sigma-Aldrich). Culture medium was replaced after 5 hours of transfection (siRNA- and
shRNA-treated wells). After 48 hours of transfection, 10 µl of MTT (5 mg/ml in PBS, 0.2 µm filter-sterilised) (Sigma-Aldrich) was added per well containing 100 µl and incubated at 37°C with 5% CO₂ in the dark for 1 hour. Thereafter, the medium was removed and the black formazan precipitate dissolved by addition of 200 µl dimethyl sulfoxide (DMSO). After mixing by gentle shaking for 2 minutes, the light absorbance of each well was measured at two wavelengths (570 nm, with 655 nm as the reference) using a Microplate reader (model 680, BioRad).

2.11 Virus production

Most of the viral work was done in the HIV Pathogenesis Research Laboratory with permission of Dr Maria Papathanasopoulos, while later work was carried out in our laboratory. For the initial challenge assay in TZM-bl cells, CAP210.2.00.E8 Env-pseudotyped virus was used (a gift from Maria Papathanasopoulos). The CAP210.2.00.E8 plasmid (NARRRP Catalogue Reagent 11317, available from Drs. L. Morris, K. Mlisana and D. Montefiori (Li et al., 2006)) expresses HIV-1 Subtype C env/rev. The env⁻/rev⁻ backbone plasmid, pSG3Δenv (NARRRP Catalogue Reagent 11051, available from Drs. John C. Kappes and Xiaoyun Wu (Wei et al., 2002; Wei et al., 2003)), contains a defective vpu gene.

For the final challenge assay, HIV-1p81A⁻ virus was used. The p81A-4 plasmid (NARRRP Catalogue Reagent 11440, available from Dr. Bruce Chesebro (Toohey et al., 1995)) is a full length HIV-1 infectious molecular clone containing the V1-V3 envelope regions of the Ba-L HIV-1 strain in an NL4-3 (Subtype B) strain background. Infectious virions were produced by transfecting the plasmid into HEK293T cells, seeded the previous day at 2 × 10⁶ cells in a 25 cm² flask (Corning® flask, Corning, NY, USA) in 5 ml medium. Four micrograms of p81A-4 plasmid was suspended in 150 µl serum-free
DMEM, mixed with 40 µl PolyFect® transfection reagent (QIAGEN, Hilden, Germany), left at room temperature for five minutes and then added to the seeded cells' medium. Cells were incubated at 37°C with 5% CO₂ and medium changed 24 hours after transfection. A further 24 hours later, virus-containing supernatant was collected, centrifuged briefly (200 × g for five minutes) to remove cellular debris and filtered through a 0.2 µm sterile filter. Serum concentration was adjusted to 20% FBS and aliquots frozen at -70°C.

2.12 Viral titration

An end-point titration technique was used for viral titering, i.e. the TCID50 (50% tissue culture infectious dose) was determined based on the protocol published by the ACTG Laboratory Technologist Committee of the HIV/AIDS Network Co-ordination (2004). Previous data from our laboratory has shown that the TCID50 determined from LTR-luciferase activity does not differ greatly from the TCID50 obtained by p24 ELISA (enzyme-linked immunosorbent assay). Thus, the TCID50 was only determined by luciferase assay in this study. A more accurate titre involves calculating the multiplicity of infection (m.o.i.), which measures the average number of viral particles present per cell. However, as this was not technically feasible in this study, the TCID50 method was used. A dosage of 500TCID50 was used in all challenge assays as this dose did not cause cell death and yielded LTR-luciferase activity values within a reliable detection range.

Cells were seeded at 5 × 10⁴ cells per well of a 96-well tissue culture plate (TPP®). The following day, the culture medium was removed and replaced with 50 µl DMEM containing 10% FBS and 15 µg/ml DEAE-dextran (GE Healthcare Life Sciences, Uppsala, Sweden). The virus stock was initially diluted in seven serial three-fold dilutions, using DEAE/dextran-containing medium (described above), a total volume of 750 µl and 250 µl transfers. To each column of three wells containing 50 µl medium, 150 µl of the
corresponding virus dilution was added per well (i.e. a 3:4 dilution). Thus the final titration consisted of seven serial four-fold dilutions (i.e. $4^{-2}; 4^{-3}; 4^{-4}; ...; 4^{-8}$). One column of wells served as a control and was treated with DEAE/dextran-containing medium without any added virus. The data obtained from the TCID50 Luciferase assay were analyzed using the Spearman-Kärber formula detailed in Appendix B.

2.13 Viral challenge assays

The viral challenge assays were performed in TZM-bl cells as they are permissive for infection with viruses that are tropic for both T-cells and macrophages. In addition, TZM-bl cells support a full viral lifecycle while their integrated reporter genes provide a convenient and simple means of assessing changes in viral transcriptional activity. The Subtype C-enveloped pseudovirus (Section 2.11) was used as an initial screen for the effects of each gene's suppression during a single round of infection. This was followed by a challenge assay consisting of two parts that was based on the approach of Brass et al., (2008) (Figure 2-4). In the first part, fully infectious virus (i.e. capable of generating virions that can infect new cells) was used to challenge cells in which SPTBN1, SP110 and the control genes had been inhibited. The virus used was Subtype B virus produced from a full-length infectious molecular clone (p81A-4) that is similar to the NL4-3 strain except for the V1-V3 envelope region, which is of the macrophage-tropic BaL strain (Section 2.11). In the second part of the assay, the culture medium from the infected cells was transferred to uninfected, untransfected cells to assess the effect of knockdown on infectivity (i.e. virion production and infection of new cells).
Figure 2-4: Schematic of two-part viral challenge assay

The schematic for the two-part viral challenge assay was based on the approach used by Brass et al. (2008), with Part 1 measuring early-stage viral replication and Part 2 measuring later stages (assembly and budding). TZM-bl cells were transfected with pooled siRNAs and, after 72 hours of gene knockdown, infected with HIV-1. After 48 hours, integrated LTR-luciferase reporter gene expression was measured (Part 1) and the virus-containing supernatant transferred to fresh TZM-bl cells. After 24 hours infection, luciferase reporter activity was again measured for Part 2 of the assay. Figure adapted from Brass et al. (2008).

2.13.1 Env-pseudotyped viral challenge

For the CAP210.2.00.E8 Env-pseudotyped viral challenge assay, TZM-bl cells were seeded at $5 \times 10^3$ cells per well of a 96-well tissue culture plate (TPP®). The RNAi effecters were transfected the following day, using 0.4 µl or 0.2 µl DharmaFECT® 4 for siRNA transfections (for 100 nM and 50 nM siRNA, respectively) and 0.4 µl ESCORT™ V for shRNA transfections (120 ng DNA : 100 ng of the relevant shRNA or control plasmid, 10 ng pCI-EGFP and 10 ng phRL-CMV). The Virus Control and No Virus control wells were not transfected. Medium (100 µl) was replenished 30 minutes to two hours prior to ESCORT™ V transfection and again 4-6 hours after DharmaFECT® 4 and ESCORT™ V
transfections. After 72 hours of transfection (when knockdown is still operative), the cells were infected with 500TCID₅₀ of pseudovirus (i.e. 50 TCID₅₀’s in 100 µl) in medium containing 5% FBS and DEAE-dextran at a final concentration of 10 µg/ml. After 48 hours of infection, a luciferase assay was carried out (see Section 2.6), measuring Firefly (LTR) luciferase relative to Renilla luciferase activity for the shRNA-transfected wells and LTR-luciferase activity only for the siRNA-transfected wells. Data were normalised to the mean of the Non-specific control (siHBV).

2.13.2 Replication-competent viral challenge

As a result of several practical issues, Part 1 in this study was measured using LTR-Luciferase activity as opposed to p24 Gag production. For Part 1, cells were seeded and transfected with siRNA as was carried out for the pseudovirus challenge assays, with untransfected infected (Virus Control) and untransfected uninfected (No Virus) control wells included. After 72 hours of transfection, cells were infected with p81 virus at a TCID₅₀ of 500/ml in 200 µl medium containing 10% FBS and 15 µg/ml DEAE-dextran. The No Virus control wells received the same medium without virus added. After two hours incubation, the medium was removed and replaced with fresh complete medium before further incubation at 37 °C with 5% CO₂. After 48 hours of infection, Part 2 of the assay was initiated by transferring 80 µl of virus-containing supernatant to the wells of a replica plate seeded the previous day with 5.4 × 10³ TZM-bl cells per well. These cells were cultured in medium containing 15 µg/ml DEAE-dextran to facilitate infection. A Dual-Luciferase reporter assay was performed (as described in Section 2.6) on the Part 1 cells after remaining supernatant had been discarded. After 20 hours of infection, the Part 2 cells were lysed for an assay of LTR-Luciferase activity (also using the Dual-Luciferase reporter assay, with only Firefly Luciferase activity measured for siRNA-treated wells).
2.14 Statistical analyzes

All data were analyzed using GraphPad Prism version 4.03 for Windows (GraphPad Software, San Diego, CA, USA, www.graphpad.com). Non-parametric one-way analysis of variance (ANOVA) was used to assess the statistical significance of data. The post-tests performed are described with each reported result in Chapter 3.
CHAPTER 3 : RESULTS

This study aimed to verify whether several genes that have recently been implicated as being novel HIV-1 host factors are required for viral infection and replication. Five genes were selected from the published dataset (Brass et al., 2008) as it was expected that some of these would be false positives (i.e. not true HIV-1 host factors) and also that some difficulties might arise during experiments to assess their effect on the virus and to inhibit their function by RNAi. Their selection was specifically made in an unbiased way, using criteria that would select for genes whose effect on HIV-1 replication was supported by good evidence and that could potentially be developed as novel drug/gene therapy targets.

3.1 Criteria-based selection of five genes

The five genes that were selected from the list of 237 novel HDFs published by Brass et al. (2008), as well as the criteria supporting their selection, are summarised in Table 3-1. A full portrayal of the selection process is shown in Appendix C. The first stage selected genes that caused a 75% or greater reduction in HIV-1 replication (see Table C-1). This cut-off value was high enough to narrow down the original gene list while ensuring a selection of genes remained that had scored in both early (Part 1) and late (Part 2) stages of the HIV-1 lifecycle.

The second stage selected for genes for which more than one siRNA caused a reduction in viral activity (Table C-2). The greater the number of individual siRNAs targeting the same gene that cause a significant decrease in viral activity, the more likely that the effect is gene-specific. From this stage onwards, the list of genes scoring in Part 1
and Part 2 were processed separately. The next stage considered the functional importance of the Part 1 genes. Six genes (see Table C-3) were excluded from the selection based on their biological function as recorded in the curated UniProt database. For example, WNT1 is a member of the Wnt signalling pathway, while NUP107 is an essential component of the nuclear pore complex. Knockdown of such genes would likely be toxic to the cell and thus not optimal target genes.

Of the five genes to be studied, three were chosen from the Part 1 list (involved in viral entry up until Gag translation) and two from the Part 2 list (involved in later stages such as assembly and budding). The final three Part 1 genes were chosen as the first three on the remaining sorted list (see Table C-4). The Part 2 genes were processed in the same way as explained above, with six genes being excluded based on functional importance (Table C-5). As the remaining list contained several more genes than the Part 1 list, an additional criterion was introduced to further focus the list: the number of gene homologues in other species (i.e. orthologues) was considered, as genes that are highly conserved across many species are more likely to have essential functions. A cut-off of five homologues was chosen that served to shorten the Part 2 gene list by ~30% (see Table C-6). The final selection (Table C-7) included one gene with three individual scoring siRNAs (PRDM14) and one with a published reference to interaction with pathogens (SP110) (see section 1.6.1.5). Thus, these two genes were selected to represent the late stage of HIV-1 replication in this study.
Table 3-1: The final five putative HDFs selected for validation in this study

<table>
<thead>
<tr>
<th>Gene</th>
<th>NCBI GeneID</th>
<th>Extent of virus knockdown</th>
<th>Number of siRNAs</th>
<th>Involvement in protein complexes and/or molecular pathways</th>
<th>Function</th>
<th>References to HIV-1 interaction</th>
</tr>
</thead>
<tbody>
<tr>
<td>SPTBN1 (spectrin, beta, non-erythrocytic 1)</td>
<td>6711</td>
<td>76%</td>
<td>3 (Part 1)</td>
<td>None</td>
<td>May be involved in calcium-dependant movement of the cytoskeleton at the membrane, through its interaction with calmodulin</td>
<td>(Shoeman et al., 2002)</td>
</tr>
<tr>
<td>TMED2 (transmembrane emp24 domain trafficking protein 2)</td>
<td>10959</td>
<td>82%</td>
<td>3 (Part 1); 1 (Part 2)</td>
<td>Interacts with ARFGAP1</td>
<td>May be involved in the budding of coatamer-coated and other coated vesicles. Could bind and collect cargo molecules into budding vesicles.</td>
<td>None</td>
</tr>
<tr>
<td>KIAA1012</td>
<td>22878</td>
<td>89%</td>
<td>4 (Part 1)</td>
<td>Part of the multisubunit TRAPP (transport protein particle) complex.</td>
<td>May be involved in vesicular transport from endoplasmic reticulum to Golgi.</td>
<td>None</td>
</tr>
<tr>
<td>PRDM14 (PR-domain containing 14)</td>
<td>63978</td>
<td>81%</td>
<td>3 (Part 2)</td>
<td>None</td>
<td>May be involved in transcriptional regulation.</td>
<td>None</td>
</tr>
<tr>
<td>SP110 (SP110 nuclear body protein)</td>
<td>3431</td>
<td>84%</td>
<td>2 (Part 2)</td>
<td>None</td>
<td>Transcription factor. May be a nuclear hormone receptor co-activator. Enhances transcription of genes with retinoic acid response elements (RARE).</td>
<td>(Izmailova et al., 2003)</td>
</tr>
</tbody>
</table>
3.2 Gene-specific knockdown by Pol III-expressed shRNAs

3.2.1 Design of shRNA expression cassettes

In order to validate the classification of SPTBN1, TMED2, KIAA1012, PRDM14 and SP110 as HDFs, as well as their role in early versus late stages of HIV-1 replication, their function was inhibited using RNAi effecters targeted to the coding region of each gene. Because gene knockdown using siRNAs is transient, shRNA expression cassettes were designed that would provide a constant supply of RNAi effecters for stabler suppressive effects. The U6 Pol III promoter is useful in this regard: shRNA-encoding sequences that have a G-residue at the 5’ end and at least four terminating T-residues are efficiently expressed to form shRNAs with two-nucleotide 3’ overhangs that allow Dicer processing into siRNAs (Paule and White 2000; Brummelkamp et al., 2002; Paddison et al., 2002; Paul et al., 2002).

Initially, shRNAs targeting three different regions of each of the five genes (SPTBN1, TMED2, KIAA1012, PRDM14 and SP110) were designed. Their design was based on the Broad Institute’s online TRC MISSION shRNA Library (Section 2.1.1). However, several candidate sequences available on this database contained three or more T-residues in the guide sequence (which cannot be altered without affecting its targeting efficacy). To avoid risking premature termination of Pol III transcription, another approach was used for the design of these shRNA sequences. The difference in thermodynamic stability between the 5’ and 3’ ends of a siRNA duplex is largely responsible for ensuring correct selection of the guide strand for loading into the RISC (Khvorova et al., 2003; Schwarz et al., 2003) and an effective predictor of siRNA efficacy (Heale et al., 2005). An online application which ranks candidate siRNA duplexes based on this difference was used to design the sequences encoding shTMED-2, shKIAA-3 and shPRDM-3 (Section 2.1.1).
The specificity of all shRNA sequences was assessed using BLAST (see Section 2.1.2) to detect off-target matches with other gene sequences. While partial matches were found in some instances (data not shown), no matches with the full seed region (nucleotides 2-8 of the guide strand) were found that might cause miRNA-like off-target translational suppression (Jackson et al., 2003; Jackson et al., 2006). Finally, the secondary structure of the shRNA sequences was assessed using the RNA Mfold program (Mathews et al., 1999; Zuker 2003) (Section 2.1.2). This algorithm uses free energy predictions to calculate the hybridization and folding of a linear RNA sequence at 37°C. In all cases, the desired hairpin structure was most thermodynamically favourable and thus most likely to form under physiological conditions. This is important for efficient Dicer processing and subsequently effective RNAi silencing (Siolas et al., 2005).

### 3.2.2 Cloning of shRNA expression cassettes and reporter targets

For each gene a set of three U6 Pol III shRNA expression cassettes and one plasmid consecutively encoding each shRNA’s minimal target sequence were designed. Cloning of the shRNA cassettes used a rapid and simple PCR-based approach that is described in Section 2.3.1. The minimal target sequences were inserted by positional cloning into the 3’ UTR of the Renilla luciferase reporter gene expressed from the psiCHECK-2.2 plasmid (see Sections 2.2 and 2.3.2). This plasmid separately expresses Firefly luciferase, providing a useful dual-reporter system whereby knockdown of the minimal target RNA can be detected by measuring Renilla luciferase activity relative to background Firefly luciferase.

Restriction digestions were performed with XbaI and HindIII, restriction enzymes whose sites flank the T/A cloning site of pTZ57R/T, to verify that individual clones had inserts of the correct size. Figure 3-1a (pg 60) shows the digestion products after gel electrophoresis: correctly inserted clones produced two bands, one of size 368 bp (the
insert) and another of 2840 bp (the backbone of pTZ57R/T). Insertion of the correct sequence was subsequently verified by sequencing the insert as shown in the partial electrophoretogram in Figure 3-1b. All five of the reporter target plasmids were correctly cloned (data not shown). Likewise, of the 15 designed shRNAs, 10 cloned correctly (verified by sequencing): shSPTBN-1, shSPTBN-2, shSPTBN-3, shTMED-2 (which was subsequently used as a positive control), shKIAA-1, shKIAA-2, shPRDM-1, shPRDM-3, shSP110-1 and shSP110-2. The five remaining shRNAs (shTMED-1, shTMED-3, shKIAA-3, shPRDM-2 and shSP110-3) proved difficult to clone, specifically at the second round of PCR-amplification. Nevertheless, sufficient shRNA expression cassettes were correctly cloned to continue.

3.2.3 Knockdown of luciferase reporter targets

The ability of each expressed shRNA to inhibit its respective target sequence was assessed by co-transfecting each shRNA (or a non-targeting control, shHBx-5 or an empty vector, Mock) with its respective reporter target plasmid into HEK293T cells. After 48 hours to allow for knockdown to occur, the cells were lysed and the relative luminescence of Renilla (reporter) versus Firefly luciferase activity measured (see Section 2.6). Figure 3-2 (pg 61) shows these ratios normalised to the Mock control, empty pTZU6+1 (see Section 2.2).
Figure 3-1: Example of restriction digest and sequence data used to verify correct clones.

a) Digestion of correctly inserted clones (lanes 2-17) with XbaI and HindIII restriction enzymes produced products of size 368 bp and 2840 bp after separation by ethidium bromide-stained agarose gel electrophoresis and visualisation under UV light. A previously sequence-verified shRNA clone, shTMED-2 was used as a positive control (when digested) (lane 19) and as a size reference for the larger digestion product when undigested (3208 bp) (lane 20). A molecular weight marker, O’GeneRuler™ 100 bp Plus DNA Ladder (Fermentas), was used as an additional band-size reference in lanes 1 and 21. b) Part of an electrophoretogram showing the sequence of the 3’ end of the U6 promoter and the shRNA sequence for correctly cloned shTMED-2.

As shown in Figure 3-2, the non-specific shRNA control (shHBx-5) did not cause unwanted off-target knockdown of any of the reporter target sequences relative to the Mock control, which serves to normalise for effects of the Pol III promoter. Moreover, all correctly cloned shRNAs yielded highly effective knockdown (75% or more, P < 0.01) of the reporter targets. This indicates that these shRNAs are being correctly expressed and processed by the RNAi pathway for recognition and action on their target sequences.
Figure 3-2: Knockdown efficiency of shRNAs against luciferase reporter targets.

Dual Luciferase Reporter Assay data compiled from several experiments. Data represent Renilla Luciferase activity relative to background Firefly Luciferase. The percentage target knockdown relative to Mock is indicated above each shRNA column. The x-axis is divided into 5 groups: one for each gene’s reporter target. Key to symbols: M is Mock; 1, 2 and 3 refer to the shRNAs targeting each gene (e.g. 2 in the TMED2 group is shTMED-2); NS is non-specific shRNA, shHBx-5. Sample values (n=3) were normalised to the Mock control, with error bars representing the standard error of the mean (SEM). ANOVA with Dunnett’s post-test comparison with Mock: ns (p > 0.05); ** (p < 0.01).
3.3 Accurate detection of mRNA by quantitative RT-PCR

The ability of the shRNAs to suppress minimal target sequences does not necessarily translate into effective inhibition of the endogenous mRNA target. Frequently, secondary structure within the mRNA prevents effective RNAi knockdown (see Section 4.6.2). To accurately detect the level of mRNA suppression by each shRNA, qRT-PCR assays were set up for specific target RNA quantification relative to a stably expressed housekeeping gene.

3.3.1 Testing of non-targeting controls

qRT-PCR is an extremely sensitive technique that requires testing with several controls to ensure amplification is occurring correctly. Three non-targeting shRNAs (shHBx-5 and shHBx-10, both targeting HBV, and shLacZ, targeting the lacZ gene of E. coli) (see Section 2.2) were tested to ensure that non-specific knockdown effects were not detected for each gene. Knockdown of SPTBN1 by shHBx-5, shHBx-10 and shLacZ is shown in Figure 3-3a. Figures 3-3b and c show knockdown of TMED2, KIAA1012 and SP110 by shHBx-5 and shLacZ. GAPDH was used as the reference gene for relative mRNA quantification in Figure 3-3a and b, while β-Actin was used in Figure 3-3c. Based on these data, shHBx-5 (from here on referred to as shHBx) was used as a non-specific control in subsequent experiments as it consistently showed no significant knockdown of the various genes tested.

Amplification of PRDM14 showed extremely low expression levels in several cell lines tested (data not shown). The cycle threshold values (i.e. the point at which the PCR amplification rate becomes exponential) for amplification of cDNA and water blank samples with PRDM14-specific primers were indistinguishable. Since normal expression levels of this gene could not be accurately detected, it is unlikely that product would be detected in RNAi-suppressed samples. For this reason, this gene was not studied further.
Figure 3-3: Testing knockdown of SPTBN1, TMED2, KIAA1012 and SP110 by non-specific shRNA controls.

Non-specific knockdown of SPTBN1: a) shHBx-5, shHBx-10 and shLacZ were tested for non-specific knockdown of SPTBN1 by qRT-PCR assay using GAPDH as housekeeping gene. Sample values (n=3, SEM) were normalised to Mock. One-way ANOVA with Dunnett's post-test comparison with Mock: Not significant, ns (p > 0.05). Non-specific knockdown of TMED2, KIAA1012 and SP110: b) shHBx-5, shHBx-10 and shLacZ were tested for non-specific knockdown of TMED2, KIAA1012 and SP110 by qRT-PCR assay. GAPDH was used as housekeeping gene with values normalised to Mock (n=2, SEM). Two-way ANOVA with Bonferroni's post-test comparison with Mock: ns (p > 0.05). c) As in b) with β-Actin used as housekeeping gene and shHBx-10 not tested.
3.3.2 Testing of positive controls

As a positive control, shHTATSF-2, a verified shRNA targeting the HTATSF1 gene involved in regulating HIV-1 transcript splicing (Section 1.2.6) was tested. As shown in Figure 3-4a, quantification of HTATSF1 mRNA levels relative to GAPDH failed to show the expected knockdown. When repeated using β-actin as the reference gene (Figure 3-4b), the expected level of knockdown (50%, $P < 0.01$) was obtained. This effect was seen in several experiments (Figure 3-3b): where GAPDH was used as the reference gene there was high variability between triplicate sample values and between successive experiments; this was overcome by using β-actin instead (Figure 3-3c). Thus, subsequent experiments quantified gene expression relative to β-actin.

**Figure 3-4: Knockdown of positive control HTATSF1.**

Knockdown of positive control HTATSF1 by qRT-PCR assay using GAPDH (a) and β-actin (b) as housekeeping genes. Sample values ($n = 3$, SEM) were normalised to Mock. ANOVA with Dunnett’s post-test comparison with Mock: ns ($p > 0.05$); ** ($p < 0.01$).
Two more positive controls were tested that target viral genetic elements. The TZM-bl cell line contains a stably integrated Firefly luciferase reporter gene that is expressed by the HIV-1 LTR promoter. By transfecting these cells with a Tat expression cassette, transactivation of the Tat-responsive LTR promoter can be induced (Section 1.1.3), which in turn increases reporter gene expression several fold. As shown in Figure 3-5a and b, verified U6 promoter-expressed shRNAs targeting the LTR and Tat RNA of HIV-1 (shLTR and shTat, respectively - Section 2.2) were tested relative to a Mock control (an empty vector containing the U6 promoter). In addition, the experiment was tested in the presence and absence of Tat (expressed from a CMV promoter), with an empty vector containing the CMV promoter added in Tat’s absence to control for promoter activity.

shLTR was designed to target a Subtype C virus LTR sequence. When this shRNA was tested previously as part of a long hairpin RNA (lhRNA) construct to knock down a Subtype C LTR-driven luciferase reporter (Barichievy et al., 2007) it yielded 90% knockdown in the presence of Tat (which upregulates LTR expression). When tested using the same conditions against a Subtype B LTR-driven reporter, it yielded only 50% knockdown in the presence of Tat. The data obtained here for shLTR is comparable: Figure 3-5a shows the effect of Tat on upregulating luciferase levels. shLTR knockdown is only significant in the presence of Tat (representing an infection), matching the expected value of 50% knockdown (P < 0.001) of this integrated Subtype B LTR-driven reporter. Figure 3-5b shows similar results for the second positive control, shTat. This shRNA yielded ~30% knockdown (P < 0.001) when its target (Tat) was added. This is lower than expected (Scherer et al., 2004), but may be due to higher target concentration in this system, where Tat was strongly expressed from a plasmid.
Knockdown of integrated LTR-luciferase expression in TZM-bl cells by shLTR (a) and shTat (b), respectively, in the presence (+) or absence (-) of the HIV-1 transcriptional Transactivator protein Tat. Luciferase expression was quantified by qRT-PCR assay using β-Actin as housekeeping gene. Values were normalised to Mock (the first column) with n=3, SEM. ANOVA with Bonferroni’s post-test comparisons (see arrows): ns (p > 0.05); *** (p <0.0001).

3.3.3 Other controls used in optimising qRT-PCR assays

Negative controls (without cDNA template) were included in every qRT-PCR assay for each primer set to control for contamination in the reaction. In addition, negative-RT controls (without Reverse transcriptase) were included regularly to check for DNA contamination during reverse transcription of RNA samples. Finally, the PCR product amplified by each set of primers (i.e. for GAPDH, SPTBN1, TMED2, KIAA1012 and SP110) was cloned into T/A vector pTZ57R/T to create target plasmid standards. A log_{10} dilution (from 0.1 pg to10 ng) of each plasmid was then amplified by qPCR to create standard curves for each gene (Figure 3-6). As 10 ng of each target plasmid was amplified with a cycle threshold (Ct) value of 12, this amount was regularly included in assays as a standard to detect inter-assay variation. The standards continued to amplify
at this Ct value (even using $\beta$-Actin as reference gene), giving support for the reliability and consistency of the assays. With the necessary controls in place, verification of shRNA knockdown of the four remaining genes could be done.

![Standard curve of plasmid standards for each primer set](image)

**Figure 3-6: Standard curve of plasmid standards for each primer set**

The PCR product amplified by each gene’s primer set was cloned into a T/A plasmid and used to create a log standard curve for each primer set. The Ct value represents the cycle threshold, the number of PCR cycles at which amplification becomes exponential.

### 3.4 shRNA-mediated knockdown of endogenous mRNA targets

The ability of the designed shRNAs to knock down their endogenous mRNA targets was verified in both 116 cells and TZM-bl cells by qRT-PCR. Representative data are shown in Figures 3-7a and b, 3-8, 3-9a and b and 3-10. Figure 3-7 shows knockdown of *SPTBN1* endogenous mRNA. Although the sample means of shSPTBN-1 and shHBx in Figure 3-7a show almost a two-fold decrease from the Mock, statistical analysis of intra- and inter-group variation made it clear that there is no significant knockdown ($p > 0.05$). Figure 3-7b (using *GAPDH* as housekeeping gene) provides support for the lack of significant knockdown of *SPTBN1*. Retrospective analysis of the sequences encoding shSPTBN-2 and shSPTBN-3 showed an error in design. As the luciferase reporter targets
were thus designed and cloned based on incorrect shRNA sequence information, knockdown of these targets could be obtained but not the endogenous mRNA.

Figure 3-7: Knockdown of endogenous SPTBN1 mRNA by shRNAs.

a) SPTBN1 knockdown in 116 cells by qRT-PCR assay using β-Actin as housekeeping gene. b) SPTBN1 knockdown in TZM-bl cells by qRT-PCR assay using GAPDH as housekeeping gene. Data normalised to Mock (n=3, SEM). ANOVA with Dunnett’s post-test comparison: ns (p > 0.05); (single data point, no statistical analysis done).

As shown in Figures 3-8 and 3-9a and b, statistically significant (i.e. P < 0.05) knockdown of TMED2 and KIAA1012 mRNA could not be achieved, though the shRNA design and sequence was correct. In all cases, the maximum amount of shRNA was transfected with good efficiency in either 116 cells or TZM-bl cells. SP110 was the only gene that could be suppressed by a shRNA. SP110 mRNA was decreased relative to controls by ~60% by shSP110-2 (Figure 3-10, pg 70). This was the only functional shRNA that could be taken forward.
Figure 3-8: Knockdown of endogenous *TMED2* mRNA by shRNAs.

*TMED2* knockdown in 116 cells by qRT-PCR assay using β-Actin as housekeeping gene. Data normalised to Mock (n=3, SEM). ANOVA with Dunnett's post-test comparison: ns (p > 0.05).

Figure 3-9: Knockdown of endogenous *KIAA1012* mRNA by shRNAs.

a) *KIAA1012* knockdown in 293T cells by qRT-PCR assay using β-actin as housekeeping gene. b) *KIAA1012* knockdown in TZM-bl cells by qRT-PCR assay using GAPDH as housekeeping gene. Data normalised to Mock (n=3, SEM). ANOVA with Dunnett's post-test comparison: ns (p > 0.05).
Figure 3-10: Knockdown of endogenous \textit{SP110} mRNA by shRNAs.

\textit{SP110} knockdown in TZM-bl cells by qRT-PCR assay using $\beta$-actin as housekeeping gene. Data normalised to Mock ($n=3$, SEM). ANOVA with Dunnett's post-test comparison: ns ($p > 0.05$); ** ($p < 0.01$).

### 3.5 Knockdown of endogenous mRNA by siRNA SMARTpools

As the majority of the designed shRNAs failed to effect significant knockdown of their endogenous targets, a different approach was required. Commercially available siRNA SMARTpools (Dharmacon, CO, USA) comprising four individual siRNAs per gene were purchased that are specially designed and modified for enhanced target specificity and are guaranteed to effect specific target knockdown of 75% or more (Section 4.6.4). Only siRNAs targeting \textit{SPTBN1} and \textit{SP110} were selected, as these genes seemed the most promising of the five originally chosen (Table 3-1). Figure 3-11 shows the knockdown efficacy of the SMARTpool siRNAs targeting \textit{SPTBN1} (a) and \textit{SP110} (b) over a range of siRNA concentrations, in comparison to a non-specific siRNA control (targeting HBV) (Section 2.1.3). The SMARTpool siRNAs successfully knocked down both genes by 90% or more (at 100 nM concentration). Moreover, knockdown efficacy decreased in proportion to the amount of siRNA transfected. This dosage response supports RNAi as being the mechanism for the observed silencing.
Figure 3-11: Knockdown of endogenous SPTBN1 mRNA by siRNA SMARTpools.

a) Knockdown of SPTBN1 (a) and SP110 (b), respectively, by qRT-PCR assay using β-actin as housekeeping gene. b) TZM-bl cells were transfected with Dharmacon® SMARTpool siRNAs at 100, 50, 25, 10 and 5 nM concentrations. Data normalised to Non-specific control siHBV (n=3, SEM). ANOVA with Dunnett’s post-test comparison: ns (p > 0.05); * (p < 0.05); ** (p < 0.01).
3.6 Knockdown efficacy of shRNAs derived from pooled siRNAs

The knockdown described above was mediated by siRNAs that were introduced into the cell cytoplasm. The activity of siRNAs is typically potent but short-lived: they are diluted during cell division and degraded over time. In contrast, shRNAs can be constitutively expressed from plasmids at a high rate (e.g. using the U6 Pol III promoter). This ensures a constant, large supply of RNAi effectors that can mediate longer-lasting knockdown than siRNAs can. For this reason, new shRNAs were designed based on the sequences of the siRNA pools (designated with the suffixes A-D). Because of the difficulties encountered during cloning of the original shRNAs, reverse primers were designed that incorporated the entire shRNA sequence upstream of the primer binding site. All the shRNAs could be cloned correctly, except for shSP110-D that had a single incorrect base in the promoter region. The knockdown efficacy of these shRNAs is shown in Figure 3-12. Though knockdown of their endogenous targets was statistically highly significant (P < 0.01), the difference in efficacy (approximately two-fold less) between the pooled siRNAs and their individual shRNA versions is striking. The inability of shSPTBN-A to knock down its target mRNA may be a result of target site accessibility (Section 4.6.2) or inefficient processing by Dicer.
Figure 3-12: Knockdown of SPTBN1 and SP110 by shRNAs derived from SMARTpool siRNAs

Knockdown of SPTBN1 (a) and SP110 (b), respectively, by shRNAs derived from pooled siRNA target sequences. Gene expression quantified by qRT-PCR assay using B-actin as housekeeping gene. Data normalised to Mock (n=3, SEM). ANOVA with Dunnett’s post-test comparison: ns (p > 0.05); ** (p < 0.01).
3.7 Delivery to SupT1 cells by electroporation

One of the objectives of this project involved optimising delivery of RNAi effecters to SupT1 cells, a T-cell line that grows in suspension. Suspension cells are notoriously difficult to transfect using conventional transfection reagents. Viral vectors offer one alternative for delivering shRNAs to T-cells. Various types of viral vectors, their advantages and drawbacks are discussed in Sections 4.6.7.2 and 4.6.7.3. Another delivery method is electroporation. This well-characterised technique is simple and rapid and widely used for introducing nucleic acids and protein molecules into primary cells, suspension cells and other cell types that are refractory to transfection. Electroporation entails delivering a pulsed electrical field across the cellular membrane, creating water-filled pores that rapidly open to allow a highly charged molecule like DNA a short period of time (~ 20 ms) to enter the cell before resealing (Langridge et al., 1987).

Electroporation of SupT1 cells was attempted with eGFP-expressing plasmid using various protocols as outlined in Section 2.8 and Tables B-1 and B-2 (Appendix B). The efficacy of each experiment is summarised in Table 3-2, with values representing the average of three data points. The number of dead cells (staining blue) and the number of transfected cells (expressing GFP) are shown as a proportion of the total number of cells. In some cases very few cells survived the procedure – even after 24 hours outgrowth their numbers were too few for accurate quantification of cell viability or electroporation efficacy. While the different protocols tested resulted in varying levels of cell toxicity, the overall efficacy of the technique was disappointingly low. Thus, this technique was not be used in this study.
Table 3-2: Results of electroporation experiments (average of data in triplicate)

<table>
<thead>
<tr>
<th>Protocol</th>
<th>Average cell death</th>
<th>Average transfection efficacy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Jurkat/1</td>
<td>Accurate quantifications not possible – cells too sparse.</td>
<td></td>
</tr>
<tr>
<td>Jurkat/2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Jurkat/3</td>
<td>17%</td>
<td>2.6%</td>
</tr>
<tr>
<td>Jurkat/4</td>
<td>24%</td>
<td>2.5%</td>
</tr>
<tr>
<td>Jurkat/5</td>
<td>0%</td>
<td>0%</td>
</tr>
<tr>
<td>Jurkat/6</td>
<td>0%</td>
<td>0%</td>
</tr>
<tr>
<td>1/1</td>
<td>91%</td>
<td>5%</td>
</tr>
<tr>
<td>1/2</td>
<td>100%</td>
<td>0%</td>
</tr>
<tr>
<td>2/1</td>
<td>47%</td>
<td>1%</td>
</tr>
<tr>
<td>2/2</td>
<td>45%</td>
<td>2%</td>
</tr>
<tr>
<td>3</td>
<td>Accurate quantifications not possible.</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

3.8 Delivery to SupT1 cells by transfection reagents

When DharmaFECT-4 transfection reagent was tested for its ability to deliver a Cy3 fluorescently-labelled siRNA to SupT1 cells at three different cell densities (Section 2.9), a moderate fluorescent signal could be observed that co-localised with individual cells (Figure 3-13b). Moreover, ESCORT-V transfection reagent could deliver GFP-expressing plasmid DNA to SupT1 cells, as shown in Figure 3-13a.
Figure 3-13: Delivery of plasmid DNA and siRNA to SupT1 cells

SupT1 cells transfected with GFP-expressing plasmid by ESCORT-V transfection reagent (a) and with Cy3-labelled siRNA by DharmaFECT-4 transfection reagent (b) were visualised by fluorescence microscopy 48 hours after transfection. The images represent 100 × magnification, except for the first picture in a) at 200 × magnification, with scale bars indicated in white for each magnification.

Though the viability of the samples was not directly measured, there were no visible detrimental effects compared to untreated cells. The functionality of the siRNAs
and shRNAs transfected into SupT1 cells was subsequently tested by qRT-PCR. Neither the siRNA SMARTpools (that had yielded impressive knockdown in TZM-bl and 293T cells), nor the two best shRNAs targeting each gene gave any significant knockdown in the SupT1 cell line (Figures 3-14 and 3-15). This indicates that the intracellular delivery of both RNAi effectors by transfection reagent is not high enough to yield functional knockdown.

**Figure 3-14: Knockdown by siRNA SMARTpools in SupT1 cells**

Knockdown of *SPTBN1* (a) and *SP110* (b), respectively, in SupT1 cells transfected with siRNA SMARTpools at 100 nM and 50 nM concentrations using DharmaFECT-4 transfection reagent. Gene expression quantified by qRT-PCR assay using *B-actin* as housekeeping gene. Data normalised to siHBV non-specific control (n=2, SEM). ANOVA with Dunnett’s post-test comparison: ns (p > 0.05).
The best shRNAs targeting *SPTBN1* (shSPTBN-D) and *SP110* (shSP110-2) were tested for their knockdown ability after delivery to SupT1 cells by ESCORT-V transfection reagent. Gene expression quantified by qRT-PCR assay using *B-actin* as housekeeping gene and normalised to Mock (n=3, SEM). ANOVA with Dunnett’s post-test comparison: ns (p > 0.05).

As the RNAi effecters could not be functionally introduced (by either electroporation or transfection reagent) into a cell line that can be infected by HIV-1 and that is also closely related to the virus’ natural host cell type (CD4+ T helper cells), the rest of this study focused on validating the data produced by Brass *et al.* (2008) in TZM-bl cells. However, before proceeding with viral challenges of the transfected cells, it was important to ascertain whether the shRNAs and siRNAs developed produced any cytotoxic effects when introduced into cells.

### 3.9 Cytotoxicity assay

All the functional RNAi effecters used in this study, as well as their respective controls, were transfected into TZM-bl cells and the resultant cytotoxicity measured relative to untreated cells (Section 2.10). In addition, cells were treated with several concentrations of TSA as a positive control. As shown in Figures 3-16, none of the shRNAs nor their controls showed significant toxicity when compared to untreated...
controls. However, the siRNA non-specific control (siHBV) showed highly significant toxicity (P < 0.01), comparable to the lowest concentration of the positive control (TSA 100 nM). As the siRNA SMARTpools do not show a significant difference in toxicity to siHBV, this implies that siRNA treatment in general is causing cytotoxicity. This is further supported by the qualitative observation that the cells transfected with siRNAs appeared to grow slower (data not shown) than those transfected with shRNAs. As the toxicity associated with siRNA transfection appeared to be slight and not specific to an individual siRNA’s effects, we decided that sufficient controls were in place to continue with viral challenge assays.

Figure 3-16: MTT assay for RNAi-related cytotoxicity

TZM-bl cells were treated with HDAC inhibitor Trichostatin A (TSA) at various concentrations as a positive control and transfected with the siRNA-derived shRNAs, siRNA pools and their respective controls. Mitochondrial Reductase activity was measured as solubilised formazan production (absorbance at 570 nm) relative to background absorbance (655 nm). Data were normalised to the mean of the untreated samples (n=3, SEM). ANOVA with Dunnett’s post-test comparisons with the untreated control (dashed lines) or with Mock and siHBV (solid lines): ns (p > 0.05), ** (p < 0.01).
The effect of SPTBN1 and SP110 knockdown on HIV-1 replication

Single-round challenge of transfected cells with pseudovirus

An initial screen was set up in TZM-bl cells after 72 hours of gene knockdown by shRNAs (Figure 3-17) and siRNAs (Figure 3-18, pg 82) targeting SPTBN1 and SP110, as well as their respective controls (Section 2.13.1). As shown in Figure 3-17a, the control shRNAs shLTR and shTat (targeting the viral LTR promoter and Tat transactivator-encoding genes) and shHTATSF-2 and shLEDGF (targeting verified HIV-1 host factors, HTATSF1 (Section 1.2.6) and LEDGF (Section 1.2.4)) inhibit the virus to their respective expected levels. shLEDGF does not show significant viral inhibition as incomplete gene knockdown ensures sufficient gene product remains for HIV-1 to replicate unhindered (Vandekerckhove et al., 2006).

The effects of shRNA-mediated knockdown on pseudovirus infection shown in Figure 3-17b are not statistically significant. This was confirmed by additional statistical analysis excluding the variable samples shSP110-B and shSP110-C (data not shown). This result is not surprising considering the weaker knockdown efficacy of the shRNAs compared to the siRNA pools. For this reason, the shRNAs were not used in the subsequent viral challenge assay. The challenge assay described above included samples transfected with siRNA (Figure 3-18). As the siRNA- and shRNA-treated cells were part of the same experiment, the shRNA controls (Figure 3-17a) provided support for the validity of the siRNA challenge data in the absence of corresponding siRNA controls. In contrast to the shRNA-treated cells, the data in Figure 3-18 show a significant decrease in viral activity when SPTBN1 is knocked down, though only at the highest siRNA concentration (100 nM). There was a significant difference between the untransfected Virus Control samples and those treated with the non-specific control siRNA targeting
HBV. This indicates some repressive effect of the siRNA transfection procedure on virus replication and/or on cell viability, as indicated by the MTT assay (Figure 3-16).

Figure 3-17: Changes in pseudovirus activity in cells transfected with shRNA

a) TZM-bl cells were transfected with shRNA controls and infected 72 hours later with HIV-1 pseudovirus (expressing a Subtype C envelope). Forty-eight hours later, LTR-Luciferase activity was measured relative to background Renilla Luciferase. Virus Control and No Virus control wells were not transfected and only assayed for LTR-Luciferase activity. All data were normalised to non-specific shRNA control, shHBx, and analyzed by ANOVA with Dunnett’s post-test comparison (n=3, SEM): ns (p > 0.05), * (p < 0.05), ** (p < 0.01). b) The same experiment as in a) showing the shRNA data.
TZM-bl cells were transfected with siRNA SMARTpools (at 100 nM and 50 nM concentrations) and siHBV control. After 72 hours of knockdown, cells were infected with pseudovirus (except for No Virus control wells). Forty-eight hours later, LTR-Luciferase activity was measured. All data (n=3, SEM) normalised to siHBV. ANOVA with Dunnett's post-test comparison with siHBV: ns (p > 0.05), ** (P < 0.01).

### 3.10.2 Two-part challenge of transfected cells with replication-competent virus

After the initial screen for the requirement of each gene for pseudovirus activity, siRNA-transfected cells were challenged with fully replicating virus. The assay was set up in two parts (Figure 3-19) to mimic as closely as possible the experiments carried out by Brass et al. (2008). The virus used, HIV-1p81A-4 (derived from the molecular clone p81A-4) is capable of a full round of replication, producing assembled virions (during Part 1) that can infect fresh cells in a second round of infection (Part 2) (Sections 2.11 and 2.13.2). As shown in Figure 3-19b, only SPTBN1 was shown to have a significant effect on viral infection. This effect was only seen at the highest siRNA concentration (100 nM) and supports the pseudovirus challenge data (Figure 3-18). In addition, the gene appears to be required during a later stage of infection as only the Part 2 data shows a statistically
significant (P < 0.01) effect. However the sample means in Part 1 and 2 follow the same trend (compare Figure 3-19a and b). SP110 knockdown did not negatively affect TZM-bl cell infection by HIV-1p81A-4 (during either Part 1 or 2). This also supports the data from the pseudovirus challenge assay. Thus, in summary, SPTBN1 knockdown decreases HIV-1 replication by 60-75% (P < 0.01), seemingly during the late stage of infection. While Brass et al. (2008) showed a similar decrease in viral activity (76%), they indicated an early-stage role for SPTBN1. In addition, this study found no indication for SP110 as a HIV-1 host factor, while Brass et al. showed that SP110 knockdown caused a 84% virus reduction during the later stages of the viral lifecycle.

![Figure 3-19: Changes in HIV-1p81A-4 activity in cells transfected with siRNAs](image)

**Figure 3-19: Changes in HIV-1p81A-4 activity in cells transfected with siRNAs**

a) TZM-bl cells transfected with siRNA SMARTpools targeting SPTBN1 and SP110, challenged with HIV-1p81A-4 72 hours post-transfection, and assayed for LTR-Luciferase activity 48 hours later (Part 1). b) Fresh TZM-bl cells infected with viral supernatant from Part 1 cells and assayed for LTR-Luciferase activity 20 hours later (Part 2). All data (n=3, SEM) normalised to siHBV. ANOVA with Dunnett’s post-test comparison with siHBV: ns (p > 0.05), * (P < 0.05).
CHAPTER 4: DISCUSSION

4.1 The possible roles of SPTBN1 in HIV-1 replication

This study has shown that SPTBN1, but not SP110, is required for infection and early-stage replication of a pseudovirus expressing Subtype C envelope in TZM-bl reporter cells. An early role in viral replication is supported by Brass et al. (2008) who showed viral suppression of 76% during Part 1 of their assay by 2 individual siRNAs. In contrast, this study’s subsequent challenge assay with fully replicating virus showed a greater requirement during the later stages (post-translation up to assembly and budding of new virions). An important difference between this study and that published by Brass et al. (2008) is that Part 1 of their study was assessed by p24 assay (i.e. from viral entry to Gag translation), while this study measured LTR-Luciferase activity (which covers viral entry to Tat-induced transcription and reporter translation). Thus post-transcriptional modifications, export and translation of viral transcripts (apart from Tat and Rev) is not included in this study’s Part 1. A “late-stage” classification in this study could translate into an “early-stage” role in terms of Brass et al.’s delineation of the lifecycle. In addition, the Part 2 data of Brass et al. (2008) was normalized to the number of Part 1 donor cells. This was not possible in this study and could affect the correlation between Part 1 and 2 samples. In spite of the above complications, the detection of this gene in all three challenge assays merits some discussion on its potential role(s) in HIV-1 replication.

As mentioned in Section 1.6.1.1, spectrins crosslink actin, organize transmembrane proteins and regulate stability and shape of the plasma membrane. β-spectrins interact with a wide array of membrane proteins to form a junctional complex and a major attachment site for the cytoskeleton to the cell membrane (Mouro-Chanteloup et al., 2003; Nicolas et al., 2003). These interactions occur via direct binding of various
forms of phospholipids (Diakowski et al., 1999). In addition, spectrins may bind calmodulin to cause calcium-dependant movement of the cytoskeleton at the membrane. As described in Sections 1.2.1, 1.2.3 and 1.2.10, multiple associations occur between the cytoskeleton and HIV-1 during both entry and exit from the cell, as well as during trafficking to and from the nucleus [reviewed by (Naghavi and Goff 2007)]. Thus, both early and late roles for SPTBN1 in HIV-1 replication are plausible.

4.2 **SP110 and HIV-1 replication**

SP110 was not detected as a host factor in either of the two screens performed here (in spite of almost-complete knockdown in the same cell line used by Brass et al., (2008). As mentioned in Section 1.6.1.5, this gene is recruited by EBV and stabilizes A. phagocytophilium gene expression. Thus a possible role in HIV-1 replication is not unlikely. As a retinoic acid receptor transcriptional co-activator, with a C-terminal region that is homologous to the Transcription Intermediary Factor 1 (TIF1) family of proteins, it's most likely role would be during viral transcription (i.e. Part 1). In contrast, Brass et al. (2008) found only a post-Gag translation requirement for SP110.

SP110 is also induced by the Interferon response (Regad and Chelbi-Alix 2001). This innate immune pathway functions to inactivate viral replication. In response, several viruses, including HIV-1 (Gunnery et al., 1990; Martinand et al., 1999) have evolved means to inhibit this cellular process. Possibly, HIV-1 interacts with SP110 to further regulate this response. Apart from these factors, high SP110 expression is restricted to peripheral blood leukocytes (PBLs) and the spleen. Although its expression in HeLa and 293T cells could be detected by qRT-PCR in this study, this gene’s involvement in the Interferon response, specifically during HIV-1 infection, may only be active in immune cells (e.g. PBLs (peripheral blood leukocytes), which interestingly are also the natural
targets of HIV-1). Thus, the inconclusive data reported for this gene in this study and by Brass et al. (2008) highlights the risks of using cellular models that do not represent the biological context of the process being studied.

4.3 Comparison of viruses used in this study and by Brass et al. (2008)

As TZM-bl reporter cells and siRNA SMARTpools were used in both studies, the discrepancy in results seems to reflect differences between the viruses. As the CAP210.2.00.E8 pseudovirus and HIV-IIIB (HTLV-IIIB/H9 from Dr Robert Gallo, NARRRP Catalogue Reagent 398, used by Brass et al., 2008) data agree on an early requirement for SPTBN1, while the HIV-1p81A-4 data indicate a later stage of involvement, we would expect the pseudovirus and HIV-IIIB to have common properties (particularly regarding cellular entry) that differ from those of HIV-1p81A-4. In contrast, HIV-IIIB is T-cell tropic (i.e. using CXCR4 as entry co-receptor) while both the Env-pseudotyped virus and HIV-1p81A-4 are macrophage-tropic (i.e. using CCR5 co-receptor), the latter due to the insertion of a BaL strain V3 loop region into the envelope of NL4-3 virus (Chesebro et al., 1992). Moreover, HIV-IIIB and HIV-1p81A-4 are both Subtype B viruses while the pseudovirus has a Subtype C envelope. Thus these aspects of the viruses do not account for the results. Other differences between the viruses include the fact that the lab-adapted strain of HIV-IIIB used by Brass et al. (2008) is deficient for the viral accessory genes nef and vpu, while a molecular clone (HXBC2, Genbank accession number K03455) derived from this strain encodes a truncated form of Vpr. It is interesting that the pseudovirus in this study is also lacking vpu, while HIV-1p81A-4 has an intact genome, except for the V1-V3 region substitution that changes its cell tropism. Since a Vpu-deficiency is common to the pseudovirus and HIV-IIIB, this loss-of-function may explain the pattern of results obtained.
Vpu has several functions in viral pathogenesis [recently reviewed by (Nomaguchi et al., 2008)]. Briefly, Vpu is required for virus maturation and release (Klimkait et al., 1990). Deletion of vpu was shown to cause retention of virus at the plasma membrane, as a result of host Tetherin that is normally degraded by Vpu (Neil et al., 2008), and increased cytopathic effects (Iwatani et al., 1997). These cytopathic effects were found to result from association between virions / envelope protein with abnormally high levels of surface CD4 receptor, which is normally degraded by Vpu (Willey et al., 1992). Down-regulation of CD4 has several functions, one being to prevent superinfection of infected cells (Wildum et al., 2006). Loss of these functions may have had significant implications on the pseudovirus data obtained in this study and the findings of Brass et al. (2008). While their study reported no change in CD4 surface expression after Rab6 gene knockdown, they did not take into account the effect of HIV-IIIB infection on CD4 levels. Tetherin-mediated membrane retention of virus (in the absence of Vpu) could mask a late-acting function of SPTBN1 (and other membrane proteins) in aiding the release of virus particles. Alternatively, superinfection (resulting from increased CD4 levels) would place a greater burden on viral entry pathways, possibly making an early requirement for SPTBN1 more pronounced than during a Vpu-regulated infection.

As Nef also functions in down-regulating CD4 surface expression, the lack of both vpu and nef in HIV-IIIB may explain the stronger requirement for SPTBN1 observed by Brass et al. (2008). Nef also directly interacts with the Actin cytoskeleton to aid post-fusion trafficking (Campbell et al., 2004). Thus the loss of a putative SPTBN1 function at the membrane during entry may not be as debilitating in the presence of Nef (as observed in the pseudovirus data). Finally, cytopathic effects induced by vpu deletion could have diverse effects on cellular functions, making the interpretation of any one gene’s role in cellular or viral processes inaccurate.
4.4 Gene selection

The criteria comprising the selection process used in this study to compile a list of individual genes representing both the early and late stages of the HIV lifecycle are logically and biologically sound. However, as three of the five selected genes could not be fully studied (PRDM14 because of low expression levels in the cell lines used; KIAA1012 and SP110 because of target mRNA inaccessibility to the shRNAs) in viral challenge assays, their status as HDFs is inconclusive. Such obstacles may be prevented in future studies by using online databases such as the NCBI Gene Expression Omnibus (GEO) Dataset that provides sequence- and microarray-based expression data in a range of cellular types. Moreover, bioinformatic measures of target site accessibility should be incorporated into the shRNA design process (see Section 4.6.2) or genes selected for which effective RNAi effecters (at the mRNA or protein level) have been published.

4.5 Broader perspectives

4.5.1 Large-scale lentiviral shRNA transduction of T-cells

A recent study by Yeung et al. (2009) used a lentiviral shRNA library to transduce Jurkat T-cells before infection with full-length NL4-3 virus. Such a study is far more valuable in representing a natural T-cell infection than siRNA screens in HeLa or 293T cells using lab-adapted viral strains. This study (Yeung et al., 2009) identified 252 host genes whose knockdown did not affect T-cell viability (over 3 weeks of selection) and that protected the cell from lytic viral replication. The candidates identified by Yeung et al. (2009) have great potential as future drug targets since cytotoxic effects resulting from long-term knockdown has already been ruled out, while the ability of each gene to completely protect the cell from virus-induced lysis (over 4 weeks of infection) has also
been shown. In addition, reversal of viral suppression could be achieved by reconstituting the inhibited gene product (for three genes), thus proving the direct and specific effects of the knockdown. A disadvantage of such an experiment is the incomplete genome coverage resulting from incomplete transduction efficiency. Also, Yeung et al. (2009) acknowledge several factors that can result in false negative and false positive results. However, by repeating such a screen several times, both coverage and accuracy can be improved. Moreover, by repeating such a screen in primary T-cells and macrophages (that are refractory to siRNA transfection but not lentiviral transduction), a shortlist of candidates can be assembled that can be confirmed by detailed studies of specific viral processes.

On comparing their data with the three siRNA genomic screen datasets (Brass et al., 2008; Konig et al., 2008; Zhou et al., 2008), Yeung et al. found that their 252 genes were mostly distinct. However, the cellular pathways used by HIV-1 seem to be conserved across the studies. This is interesting, considering the differences in cell lines, viral strains, study design and detection assays used. They suggest that further study of these common cellular pathways may be more informative than focusing on individual genes identified. This confirms the need for using microarray analyzes of total gene expression profile changes, as well as robust and powerful bioinformatic algorithms that take into account protein interactions and molecular pathway networks in order to detect overlaps and patterns that might easily be missed when dealing with such large datasets. This study has shown that there are many reasons for which study of an individual gene might be discontinued. By studying the interacting proteins and functional gene networks of such a gene in parallel, such obstacles would have little impact on the investigative potential of the study. In addition, the results would be more robust and informative, with data on multiple interacting genes providing a biological context that does not exist in single gene studies.
4.5.2 Alternative techniques for characterizing the viral/cellular interactome

The task of documenting the viral/cellular interactome is vast and complex. By combining the data obtained from a wide variety of approaches, there is a greater likelihood that the final picture will be a true reflection of such a dynamic process. Moreover, because these interactions are dynamic, using only two time-points for measurements (as used in this study and in all four genomic RNAi screens mentioned) would not detect subtle and transient changes. Thus kinetic analyzes of several viral markers (for example, LTR-expression, p24-Gag levels, CD4 receptor levels and total viral load) over time would be more informative.

Several novel technologies have been developed for high-throughput analysis of protein interactions that are useful in biochemically validating the interaction data from standard yeast-two-hybrid (Y2H) screens. These technologies [reviewed by (Bailer and Haas 2009)] include: nucleic acid programmable protein arrays, NAPPA (Ramachandran et al., 2004), that transcribe and translate DNA printed on array chips for co-immunoprecipitation (CoIP) assays; luminescence-based mammalian interactome mapping, LUMIER (Barrios-Rodiles et al., 2005), that uses luciferase-fusion proteins to detect protein interactions; protein fragment complementation assays, PCA (Nyfeler et al., 2005), based on the reconstitution of a fluorescent protein when its domains (fused to the proteins of interest) interact and finally, mammalian protein-protein interaction trapping, MAPPIT (Eyckerman et al., 2001), to detect the transcriptional activity of an endogenous reporter gene that is activated by a signaling cascade when the proteins of interest (fused to the membrane receptor and signal molecule, respectively) interact.

Other approaches to mapping the viral/cellular interactome include microarray analyzes of HIV-induced changes in the host transcriptome and proteome (Geiss et al., 2000), shotgun liquid chromatography-tandem mass spectrometry (Chan et al., 2009), as
well as high-resolution microscopy techniques. The many applications of the latter are particularly exciting for viral research (Arhel et al., 2006; Brandenburg et al., 2007). PALM/STORM (photoactivation localization microscopy/stochastic optical reconstruction microscopy) combines super-resolution microscopy with superior algorithms for visualizing molecules at the nanometer scale and detecting individual fluorescent signals in three dimensions over time (Bates et al., 2007; Huang et al., 2008). In addition, single particle tracking and new molecular beacons have been used to study RNA localization and trafficking as effectively as the GFP reporter allows protein visualization (Bratu et al., 2003; Mhlanga et al., 2005). The application of such technologies to future studies of viral replication and infection will enable the real-time detection of individual reporter signals in single live cells. This will greatly reduce the inaccuracies of previous experiments where averaged static signals across cell populations were used, potentially masking multiple subtle variations that can now be more fully explored.

4.6 Factors of RNAi experiments highlighted by this study

4.6.1 Primer design in PCR-based cloning of shRNAs

The variable ease with which each shRNA could be cloned (using the two-step PCR approach) may be because of specific sequence combinations that make hairpin formation thermodynamically more favorable than template/primer binding. The second-round PCR reverse primers were designed to bind the loop region as well as part of the sense strand of the first-round product, allowing hairpin formation to occur in both the primer and the template. In retrospect, the original primer design suggested by (Castanotto et al., 2002), using reverse primers that are complementary to each other over the loop region only, would have avoided this hairpin formation. The comparatively greater success at cloning the siRNA-based shRNAs (using single-round PCR with full-
length primers, where the template did not contain hairpin-forming sequences) gives support for using this approach in future, as no optimization or additional reagents was required for efficient PCR amplification.

4.6.2 Target mRNA accessibility to the RNAi machinery

The ability of the original set of shRNAs to knock down Luciferase reporter targets by at least 75% (with most affording over 90% knockdown) indicates efficient Dicer processing of the guide strand and RISC activation. In addition, the fact that a non-targeting shRNA showed no effect on each reporter target supports this knockdown as being sequence-specific. The fact that such efficiently processed and highly functional shRNAs had no significant effect on their endogenous mRNA targets (excepting shSP110-2) is likely to be a result of the inherent properties of the target mRNA. Both bioinformatic modeling and experimental studies have shown that the local secondary structure of target mRNAs affects the biological activity of siRNAs by preventing loaded RISCs from accessing the site (Bohula et al., 2003; Kretschmer-Kazemi Far and Sczakiel 2003; Overhoff et al., 2005). Indeed, one way in which HIV-1 can evade RNAi is by altering its secondary structure to hinder target site accessibility, which in turn lowers the stability of the siRNA/target duplex (Westerhout et al., 2005). Moreover, a bioinformatic analysis of 100 siRNAs and 101 shRNAs targeted to 103 endogenous human genes (Shao et al., 2007) predicted that siRNA/shRNA efficacy could be improved by 40% just by ensuring target accessibility. Shao et al. (2007) described a measure of target structural accessibility (termed ΔG_{disruption}) that was found to be the most important predictor of RNAi efficacy. Such factors would be given more attention in future shRNA designs.
4.6.3 The importance of reliable reference genes and controls in qRT-PCR assays

Target site accessibility cannot be predicted by assays like the Dual Luciferase Reporter Assay used in this study. The oligodeoxyribonucleotide/RNase H method has been used effectively to give a direct indication of target accessibility (Scherr and Rossi 1998; Lee et al., 2002), though it cannot measure overall knockdown efficacy. A simpler alternative to assess the RNAi activity on the endogenous mRNA is to use qRT-PCR (used here to address the second objective of the study). This technique (if set up correctly) is both sensitive and accurate, allowing the detection of changes in RNA levels over a large dynamic range.

The qRT-PCR data obtained in this study clearly shows the importance of having the correct controls in place. In this study, several negative and positive controls were tested: the no-template and negative-RT controls, target plasmid standards, as well as melting curve analysis, all served to ensure that the correct template (cDNA) was being amplified during each assay, with no contamination and with the same efficiency. Apart from these, two different housekeeping genes (β-actin and GAPDH) were compared for their reliability as references for normalizing mRNA levels. β-actin and GAPDH have been used as single reference genes in more than 90% of studies (Vandesompele et al., 2002). In spite of this, several studies on the reliability of commonly used housekeeping genes have shown that neither are in fact optimal reference genes (Zhong and Simons 1999; Selvey et al., 2001; Glare et al., 2002; Watson et al., 2007). Their expression varies across several tissue types tested and are affected by cell cycle activation (Radonic et al., 2004). This is of concern as the requirement for a reference gene in mRNA detection is that the gene’s expression is as constant as biologically possible, without being affected by experimental procedures.

In this study, where RNA was amplified by qRT-PCR from two cell lines, β-actin consistently proved to be more reliable than GAPDH, with the data normalized to β-actin
showing very little variation between replicate samples and assays. The suitability of $\beta$-actin is further supported by the fact that the knockdown efficacies of the positive control shRNAs (that were previously determined by qRT-PCR) matched the data obtained in this study. The only discrepancy observed (for shTat) is due to differences in experimental setup, with the weaker knockdown observed here being attributable to the far higher target concentration (expressed from a plasmid) than was present when originally tested (expressed endogenously from a stable cell line). Moreover, the levels of LTR-driven reporter gene expression in the absence of Tat are so low that any knockdown would not be detected in comparison to the decrease detected when Tat was added.

4.6.4 Off-target effects and the benefits of Dharmacon’s ON-TARGETplus SMARTpool technology

An important aspect of any RNAi study is non-specific or off-target effects. The only comprehensive way to rule this out is to use microarrays to assess changes in the gene expression profile before and after RNAi treatment. As this option was not feasible in this project, non-specific control shRNAs (targeting the HBV genome and a host gene, $\text{lacZ}$) and a control siRNA (also targeting HBV) were tested for their effects against each of the genes studied. Once their reliability had been ascertained, these were included in every qRT-PCR assay to indicate whether knockdown effects were shRNA/siRNA-specific or not. Finally, the online BLAST tool was useful in screening each shRNA for non-specific matches between the guide sequence and other human gene sequences.

Several aspects of Dharmacon’s ON-TARGETplus SMARTpool technology ensure specific silencing by the siRNAs used in this study. Firstly, chemical modifications in both strands (Jackson et al., 2006; Chen et al., 2008) serve to ensure that the antisense (guide) strand is processed and loaded into RISC and to stabilize target specificity. Secondly, the effect of pooling not only increases knockdown efficacy but also minimizes
off-target silencing by decreasing the concentration of each individual siRNA. Seed region filters and seed frequency analyzes (Birmingham et al., 2006; Anderson et al., 2008) have also been incorporated into the design of each siRNA to prevent miRNA-like off-target silencing of genes with seed sequence matches in their 3’-UTRs. Finally, microarray analyzes of the number of genes’ expression affected by these siRNAs (Anderson et al., 2008) show a marked decrease in off-target effects.

The SMARTselection™ design algorithm, combined with the dual-strand modifications and pooling, not only decrease off-target silencing but are also responsible for the impressive knockdown efficacy of the siRNAs. The increase in knockdown of both genes at the lowest concentration (5 nM - Figure 3-11) is unexpected. A possible explanation is that the siRNA/lipid charge ratio at such a low concentration favours lipoplex formation (lipid molecules complexed with DNA or RNA), enhancing transfection and thus knockdown efficacy. The comparatively weaker knockdown efficacy of the shRNAs derived from the siRNA pools is unsurprising. Not only is the pooling effect lost, as well as the chemical modifications that enhanced both specificity and efficiency in the SMARTpools, but potential variability in Dicer cleavage sites (Vermeulen et al., 2005) could produce guides sequences of slightly different lengths, affecting the ultimate seed sequence of the guide and thus its targets for knockdown. While shRNAs provide stabler, longer-term silencing, with Dicer processing increasing the efficiency of RISC loading (Kim et al., 2005; Siolas et al., 2005), the siRNA ON-TARGETplus SMARTpool technology provides more potent knockdown with fewer off-target effects, which was important for achieving the viral knockdown objective of the study.

4.6.5 siRNA-related cytotoxicity

The third aim of this project was to confirm that no significant toxicity was induced in the treated cells. Trypan-blue exclusion (the classical standard for measuring the
proportion of viable cells in a sample) detects only late apoptotic and necrotic cells where the plasma membrane is sufficiently damaged to allow uptake of the dye, whereas the MTT assay used in this study can detect apoptotic cells and distinguish them from viable and dead cells (Feng et al., 2002). The MTT assay assesses metabolic activity (and thus cell viability) by measuring mitochondrial enzyme activity and cytoplasmic NADH reduction. It is important to ensure that the cells to be tested have a uniform baseline metabolic rate. The cells tested in this study were healthy, proliferating well and not subjected to any extreme environmental changes before the test. They were seeded uniformly from one population and transfected in the same way as during previous knockdown assays. The only observed toxicity (relating to siRNA transfection) may be related to siRNA-induced activation of the innate immune response when certain sequence motifs are recognized by Toll-like Receptor 7/8 (Hornung et al., 2005; Judge et al., 2005). This ultimately leads to sequence-independent suppression of protein synthesis and enhanced mRNA degradation through the activation of many host defense genes (Kariko et al., 2004). However, this is unlikely as siHBV has previously tested negative for Interferon induction (data not shown) while the design and construction of the SMARTpools ensures that such motifs are excluded while including chemical modifications that can avoid such an effect (Sioud 2008).

Alternatively off-target silencing may have down-regulated essential gene functions. As described previously, several preventative measures serve to significantly decrease the number of off-target effects in the ON-TARGETplus SMARTpool siRNAs. Another possibility is that the targeted genes’ loss of function is responsible for the toxicity. In spite of the care taken during selection of these genes, very few of the protein interactions and molecular pathway interactions of the selected genes have been described. As with any study where relatively unknown human genes are targeted, in-depth characterization of the gene in several cell lines would need to be done and more sensitive techniques than the MTT assay used to assess the exact cellular system that is
affected by the knockdown. However, in light of the fact that both siRNA pools and the siRNA control affected cell viability to a similar degree suggests a more general cause of toxicity.

The transfection procedure is a likely candidate for siRNA-associated toxicity. One study, comparing a number of commercially available cationic liposome/lipid-based systems used for siRNA delivery, found that the doses required for maximum transfection efficiency and target knockdown also caused non-specific down-regulation of cellular protein expression (Spagnou et al., 2004). One of these transfection reagents is widely used in RNAi studies because of its high transfection efficiencies across a broad range of cells. In contrast, the siRNA transfection reagent used in this study (DharmaFECT 4) claims to have very low toxicity. Alamar blue viability assays on transfected cells (matching the exact conditions used in this study excepting the cell line) indicated 80% viability post-transfection (data unpublished, Dharmacon). The alamar blue assay is comparable in sensitivity (if not slightly more sensitive) to the MTT assay used here (Hamid et al., 2004).

The only remaining explanation for the observed toxicity is that the untransfected, untreated control sample wells contained more cells than the siRNA-transfected wells because of cells lifting off during the two changes of culture medium in the latter. For the shRNA-transfected wells, the transfection mix was added directly to the medium, which is less disruptive to the cell monolayer. This explanation is supported by the difference in LTR-activity observed between the Virus Control and siHBV-treated samples in the viral challenge assays (Figure 3-18 for an example). In summary, it is unlikely that the siRNA transfection conditions are cytotoxic, with the data probably reflecting unequal baseline signals as a result of the transfection process.
4.6.6 Delivery of RNAi effecters to a T-cell line by electroporation and lipid-based reagents

Another objective of this study was to determine whether shRNAs or siRNAs could be delivered to a T-cell line (that is permissive to natural infection by HIV) using reagents and tools available to us. Unfortunately, electroporation is clearly too harsh a delivery technique, with insufficient cells remaining for several of the tested protocols. If delivery of the RNAi effectors significantly affects cell viability (altering several cellular pathways’ normal functioning), this would severely limit our ability to accurately detect these pathways’ interactions with HIV. Apart from the adverse effects on viability, the delivery efficiency of the protocols tested here was also disappointing. In contrast, the DharmaFECT 4 and ESCORT-V transfection reagents showed better delivery (of siRNA and DNA, respectively) with no visible effects on cell viability. However, delivery was still inadequate for achieving the knockdown potential of the RNAi effecters. To infer the involvement of a gene product in viral replication very good knockdown of gene expression must be achieved and thus highly efficient delivery of the RNAi effecters is an absolute requirement.

4.6.7 Alternative delivery methods

Two alternatives to the delivery methods tested in this study are nucleofection and lentiviral transduction. Both have been reported to be successfully applied (Schroers et al., 2002; Zhou et al., 2003; Maasho et al., 2004; Anderson et al., 2007; ter Brake et al., 2008), however they are considerably expensive. Generation of recombinant lentiviral vectors is also time-consuming and requires some experience.
4.6.7.1 Nucleofection

Nucleofection™ is an adaptation of electroporation that delivers any nucleic acid directly to the nucleus (and cytoplasm) without requiring cells to be dividing. This technique, developed by the Amaxa biotechnology company (Lonza, Cologne, Germany), specializes in delivery to difficult-to-transfect cell types, including primary cells, stem cells, suspension cells and non-dividing cells. Nucleofection™ uses both electrical parameters (which can only be used on the Nucleofector™ device) and a buffer specially formulated for each cell type. The company’s website (www.lonzabio.com) has reports of 5 independent researchers that have successfully transfected SupT1 cells using their technology, with delivery ranging from 30 to 65% and 75-98% cell viability. This technique is promising for transient transfection as it is simple and rapid, with a cell-type specific buffer and existing protocols that have been proven to work. However, in the HIV-1 gene therapy setting stable transfection of cells is important, thus limiting the application of this technique.

4.6.7.2 Lentiviral transduction

Lentiviral transduction has been developing for many years and has been used in a wide variety of settings including delivery of potential gene therapies to inhibit HIV-1 [reviewed by (Escors and Breckpot 2010)]. Lentiviral vectors are attractive for gene therapy because stable integration of the transgene into the genome occurs, allowing permanent knockdown of targeted genes. Moreover, as with nucleofection, transduction of non-dividing, primary and stem cells is possible (Hamilton et al., 2009). Finally, these vectors have low immunogenic properties, making them useful for long-term challenge assays to detect viral resistance and changes in replication kinetics over time [reviewed by (Nayak and Herzog 2009)].
Their great potential in the gene therapy field is illustrated by a triple-therapy vector comprising a TAR decoy, a ribozyme targeted to CCR5 and a shRNA against rev/tat. This gene therapy has proved to be both safe and highly effective against HIV-1 in T-cells/PBMCs (peripheral blood mononuclear cells) (Li et al., 2003), hematopoietic stem cells, HSCs (Li et al., 2005) and in SCID-hu (severe combined immune deficient human) mouse-derived T-cells (Anderson et al., 2007). Effective delivery of this vector to the HSCs is particularly exciting, as these cells give rise to the lymphocytes and other cells of the humoral immune system, many of which are targets of HIV-1 infection. This therapy has been approved by the FDA for participation in two planned human clinical trials.

The drawback of using any integrating viral vector is the risk of insertional mutagenesis occurring - a concern enhanced after three X-SCID (X-linked severe combined immunodeficiency) patients developed leukemia during a gene therapy trial due to onco-retroviral vector integration adjacent to a proto-oncogene promoter (Hacein-Bey-Abina et al., 2003). Since then most studies use lentiviral vectors, as they tend to integrate distally from promoters in introns, lowering the oncogenic risk (Wu et al., 2003) or non-integrating lentiviral vectors [reviewed by (Philpott and Thrasher 2007)] that remove this risk entirely while retaining their high transduction efficiency (particularly in non-dividing cells). Lentiviral vectors have been extensively developed for increased safety (Pauwels et al., 2009). However, in spite of these safety precautions there is always a chance that recombination between the vector and wildtype virus may occur, potentially mobilizing the vector and rendering it replication-competent. However, this risk may be used to our advantage by ensuring that the vector is spread to every cell that becomes infected (Morris and Rossi 2004).
4.6.7.3 Other viral delivery vectors

Apart from lentiviral vectors, other popular viral vectors have been developed from Adenovirus (AdV), and Adeno-associated virus (AAV). AdV vectors (usually derived from serotypes 2 or 5) have a broad tissue-tropism [reviewed by (Lipiec et al., 2009)], can transduce quiescent and proliferating cells and rapidly induce strong transgene expression – important in infectious disease settings. Vector toxicity and immunogenicity (leading to rapid clearance of transduced cells) has been overcome by the generation of “gutless” AdV vectors, from which all adenovirus coding sequences (excepting the inverted terminal repeats, ITRs, and packaging signal, ψ) are deleted (McCaffrey et al., 2008).

AAVs have long-term in vivo stability, up to 6 years (Bankiewicz et al., 2006) in spite of mostly remaining episomal (i.e. non-integrating), which lowers concerns of insertional mutagenesis. The drawback of slow onset of transgene expression by single-stranded AAVs was overcome by generation of self-complementary (sc) double-stranded AAVs that reach maximal expression within several days (McCarty et al., 2003; Wang et al., 2003). The maximum insert size of scAAV vectors (~2.2 Kb), while being prohibitive for most applications, is ideal for expression of RNAi triggers like shRNAs.

In brief, while cloning and packaging high-titer viral vector preparations is a relatively lengthy process requiring specific resources and skills, the many developments and options in this field allow highly efficient transgene delivery to almost any target cell desired, allowing stable expression with few side effects. Moreover, while there are significant concerns relating to safety of these vectors, these continue to be addressed in each new generation of vector development. As a result, viral vectors are the most promising and widely used vectors for delivery of gene therapies.
CHAPTER 5 : CONCLUSIONS

In summary, this study has verified that SPTBN1 is required for HIV-1 replication in a HeLa-derived cell line. However, unlike Brass et al. (2008), we have shown a later (post-transcriptional) stage of involvement. In addition, SP110 (implicated by Brass et al. (2008) as a late-acting HDF) does not appear to have any involvement in the viral lifecycle. These discrepancies between our findings and those of Brass et al. (2008) may reflect the absence of several important viral accessory genes from the virus used in the latter study. As both studies used a HeLa-derived reporter cell line rather than CD4+ T-cells, the need for using physiologically representative viral and cellular models when using RNAi to study virus-host interactions has been highlighted. Moreover, a set of viral markers that represent the viral replication cycle should be measured over several time-points, to ensure that experiments are comprehensive and maximally informative.

This study has accentuated several factors regarding RNAi-based knockdown experiments. Firstly, accessibility of the mRNA target site to RNAi activity is essential. Secondly, off-target knockdown effects, as well as cytotoxicity arising from the RNAi process, the transfection conditions or target knockdown must be carefully considered. Thirdly, while qRT-PCR is an immensely sensitive and powerful technique, it can give misleading results if reliable reference genes and controls are not used. Finally, while T-cells appear to be refractory to efficient transfection by lipid-based reagents and electroporation, advances in viral delivery systems, as well as the development of newer approaches (like Nucleofection), have made RNAi-based research in difficult-to-transfect cells possible.

Future prospects are promising as there have been major advances in our ability to directly visualize any host/viral molecular interaction in high resolution and in real-time within single live cells. These technologies are continually improving, enhancing our
understanding of the complex and dynamic processes that occur during a cellular HIV-1 infection. By complementing existing approaches to confirm and characterize the hundreds of newly identified candidate host factors, the goal of developing novel effective antiviral therapies will be more easily reached.
APPENDIX A: Sequence information

Table A-1: Genomic location of shRNA target sequences

<table>
<thead>
<tr>
<th>Gene targeted</th>
<th>RNAi effector</th>
<th>Transcript code</th>
<th>Target location</th>
</tr>
</thead>
<tbody>
<tr>
<td>SPTBN1</td>
<td>shSPTBN-1</td>
<td>NM_003128.2 (transcript variant 1, TV1)</td>
<td>4864-4884</td>
</tr>
<tr>
<td></td>
<td></td>
<td>NM_178313.2 (transcript variant 2, TV2)</td>
<td>4961-4981</td>
</tr>
<tr>
<td></td>
<td>shSPTBN-2</td>
<td>TV1</td>
<td>6121-6141*</td>
</tr>
<tr>
<td></td>
<td></td>
<td>TV2</td>
<td>6218-6238*</td>
</tr>
<tr>
<td></td>
<td>shSPTBN-3</td>
<td>TV1</td>
<td>1896-1916*</td>
</tr>
<tr>
<td></td>
<td></td>
<td>TV2</td>
<td>1993-2013*</td>
</tr>
<tr>
<td>TMED2</td>
<td>shTMED-1</td>
<td>NM_006815.3</td>
<td>584-604</td>
</tr>
<tr>
<td></td>
<td>shTMED-2</td>
<td></td>
<td>153-173</td>
</tr>
<tr>
<td></td>
<td>shTMED-3</td>
<td></td>
<td>162-182</td>
</tr>
<tr>
<td>KIAA1012</td>
<td>shKIAA-1</td>
<td>NM_014939.3</td>
<td>4394-4414</td>
</tr>
<tr>
<td></td>
<td>shKIAA-2</td>
<td></td>
<td>536-556</td>
</tr>
<tr>
<td></td>
<td>shKIAA-3</td>
<td></td>
<td>1151-1171</td>
</tr>
<tr>
<td>PRDM14</td>
<td>shPRDM-1</td>
<td>NM_024504.2</td>
<td>303-323</td>
</tr>
<tr>
<td></td>
<td>shPRDM-2</td>
<td></td>
<td>1544-1564</td>
</tr>
<tr>
<td></td>
<td>shPRDM-3</td>
<td></td>
<td>691-711</td>
</tr>
<tr>
<td>SP110</td>
<td>shSP110-1</td>
<td>NM_004509.3 (TV1)</td>
<td>2162-2182</td>
</tr>
<tr>
<td></td>
<td></td>
<td>NM_080424.2 (TV3)</td>
<td>2234-2254</td>
</tr>
<tr>
<td></td>
<td>shSP110-2</td>
<td>TV1 + NM_004509.3(TV1)</td>
<td>348-368</td>
</tr>
<tr>
<td></td>
<td></td>
<td>TV1 + TV2 + TV3</td>
<td>471-491</td>
</tr>
<tr>
<td></td>
<td>shSP110-3</td>
<td>TV1 + TV2 + TV3</td>
<td></td>
</tr>
</tbody>
</table>

* These sequences were the intended targets but were not actually targeted due to incorrect insertion of the sense (anti-guide) sequence in the position of the guide sequence during design of shRNAs shSPTBN-2 and -3.
<table>
<thead>
<tr>
<th>Gene targeted</th>
<th>RNAi effector</th>
<th>Transcript code</th>
<th>Target location</th>
</tr>
</thead>
<tbody>
<tr>
<td>SPTBN1</td>
<td>shSPTBN-A</td>
<td>NM_003128.2 (transcript variant 1, TV1)</td>
<td>5239-5257</td>
</tr>
<tr>
<td></td>
<td></td>
<td>NM_178313.2 (transcript variant 2, TV2)</td>
<td>5336-5354</td>
</tr>
<tr>
<td></td>
<td>shSPTBN-B</td>
<td>TV1</td>
<td>1664-1682</td>
</tr>
<tr>
<td></td>
<td></td>
<td>TV2</td>
<td>1761-1779</td>
</tr>
<tr>
<td></td>
<td>shSPTBN-C</td>
<td>TV1</td>
<td>4146-4164</td>
</tr>
<tr>
<td></td>
<td></td>
<td>TV2</td>
<td>4243-4261</td>
</tr>
<tr>
<td></td>
<td>shSPTBN-D</td>
<td>TV1</td>
<td>5876-5894</td>
</tr>
<tr>
<td></td>
<td></td>
<td>TV2</td>
<td>5973-5991</td>
</tr>
<tr>
<td>SP110</td>
<td>shSP110-A</td>
<td>NM_004509.3 (TV1)</td>
<td>1530-1548</td>
</tr>
<tr>
<td></td>
<td></td>
<td>NM_004510.3 (TV2)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>NM_080424.2 (TV3)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>shSP110-B</td>
<td>TV1 + TV2 + TV3</td>
<td>499-517</td>
</tr>
<tr>
<td></td>
<td></td>
<td>TV1 + TV2 + TV3</td>
<td>1714-1732</td>
</tr>
<tr>
<td></td>
<td>shSP110-C</td>
<td>TV1 + TV2 + TV3</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>TV1 + TV2 + TV3</td>
<td>1204-1222</td>
</tr>
<tr>
<td></td>
<td>shSP110-D</td>
<td>TV1 + TV2 + TV3</td>
<td></td>
</tr>
</tbody>
</table>
Table A-3: Primer sequences used in two-step PCR-amplification of shRNA expression cassettes for SPTBN1, TMED2 and KIAA1012

<table>
<thead>
<tr>
<th>Primer Name</th>
<th>Sequence Information</th>
</tr>
</thead>
<tbody>
<tr>
<td>Universal U6F</td>
<td><strong>GGATCC</strong>GATTCAAGGTCGGGCAGGAAGAGGGG<strong>CT</strong></td>
</tr>
<tr>
<td>SPTBN_1 R1</td>
<td>CTTT<strong>GGGTCAGG</strong>AGAGATGTGCTCAATGCAGGCGGTGTTTCTGCTTCCACAA</td>
</tr>
<tr>
<td>SPTBN_1 R2</td>
<td>AAAAAACCTCGCATTTGACAGCATCTTTT<strong>GGGTCAGG</strong>AGAGATG</td>
</tr>
<tr>
<td>SPTBN_2 R1</td>
<td>ATTAT<strong>GGGTCAGG</strong>TAATGCAGTGGTGAATGTCGCGGTGTTTCTGCTTCCACAA</td>
</tr>
<tr>
<td>SPTBN_2 R2</td>
<td>AAAAAACGACTTTAACTCGTGACATT<strong>GGGTCAGG</strong>TAATGCA</td>
</tr>
<tr>
<td>SPTBN_3 R1</td>
<td>TGAT<strong>GGGTCAGG</strong>CTGAGAACCCTGTTATGACTCGCTGCGGTGTTTCTGCTTCCACAA</td>
</tr>
<tr>
<td>SPTBN_3 R2</td>
<td>AAAAAATAGTCTTGGACAGAAATACAG<strong>GGGTCUGG</strong>ACATGAA</td>
</tr>
<tr>
<td>TMED_1 R1</td>
<td>CACAA<strong>GGGTCAGG</strong>TGTGTTGGTCTGAGTGGCTCGGTGTTTCTGCTTCCACAA</td>
</tr>
<tr>
<td>TMED_1 R2</td>
<td>AAAAAAGAGCCATCAACGACACACACAC<strong>GGGTCAGG</strong>TGTTGTT</td>
</tr>
<tr>
<td>TMED_2 R1</td>
<td>TATTT<strong>GGGTCAGG</strong>ACATAGCAGGACACGCTCGGCGGTGTTTCTGCTTCCACAA</td>
</tr>
<tr>
<td>TMED_2 R2</td>
<td>AAAAAAGGCTTGGGTGTTTGGTCTGACTAG<strong>GGGTCAGG</strong>ACATG</td>
</tr>
<tr>
<td>TMED_3 R1</td>
<td>GCATT<strong>GGGTCAGG</strong>ATGCTAAGCAGCAGCAGCAGCGGTGTTTCTGCTTCCACAA</td>
</tr>
<tr>
<td>TMED_3 R2</td>
<td>AAAAAACTCGGGCTATTTGGTACGTACATG<strong>GGGTCAGG</strong>TGCTAA</td>
</tr>
<tr>
<td>KIAA_1 R1</td>
<td>TAGTT<strong>GGGTCAGG</strong>AATCGACAGCTACATCAGCAGCGGTGTTTCTGCTTCCACAA</td>
</tr>
<tr>
<td>KIAA_1 R2</td>
<td>AAAAAAGCTGATAGTGATGCTAGT<strong>GGGTCAGG</strong>ACATAT</td>
</tr>
<tr>
<td>KIAA_2 R1</td>
<td>CGTAT<strong>GGGTCAGG</strong>TACGTGAGTGTAGTTAGGCGGTGTTTCTGCTTCCACAA</td>
</tr>
<tr>
<td>KIAA_2 R2</td>
<td>AAAAAACGCTAATAACCTACGCTAC<strong>GGGTCAGG</strong>TACGTGA</td>
</tr>
<tr>
<td>KIAA_3 R1</td>
<td>CTGAT<strong>GGGTCAGG</strong>TGAAGTTATAGTACAGAGCGCGGTGTTTCTGCTTCCACAA</td>
</tr>
<tr>
<td>KIAA_3 R2</td>
<td>AAAAAACGATCATAACACTAC<strong>GGGTCAGG</strong>TGAGT</td>
</tr>
</tbody>
</table>

**Key:**  
- **Bold:** BamHI;  
- **Italic:** EcoRI;  
- **Bold Italic:** Loop;  
- **Red:** Guide sequence;  
- **Blue:** U6 promoter priming;  
- **Underlined:** Termination signal
**Table A-4: Primer sequences used in two-step PCR-amplification of shRNA expression cassettes for PRDM14 and SP110**

| PRDM_1 R1 | AGAAAAAGGGTCAGGTTCTGTAAGGTCCATAGGACGCGGTGTTTCGCTCTTTTCCACAA |
| PRDM_1 R2 | AGAAAAACGTCCTATGGACACTACAGAATTGGTCTCGTTCTGTA |
| PRDM_2 R1 | TGTTGAGGTTCAGGAACATGAAGATGGATCCCGCGGTGTTCGTCTTTTCCACAA |
| PRDM_2 R2 | AGAAAAACGATCCACCTTTCTCATGGTTGAGGTTCAGGAACATGAAGATGGATCCCGCGGTGTTCGTCTTTTCCACAA |
| PRDM_3 R1 | TTATTGAGGTTCAGGAATAACTGAGGTTCCTCAGCGGTGTTTCGCTCTTTTCCACAA |
| PRDM_3 R2 | AGAAAAACGATCCACCTTTCTCATGGTTGAGGTTCAGGAATAACTGAGGTTCCTCAGCGGTGTTTCGCTCTTTTCCACAA |
| SP110_1 R1 | CCATAGGGTCAGGATGGGTTGCGAGACATCAGCGGTGTTTCGCTCTTTTCCACAA |
| SP110_1 R2 | AGAAAAACGATCCACCTTTCTCATGGTTGCGAGACATCAGCGGTGTTTCGCTCTTTTCCACAA |
| SP110_2 R1 | TCATTGGGTCAAGTATGGGTTGCGAGACATCAGCGGTGTTTCGCTCTTTTCCACAA |
| SP110_2 R2 | AGAAAAACGATCCACCTTTCTCATGGTTGCGAGACATCAGCGGTGTTTCGCTCTTTTCCACAA |
| SP110_3 R1 | CATTGGGTCAAGTATGGGTTGCGAGACATCAGCGGTGTTTCGCTCTTTTCCACAA |
| SP110_3 R2 | AGAAAAACGATCCACCTTTCTCATGGTTGCGAGACATCAGCGGTGTTTCGCTCTTTTCCACAA |

**Key:** **Bold:** BamHI; **Italic:** EcoRI; **Bold Italic:** Loop; **Red:** Guide sequence; **Blue:** U6 promoter priming; **Underlined:** Termination signal
### Table A-5: Primer sequences used in single-round PCR-amplification of siRNA-derived shRNA expression cassettes

<table>
<thead>
<tr>
<th>Primer Set</th>
<th>Primer Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>Universal U6F</td>
<td>5' - GGATCCCTAGAGCGGCGCCTAGCAAGGTCGGCGAGGAAGGAGGCC - 3'</td>
</tr>
<tr>
<td>SPTBN_A R</td>
<td>5' – AAAAAACGCTGAARGTGAGCGCATTAT\textcolor{Orange}{TGGGTCAGG}ATAATACGCTCAGTTCAGGCA CGGTGTTTCGCTCCTTCCACAA – 3'</td>
</tr>
<tr>
<td>SPTBN_B R</td>
<td>5' – AAAAAACGCGATACGAGGGCAGCTGATATT\textcolor{Orange}{TGGGTCAGG}ATAATACGCTCCTCAGGCGGC CGGTGTTTCGCTCCTTCCACAA – 3'</td>
</tr>
<tr>
<td>SPTBN_C R</td>
<td>5' – AAAAAAGGACATGTCTTAAGATGATATT\textcolor{Orange}{TGGGTCAGG}AATCCATCGTAAGACATGTCC CGGTGTTTCGCTCCTTCCACAA – 3'</td>
</tr>
<tr>
<td>SPTBN_D R</td>
<td>5' – AAAAAAGTGACAGGCCAGCAGTATATT\textcolor{Orange}{TGGGTCAGG}ATAATACGCTCCTCAGGCGGC CGGTGTTTCGCTCCTTCCACAA – 3'</td>
</tr>
<tr>
<td>SP110_A R</td>
<td>5' – AAAAAAAGGAGAARGATATCTGTTCATTT\textcolor{Orange}{TGGGTCAGG}AATGAACAGATATCTTTCTCC CGGTGTTTCGCTCCTTCCACAA – 3'</td>
</tr>
<tr>
<td>SP110_B R</td>
<td>5' – AAAAAAAGAGAARGATATCTGTTCATTT\textcolor{Orange}{TGGGTCAGG}AATGAACAGATATCTTTCTCC CGGTGTTTCGCTCCTTCCACAA – 3'</td>
</tr>
<tr>
<td>SP110_C R</td>
<td>5' – AAAAAAGTAGATGGAACTTGGTTAATT\textcolor{Orange}{TGGGTCAGG}AATGAACAGATATCTTTCTCC CGGTGTTTCGCTCCTTCCACAA – 3'</td>
</tr>
<tr>
<td>SP110_D R</td>
<td>5' – AAAAAAGAGGATGACTCAACTTGTATT\textcolor{Orange}{TGGGTCAGG}AATGAACAGATATCTTTCTCC CGGTGTTTCGCTCCTTCCACAA – 3'</td>
</tr>
</tbody>
</table>

**Key:** Orange: \textit{Bam}H1; Purple: \textit{Xho}I; Dark Red: \textit{Nof}I; Green: \textit{Nhe}I; Bold: Overlapping restriction sites; \textbf{Bold Italic}: Loop; Red: Guide sequence; Blue: U6 promoter priming; Underlined: Termination signal; R: Purine (A or G); Y: Pyrimidine (C or T)
**Table A-6: shRNA Target Oligonucleotide Sequences**

<p>| | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>SPTBN1 F1</td>
<td>5' - TCAGATATCCCTCGCATTTGACGACATCTTTGCAGACT - 3'</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SPTBN1 F2</td>
<td>5' - TTAATCTCGACATTAGCTAATACTGAGTCTGATGC - 3'</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SPTBN1 R1</td>
<td>5' - GTTAAAGTCCAGAGATGTGCTAATGGCGAGGATATC - 3'</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SPTBN1 R2</td>
<td>5' - GGCCGCATCGAAGCTCCTGATGATGACGTA - 3'</td>
<td></td>
<td></td>
</tr>
<tr>
<td>TMED2 F1</td>
<td>5' - TCAGATATCGAGCATCAAACAGAAAGCCCA - 3'</td>
<td></td>
<td></td>
</tr>
<tr>
<td>TMED2 F2</td>
<td>5' - CCGTCTCGGGCTTATCTCGGGCTATTCGTTAGCATGC - 3'</td>
<td></td>
<td></td>
</tr>
<tr>
<td>TMED2 R1</td>
<td>5' - GACCGTGGCCTTTGTTGCTGATGGCTGATGCATATC - 3'</td>
<td></td>
<td></td>
</tr>
<tr>
<td>TMED2 R2</td>
<td>5' - GGCCGCATGCTAACGAAATAGCCCGA - 3'</td>
<td></td>
<td></td>
</tr>
<tr>
<td>KIAA1012 F1</td>
<td>5' - TCAGATATCGCTGATGGTACGATCGATAGTCTAA - 3'</td>
<td></td>
<td></td>
</tr>
<tr>
<td>KIAA1012 F2</td>
<td>5' - TAATCAACCTCAGTGGCCTTGTACTATAACTTCA - 3'</td>
<td></td>
<td></td>
</tr>
<tr>
<td>KIAA1012 R1</td>
<td>5' - GATTATTAGGAATATGACATCTACATCAGC - 3'</td>
<td></td>
<td></td>
</tr>
<tr>
<td>KIAA1012 R2</td>
<td>5' - GGCCGCCTGAAGTATATGACAAAGGGCTACGTGAAGTT - 3'</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PRDM14 F1</td>
<td>5' - TCAGATATCCGTCTATGGGACACTACAGAAACGGAT - 3'</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PRDM14 F2</td>
<td>5' - CCACATTCTCTATGGTGCTGAGGACCTCCAGTTATG - 3'</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PRDM14 R1</td>
<td>5' - TGTTGATCGCTTCTGTAGTGTCATAGGACGATATC - 3'</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PRDM14 R2</td>
<td>5' - GGCCGCATAAATCAGGAGGTCTCTCAGCAACATGAAAGA - 3'</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SP110 F1</td>
<td>5' - TCAGATATCGCCTGATTTCGCAACACTACCTCCC - 3'</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SP110 F2</td>
<td>5' - TAGACAACCTCCATCTCTGCTCTTCTAGGTGACATTGC - 3'</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SP110 R1</td>
<td>5' - GTCTAGAGGTATGTTGTCAGAAATCGGCGAATCGATCGATTT - 3'</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SP110 R2</td>
<td>5' - GGCCGAATGCTACAGGAGGAGCAGAGGATCGGAGTT - 3'</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Key:** Underlined: **Xhoi** site; **Bold:** EcoRV; **Bold, italic:** NotI
<table>
<thead>
<tr>
<th>GENE</th>
<th>PRIMER</th>
<th>SEQUENCE (5’-3’)</th>
</tr>
</thead>
<tbody>
<tr>
<td>GAPDH</td>
<td>Forward (F)</td>
<td>CAT GAG AAG TAT GAC AAC AGC CT</td>
</tr>
<tr>
<td></td>
<td>Reverse (R)</td>
<td>AGT CCT TCC ACG ATA CCA AAG T</td>
</tr>
<tr>
<td>β-ACTIN</td>
<td>F</td>
<td>AGG TCA TCA CCA TTG GCA ATG AG</td>
</tr>
<tr>
<td></td>
<td>R</td>
<td>TCT TTG CGG ATG TCC ACG TCA</td>
</tr>
<tr>
<td>SPTBN1</td>
<td>F</td>
<td>CCT CTG ATC GTA AAG CCA AGA C</td>
</tr>
<tr>
<td></td>
<td>R</td>
<td>CCA CTC GTG TTT CCG ATT GAG</td>
</tr>
<tr>
<td>TMED2</td>
<td>F</td>
<td>AGC TCA CCA GAA CAA GCT AG</td>
</tr>
<tr>
<td></td>
<td>R</td>
<td>GTC GTT GAT GGC TCT GTG TAT T</td>
</tr>
<tr>
<td>KIAA1012</td>
<td>F</td>
<td>ACT TTC GGT GAT AAT ACT GCT G</td>
</tr>
<tr>
<td></td>
<td>R</td>
<td>GAA GCT GAT GAT ATG GCT GT</td>
</tr>
<tr>
<td>PRDM14</td>
<td>F</td>
<td>CAG GAC AAG GTG TGC TAC CC</td>
</tr>
<tr>
<td></td>
<td>R</td>
<td>GCT GCC GGA AAG GTT GGA A</td>
</tr>
<tr>
<td>SP110</td>
<td>F</td>
<td>TCG GAA TGA GGA TGG AAC TTG G</td>
</tr>
<tr>
<td></td>
<td>R</td>
<td>CAG AGC AAA AGT CCA CTC TTC AG</td>
</tr>
</tbody>
</table>
This represents the antisense sequence (guide) of the siRNA and was labeled at the 3’ end with the Cy3 fluorescent dye. The sense sequence is as follows: 5’-ACCUUGAAGCAUACUUCAAdTdT-3’ (d: deoxy).

<table>
<thead>
<tr>
<th>siRNA/shRNA GUIDE SEQUENCE (5’-3’)</th>
<th>TARGET Q-PCR PRIMERS (5’-3’)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>siHBV:</strong> UUGAAGUAUGCCUCAAGGUCdG</td>
<td>N/A</td>
</tr>
<tr>
<td><strong>shHBx-5:</strong> CAG AGG TGA AGC GAA GTG CAC ACG G</td>
<td>N/A</td>
</tr>
<tr>
<td><strong>shHBx-10:</strong> CCA AGG CAC AGC TTG GAG GCT TGA A</td>
<td>N/A</td>
</tr>
<tr>
<td><strong>shHTATSF-2:</strong> TTG CGG TAG AAC TAG ATG CGC</td>
<td><strong>HTATSF1 F:</strong> AGT GGG ACC TGG ACA AAA AGG</td>
</tr>
<tr>
<td><strong>shLEDGF:</strong> TCG CTT CCT CAT GCT GTC T</td>
<td><strong>HTATSF1 R:</strong> GTT CCG GGG CTT TTT CTT GTG</td>
</tr>
<tr>
<td><strong>shLTLR:</strong> TCT GAG GGA TCT CTA GTT ACT T</td>
<td><strong>LEDGF:</strong> N/A</td>
</tr>
<tr>
<td><strong>shTat:</strong> GCT CTT CGT CGC TGT CTC CGC</td>
<td><strong>Firefly Luc F:</strong> (EcoRI site in bold) ATT AGA ATT CAT GGA AGA CGC CAA AAA C</td>
</tr>
<tr>
<td></td>
<td><strong>Firefly Luc R:</strong> (HindIII site in bold) ATT AGG GCC CAC TAG <strong>TAA GCT TAC</strong> CAC ATT TGT AGA GGT TTT AC</td>
</tr>
</tbody>
</table>
APPENDIX B: Experimental procedures

1) Ethanol/EDTA precipitation

After addition of 2.5 µl 125 mM EDTA and 30 µl of absolute ethanol, the reaction was incubated for 15 minutes in the dark at room temperature. Centrifugation at 12 000 × g for 15 minutes ensued, after which the supernatant was discarded. Finally, 30 µl of 70% ethanol was added before centrifugation at 12 000 × g for 10 minutes, after which the supernatant was again discarded.

2) Bacterial cell culturing

Using aseptic technique and working near a gas flame, a scraping of frozen bacterial cells (e.g. DH5α E. coli cells) were inoculated into Luria Bertani (LB) broth (10 g/L bacto-tryptone, 5 g/L yeast extract, 5 g/L sodium chloride in distilled water and autoclaved), such that the final volume did not exceed 1/10th of the flask's volume. The flask was incubated overnight (approximately 16 hours or until growth was in the log phase) at 37°C in a shaking incubator. To control for contamination, a flask was inoculated as described above but with 1 × ampicillin: any turbidity after overnight culture indicated growth of contaminants (i.e. ampicillin-resistant bacteria or fungi).

3) Agar bacterial plates

Agar plates were made by adding 1% w/v agar to LB broth (see Bacterial cell culturing (section 2)) and autoclaving. After cooling (without setting), 100 µg/ml ampicillin was added if necessary and 30-50 ml agar poured into each sterile bacterial plate (Greiner Bio-One, GmbH, Kremsmunster, Austria) using aseptic technique near a gas flame. After setting, the plates were stored at 4°C.
4) **Chemically competent bacterial cells**

A 1 in 50 dilution of overnight bacterial culture in fresh LB broth (see Bacterial cell culturing (section 2)) was made and allowed to incubate at 37°C with moderate shaking. During incubation, the transformation buffer (100 mM CaCl₂, 10 mM PIPES HCl, 15% v/v Glycerol) was allowed to thaw on ice. After two hours, growth was arrested by placing the cultures on ice for five minutes. The cultured cells were then pelleted by centrifugation at 1200 × g for 15 minutes. Working aseptically and on ice, the supernatant was discarded and the pellet gently resuspended in 1 ml of transformation buffer. The volume was then made up to 20 ml with transformation buffer and incubated on ice for 30 minutes. Thereafter, the cells were pelleted by centrifugation at 500 – 700 × g for 10 minutes. The supernatant was carefully removed, the pellet gently resuspended in 1.5 – 2 ml transformation buffer and aliquots of 100 µl stored at -70°C.

5) **Small-scale preparations of plasmid DNA**

Bacterial cells transformed with plasmid DNA were cultured overnight (12-16 hours) in 4 ml LB (see Bacterial cell culturing (section 2)) at 37°C before DNA purification using the High Pure Plasmid Isolation kit (Roche, Mannheim, Germany). The culture was pelleted at 13 000 × g for 30 seconds and the pellet resuspended in 250 µl Suspension Buffer/RNase (50 mM Tris-HCl, 10 mM EDTA, pH 8, 0.1 mg/ml RNase A). Thereafter, 250 µl Lysis Buffer was added, mixed gently and incubated for five minutes at room temperature. Next 350 µl chilled Binding Buffer was added, gently mixed and incubated for 5 minutes on ice before centrifugation at 13 000 × g for 10 minutes. The supernatant containing the DNA was transferred to a High Pure filter tube, centrifuged at 13 000 × g for one minute and the flow-through discarded. The resin/filter binding the DNA was washed of contaminants using 700 µl Wash Buffer II and centrifuged at 13 000 × g for one minute and an additional minute after discarding the flow-through. Finally the pure DNA was eluted by applying 100 µl Elution Buffer to the centre of the filter and centrifuging after one minute at 13 000 × g for a minute. The quantity and quality of the DNA was then measured at 260 nm relative to 280 nm using the NanoDrop™ 1000 spectrophotometer (Thermo Scientific).
6) Bulk preparations of plasmid DNA

A scraping was taken from a 60% glycerol stock of bacterial cells transformed with the desired plasmid and used to inoculate 200 ml of LB supplemented with 100 µg/ml ampicillin for overnight culture. After approximately 16 hours of growth, the cells were pelleted by centrifugation at 3 040 × g for 30 minutes. During the centrifugation step, Qiagen HiSpeed Maxi Tips (one for each plasmid preparation) were equilibrated by filling with dH$_2$O, followed by 20 ml equilibration buffer (QBT) (750 mM NaCl, 50 mM MOPS (pH 7.0), 15% v/v isopropanol, 0.15% v/v Triton® X-100) and allowing to empty by gravity flow through the resin into a waste container. After centrifugation, the supernatant was discarded and the pellet gently resuspended in 10 ml chilled resuspension buffer (P1) (50 mM Tris-chloride (pH 8.0), 10 mM EDTA, 100 µg/ml RNase A). The cells were lysed by addition of 10 ml lysis buffer (P2) (200 mM NaOH, 1% w/v SDS), mixed by inversion and incubated for 5 minutes at room temperature. Thereafter, 10 ml chilled neutralization buffer (P3) (3 M potassium acetate, pH 5.5) was added, mixed by inversion and incubated for 5 minutes on ice. Centrifugation at 3 040 × g for 20 minutes followed, after which the supernatant (free of cellular debris) was carefully poured into the equilibrated HiSpeed (Qiagen) tips and allowed to filter through the resin. The low salt and pH conditions allow the DNA to bind to the anion-exchange resin. The HiSpeed tips were washed with 60 ml wash buffer (QC) (1 M NaCl, 50 mM MOPS (pH 7.0), 15% v/v isopropanol). The DNA was then eluted into a clean tube containing 10 ml isopropanol by addition of 15 ml elution buffer (QF) (1.25 M NaCl, 50 mM Tris-chloride (pH 8.5), 15% v/v isopropanol), mixed and stored at -20°C overnight. The following day, the precipitated DNA was pelleted (by centrifugation at 10 000 × g for 1 hour), washed by resuspension in 70% v/v ethanol and again centrifuged for 15 minutes. The final pellet was air-dried before resuspension in 200 µl dH$_2$O.

7) Mammalian cell culture

All mammalian cell culture (including transfections) was done in a separate tissue culture room using sterile techniques. Handling of cultures was performed inside a biological laminar flow cabinet, decontaminated with Virkon® disinfectant and 70% ethanol, before incubation in sterile incubators fitted with HEPA (high efficiency particulate air) filters, warmed to 37 °C and ventilated with 5% CO$_2$ gas.
i. **Passage of 116, HEK293 and 293T cells**

When cells reached 80% confluency (i.e. ~80% of the culture vessel's surface area is covered with an even monolayer of cells) they were passaged. This involved lifting the adherent cells from the surface of the culture vessel by gently washing with pre-warmed phosphate buffered saline, PBS (137 mM NaCl, 2.7 mM KCl, 10 mM Na$_2$HPO$_4$, 1.76 mM KH$_2$PO$_4$, pH 7.4) supplemented with 0.01% w/v EDTA. Once all the cells were in suspension a fraction of the volume (generally one quarter) was returned to the culture vessel (or transferred to a fresh one) and the appropriate culture volume made up with fresh warmed complete medium.

ii. **Passage of TZM-bl cells**

These cells were also passaged at 80% confluency. After removing culture medium, the cells were gently rinsed of residual serum with pre-warmed PBS/EDTA, being careful not to lift off any cells, and the rinse fluid discarded. After rinsing, sufficient 0.25% Trypsin / 1 mM EDTA (GIBCO, Invitrogen, CA, USA) was added to cover the monolayer of cells. The cells were returned to the 37 °C incubator for 4 minutes or until the cells easily lifted from the culture vessel surface. Thereafter, fresh medium was added to the cell suspension (at a 3 : 1 volume ratio of medium to Trypsin), before returning one quarter to the culture vessel with fresh warmed complete medium.

iii. **Passage of SupT1 cells**

Unlike the 116 and TZM-bl adherent cell lines, SupT1 cells grow in suspension and were passaged when cell density reached ~5 × 10$^5$ cells/ml. The cell suspension was pelleted by centrifugation at 500 × g for 5 minutes, after which the supernatant was removed and the pelleted cells washed by resuspending in 1 × PBS (no EDTA added). The washed cells were again pelleted by centrifugation and resuspended in PBS. Thereafter 1/10$^6$ of the cell suspension was returned to the culture flask with fresh warmed complete medium.
8) **Cell counting with a haemocytometer**

Cells were resuspended in PBS (4-5 ml) and a 50 µl sample diluted 1:2 with Trypan-blue. After gently mixing, the sample was loaded by capillary action into both chambers of a haemocytometer (Hausser Scientific, Horsham, PA) covered with a 22 × 26 mm glass coverslip (Marienfeld, Germany). After focusing on the grid under a light microscope, the number of white cells in four quadrants and the centre square were counted. The total number of cells was divided by the number of quadrants counted (i.e. 5), multiplied by the dilution factor (i.e. 2) and by 10⁴. This calculation gives the approximate number of cells per ml and can be used to calculate the dilutions needed to seed the required number of cells per well.

9) **Algorithm for TCID50 calculation**

The Spearman-Kärber formula for calculating 50% tissue culture infectious dose (TCID50) is given as:

\[ M = x_k + d[0.5 - (1/n)(r)] \]

where \( x_k \) is the dose of the highest dilution, \( r \) is the sum of negative responses, \( d \) is the spacing between dilutions and \( n \) represents the number of wells per dilution (Schmidt and Emmons 1989).

**Example of TCID50 calculation**

The following viral dilutions were used in triplicate: \( 4^{-2}, 4^{-3}, 4^{-4}, 4^{-5}, 4^{-6}, 4^{-7}, 4^{-8} \). There are 9 negative responses out of the 21 total responses (i.e. nine samples’ means are statistically significant (P < 0.05) from the mean of the No Virus control). The volume per well is 200 µl.

Then: \( x_k = 8 \) (because \( 4^{-8} \) was the highest dilution); \( r = 9 \); \( d = 1 \) and \( n = 3 \)

Thus: \( M = 8 + 1[0.5 - (1/3)(9)] = -5.5 \)

The 50% endpoint is \( 4^M \), thus \( 10^x = 4^{-5.5} \)

Thus the titre, \( x \), is 3.31 TCID50s per 200 µl.

The TCID50/ml is the above value multiplied by 5: \( 1.02 \times 10^4 / \text{ml} \)
10) Electroporation protocols (two tables)

Table B-1: Preset electroporation protocols

<table>
<thead>
<tr>
<th>PROTOCOL</th>
<th>Preset – Mammalian 11 (Jurkat)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Volume (µl)</td>
<td>400</td>
</tr>
<tr>
<td></td>
<td>200</td>
</tr>
<tr>
<td></td>
<td>100</td>
</tr>
<tr>
<td>Number of Cells</td>
<td>$1 \times 10^5$</td>
</tr>
<tr>
<td>Buffer</td>
<td>OptiMEM</td>
</tr>
<tr>
<td></td>
<td>DMEM</td>
</tr>
<tr>
<td></td>
<td>OptiMEM</td>
</tr>
<tr>
<td></td>
<td>DMEM</td>
</tr>
<tr>
<td></td>
<td>OptiMEM</td>
</tr>
<tr>
<td></td>
<td>DMEM</td>
</tr>
<tr>
<td>DNA (ng)</td>
<td>40</td>
</tr>
<tr>
<td>Cuvette size (cm)</td>
<td>0.4</td>
</tr>
<tr>
<td></td>
<td>0.2</td>
</tr>
<tr>
<td>Temperature (°C)</td>
<td>25 (room temperature)</td>
</tr>
<tr>
<td>Outgrowth</td>
<td>2 ml medium (10% FBS)</td>
</tr>
<tr>
<td></td>
<td>500 µl medium (10% FBS)</td>
</tr>
<tr>
<td>Waveform</td>
<td>Exponential</td>
</tr>
<tr>
<td>Voltage (V)</td>
<td>140</td>
</tr>
<tr>
<td>Capacitance (µF)</td>
<td>1000</td>
</tr>
<tr>
<td>Number of pulses</td>
<td>Single</td>
</tr>
</tbody>
</table>
Table B-2: Electroporation protocols from the literature

<table>
<thead>
<tr>
<th>PROTOCOL</th>
<th>1&lt;sup&gt;a&lt;/sup&gt;</th>
<th>2&lt;sup&gt;b&lt;/sup&gt;</th>
<th>3&lt;sup&gt;c&lt;/sup&gt;</th>
<th>4&lt;sup&gt;d&lt;/sup&gt;</th>
<th>5&lt;sup&gt;e&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>Volume (µl)</td>
<td>300</td>
<td>400</td>
<td>200</td>
<td>250</td>
<td></td>
</tr>
<tr>
<td>Number of Cells</td>
<td>7 x 10&lt;sup&gt;6&lt;/sup&gt;</td>
<td>5 x 10&lt;sup&gt;6&lt;/sup&gt;</td>
<td>2 x 10&lt;sup&gt;6&lt;/sup&gt;</td>
<td>1 x 10&lt;sup&gt;6&lt;/sup&gt;</td>
<td>2 x 10&lt;sup&gt;6&lt;/sup&gt;</td>
</tr>
<tr>
<td>Buffer</td>
<td>DMEM + 10% FBS</td>
<td>DMEM + 10% FBS + 1.25% DMSO</td>
<td>GenePulser</td>
<td>Phosphate-buffered sucrose</td>
<td>OptiMEM</td>
</tr>
<tr>
<td>DNA (µg)</td>
<td>30</td>
<td>2</td>
<td>5</td>
<td>25</td>
<td>5</td>
</tr>
<tr>
<td>Cuvette size (cm)</td>
<td>0.2</td>
<td>0.4</td>
<td>0.2</td>
<td>0.4</td>
<td></td>
</tr>
<tr>
<td>Temperature (°C)</td>
<td>25 (room temperature)</td>
<td></td>
<td></td>
<td>0 (ice)</td>
<td></td>
</tr>
<tr>
<td>Outgrowth</td>
<td>10 ml DMEM + 10% FBS</td>
<td>10 ml DMEM + 10% FBS + 1.25% DMSO</td>
<td>500,000 fresh cells in 27.5 ml medium</td>
<td>1 ml DMEM + 10% FBS</td>
<td></td>
</tr>
<tr>
<td>Waveform</td>
<td>Exponential</td>
<td>Square</td>
<td>Exponential</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Voltage (V)</td>
<td>250</td>
<td>100</td>
<td>250</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Capacitance (µF)</td>
<td>960</td>
<td>N/A</td>
<td>950</td>
<td>900</td>
<td></td>
</tr>
<tr>
<td>Number of pulses</td>
<td>Single</td>
<td>10</td>
<td>Single</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pulse duration (ms)</td>
<td>N/A</td>
<td>2</td>
<td>N/A</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pulse interval (ms)</td>
<td>N/A</td>
<td>100</td>
<td>N/A</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

11) Safety precautions when handling Ethidium Bromide

Avoid inhalation or ingestion, as well as contact with the eyes or skin. Use gloves, separate pipettes and glassware and dispose of waste appropriately. This dye intercalates into nucleic acid bases and may be mutagenic and thus carcinogenic.
APPENDIX C: Selection of genes

This appendix contains tables listing the genes that were excluded or selected during selection of the genes to be studied. The title of each table describes each stage of the process, where Part 1 and 2 refers to the early and late stages, respectively, of the HIV-1 lifecycle as defined by Brass et al. (2008). In addition, the tables themselves are adapted from Supplementary Table 2 published by Brass et al. (2008). The original table of genes was successively sorted according to criteria explained in Section 3-1.
Table C-1: Genes that decreased HIV-1 activity by 75% or more

<table>
<thead>
<tr>
<th>Symbol</th>
<th>SMARTpools, percent versus controls (part one, p24, or part two, RLU)</th>
</tr>
</thead>
<tbody>
<tr>
<td>KIAA1012</td>
<td>11% p24</td>
</tr>
<tr>
<td>ATP6V0A1</td>
<td>12% RLU</td>
</tr>
<tr>
<td>FLJ90680</td>
<td>12% RLU</td>
</tr>
<tr>
<td>STAC2</td>
<td>12% RLU</td>
</tr>
<tr>
<td>ZNF512B</td>
<td>13% RLU</td>
</tr>
<tr>
<td>CXorf50</td>
<td>14% RLU</td>
</tr>
<tr>
<td>DIMIT1L</td>
<td>15% p24</td>
</tr>
<tr>
<td>ASXL2</td>
<td>15% RLU</td>
</tr>
<tr>
<td>DDX55</td>
<td>15% RLU</td>
</tr>
<tr>
<td>CXC4R4</td>
<td>16% p24</td>
</tr>
<tr>
<td>CACNG1</td>
<td>16% RLU</td>
</tr>
<tr>
<td>FLJ10154</td>
<td>16% RLU</td>
</tr>
<tr>
<td>SP110</td>
<td>16% RLU</td>
</tr>
<tr>
<td>GABARA PL2 (ATG8)</td>
<td>17% RLU</td>
</tr>
<tr>
<td>STT3A</td>
<td>17% RLU</td>
</tr>
<tr>
<td>NUP107</td>
<td>18% p24</td>
</tr>
<tr>
<td>THOC2</td>
<td>18% p24</td>
</tr>
<tr>
<td>TMED2</td>
<td>18% p24</td>
</tr>
<tr>
<td>SPCS3</td>
<td>18% RLU</td>
</tr>
<tr>
<td>TUBAL3</td>
<td>18% RLU</td>
</tr>
<tr>
<td>C20orf174</td>
<td>19% p24</td>
</tr>
<tr>
<td>C9orf131</td>
<td>19% RLU</td>
</tr>
<tr>
<td>PRDM14</td>
<td>19% RLU</td>
</tr>
<tr>
<td>AGBL5</td>
<td>20% p24</td>
</tr>
<tr>
<td>PNRC1</td>
<td>20% p24</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Symbol</th>
<th>SMARTpools, percent versus controls (part one, p24, or part two, RLU)</th>
</tr>
</thead>
<tbody>
<tr>
<td>FAM76B</td>
<td>20% RLU</td>
</tr>
<tr>
<td>NDUFB7</td>
<td>20% RLU</td>
</tr>
<tr>
<td>NMT1</td>
<td>20% RLU</td>
</tr>
<tr>
<td>ADAM10</td>
<td>21% p24</td>
</tr>
<tr>
<td>RICS</td>
<td>22% p24</td>
</tr>
<tr>
<td>EDNRA</td>
<td>22% RLU</td>
</tr>
<tr>
<td>MPHOSPH6</td>
<td>22% RLU</td>
</tr>
<tr>
<td>OST48 (DDOST)</td>
<td>22% RLU</td>
</tr>
<tr>
<td>PDIA6</td>
<td>22% RLU</td>
</tr>
<tr>
<td>C3orf56</td>
<td>23% p24</td>
</tr>
<tr>
<td>TNPO3</td>
<td>23% p24</td>
</tr>
<tr>
<td>ATG16L2</td>
<td>23% RLU</td>
</tr>
<tr>
<td>JHDM1D</td>
<td>23% RLU</td>
</tr>
<tr>
<td>DKFZp686024</td>
<td>24% p24</td>
</tr>
<tr>
<td>SPTBN1</td>
<td>24% p24</td>
</tr>
<tr>
<td>WNT1</td>
<td>24% p24</td>
</tr>
<tr>
<td>LOC285550</td>
<td>24% RLU</td>
</tr>
<tr>
<td>PPP2R2A</td>
<td>24% RLU</td>
</tr>
<tr>
<td>C8orf14</td>
<td>25% p24</td>
</tr>
<tr>
<td>GCN5L2</td>
<td>25% p24</td>
</tr>
<tr>
<td>TIMM8A</td>
<td>25% RLU</td>
</tr>
<tr>
<td>UBQLN4</td>
<td>25% RLU</td>
</tr>
</tbody>
</table>
Table C-2: Genes that scored with two or more siRNAs in Part 1 (left) and Part 2 (right)

<table>
<thead>
<tr>
<th>Symbol</th>
<th>p24 siRNAs (part one)</th>
<th>Symbol</th>
<th>Beta-gal siRNAs (scored in part two only)</th>
</tr>
</thead>
<tbody>
<tr>
<td>KIAA1012</td>
<td>4</td>
<td>DDX55</td>
<td>3</td>
</tr>
<tr>
<td>DIMT1L</td>
<td>4</td>
<td>PRDM14</td>
<td>3</td>
</tr>
<tr>
<td>TNPO3</td>
<td>4</td>
<td>NMT1</td>
<td>3</td>
</tr>
<tr>
<td>TMED2</td>
<td>3</td>
<td>OST48 (DDOST)</td>
<td>3</td>
</tr>
<tr>
<td>SPTBN1</td>
<td>3</td>
<td>DKFZp686O24166</td>
<td>2</td>
</tr>
<tr>
<td>WNT1</td>
<td>3</td>
<td>ATP6V0A1</td>
<td>2</td>
</tr>
<tr>
<td>C8orf14</td>
<td>3</td>
<td>FLJ90680</td>
<td>2</td>
</tr>
<tr>
<td>CXCR4</td>
<td>2</td>
<td>STAC2</td>
<td>2</td>
</tr>
<tr>
<td>NUP107</td>
<td>2</td>
<td>CXorf50</td>
<td>2</td>
</tr>
<tr>
<td>AGBL5</td>
<td>2</td>
<td>FLJ10154</td>
<td>2</td>
</tr>
<tr>
<td>ADAM10</td>
<td>2</td>
<td>SP110</td>
<td>2</td>
</tr>
<tr>
<td>RICS</td>
<td>2</td>
<td>STT3A</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>SPCS3</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>EDNRA</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>PDIA6</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>ATG16L2</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>PPP2R2A</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>TIMM8A</td>
<td>2</td>
</tr>
<tr>
<td>Symbol</td>
<td>UniProt function</td>
<td></td>
<td></td>
</tr>
<tr>
<td>--------------</td>
<td>-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------</td>
<td></td>
<td></td>
</tr>
<tr>
<td>OST48 (DDOST)</td>
<td>Essential subunit of N-oligosaccharyl transferase enzyme which catalyzes the transfer of a high mannose oligosaccharide to an asparagine residue within an Asn-X-Ser/Thr consensus motif in nascent polypeptide chains.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ATP6V0A1</td>
<td>Required for assembly and activity of the vacuolar ATPase. Potential role in differential targeting and regulation of the enzyme for a specific organelle (By similarity).</td>
<td></td>
<td></td>
</tr>
<tr>
<td>STT3A</td>
<td>Component of the N-oligosaccharyl transferase enzyme which catalyzes the transfer of a high mannose oligosaccharide from a lipid-linked oligosaccharide donor to an asparagine residue within an Asn-X-Ser/Thr consensus motif in nascent polypeptide chains. N-glycosylation occurs cotranslational and the complex associates with the Sec61 complex at the channel-forming translocon complex that mediates protein translocation across the endoplasmic reticulum (ER). SST3A seems to be involved in complex substrate specificity (By similarity).</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SPCS3</td>
<td>Component of the microsomal signal peptidase complex which removes signal peptides from nascent proteins as they are translocated into the lumen of the endoplasmic reticulum (By similarity).</td>
<td></td>
<td></td>
</tr>
<tr>
<td>EDNRA</td>
<td>Receptor for endothelin-1. Mediates its action by association with G proteins that activate a phosphatidylinositol- calcium second messenger system. The rank order of binding affinities for ET-A is: ET1 &gt; ET2 &gt;&gt; ET3.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>TIMM8A</td>
<td>Mitochondrial intermembrane chaperone that participates in the import and insertion of some multi-pass transmembrane proteins into the mitochondrial inner membrane. Also required for the transfer of beta-barrel precursors from the TOM complex to the sorting and assembly machinery (SAM complex) of the outer membrane. Acts as a chaperone-like protein that protects the hydrophobic precursors from aggregation and guide them through the mitochondrial intermembrane space. The TIMM8-TIMM13 complex mediates the import of proteins such as TIMM23, SLC25A12/ARALAR1 and SLC25A13/ARALAR2, while the predominant TIMM9-TIMM10 70 kDa complex mediates the import of much more proteins. Probably necessary for normal neurologic development.</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table C-3: Part 1 genes excluded due to important cellular functions
<table>
<thead>
<tr>
<th>Symbol</th>
<th>p24 siRNAs (part one)</th>
<th>Beta-gal siRNAs (scored in part two only)</th>
<th>SMARTpools, percent versus controls (part one, p24, or part two, RLU)</th>
<th>Description</th>
<th>UniProt function</th>
</tr>
</thead>
<tbody>
<tr>
<td>KIAA1012</td>
<td>4</td>
<td>11% p24</td>
<td>KIAA1012</td>
<td>May play a role in vesicular transport from endoplasmic reticulum to Golgi.</td>
<td></td>
</tr>
<tr>
<td>TMED2</td>
<td>3</td>
<td>18% p24</td>
<td>transmembrane emp24 domain trafficking protein 2</td>
<td>Could have a role in the budding of coatomer-coated and other species of coated vesicles. Could bind cargo molecules to collect them into budding vesicles.</td>
<td></td>
</tr>
<tr>
<td>SPTBN1</td>
<td>3</td>
<td>24% p24</td>
<td>spectrin, beta, non-erythrocytic 1</td>
<td>Fodrin, which seems to be involved in secretion, interacts with calmodulin in a calcium-dependent manner and is thus candidate for the calcium-dependent movement of the cytoskeleton at the membrane.</td>
<td></td>
</tr>
<tr>
<td>C8orf14</td>
<td>3</td>
<td>25% p24</td>
<td>chromosome 8 open reading frame 14</td>
<td></td>
<td></td>
</tr>
<tr>
<td>AGBL5</td>
<td>2</td>
<td>20% p24</td>
<td>ATP/GTP binding protein-like 5 non TM</td>
<td></td>
<td></td>
</tr>
<tr>
<td>RICS</td>
<td>2</td>
<td>22% p24</td>
<td>Rho GTPase-activating protein</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
### Table C-5: Part 2 genes excluded due to important cellular functions

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Homologs</th>
</tr>
</thead>
<tbody>
<tr>
<td>FLJ90680</td>
<td>Stac2 (Mus musculus); Stac2_predicted (Rattus norvegicus); STAC2 (Canis lupus familiaris); STAC2 (Homo sapiens)</td>
</tr>
<tr>
<td>STAC2</td>
<td>CXorf50 (Homo sapiens); LOC653687 (Homo sapiens)</td>
</tr>
<tr>
<td>CXorf50</td>
<td>9430010003Rik (Mus musculus); FLJ10154 (Homo sapiens)</td>
</tr>
<tr>
<td>FLJ10154</td>
<td>Sp110 (Rattus norvegicus); Sp110 (Mus musculus); SP110 (Pan troglodytes); SP110 (Homo sapiens)</td>
</tr>
<tr>
<td>SP110</td>
<td>Prdm14 (Mus musculus); LOC563151 (Danio rerio); PRDM14 (Canis lupus familiaris); LOC464224 (Pan troglodytes); PRDM14 (Homo sapiens)</td>
</tr>
<tr>
<td>PRDM14</td>
<td>Atg16L2 (Mus musculus); LOC100007282 (Danio rerio); ATG16L2 (Canis lupus familiaris); ATG16L2 (Homo sapiens)</td>
</tr>
<tr>
<td>ATG16L2</td>
<td>DKFZp686O24166</td>
</tr>
</tbody>
</table>

Part 2 genes with 5 or fewer homologs
<table>
<thead>
<tr>
<th>Symbol</th>
<th>UniProt function</th>
</tr>
</thead>
<tbody>
<tr>
<td>DIMT1L</td>
<td>Specifically dimethylates two adjacent adenosines in the loop of a conserved hairpin near the 3'-end of 18S rRNA in the 40S particle (By similarity).</td>
</tr>
<tr>
<td>WNT1</td>
<td>Ligand for members of the frizzled family of seven transmembrane receptors. Probable developmental protein. May be a signaling molecule important in CNS development. Is likely to signal over only few cell diameters.</td>
</tr>
<tr>
<td>CXCR4</td>
<td>Receptor for the C-X-C chemokine CXCL12/SDF-1. Transduces a signal by increasing the intracellular calcium ions level. Involved in haematopoiesis and in cardiac ventricular septum formation. Plays also an essential role in vascularization of the gastrointestinal tract, probably by regulating vascular branching and/or remodeling processes in endothelial cells. Could be involved in cerebellar development. In the CNS, could mediate hippocampal-neuron survival. Acts as a coreceptor (CD4 being the primary receptor) for HIV-1 X4 isolates and as a primary receptor for some HIV-2 isolates. Promotes Env-mediated fusion of the virus.</td>
</tr>
<tr>
<td>NUP107</td>
<td>Essential component of nuclear pore complex. Required for the assembly of peripheral proteins into the nuclear pore complex.</td>
</tr>
<tr>
<td>ADAM10</td>
<td>Cleaves the membrane-bound precursor of TNF-alpha at '76-Ala-Val-77' to its mature soluble form. Responsible for the proteolytic release of several other cell-surface proteins, including heparin-binding epidermal growth-like factor, ephrin-A2 and for constitutive and regulated alpha-secretase cleavage of amyloid precursor protein (APP). Contributes to the normal cleavage of the cellular prion protein. Involved in the cleavage of the adhesion molecule L1 at the cell surface and in released membrane vesicles, suggesting a vesicle-based protease activity. Controls also the proteolytic processing of Notch and mediates lateral inhibition during neurogenesis (By similarity).</td>
</tr>
<tr>
<td>TNPO3</td>
<td>Seems to function in nuclear protein import as nuclear transport receptor. In vitro, mediates the nuclear import of splicing factor SR proteins SFRS1 and SFRS2, by recognizing phosphorylated RS domains.</td>
</tr>
<tr>
<td>Symbol</td>
<td>p24 siRNAs (part one)</td>
</tr>
<tr>
<td>--------</td>
<td>------------------------</td>
</tr>
<tr>
<td>FU90680</td>
<td>2</td>
</tr>
<tr>
<td>STAC2</td>
<td>2</td>
</tr>
<tr>
<td>CXorf50</td>
<td>2</td>
</tr>
<tr>
<td>FU10154</td>
<td>2</td>
</tr>
<tr>
<td>SP110</td>
<td>2</td>
</tr>
<tr>
<td>PRDM14</td>
<td>3</td>
</tr>
<tr>
<td>ATG16L2</td>
<td>2</td>
</tr>
<tr>
<td>DKFZp686O24166</td>
<td>1</td>
</tr>
</tbody>
</table>
REFERENCES


